

Perspective: Quantum mechanical methods in biochemistry and biophysics

Qiang Cui

Citation: [The Journal of Chemical Physics](#) **145**, 140901 (2016);

View online: <https://doi.org/10.1063/1.4964410>

View Table of Contents: <http://aip.scitation.org/toc/jcp/145/14>

Published by the [American Institute of Physics](#)

Articles you may be interested in

[Perspective: Kohn-Sham density functional theory descending a staircase](#)

The Journal of Chemical Physics **145**, 130901 (2016); 10.1063/1.4963168

[Perspective: How good is DFT for water?](#)

The Journal of Chemical Physics **144**, 130901 (2016); 10.1063/1.4944633

[Perspective: Machine learning potentials for atomistic simulations](#)

The Journal of Chemical Physics **145**, 170901 (2016); 10.1063/1.4966192

[Perspective: Fundamental aspects of time-dependent density functional theory](#)

The Journal of Chemical Physics **144**, 220901 (2016); 10.1063/1.4953039

[Perspective: Found in translation: Quantum chemical tools for grasping non-covalent interactions](#)

The Journal of Chemical Physics **146**, 120901 (2017); 10.1063/1.4978951

[Perspective: Defining and quantifying the role of dynamics in enzyme catalysis](#)

The Journal of Chemical Physics **144**, 180901 (2016); 10.1063/1.4947037



Perspective: Quantum mechanical methods in biochemistry and biophysics

Qiang Cui^{a)}

Department of Chemistry and Theoretical Chemistry Institute, University of Wisconsin-Madison,
1101 University Avenue, Madison, Wisconsin 53706, USA

(Received 16 July 2016; accepted 22 September 2016; published online 10 October 2016)

In this perspective article, I discuss several research topics relevant to quantum mechanical (QM) methods in biophysical and biochemical applications. Due to the immense complexity of biological problems, the key is to develop methods that are able to strike the proper balance of computational efficiency and accuracy for the problem of interest. Therefore, in addition to the development of novel *ab initio* and density functional theory based QM methods for the study of reactive events that involve complex motifs such as transition metal clusters in metalloenzymes, it is equally important to develop inexpensive QM methods and advanced classical or quantal force fields to describe different physicochemical properties of biomolecules and their behaviors in complex environments. Maintaining a solid connection of these more approximate methods with rigorous QM methods is essential to their transferability and robustness. Comparison to diverse experimental observables helps validate computational models and mechanistic hypotheses as well as driving further development of computational methodologies. Published by AIP Publishing. [<http://dx.doi.org/10.1063/1.4964410>]

I. INTRODUCTION

Since the initial development of quantum mechanical (QM) theories of chemical bonding and reactivity, researchers have applied these methods to biochemical and biophysical problems.¹ Understanding the fundamental physical basis of enzyme catalysis² and of the vision process^{3,4} using the language of quantum mechanics served as a major motivation for Karplus, Warshel, and Levitt to develop effective computational methods that were recognized by the 2013 Nobel Prize in Chemistry. This richly deserved honor underlines the tremendous power and promise of QM based methods for the analysis of biological phenomena. Indeed, there is now a rich literature that demonstrates the value of QM⁵ or hybrid QM/MM^{6–11} (or ONIOM¹²) for the analysis of biological problems that range from enzyme catalysis,¹³ bioenergy transduction,¹⁴ biomimetics, and ligand design¹⁵ to systems biology applications.¹⁶ On the other hand, there remain ample opportunities and challenges for the development of more quantitative and efficient QM based methods to allow the investigation of increasingly complex biological problems at an affordable cost.

In this perspective, I reflect on several relevant topics that have been at the forefront of recent pursuits by the theoretical chemistry community. The goal is to highlight key challenges and opportunities so as to attract additional brilliant minds to join the development of the next generation of QM based tools to tackle some of the most exciting problems in biochemistry and biophysics. The choices made here are clearly personal, and we focus on developments which are most relevant to biological applications such that the scope does not significantly overlap several recent Perspective articles (for example, Refs. 17–20), which discussed QM

methods (wavefunction and density functional theories) for more general chemical and physical applications. Although multi-scale methods such as QM/MM are clearly important to many biological applications, we will focus the discussion on the QM component since several excellent recent reviews,^{8,11} including a recent Perspective article,²¹ are available on multi-scale methodologies. Also not covered in detail are topics related to nuclear quantum effects and non-adiabatic phenomena, which are particularly important in the study of kinetic isotope effects, spectroscopy and photobiology. Exciting progress has been made in the treatment of these effects and interested readers are referred to several recent review articles.^{22–25}

II. QUANTUM MECHANICAL METHODS FOR BIOMOLECULAR FORCE FIELD DEVELOPMENTS

The function of biomolecules depends crucially on their structure and dynamics. Therefore, properly describing the energy landscape of biomolecules is a prerequisite to any meaningful mechanistic analysis or rational design work. Although the basic “alphabet” of biomolecular composition is rather limited (not to consider the rich set of post-translational modifications), subtleties of intra- and intermolecular interactions make an accurate description of the energy landscape a highly non-trivial task. Accurate QM based methods are important for either guiding the development of classical force fields or serving directly as the underlying potential energy function.

A. Development of classical force fields

Due to the importance of adequately sampling the conformational space to biomolecular applications, classical force fields will likely remain the dominant form of potential

^{a)}Author to whom correspondence should be addressed. Electronic mail: cui@chem.wisc.edu

function for the near future due to their computational efficiency. The parameterization of an accurate classical force field, however, has not been straightforward and often considered as “an art” in the literature. There are several reasons for such complexity. Foremost, the demand on the accuracy of a force field is high for a reliable description of the energy landscape; an error of 0.6 kcal/mol in the relative energy of two conformational basins, for example, will lead to an error of 3-fold in population at 300 K due to the Boltzmann distribution ($P \propto e^{-\beta \Delta E}$). In this context, it is been argued that naturally evolved sequences have been optimized through evolution such that the folding landscape is sufficiently funneled in shape and therefore capturing the key structural features does not necessarily require a highly detailed potential energy function.²⁶ Although there is extensive literature and discussion motivated by such “principle of minimal frustration,”²⁷ which led to the development of valuable coarse-grained protein models,^{28,29} a highly accurate potential function is crucial to an atomic level description of biological function. Second, due to efficiency consideration, most popular biomolecular force fields employ rather simple functional forms and explicitly describe a limited range of physical phenomena associated with intra- and inter-molecular interactions. For example, even recently reported polarizable force fields for biomolecules^{30–34} do not uniformly consider hyperconjugation, charge-penetration, anisotropy in exchange-repulsion and charge transfer^{35,36} in

an explicit manner. As a result, the components of popular biomolecular force fields cannot be directly compared to any meaningful decomposition from a high-level quantum mechanical calculation (e.g., Symmetry Adapted Perturbation Theory,³⁷ see additional discussion below), and the accuracy of force field calculations relies to a significant degree on error compensation among different components. This not only limits the degree of transferability but also requires careful “tweaking” of force field terms (for example, protein backbone potential) to match either QM calculations or experimental observables. Along this line, another complication is that biomolecules are usually in a water environment, thus any error in the water potential^{38,39} needs to be compensated by the biomolecule-water interaction, further making the physical meaning of individual force field terms less transparent.

As an example from our recent work,⁴⁰ Fig. 1 compares the structural response of a small protein (Staphyl Nuclease) to the titration of an internal residue predicted from MD simulations using two popular non-polarizable force fields (AMBERff99SB-nmr^{41,42} CHARMM36⁴³) and a recently reported polarizable field based on Drude oscillators.³¹ There are clear discrepancies among the different simulations, although many trends are consistent among the different force fields.⁴⁰ This is perhaps a somewhat extreme example; nevertheless, it serves as a reminder that further developments in biomolecular force fields are still required despite the tremendous progress made in the last few decades. For

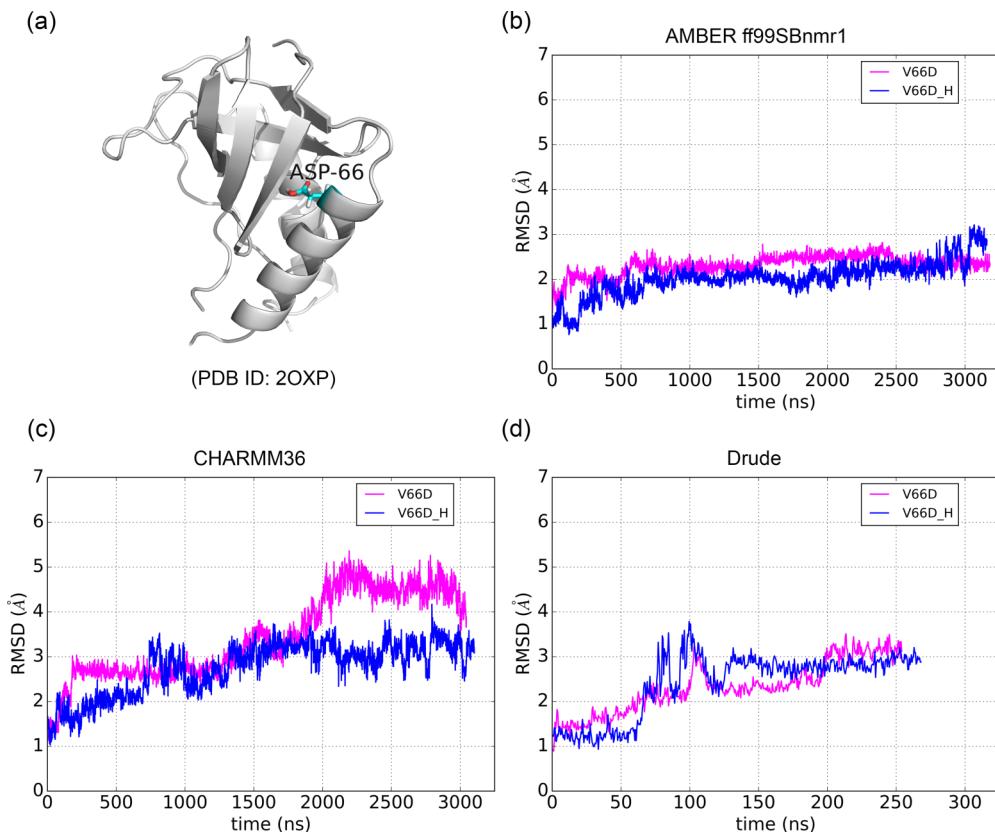


FIG. 1. The structural properties of the V66D mutant of Staphyl Nuclease with a buried titratable residue (Asp66) in different protonation states (V66D/V66D_H) depend on the force field used in molecular dynamics simulations. (a) The overall structure of Staphyl Nuclease and location of Asp66. ((b)-(d)) Root-Mean-Square-Deviation (RMSD) of the structure relative to the initial structure (based on crystal structure, PDB code: 2OXP) as a function of the simulation time using two fixed-charge force fields and a polarizable force field. Note that considerable structural transition occurs after several hundreds of nanoseconds.

example, recent discussion⁴⁴ critically revisited the issue of identifying the proper models for the representation of electronic polarization in force field development and the response model used to quantify polarization within a chosen representation. Regarding the significance of explicitly including polarization, it is interesting to note a recent study of Bradshaw and Essex,⁴⁵ who found that solvation free energies computed using the more advanced AMOEBA force field, even after further parameter optimization, do not differ statistically significantly from those computed using the AMBER-GAFF fixed-charge model.

Bearing the above discussion in mind, an enticing direction is to develop physically motivated classical force fields whose components are individually fitted against QM calculations. The idea itself is hardly new and goes back to the early days of potential energy function development.³⁵ What is required is a robust and systematic QM approach that decomposes intra- and inter-molecular interactions in a physically transparent manner. Various “energy decomposition” schemes have been proposed in the field of quantum chemistry over the years,^{37,46–53} and some of them indeed have been used to guide the development of force fields. A pioneering example in the context of biomolecular force field development is the SIBFA model of Gresh and co-workers,⁵⁴ who carefully considered the consistency between various force field terms, especially polarization, with decomposition results from Reduced Variational Space (RVS) analysis.⁴⁷ Most of the early energy decomposition schemes were limited to somewhat low level of theory, such as Hartree-Fock, Density Functional Theory (DFT), or Møller-Plesset perturbation theory. As a consequence, the decomposition results are instructive but not quantitatively accurate; directly fitting to these decomposition results is unlikely to lead to a highly accurate force field.

In the last decade or so, an increasingly popular approach for computing intermolecular interactions is the Symmetry Adapted Perturbation Theory (SAPT), thanks to more efficient numerical implementation and integration with a DFT description for the monomers (DFT-SAPT).^{55,56} The accuracy of SAPT truncated at different orders has been carefully benchmarked by several groups^{36,57,58} for biologically relevant systems. For example, Parker *et al.*⁵⁸ recommended gold, silver, and bronze standard of SAPT, which has mean absolute errors for intermolecular interactions on the order of 0.15, 0.30, and 0.49 kcal/mol, respectively, based on benchmark calculations for several popular databases. Herbert and co-workers carefully examined the problems by using SAPT to treat interactions involving ions and proposed corrections to remove artifacts associated with the single exchange approximation.⁵⁹ In addition to good accuracy, an attractive feature of SAPT is that the intermolecular interactions are given in terms of physically meaningful components, such as electrostatics, exchange-repulsion, induction, and dispersion, although higher-order contributions involve nominally “cross-terms” such as exchange-dispersion and exchange-induction.⁶⁰

Several authors have successfully used SAPT results to parameterize physically motivated force fields for condensed phase applications. For example, Schmidt and co-workers

have pioneered a set of protocols for developing force fields based on SAPT and response property calculations;⁶¹ they demonstrated the remarkable accuracy and transferability of such SAPT based force fields for several systems across a broad range of conditions.⁶² By combining SAPT based force fields and enhanced sampling techniques, Tuckerman and co-workers⁶³ have shown that small free energy difference between crystal polymorphs can be reliably predicted. Tafipolsky *et al.*⁶⁴ commented on the form of potential function in force field based on a SAPT decomposition of water-water interactions. For an excellent recent review of SAPT based force field developments, see Ref. 65.

To extend the energy decomposition based protocols to the parameterization of biomolecular force fields will be an exciting yet non-trivial endeavor. Intermolecular and intramolecular interactions along the peptide chain need to be decoupled in a physical fashion; along this line, the development of atomic SAPT partitioning analysis by Parrish and Sherrill can be particularly useful.^{66,67} Equally important are that the functional forms in the classical force fields need to be extended to describe effects such as hyperconjugation, charge-penetration, and anisotropy in exchange-repulsion in an explicit manner; charge-transfer has been proposed to be particularly important for interactions involving metal ions.^{68–71} Another interesting issue concerns the analysis of interactions beyond a pair of molecules,⁷² which can be done in the framework of Morokuma-Kitaura⁴⁶ or RVS-self-consistent-field analysis^{47,48} but not (yet) straightforward with SAPT; going beyond dimers for the analysis of non-covalent interactions is crucial because the balance of different physical terms, such as that between induction and exchange-repulsion, is different in the condensed phase from the gas phase.⁷³ Many-body dispersion interactions have also been shown to be important in extended systems.^{74,75} Finally, many biomolecules contain important co-factors such as metal ions; in some studies (e.g., biomineralization⁷⁶ and nano/bio interface⁷⁷), it is essential to calibrate the interaction between biomolecules and inorganic materials. Developing robust ways to decompose the interactions between biomolecular motifs with metal ions/complexes to guide the development of transferable and physical force fields is an important challenge.

B. Development of quantum mechanical force fields

As mentioned above, there is a concern that with limited sets of functional forms, classical force fields may not be able to capture all the subtleties of intermolecular and intramolecular interactions in complex systems like biomolecules. Therefore, an alternative is to adopt a QM description. The challenge, clearly, is to make the computation sufficiently efficient such that the gain in accuracy is not out-weighted by the loss in computational speed relative to classical force fields. As shown in Fig. 1, for example, many biologically relevant processes require simulations on the order of at least hundreds of nanoseconds, although shorter simulations with enhanced sampling techniques⁷⁸ can be applied for equilibrium properties.

There are two general directions in this regard. One approach is to develop efficient, low-scaling QM

methods based on various techniques such as divide-and-conquer,⁷⁹ fragmentation,^{80,81} and many-body expansion.⁸² For protein systems, most realistic applications have employed semi-empirical type of QM.⁸³ Integrating methodology developments with novel computational hardwares such as graphics processing units promises to further enhance the computational efficiency; it is now feasible to carry out, for example, full DFT optimizations for small solvated proteins.⁸⁴ In practice, linear-scaling QM methods are yet well established for widespread biological applications because of two considerations. First, the computational speed is still generally too low for extensive sampling, which is crucial for most biomolecular applications (see, for example, Fig. 1). Second, it is important to recognize that full QM calculations accessible to routine molecular dynamics are not necessarily highly accurate. Basis set superposition errors⁸⁵ and limitations in typical functionals in DFT⁸⁶ (e.g., missing dispersion) limit the accuracy of DFT based molecular dynamics calculations of biomolecules (including water⁸⁷), although empirical corrections have been developed.^{85,88} Therefore, in addition to continuing pushing forward the efficiency and accuracy of *ab initio* QM methods, it is also worthwhile exploring approximate QM models to balance computational speed and accuracy;⁸⁹ we discuss this further in Sec. III B.

An alternative approach is to develop a QM based force field, in the sense that a QM model is used to describe the fundamental monomer units of a biomolecule (e.g., amino acids) and more approximate (e.g., classical) models are used to describe the interactions among the monomers. Pioneering effort along this line is the Xpol model of Gao and co-workers,⁹⁰ in which the monomer units interact with each other through QM/MM type of energy terms (i.e., point charge electrostatics and Lennard-Jones for van der Waals interactions); the Xpol model has been implemented for both semi-empirical and density functional QM methods, and the semi-empirical Xpol has been used to describe a fully solvated small protein (BPTI) in molecular dynamics simulations.⁹¹ Herbert and co-workers have extended the Xpol approach by treating intermolecular interactions with pairwise-additive SAPT and an empirical dispersion model.⁹² In more recent work, York and co-workers developed a related scheme referred to as modified divide and conquer (mDC),⁹³ in which the monomer units are described with an approximate DFT model (DFTB3⁹⁴); the electrostatic component of inter-monomer interaction is described with a multipolar model, and Lennard-Jones is used to describe the van der Waals component. Compared to linear-scaling full QM models, Xpol and mDC are “hybrid” models in the sense that classical potential functions are added to the QM description; although this might seem less satisfying to purists, the practical advantage is that the accuracy of these QM based force fields can be tuned even if the QM model itself has intrinsic limitations. Further developments to ensure that interactions among monomer units are accurate and involve only physically motivated (or computed rather than fitted) parameters will help make such QM based force fields more robust and transferable. Along this line, the Effective Fragment Potential (EFP) approach⁹⁵ can be considered as a special case

of a QM based force field in which the fragment potentials involve terms that are quantum mechanical in nature, such as Fock matrix elements and overlap integrals.

Before concluding Sec. II, it is useful to remind ourselves that, for computing observables and comparing to experiments, a force field developed entirely based on first principles calculations (i.e., not fitted to condensed phase properties), in principle, should only be used in simulations that explicitly include nuclear quantum effects. In practice, whether zero-point energy conservation and other nuclear quantum effects are important depends on the properties of interest.⁹⁶

III. QUANTUM MECHANICAL METHODS FOR STUDYING REACTION MECHANISMS

The most unique and powerful application of QM methods is perhaps the study of chemical reactions. Compared to reactive force field models,^{97–99} QM methods are more general and therefore most valuable when the reaction mechanism cannot be described by a superposition of *pre-conceived* diabatic states. For many practical applications, DFT is remarkably effective; this is because although there remain many limitations in popular DFT methods,⁸⁶ most mechanistic studies require relative energetics, which tend to be less sensitive to the choice of functionals than absolute energetics (barriers/exothermicities). Nevertheless, there are several areas that will benefit from major methodological advancements and we highlight those in the following discussion.

A. Quantum mechanical methods for complex co-factors

While the potential energy surface for closed-shell species can be well described by many DFT and wavefunction based methods, open-shell species are more difficult to treat in general. Examples relevant to biological systems include transition metal ions, especially transition metal clusters, radical species, and electronically excited states for chromophores found in various light-dependent proteins. Even for relatively “simple” metal ions such as zinc and copper, popular DFT methods often have substantial errors in not only ligand-binding energies but also geometrical parameters;¹⁰⁰ even the “golden standard” single-reference wavefunction method, CCSD(T), with fairly large basis (aug-cc-pVTZ) and relativistic correction, may have errors in ligand-binding energies in the 5–10 kcal/mol range compared to experimental data when the system exhibits significant multi-reference character.^{101–103} In fact, one major challenge for identifying or refining approximate QM methods for transition metal ions is the very limited amount of accurate experimental data in the gas-phase; it is far less straightforward to unambiguously compare computed and measured data for metal ions in solution. Therefore, one important research topic is the development of highly accurate benchmark data for transition metal ions based on highly correlated computations.^{104,105}

Multi-metal clusters serve as unique catalytic co-factors in many metalloenzymes;¹⁰⁶ the most prominent examples

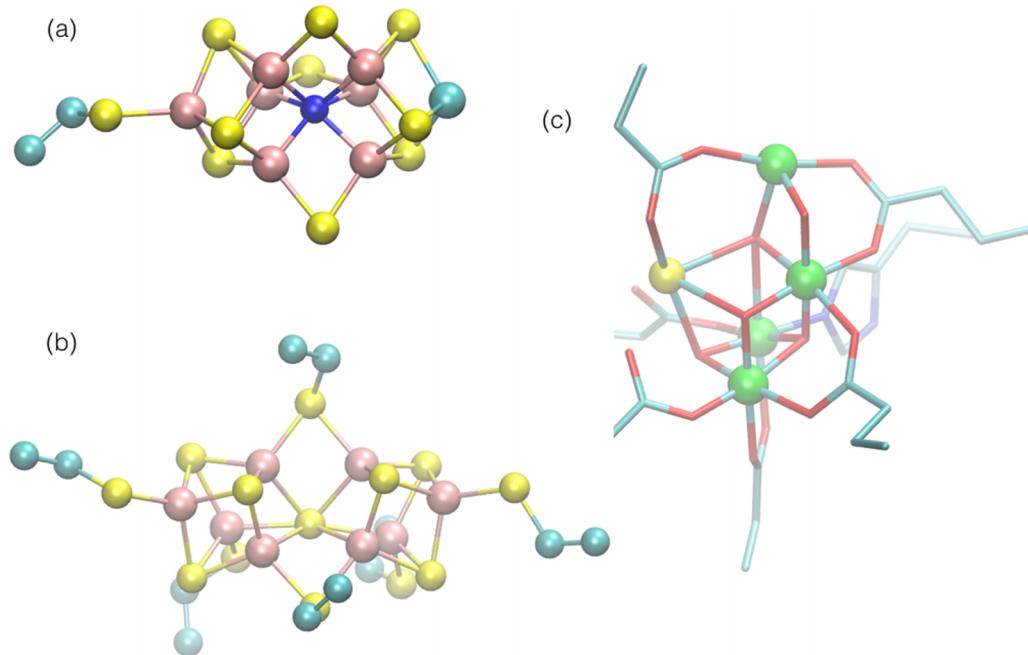


FIG. 2. Examples of metal clusters in proteins. (a) The FeMo cluster bound to a nitrogen in nitrogenase. (b) The Fe₈S₇ cluster in nitrogenase. (c) The manganese-calcium cluster in photosystem II. The structures are based on PDB files 1M1N and 3WU2, respectively.

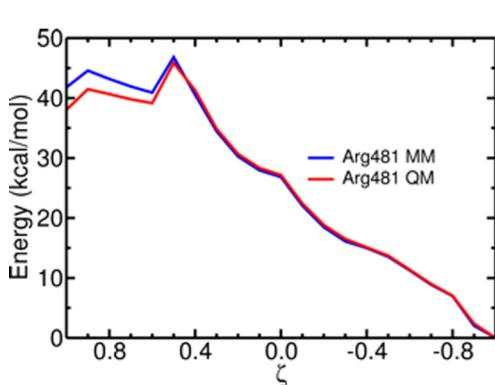
(Fig. 2) include the inorganic cubane clusters in iron-sulfur proteins, FeMo clusters in nitrogenase, and the manganese cluster in photosystem II (PSII). The complex coupling of multiple spins on the different metal centers leads to densely spaced electronic states that are difficult to describe using single-determinant (reference) wavefunction. Therefore, although symmetry-broken DFT calculations have been used effectively to probe many mechanistic issues in these systems, perhaps most notably in PSII,⁵ it is difficult to evaluate the quantitative reliability of the results. Recent development in methodologies that target strongly correlated systems such as density matrix renormalization group (DMRG),¹⁰⁷ and canonical transform theory¹⁰⁸ has made it possible to treat metal clusters with a more rigorous description of spin coupling and near-degeneracy effects, providing novel insights into the energy levels of low-lying electronic states.^{109,110} These methods, however, remain far from being “black box” and their computational cost is very high; removing these bottlenecks is essential before these methods can be routinely applied to analyze the pathway and energetics of (bio)chemical processes.

Electronically excited states are not routinely present in biological processes, except for a number of specific light-activated systems involved in bioenergetics (e.g., photoreaction center and the light-harvesting complex), sensory transduction (e.g., rhodopsin), and DNA repair (e.g., photolyase); several systems motivated by biotechnology, such as the green fluorescence protein and channel rhodopsin, have also attracted much attention. Time-dependent DFT is promising in many applications, although extensive studies on model and realistic chromophores have also established scenarios under which TD-DFT with adiabatic approximation is expected to fail, such as for the description of double excitations, charge-transfer excitations^{111–113} and conical intersection.¹¹⁴ Further

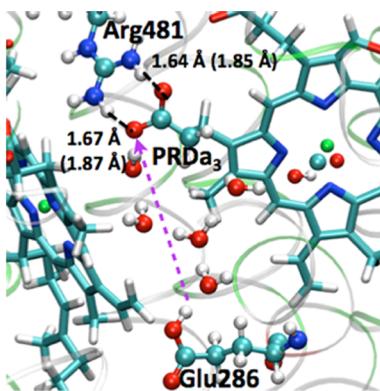
improvements in TD-DFT²⁰ or alternative methodologies (e.g., Random Phase Approximation¹¹⁵ or Multi-state DFT¹¹⁶) are essential to improve the reliability of computational analysis of these systems, especially their dynamical evolution after photon absorption.

B. The role of inexpensive quantum mechanical methods

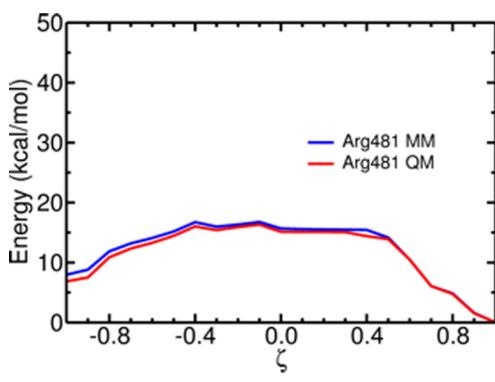
In addition to pushing forward highly accurate QM methods, further development of inexpensive QM approaches such as semi-empirical QM methods is also important to the analysis of chemical reactions in “soft matter” such as biological systems. This is because biomolecules are rich in motions that span multiple spatial and temporal scales,¹¹⁷ and it is often difficult to completely decouple them from the reactive process, especially when the reaction involves significant charge redistribution, such as long-range proton/electron transfers.¹¹⁸ Therefore, it is essential to conduct free energy sampling for mechanistic studies, while minimum energy path (MEP) results may be highly sensitive to the starting protein conformation, making it difficult to draw firm conclusions (see Fig. 3). Along this line, the minimum free energy path approach pioneered by Yang and co-workers¹¹⁹ (also see Refs. 120 and 121) in which the QM and MM fluctuations are decoupled remains arguably the most practical method for studying biochemical reactions using *ab initio* QM method. Nevertheless, it is worthwhile developing low-cost QM methods so that the validity of decoupled QM and MM fluctuations can be examined explicitly for the specific system in hand; moreover, as discussed in the Subsection III C, there are multiple ways that different levels of QM/MM calculations are integrated to obtain the most robust description of chemical reactions in biomolecules.



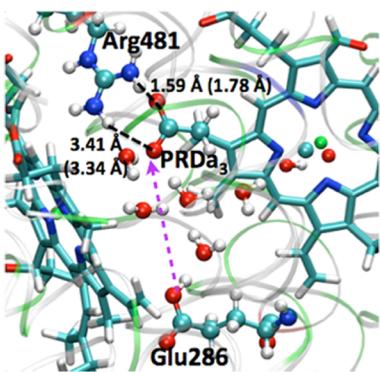
(a)



(b)



(c)



(d)

The most accessible inexpensive QM methods to biomolecular studies are semi-empirical neglect of diatomic differential overlap (NDDO) methodologies¹²³ and the self-consistent-charge density-functional-tight-binding (SCC-DFTB)¹²⁴ method. For an up-to-date review and comparison of NDDO and DFTB methodologies for the treatment of non-covalent interactions, we refer the readers to Ref. 125. Briefly, with both empirical and physically motivated corrections, NDDO and DFTB methods can be remarkably accurate for non-covalent interactions, although care must be exercised to ensure the transferability of the correction;¹²⁶ similar to the force field discussion above, it remains to investigate the reliability of these corrections to condensed phase applications.⁸⁹ For bond breaking and formation processes, several benchmark studies indicated that the latest variations of these methods, OMx¹²⁷ for NDDO and DFTB3^{94,128} for SCC-DFTB, have comparable performance in general. For several elements such as P, S,¹²⁹ and metal ions (e.g., Zn¹³⁰ and Cu¹⁰³), the DFTB3 model has advantages over published NDDO methods by including *d* orbitals and basing on DFT rather than on the uncorrelated Hartree-Fock approach. As a result, for reactions involving these elements, the DFTB3 method often exhibits substantially reduced errors, suggesting that it is worthwhile extending the DFTB3 methodology to more transition metal ions; an alternative avenue is to parameterize NDDO methods that explicitly include *d* orbitals, such as AM1/d¹³¹ and MNDO(d)¹³² methods. Nevertheless, it is important to recognize that semi-empirical methods have numerous

fundamental approximations in the basis set, integrals, and matrix elements that limit their inherent accuracy. The improvement of these methods towards quantitative accuracy for complex chemical transformations, such as phosphoryl transfers,^{129,131,133,134} and the description of water in different protonation states in different environments,⁸⁹ has not been straightforward. Along this line, one potentially productive avenue is to analyze the physical components of the underlying electronic structure using self-consistent frameworks such as NBO¹³⁵ and MOVB.¹³⁶ Similar to force field developments, by comparing physical components from semi-empirical and high-level QM calculations, one is able to identify specific aspects that require the most systematic improvements.

There are other avenues to develop inexpensive methods for studying chemical processes. Grimme and co-workers developed the “3c” set of approaches^{137,138} that correct for the basis set superposition error, missing dispersion and short-ranged basis set effects for HF/DFT calculations with small basis sets. These methods were shown to give reliable structural properties and non-covalent interactions, although there are also challenging cases;¹³⁹ their applicability to chemical reactions remains to be thoroughly tested. Moreover, these methods may still be too costly to allow extensive conformational sampling. Another framework that has received a revived interest¹⁴⁰ is the effective Hamiltonian approach.¹⁴¹ Alternatively, one may parameterize multi-state valence bond (MS-VB) models, which are very effective when the reactive species can be written in terms of a linear combination of diabatic states. The diabatic states and their

couplings can be either computed quantum mechanically in approaches such as MOVB¹³⁶ or parameterized using molecular mechanics, as in (MS)-Empirical Valence Bond (EVB);^{99,142,143} parameterization of MM based EVB model is not always straightforward, as reflected by the need of multiple sets of parameters for the same reactive group in different proton transfer events in the same enzyme.¹⁴⁴ Nevertheless, once parameterized, the (MS-)EVB model is computationally efficient and provides a framework for analyzing how the chemical environment (e.g., the level of local hydration¹⁴⁴) dictates reaction energetics and preferred pathways,^{142,144} including the decomposition of reactive free energy into enthalpic and entropic components,¹⁴⁵ which requires extensive sampling. On the other hand, the number of diabatic states included in the expansion limits the types of reaction intermediates and therefore reaction mechanisms; for example, a “proton hole” pathway¹⁴⁶ for proton transfers in carbonic anhydrase is not accessible to a MS-EVB model that includes only water and protonated water as diabatic states. Finally, one somewhat special case of inexpensive QM model is a coarse-grained model for charge transport based on fragment molecular orbitals;¹⁴⁷ it has been found effective for studying several systems including DNA¹⁴⁸ and photolyase/cryptochrome.¹⁴⁹

C. Multi-level quantum mechanical methods: In parallel and in serial

The need of combining different levels of QM methods has been recognized for many years. The most prevalent approach combines different QM methods “in parallel,” i.e., different QM methods are applied to different regions of the system in a concurrent fashion. This forms the basis of various ONIOM schemes¹⁵⁰ and embedding methodologies.^{151–154} Several recent review articles have summarized the status and application of these methods.^{8,11,12,155} Generally speaking, most technical issues related to these embedding approaches have been worked out for biological applications,⁸ although care needs to be exercised when very different QM methods are combined together¹⁵⁰ or when partitioning is made across strongly correlated electron pairs.^{156,157}

The other approach is to apply different levels of QM/MM calculations “in serial,” i.e., the goal is to use an inexpensive QM/MM method to conduct most sampling, then using more limited amount of calculations with a high-level QM/MM method to improve the energetics. Although the idea sounds intuitive, whether it works in realistic systems at a quantitative level is not always straightforward to evaluate. For example, one protocol proposed by a few authors^{158,159} employs the low-level QM/MM to compute a free energy profile using, for example, umbrella sampling; the potential energy difference between low- and high-level QM/MM results along the minimum energy path (MEP) is then taken as the correction to the low-level QM/MM method and added to the free energy profile computed at that level. This type of scheme is effective when the active site is rigid and therefore the MEPs remain close to the minimum free energy path; otherwise, the environments sampled in free energy and MEP calculations are very different, thus the energy

corrections computed along MEPs have very limited degree of transferability.

Another “serial” multi-level approach employs a thermodynamic cycle similar to that shown in Fig. 4(a). The hope is that one can gain a good estimate for the “vertical” free energy differences, $\Delta F_{s_0}^{L,H}$ and $\Delta F_{s_n}^{L,H}$, by doing most of the sampling at the low-level while using free energy perturbation (FEP) to connect the different levels (L, H) of theory;^{89,161,163,164} an alternative approach that connects different levels of theory is the Non-Boltzmann Bennett (NBB) method,¹⁶¹ which treats distribution from low-level sampling as the reweighted distribution from high-level sampling. In practice, it is clear that if the two levels of theory are very different, such single-step FEP approach is unlikely to work effectively due to the poor (or practically non-existent) overlap between the corresponding distributions. As discussed in several recent analyses, effectiveness of the FEP scheme can be improved in two practical ways. First, the low-level theory should be improved to have significant distribution overlap with the high-level theory. This can be done by generally improving the low-level QM methodology, as we and others do for DFTB and OMx, respectively; alternatively, the low-level QM can be improved “on-the-fly” for the specific system of interest using protocols such as force-matching,¹⁶⁵ para-dynamics¹⁶⁶ or machine-learning.¹⁶⁷ Second, it is beneficial to conduct a minimal level of sampling using the high-level QM/MM theory and then use a linear response approximation (LRA) to estimate the “vertical” free energy differences.¹⁶⁸ For example, as we discussed recently,¹⁶⁰ the energy gap between two levels of theory ($\Delta U^{L,H}$) along the trajectory sampled around stationary points on the free energy surface appears to closely follow Gaussian statistics for at least simple solution reactions. Under such circumstances, a LRA type model is indeed expected to work well since the cumulant expansion of the FEP expression truncates at the second-order; alternatively, LRA can be considered as a special case (infinitely fast switch) of the symmetric formula for the fast-growth method of Hummer,¹⁶⁹ who showed that the expression is exact as fourth order in the cumulant expansion.

The most formally rigorous way of integrating low- and high-level potential functions is perhaps through the multiple-time-step (MTS) framework.^{170,171} Here the low-level theory is used to propagate the dynamics at the shortest time step, while the difference between the high- and low-level potential functions is propagated at the longer time step. Assuming that the low-level calculation is negligible in cost compared to the high-level calculation, one can gain a speed up as the ratio of the long and short integration time steps; this is in the range of 5–10 for practical applications, provided that the low- and high-level energies/forces are not too different, thus their difference is indeed slowly varying. The approach has been used to integrate HF and correlated methods,¹⁷² GGA and hybrid DFT methods,¹⁷⁰ and semi-empirical QM/MM and *ab initio* QM/MM simulations.¹⁷³ More recently, the approach has been adapted in the context of path-integral MD simulations, in which DFTB was used as the low-level approach and DFT as the high-level method.¹⁷⁴

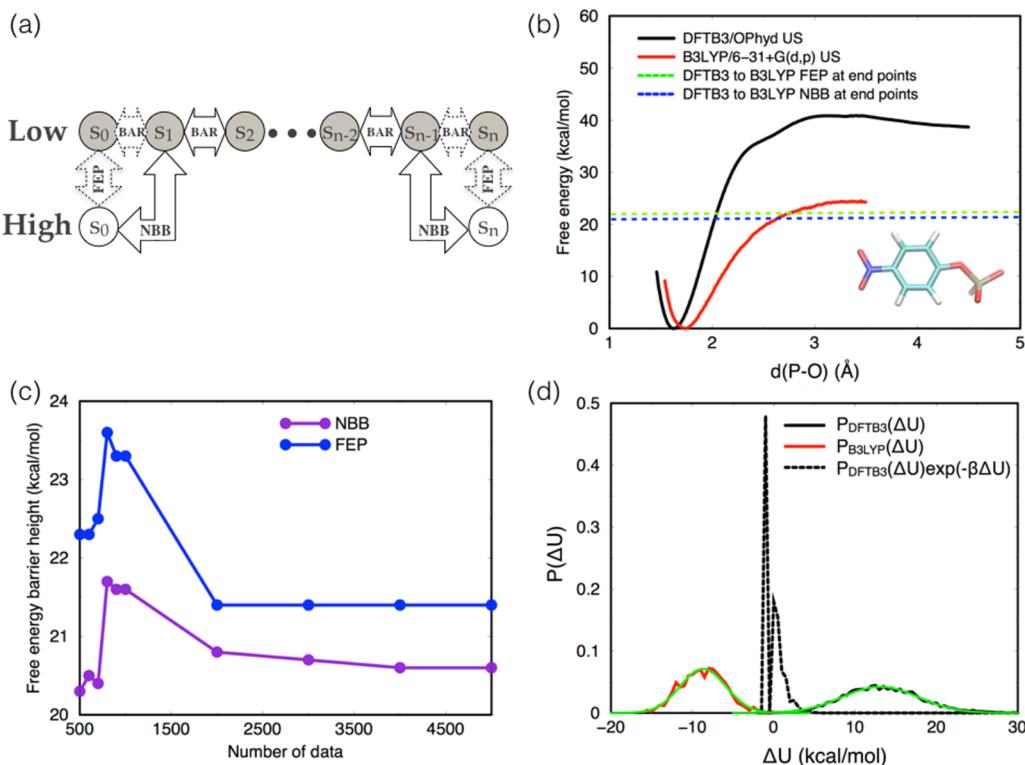


FIG. 4. Multi-level simulations that integrate low-level (L) and high-level (H) QM/MM simulations for free energies. For additional discussion and detail, see Ref. 160. (a) The qualitative scheme for combining sampling at the low-level and then using either FEP or NBB¹⁶¹ to connect to the high-level theory. s_i indicates different windows, which can be a coupling parameter in alchemical free energy simulations or an order parameter for a chemical reaction. (b) Potential of mean force (PMF) computed via standard umbrella sampling (US), and perturbation of the PMF following the NBB and FEP routes in panel (a) for p-nitrophenyl phosphate (pNPP²⁻) dissociation in solution. The low-level method is DFTB3/MM and the high-level method is B3LYP/MM. For the NBB/FEP perturbations at the end-points, only reactant and transition state windows are used, leading to corrected barrier heights indicated by dashed lines. (c) Corrected free energy barrier height via NBB/FEP using different numbers of data points per window from the low-level sampling; further improved results can be obtained by conducting sampling also at the high-level.¹⁶⁰ (d) Distribution of energy difference between the low- and high-levels of theory, ΔU^{LH} , using trajectories sampled with the two levels of theory; also plotted is the integrand of the FEP expression using the low-level trajectory (i.e., $\exp(-\beta\Delta U^{LH})P_L(\Delta U^{LH})$).¹⁶² The plots are for the transition state region in pNPP²⁻ dissociation. The ΔU^{LH} distributions are broad and data at different levels have almost vanishing overlap, thus it seems challenging for FEP or NBB to obtain reliable free energy barrier corrections, although the Gaussian nature of the ΔU^{LH} distributions (fit shown explicitly in green) suggests that a linear response model may work well.

IV. QUANTUM MECHANICAL METHODS FOR EXPERIMENTAL OBSERVABLES

In addition to intermolecular interactions and reaction energies/pathways, another important application of QM methods is to compute observables. This is not only important for validation of computational models but also essential to the molecular interpretation of experimental measurements. Such tight connection between theory/computation and experiment is by no means unique to biological studies; however, considering the complexity of biological systems, the value of theory and computation cannot be overemphasized.

For high-resolution structural information, the most popular experimental techniques for biomolecules are X-ray crystallography and NMR; major advances have been made in the area of electron microscopy, although it is mainly for near-atomic resolution structures of very large complexes. QM calculations have been used in both X-ray crystallography and NMR structural determination. For X-ray crystallography, QM method is particularly useful for co-factors that cannot be described with simple MM force fields in popular refinement programs; good examples include metal ions^{175,176} and chromophore;¹⁷⁷ we anticipate

such technique to be particularly powerful when combined with time-resolved X-ray diffraction.^{178,179} For NMR, QM calculations have been used to compute chemical shifts and spin-spin coupling constants, which help improve structural determinations.^{83,180,181} Further developments in the accuracy and efficiency of QM and QM/MM methods will greatly facilitate the refinement of difficult cases, as evidenced by the impact of using polarizable force fields in X-ray crystallography refinement.¹⁸² For metal ions and radicals, integrated spectroscopic measurement (e.g., EPR, MCD, and Mössbauer) and computation has been shown invaluable to the characterization of unusual species.^{183–185} Development of QM methods that allow a quantitative computation of spectra for complex metal centers,¹⁸⁶ such as those in Fig. 2, remains an important area of research.

Vibrational spectroscopy has been used to extract important structural information for relatively small systems, or for orientational information at (membrane) surfaces. QM based calculations have been used to parameterize frequency and coupling maps^{187–189} that aid the interpretation of vibrational spectra, especially those for non-linear spectroscopies.¹⁹⁰ For dynamical information, various time-resolved spectroscopies can be used to access different spatial and

temporal scales. Excellent review articles^{191–193} are available regarding the critical roles of theoretical/computational studies in interpreting such measurements for liquids, gas/liquid interface, and proteins. As these experimental techniques continue to evolve, we can only expect the increasing involvement of theoretical/computational studies.

To highlight the demand on computational approach, we note our recent experience in dissecting the molecular basis of continuum infrared bands associated with water molecules embedded in proteins.^{194–196} These “functional” water molecules¹⁹⁷ are essential for mediating proton transfers, and sometimes they reorganize to form proton conducting pathways only in specific kinetic states, making their observation and characterization difficult using experimental techniques such as X-ray diffraction and NMR, which target ground (resting) state of the system. Time-resolved vibrational spectroscopy is potentially powerful for characterizing the structure and dynamics of these “transient” water clusters.¹⁹⁸ For this purpose, it is essential to establish the spectroscopic signature of these water clusters, which may experience rather unusual environments that substantially perturb their spectra. A case in point concerns the water cluster between Asp96 and retinal in bacteriorhodopsin (Fig. 5(a)). The three-water cluster is formed in the N-state of the functional cycle¹⁹⁹ and sandwiched between the negatively charged

Asp96 and positively charged Schiff base; the cluster is otherwise surrounded by hydrophobic residues. Therefore, the water cluster is under major electronic polarization and was assigned to be responsible for the continuum band around 2700–2800 cm⁻¹. DFTB3/MM spectra simulations indeed reproduced this dramatic red-shift and qualitatively captured the line-shape¹⁹⁶ (Fig. 5(b)). To firmly establish this assignment, however, higher-level calculations are worthwhile because delocalization errors in DFT are known to over-estimate vibrational shifts associated with strong hydrogen-bonding interactions;²⁰⁰ along this line, adopting range-separated hybrid functionals²⁰¹ or a DFT+U approach²⁰² in the DFTB framework²⁰³ or other minimal basis QM methods²⁰⁴ is a promising direction to reduce delocalization errors. Similarly, the continuum band near 1800–2000 cm⁻¹ measured for the proton release group in bacteriorhodopsin (Figs. 5(c) and 5(d)) has also triggered debates about its assignment at the molecular level; models include a protonated water cluster^{197,205} and a pair of glutamates bridged by an “intermolecular proton bond.”^{194,195} Both models feature the excess proton being delocalized among groups in the same region, although they differ in important details regarding the major location where the proton resides. These unusual spectral features in bacteriorhodopsin highlight the need of establishing efficient and accurate QM based methods for quantitatively capturing

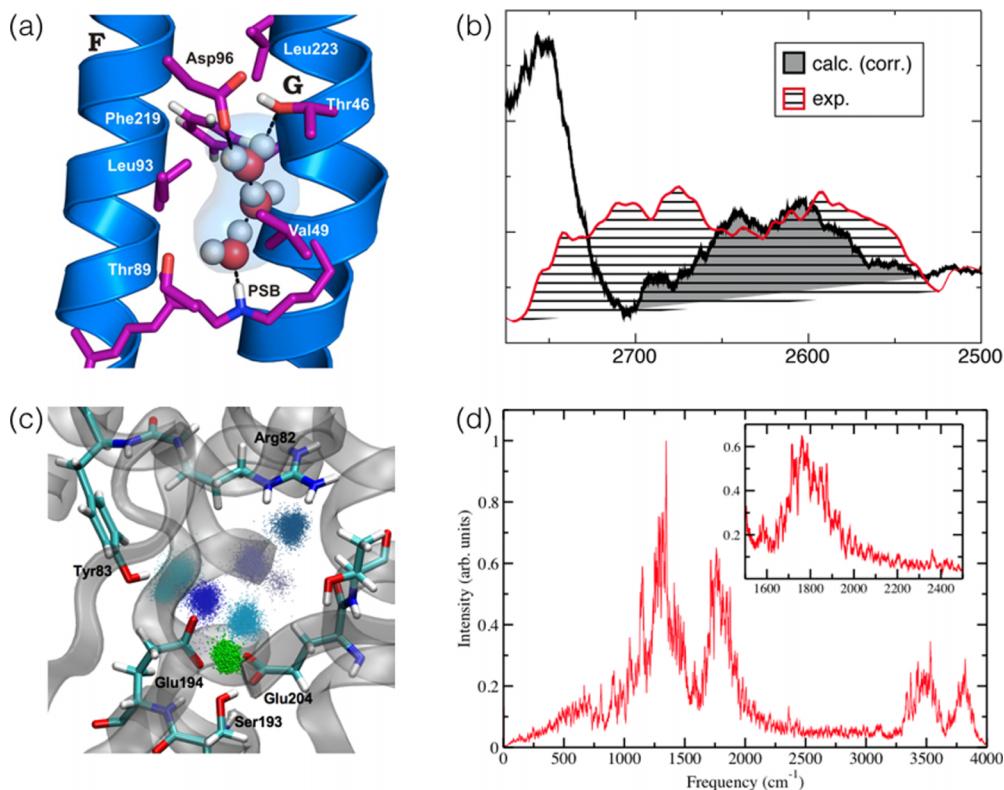


FIG. 5. Unusual infrared features associated with internal water molecules in bacteriorhodopsin represent stringent benchmarks for molecular simulations. (a) The water wire formed in the N intermediate;¹⁹⁹ the water molecules are polarized by Asp96 and the retinal, leading to significantly red-shifted infrared feature shown in panel (b), which compares computed¹⁹⁶ and experimental spectra. Panel (c) shows the distribution of water and the excess proton in the region of the proton release group from DFTB/MM simulations;¹⁹⁵ the excess proton (green dots) is largely delocalized between the pair of Glu residues, although also see Ref. 205. (d) Computed infrared spectra from DFTB/MM simulations for the proton release group; the inset shows the continuum band close to 1800–2000 cm⁻¹, which was measured experimentally.^{197,205} Panel (a) is adapted with permission from E. Freier, S. Wolf, and K. Gerwert, Proc. Natl. Acad. Sci. U. S. A. **108**, 11435–11439 (2011). Copyright 2011 PNAS. Panels (c) and (d) are adapted with permission from Goyal *et al.*, J. Am. Chem. Soc. **133**, 14981–14997 (2011). Copyright 2011 American Chemical Society.

the line shape and their variation with respect to perturbations (e.g., mutation, pH, and temperature) for reliable assignment.

In addition to spectra, many other experimental observables also help to put important constraints on a mechanistic model. Examples include kinetic isotope effects (KIEs), free energy relations, and pH dependence of enzyme reactivity.²⁰⁶ To systematically evaluate a mechanistic model that emerges from computational study, it is essential to compute multiple observables that can be either compared to available experimental data or, better yet, recorded to serve as predictions.^{207,208} Since KIEs are ratios of rate constants, they are less affected by systematic errors in the potential functions; however, KIEs depend on a proper description of variations in vibrational frequencies along the reaction coordinate, which are not straightforward to capture using empirical models such as EVB. The treatment of nuclear quantum effects is also essential, for which path-integral,^{209,210} semi-classical transition state theory,²¹¹ and extended Marcus theory²¹² have been shown to be effective. Coupling high-level QM with path-integral simulations is becoming more feasible due to recent developments of efficient contraction schemes²² and the multiple-time-step algorithm¹⁷⁴ mentioned above.

V. ENVOI

Quantum mechanical methods are important to any problem that involves complex interactions and/or chemical rearrangements. What makes biological problems particularly challenging—and fascinating—is the large number of degrees of freedom involved; these degrees of freedom typically involve configurations of atoms, although specific applications also involve sampling in the sequence space (e.g., enzyme design,²¹³ analysis of enzyme evolution,²¹⁴ and enzyme functional annotation²¹⁵). Therefore, the key to the development of effective computational methods for biological problems is to strike the proper balance between accuracy and sampling efficiency. In this perspective article, I have reflected on several areas for which further developments along this line are sorely needed. These include development of advanced force fields at both classical and quantum mechanical levels, *ab initio* and semi-empirical methods for the analysis of reactive events and computation of experimental observables. For those QM based methods to be most effective at solving important biophysical and biochemical problems, they need to be tightly coupled with advances in other fields, such as enhanced sampling methods in statistical mechanics, novel algorithms and hardwares in computer sciences, and new experimental techniques. Only with such tight integration and continuous innovation can the computer truly be the “Virgil in the world of atoms” (press release of 2013 Nobel Chemistry Prize) which can guide researchers through the *Inferno* of mechanistic uncertainty towards the *Paradiso* of truth.

ACKNOWLEDGMENTS

The work has been supported in part by grants from the National Science Foundation No. CHE-1300209 and by the National Institutes of Health No. R01GM106443. I thank

Mr. Y. Q. Zheng and Dr. P. Goyal for assistance in preparing figures in the article.

- ¹L. Pauling, *The Nature of the Chemical Bond and the Structure of Molecules and Crystals: An Introduction to Modern Structural Chemistry*, 3rd ed. (Cornell University Press, Ithaca, 1960).
- ²A. Warshel and M. Levitt, *J. Mol. Biol.* **103**, 227–249 (1976).
- ³B. Honig and M. Karplus, *Nature* **229**, 558 (1971).
- ⁴A. Warshel and M. Karplus, *J. Am. Chem. Soc.* **96**, 5677–5689 (1974).
- ⁵M. R. A. Blomberg, T. Borowski, F. Himo, R. Z. Liao, and P. E. M. Siegbahn, *Chem. Rev.* **114**, 3601–3658 (2014).
- ⁶G. Monard and K. M. Merz, *J. Acc. Chem. Res.* **32**, 904–911 (1999).
- ⁷R. A. Friesner and V. Guallar, *Annu. Rev. Phys. Chem.* **56**, 389–427 (2005).
- ⁸H. M. Senn and W. Thiel, *Angew. Chem., Int. Ed.* **48**, 1198–1229 (2009).
- ⁹J. L. Gao, S. H. Ma, D. T. Major, K. Nam, J. Z. Pu, and D. G. Truhlar, *Chem. Rev.* **106**, 3188–3209 (2006).
- ¹⁰M. P. Fruschicheva, M. J. L. Mills, P. Schopf, M. K. Singh, R. B. Prasad, and A. Warshel, *Curr. Opin. Chem. Biol.* **21**, 56–62 (2014).
- ¹¹E. Brunk and U. Rothlisberger, *Chem. Rev.* **115**, 6217–6263 (2015).
- ¹²L. W. Chung, W. M. C. Sameera, R. Ramozzi, A. J. Page, M. Hatanaka, G. P. Petrova, T. V. Harris, X. Li, Z. F. Ke, F. Y. Liu, H. B. Li, L. N. Ding, and K. Morokuma, *Chem. Rev.* **115**, 5678–5769 (2015).
- ¹³M. W. van der Kamp and A. J. Mulholland, *Biochem.* **52**, 2708–2728 (2013).
- ¹⁴S. Hayashi, H. Ueno, A. R. Shaikh, M. Umemura, M. Kamiya, Y. Ito, M. Ikeguchi, Y. Komoriya, R. Iino, and H. Noji, *J. Am. Chem. Soc.* **134**, 8447–8454 (2012).
- ¹⁵E. Bozkurt, N. Ashari, N. Browning, E. Brunk, P. Campomanes, M. A. S. Perez, and U. Rothlisberger, *Chimia* **68**, 642–647 (2014).
- ¹⁶E. Brunk, M. Neri, I. Tavernelli, V. Hatzimanikatis, and U. Rothlisberger, *Biotech. Bioeng.* **109**, 572–582 (2012).
- ¹⁷A. D. Becke, *J. Chem. Phys.* **140**, 18A301 (2014).
- ¹⁸C. D. Sherrill, *J. Chem. Phys.* **132**, 110902 (2010).
- ¹⁹K. Burke, *J. Chem. Phys.* **136**, 150901 (2012).
- ²⁰N. T. Maitra, *J. Chem. Phys.* **144**, 220901 (2016).
- ²¹A. Warshel and R. P. Bora, *J. Chem. Phys.* **144**, 180901 (2016).
- ²²S. Habershon, D. E. Manolopoulos, T. E. Markland, and T. F. Miller III, *Annu. Rev. Phys. Chem.* **64**, 387–413 (2013).
- ²³R. Kapral, *Annu. Rev. Phys. Chem.* **57**, 129–157 (2006).
- ²⁴W. H. Miller, *Annu. Rev. Phys. Chem.* **65**, 1–19 (2014).
- ²⁵S. Hammes-Schiffer and A. A. Stuchebrukhov, *Chem. Rev.* **110**, 6939–6960 (2010).
- ²⁶C. Clementi, H. Nymeyer, and J. N. Onuchic, *J. Mol. Biol.* **298**, 937–953 (2000).
- ²⁷J. N. Onuchic, Z. Luthey-Schulten, and P. G. Wolynes, *Annu. Rev. Phys. Chem.* **48**, 545–600 (1997).
- ²⁸P. C. Whitford, J. K. Noel, S. Gosavi, A. Schug, K. Y. Sanbonmatsu, and J. N. Onuchic, *Proteins: Struct., Funct., Bioinf.* **75**, 430–441 (2009).
- ²⁹C. Clementi, *Curr. Opin. Struct. Biol.* **18**, 10–15 (2008).
- ³⁰J. W. Ponder, C. J. Wu, P. Y. Ren, V. S. Pande, J. D. Chodera, M. J. Schnieders, I. Haque, D. L. Mobley, D. S. Lambrecht, R. A. DiStasio, M. Head-Gordon, G. N. I. Clark, M. E. Johnson, and T. Head-Gordon, *J. Phys. Chem. B* **114**, 2549–2564 (2010).
- ³¹P. E. M. Lopes, J. Huang, J. Shim, Y. Luo, H. Li, B. Roux, and A. D. MacKerell, Jr., *J. Chem. Theory Comput.* **9**, 5430–5449 (2013).
- ³²A. Albaugh, H. A. Boateng, R. T. Bradshaw, O. Demerdash, J. Dziedzic, Y. Z. Mo, D. T. Margul, J. Swails, Q. Zeng, D. A. Case, P. Eastman, J. W. Essex, M. Head-Gordon, V. S. Pande, J. W. Pnder, Y. H. Shao, C.-K. Skylaris, I. T. Todorov, M. E. Tuckerman, and T. Head-Gordon, *J. Phys. Chem. B* **120**, 9811–9832 (2016).
- ³³L. Lagardere, F. Lipparini, E. Polack, B. Stamm, E. Chances, M. Schnieders, P. Y. Ren, Y. Maday, and J. P. Piquemal, *J. Chem. Theory Comput.* **11**, 2589–2599 (2015).
- ³⁴R. E. Duke, O. N. Starovoytov, J. P. Piquemal, and G. A. Cisneros, *J. Chem. Theory Comput.* **10**, 1361–1365 (2014).
- ³⁵A. J. Stone, *The Theory of Intermolecular Forces* (Oxford University Press, Oxford, UK, 1997).
- ³⁶P. Hobza and J. Rezac, *Chem. Rev.* **116**, 4911–4912 (2016).
- ³⁷B. Jeziorski, R. Moszynski, and K. Szalewicz, *Chem. Rev.* **94**, 1887–1930 (1994).
- ³⁸C. Vega and J. L. F. Abascal, *Phys. Chem. Chem. Phys.* **13**, 19663–19688 (2011).
- ³⁹G. A. Cisneros, K. T. Wikfeldt, L. Ojamäe, J. B. Lu, Y. Xu, H. Torabifard, A. P. Bartok, G. Csanyi, V. Molinero, and F. Paesani, *Chem. Rev.* **116**, 7501–7528 (2016).

- ⁴⁰Y. Q. Zheng and Q. Cui, "Microscopic mechanisms that govern the titration response and pK_a values of buried residues in staphylococcal nuclease mutants," *Proteins: Struct., Funct., Bioinf.* (submitted).
- ⁴¹V. Hornak, R. Abel, A. Okur, B. Strockbine, A. Roitberg, and C. Simmerling, *Proteins: Struct., Funct., Bioinf.* **65**, 712–725 (2006).
- ⁴²D.-W. Li and R. Brüschweiler, *Angew. Chem., Int. Ed.* **49**, 6778–6780 (2010).
- ⁴³J. Huang and A. D. MacKerell, Jr., *J. Comput. Chem.* **34**, 2135–2145 (2013).
- ⁴⁴A. Li, A. Voronin, A. T. Fenley, and M. K. Gilson, *J. Phys. Chem. B* **120**, 8668–8684 (2016).
- ⁴⁵R. T. Bradshaw and J. W. Essex, *J. Chem. Theory Comput.* **12**, 3871–3883 (2016).
- ⁴⁶K. Kitaura and K. Morokuma, *Int. J. Quantum Chem.* **10**, 325–340 (1976).
- ⁴⁷P. S. Bagus, K. Hermann, and C. W. Bauschlicher, Jr., *J. Chem. Phys.* **80**, 4378–4386 (1984).
- ⁴⁸W. J. Stevens and W. H. Fink, *Chem. Phys. Lett.* **139**, 15–22 (1987).
- ⁴⁹A. E. Reed, L. A. Curtiss, and F. Weinhold, *Chem. Rev.* **88**, 899–926 (1988).
- ⁵⁰Y. R. Mo, J. L. Gao, and S. D. Peyerimhoff, *J. Chem. Phys.* **112**, 5530–5538 (2000).
- ⁵¹R. Z. Khalilullin, E. A. Cobar, R. C. Lochan, A. T. Bell, and M. Head-Gordon, *J. Phys. Chem. A* **111**, 8753–8765 (2007).
- ⁵²Q. Wu, P. W. Ayers, and Y. K. Zhang, *J. Chem. Phys.* **131**, 164112 (2009).
- ⁵³K. U. Lao and J. M. Herbert, *J. Chem. Theory Comput.* **12**, 2569–2582 (2016).
- ⁵⁴N. Gresh, G. A. Cisneros, T. A. Darden, and J.-P. Piquemal, *J. Chem. Theory Comput.* **3**, 1960–1986 (2007).
- ⁵⁵A. Hesselmann, G. Jansen, and M. Schütz, *J. Chem. Phys.* **122**, 014103 (2005).
- ⁵⁶E. G. Hohenstein, R. M. Parrish, C. D. Sherrill, J. M. Turney, and H. F. Schaefer III, *J. Chem. Phys.* **135**, 174107 (2011).
- ⁵⁷T. M. Parker, L. A. Burns, R. M. Parrish, A. G. Ryno, and C. D. Sherrill, *J. Chem. Phys.* **140**, 094106 (2014).
- ⁵⁸L. A. Burns, M. S. Marshall, and C. D. Sherrill, *J. Chem. Phys.* **141**, 234111 (2014).
- ⁵⁹K. U. Lao, R. Schaffer, G. Jansen, and J. M. Herbert, *J. Chem. Theory Comput.* **11**, 2473–2486 (2015).
- ⁶⁰H. L. Williams, K. Szalewicz, R. Moszynski, and B. Jeziorski, *J. Chem. Phys.* **103**, 4586–4599 (1995).
- ⁶¹J. R. Schmidt, K. Yu, and J. G. McDaniel, *Acc. Chem. Res.* **48**, 548–556 (2015).
- ⁶²J. G. McDaniel and J. R. Schmidt, *J. Phys. Chem. A* **117**, 2053–2066 (2013).
- ⁶³T. Q. Yu and M. E. Tuckerman, *Phys. Rev. Lett.* **107**, 015701 (2011).
- ⁶⁴M. Tafipolsky and K. Ansog, *J. Chem. Theory Comput.* **12**, 1267–1279 (2016).
- ⁶⁵J. G. McDaniel and J. R. Schmidt, *Annu. Rev. Phys. Chem.* **67**, 467–488 (2016).
- ⁶⁶R. M. Parrish and C. D. Sherrill, *J. Chem. Phys.* **141**, 044115 (2014).
- ⁶⁷R. M. Parrish, T. M. Parker, and C. D. Sherrill, *J. Chem. Theory Comput.* **10**, 4417–4431 (2014).
- ⁶⁸D. V. Sakharov and C. Lim, *J. Comput. Chem.* **30**, 191–202 (2008).
- ⁶⁹R. Chaudret, N. Gresh, O. Parisel, and J.-P. Piquemal, *J. Comput. Chem.* **32**, 2949–2957 (2011).
- ⁷⁰T. Dudev, M. Devereux, M. Meuwly, C. Lim, J. P. Piquemal, and N. Gresh, *J. Comput. Chem.* **36**, 285–302 (2015).
- ⁷¹Y. Yao, Y. Kanai, and M. L. Berkowitz, *J. Phys. Chem. Lett.* **5**, 2711–2716 (2014).
- ⁷²W. Chen and M. S. Gordon, *J. Phys. Chem.* **100**, 14316–14328 (1996).
- ⁷³T. J. Giese and D. M. York, *J. Chem. Phys.* **120**, 9903–9906 (2004).
- ⁷⁴O. A. von Lilienfeld and A. Tkatchenko, *J. Chem. Phys.* **132**, 234109 (2010).
- ⁷⁵A. Ambrosetti, N. Ferri, R. A. Distasio, Jr., and A. Tkatchenko, *Science* **351**, 1171–1176 (2016).
- ⁷⁶C. L. Freeman, J. H. Harding, D. J. Cooke, J. A. Elliott, J. S. Lardge, and D. M. Duffy, *J. Phys. Chem. C* **111**, 11943–11951 (2007).
- ⁷⁷Q. Cui, R. Hernandez, S. Mason, T. Frauenheim, J. Pedersen, and F. M. Geiger, *J. Phys. Chem. B* **120**, 7297–7306 (2016).
- ⁷⁸D. M. Zuckerman, *Annu. Rev. Biophys.* **40**, 41–62 (2011).
- ⁷⁹W. T. Yang, *Phys. Rev. Lett.* **66**, 1438–1441 (1991).
- ⁸⁰K. Kitaura, E. Ikeno, T. Asada, T. Nakano, and M. Uebayashi, *Chem. Phys. Lett.* **313**, 701–706 (1999).
- ⁸¹D. G. Fedorov and K. Kitaura, *J. Phys. Chem. A* **111**, 6904–6914 (2007).
- ⁸²R. M. Richard, K. U. Lao, and J. M. Herbert, *Acc. Chem. Res.* **47**, 2828–2836 (2014).
- ⁸³K. M. Merz, Jr., *Acc. Chem. Res.* **47**, 2804–2811 (2014).
- ⁸⁴H. J. Kulik, N. Luehr, I. S. Ufimtsev, and T. J. Martinez, *J. Phys. Chem. B* **116**, 12501–12509 (2012).
- ⁸⁵H. Kruse and S. Grimme, *J. Chem. Phys.* **136**, 154101 (2012).
- ⁸⁶A. J. Cohen, P. Mori-Sánchez, and W. T. Yang, *Chem. Rev.* **112**, 289–320 (2012).
- ⁸⁷M. J. Gillan, D. Alfe, and A. Michaelides, *J. Chem. Phys.* **144**, 130901 (2016).
- ⁸⁸S. Grimme, J. Antony, S. Ehrlich, and H. Krieg, *J. Chem. Phys.* **132**, 154104 (2010).
- ⁸⁹P. Goyal, H. J. Qian, S. Irle, X. Lu, D. Roston, T. Mori, M. Elstner, and Q. Cui, *J. Phys. Chem. B* **118**, 11007–11027 (2014).
- ⁹⁰J. L. Gao, D. G. Truhlar, Y. J. Wang, M. J. M. Mazack, P. Loffler, M. R. Provorse, and P. Rehak, *Acc. Chem. Res.* **47**, 2837–2845 (2014).
- ⁹¹W. Xie, M. Orozco, D. G. Truhlar, and J. L. Gao, *J. Chem. Theory Comput.* **5**, 459–467 (2009).
- ⁹²K. U. Lao and J. M. Herbert, *J. Chem. Phys.* **139**, 034107 (2013).
- ⁹³T. J. Giese, H. Chen, T. Dissanayake, G. M. Giambasu, H. Heldenbrand, M. Huang, E. R. Kuechler, T. S. Lee, M. T. Panteva, B. K. Radak, and D. M. York, *J. Chem. Theory Comput.* **9**, 1417–1427 (2013).
- ⁹⁴M. Gaus, Q. Cui, and M. Elstner, *J. Chem. Theory Comput.* **7**, 931–948 (2011).
- ⁹⁵M. S. Gordon, Q. A. Smith, P. Xu, and L. V. Slipchenko, *Annu. Rev. Phys. Chem.* **64**, 553–578 (2013).
- ⁹⁶M. Ceriotti, W. Fang, P. G. Kusalik, R. H. McKenzie, A. Michaelides, M. A. Morales, and T. E. Markland, *Chem. Rev.* **116**, 7529–7550 (2016).
- ⁹⁷A. Warshel, *Computer Modeling of Chemical Reactions in Enzymes and Solution* (Wiley, New York, 1991).
- ⁹⁸A. C. T. van Duin, S. Dasgupta, F. Lorant, and W. A. Goddard III, *J. Phys. Chem. A* **105**, 9396–9409 (2001).
- ⁹⁹G. A. Voth, *Acc. Chem. Res.* **39**, 143–150 (2006).
- ¹⁰⁰C. J. Cramer and D. G. Truhlar, *Phys. Chem. Chem. Phys.* **11**, 10757–10816 (2009).
- ¹⁰¹W. Jiang, N. J. DeYonker, J. J. Determan, and A. K. Wilson, *J. Phys. Chem. A* **116**, 870–885 (2012).
- ¹⁰²M. N. Weaver, K. M. Merz, Jr., D. X. Ma, H. J. Kim, and L. Gagliardi, *J. Chem. Theory Comput.* **9**, 5277–5285 (2013).
- ¹⁰³M. Gaus, H. Jin, D. Demapan, A. S. Christensen, P. Goyal, M. Elstner, and Q. Cui, *J. Chem. Theory Comput.* **11**, 4205–4219 (2015).
- ¹⁰⁴C. Hättig, W. Klopper, A. Köhn, and D. P. Tew, *Chem. Rev.* **112**, 4–74 (2012).
- ¹⁰⁵B. M. Austin, D. Y. Zubarev, and W. A. Lester, *Chem. Rev.* **112**, 263–288 (2012).
- ¹⁰⁶D. C. Rees, *Annu. Rev. Biochem.* **71**, 221–246 (2002).
- ¹⁰⁷G. K. L. Chan and S. Sharma, *Annu. Rev. Phys. Chem.* **62**, 465–481 (2011).
- ¹⁰⁸T. Yanai, Y. Kurashige, E. Neuscamman, and G. K. L. Chan, *J. Chem. Phys.* **132**, 024105 (2010).
- ¹⁰⁹S. Sharma, K. Sivalingam, F. Neese, and G. K. L. Chan, *Nat. Chem.* **6**, 927–933 (2014).
- ¹¹⁰Y. Kurashige, G. K. L. Chan, and T. Yanai, *Nat. Chem.* **5**, 660–666 (2013).
- ¹¹¹M. Wanko, M. Hoffmann, P. Strodel, A. Koslowski, W. Thiel, F. Neese, T. Frauenheim, and M. Elstner, *J. Phys. Chem. B* **109**, 3606–3615 (2005).
- ¹¹²M. A. Rohrdanz, K. M. Martins, and J. M. Herbert, *J. Chem. Phys.* **130**, 054112 (2009).
- ¹¹³O. Valsson, P. Campomanes, I. Tavernelli, U. Rothlisberger, and C. Filippi, *J. Chem. Theory Comput.* **9**, 2441–2454 (2013).
- ¹¹⁴B. G. Levine, C. Ko, J. Quenneville, and T. J. Martinez, *Mol. Phys.* **104**, 1039–1051 (2006).
- ¹¹⁵Y. Yang, H. van Aggelen, and W. T. Yang, *J. Chem. Phys.* **139**, 224105 (2013).
- ¹¹⁶Z. X. Qu and J. L. Gao, *Chem. J. Chin. Univ.* **36**, 2236–2240 (2015).
- ¹¹⁷G. Bhabha, J. T. Biel, and J. S. Fraser, *Acc. Chem. Res.* **48**, 423–430 (2015).
- ¹¹⁸D. Riccardi, P. Schaefer, Y. Yang, H. Yu, N. Ghosh, X. Prat-Resina, P. Konig, G. Li, D. Xu, H. Guo, M. Elstner, and Q. Cui, *J. Phys. Chem. B* **110**, 6458–6469 (2006).
- ¹¹⁹H. Hu, Z. Y. Lu, and W. T. Yang, *J. Chem. Theory Comput.* **3**, 390–406 (2007).
- ¹²⁰T. H. Rod and U. Ryde, *J. Chem. Theory Comput.* **1**, 1240–1251 (2005).
- ¹²¹T. Kosugi and S. Hayashi, *J. Chem. Theory Comput.* **8**, 322–334 (2012).
- ¹²²P. Goyal, S. Yang, and Q. Cui, *Chem. Sci.* **6**, 826–841 (2015).
- ¹²³W. Thiel, *Adv. Chem. Phys.* **93**, 703–757 (1996).
- ¹²⁴M. Elstner, D. Porezag, G. Jungnickel, J. Elsner, M. Haugk, T. Frauenheim, S. Suhai, and G. Seifert, *Phys. Rev. B* **58**, 7260–7268 (1998).
- ¹²⁵A. S. Christensen, T. Kubar, Q. Cui, and M. Elstner, *Chem. Rev.* **116**, 5301–5337 (2016).

- ¹²⁶A. S. Christensen, M. Elstner, and Q. Cui, *J. Chem. Phys.* **143**, 084123 (2015).
- ¹²⁷P. O. Dral, X. Wu, L. Sporkel, A. Koslowski, and W. Thiel, *J. Chem. Theory Comput.* **12**, 1097–1120 (2016).
- ¹²⁸M. Gaus, A. Goez, and M. Elstner, *J. Chem. Theory Comput.* **9**, 338–354 (2012).
- ¹²⁹M. Gaus, X. Lu, M. Elstner, and Q. Cui, *J. Chem. Theory Comput.* **10**, 1518–1537 (2014).
- ¹³⁰X. Lu, M. Gaus, M. Elstner, and Q. Cui, *J. Phys. Chem. B* **119**, 1062–1082 (2015).
- ¹³¹K. Nam, Q. Cui, J. Gao, and D. M. York, *J. Chem. Theory Comput.* **3**, 486–504 (2007).
- ¹³²W. Thiel and A. A. Voityuk, *J. Phys. Chem.* **100**, 616–626 (1996).
- ¹³³Y. Yang, H. Yu, D. York, M. Elstner, and Q. Cui, *J. Chem. Theory Comput.* **4**, 2067–2084 (2008).
- ¹³⁴V. Mlynsky, P. Banas, J. Sponer, M. W. van der Kamp, A. J. Mulholland, and M. Otyepka, *J. Chem. Theory Comput.* **10**, 1608–1622 (2014).
- ¹³⁵F. Weinhold and C. R. Landis, *Valency and Bonding* (Cambridge University Press, 2005).
- ¹³⁶Y. R. Mo and J. L. Gao, *J. Comput. Chem.* **21**, 1458–1469 (2000).
- ¹³⁷R. Sure and S. Grimme, *J. Comput. Chem.* **34**, 1672–1685 (2013).
- ¹³⁸S. Grimme, J. G. Brandenburg, C. Bannwarth, and A. Hansen, *J. Chem. Phys.* **143**, 054107 (2015).
- ¹³⁹H. Kruse, A. Mladek, K. Gkionis, A. Hansen, S. Grimme, and J. Sponer, *J. Chem. Theory Comput.* **11**, 4972–4991 (2015).
- ¹⁴⁰T. J. Watson, Jr. and G. K. L. Chan, *J. Chem. Theory Comput.* **12**, 512–522 (2016).
- ¹⁴¹H. Sun, K. F. Freed, M. E. Herman, and D. L. Yeager, *J. Chem. Phys.* **72**, 4158–4173 (1980).
- ¹⁴²S. C. L. Kamerlin and A. Warshel, *Faraday Discuss.* **145**, 71–106 (2010).
- ¹⁴³J. Danielsson and M. Meuwly, *J. Chem. Theory Comput.* **4**, 1083–1093 (2008).
- ¹⁴⁴R. B. Liang, J. M. J. Swanson, Y. X. Peng, M. Wikström, and G. A. Voth, *Proc. Natl. Acad. Sci. U. S. A.* **113**, 7420–7425 (2016).
- ¹⁴⁵J. Åqvist and S. C. L. Kamerlin, *ACS Catal.* **6**, 1737–1743 (2016).
- ¹⁴⁶D. Riccardi, P. König, X. Prat-Resina, H. Yu, M. Elstner, T. Frauenheim, and Q. Cui, *J. Am. Chem. Soc.* **128**, 16302–16311 (2006).
- ¹⁴⁷T. Kubar, P. B. Woiczikowski, G. Cuniberti, and M. Elstner, *J. Phys. Chem. B* **112**, 7937–7947 (2008).
- ¹⁴⁸T. Kubar and M. Elstner, *J. Phys. Chem. B* **112**, 8788–8798 (2008).
- ¹⁴⁹G. Ludemann, I. A. Solov'yov, T. Kubar, and M. Elstner, *J. Am. Chem. Soc.* **137**, 1147–1156 (2015).
- ¹⁵⁰T. Vreven and K. Morokuma, *J. Comput. Chem.* **21**, 1419–1432 (2000).
- ¹⁵¹J. Gao, in *Reviews in Computational Chemistry VII*, edited by K. B. Lipkowitz and D. B. Boyd (VCH, New York, 1995), p. 119.
- ¹⁵²Q. Cui, H. Guo, and M. Karplus, *J. Chem. Phys.* **117**, 5617–5631 (2002).
- ¹⁵³P. Huang and E. A. Carter, *Annu. Rev. Phys. Chem.* **59**, 261–290 (2008).
- ¹⁵⁴F. R. Manby, M. Stella, J. D. Goodpaster, and T. F. Miller III, *J. Chem. Theory Comput.* **8**, 2564–2568 (2012).
- ¹⁵⁵R. Lonsdale and A. J. Mulholland, *Curr. Top. Med. Chem.* **14**, 1339–1347 (2014).
- ¹⁵⁶G. Knizia and G. K. L. Chan, *J. Chem. Theory Comput.* **9**, 1428–1432 (2013).
- ¹⁵⁷J. D. Goodpaster, T. A. Barnes, F. R. Manby, and T. F. Miller III, *J. Chem. Phys.* **140**, 18A507 (2014).
- ¹⁵⁸S. Martí, V. Moliner, and I. Tuñón, *J. Chem. Theory Comput.* **1**, 1008–1016 (2005).
- ¹⁵⁹F. Claeyssens, J. N. Harvey, F. R. Manby, R. A. Mata, A. J. Mulholland, K. E. Ranaghan, M. Schutz, S. Thiel, W. Thiel, and H. J. Werner, *Angew. Chem., Int. Ed.* **45**, 6856–6859 (2006).
- ¹⁶⁰X. Lu, D. Fang, S. Ito, Y. Okamoto, V. Ovchinnikov, and Q. Cui, *Mol. Simul.* **42**, 1056–1078 (2016).
- ¹⁶¹G. König, P. S. Hudson, S. Boresch, and H. L. Woodcock, *J. Chem. Theory Comput.* **10**, 1406–1419 (2014).
- ¹⁶²A. Pohorille, C. Jarzynski, and C. Chipot, *J. Phys. Chem. B* **114**, 10235–10253 (2010).
- ¹⁶³M. Retegan, M. Martins-Costa, and M. F. Ruiz-Lopez, *J. Chem. Phys.* **133**, 064103 (2010).
- ¹⁶⁴I. Polyak, T. Benighaus, E. Boulanger, and W. Thiel, *J. Chem. Phys.* **139**, 064105 (2013).
- ¹⁶⁵Y. Zhou and J. Z. Pu, *J. Chem. Theory Comput.* **10**, 3038–3054 (2014).
- ¹⁶⁶N. V. Plotnikov, S. C. L. Kamerlin, and A. Warshel, *J. Phys. Chem. B* **115**, 7950–7962 (2011).
- ¹⁶⁷P. O. Dral, O. A. von Lilienfeld, and W. Thiel, *J. Chem. Theory Comput.* **11**, 2120–2125 (2015).
- ¹⁶⁸S. C. L. Kamerlin, M. Haranczyk, and A. Warshel, *J. Phys. Chem. B* **113**, 1253–1272 (2009).
- ¹⁶⁹G. Hummer, *J. Chem. Phys.* **114**, 7330–7337 (2001).
- ¹⁷⁰M. Guidon, F. Schiffmann, J. Hutter, and J. VandeVondele, *J. Chem. Phys.* **128**, 214104 (2008).
- ¹⁷¹N. Luehr, T. E. Markland, and T. J. Martinez, *J. Chem. Phys.* **140**, 084116 (2014).
- ¹⁷²R. P. Steele, *J. Chem. Phys.* **139**, 011102 (2013).
- ¹⁷³K. Nam, *J. Chem. Theory Comput.* **10**, 4175–4183 (2014).
- ¹⁷⁴O. Marsalek and T. E. Markland, *J. Chem. Phys.* **144**, 054112 (2016).
- ¹⁷⁵U. Ryde and K. Nilsson, *J. Am. Chem. Soc.* **125**, 14232–14233 (2003).
- ¹⁷⁶U. Ryde, C. Greco, and L. De Gioia, *J. Am. Chem. Soc.* **132**, 4512–4513 (2010).
- ¹⁷⁷Y. W. Hsiao, E. Sanchez-Garcia, M. Doerr, and W. Thiel, *J. Phys. Chem. B* **114**, 15413–15423 (2010).
- ¹⁷⁸J. Tenboer *et al.*, *Science* **346**, 1242–1246 (2014).
- ¹⁷⁹T. R. M. Barends *et al.*, *Science* **350**, 445–450 (2015).
- ¹⁸⁰X. He, B. Wang, and K. M. Merz, Jr., *J. Phys. Chem. B* **113**, 10380–10388 (2009).
- ¹⁸¹J. A. Vila, J. M. Aramini, P. Rossi, A. Kuzin, M. Su, J. Seetharaman, R. Xiao, L. Tong, G. T. Montelione, and H. A. Scheraga, *Proc. Natl. Acad. Sci. U. S. A.* **105**, 14389–14394 (2008).
- ¹⁸²M. J. Schnieders, T. D. Fenn, V. S. Pande, and A. T. Brunger, *Acta Crystallogr., Sect. D* **65**, 952–965 (2009).
- ¹⁸³K. M. Lancaster, M. Roemelt, P. Ettenhuber, Y. L. Hu, M. W. Ribbe, F. Neese, U. Bergmann, and S. DeBeer, *Science* **334**, 974–977 (2011).
- ¹⁸⁴R. Bjornsson, F. Neese, R. R. Schrock, O. Einsle, and S. DeBeer, *J. Biol. Inorg. Chem.* **20**, 447–460 (2015).
- ¹⁸⁵N. Cox, D. A. Pantazis, F. Neese, and W. Lubitz, *Acc. Chem. Res.* **46**, 1588–1596 (2013).
- ¹⁸⁶F. Neese, *Coord. Chem. Rev.* **253**, 526–563 (2009).
- ¹⁸⁷L. Wang, C. T. Middleton, M. T. Zanni, and J. L. Skinner, *J. Phys. Chem. B* **115**, 3713–3724 (2011).
- ¹⁸⁸T. L. Jansen and J. Knoester, *J. Chem. Phys.* **124**, 044502 (2006).
- ¹⁸⁹W. Zhuang, D. Abramavicius, T. Hayashi, and S. Mukamel, *J. Phys. Chem. B* **110**, 3362–3374 (2006).
- ¹⁹⁰S. Mukamel, *Annu. Rev. Phys. Chem.* **51**, 691–729 (2000).
- ¹⁹¹H. J. Bakker and J. L. Skinner, *Chem. Rev.* **110**, 1498–1517 (2010).
- ¹⁹²G. R. Medders and F. Paesani, *J. Chem. Theory Comput.* **11**, 1145–1154 (2015).
- ¹⁹³C. Kotting and K. Gerwert, *Biol. Chem.* **396**, 131–144 (2015).
- ¹⁹⁴P. Phatak, N. Ghosh, H. Yu, Q. Cui, and M. Elstner, *Proc. Natl. Acad. Sci. U. S. A.* **105**, 19672–19677 (2008).
- ¹⁹⁵P. Goyal, N. Ghosh, P. Phatak, M. Clemens, M. Gaus, M. Elstner, and Q. Cui, *J. Am. Chem. Soc.* **133**, 14981–14997 (2011).
- ¹⁹⁶S. Wolf, E. Freier, Q. Cui, and K. Gerwert, *J. Chem. Phys.* **141**, 22D524 (2014).
- ¹⁹⁷F. Garczarek and K. Gerwert, *Nature* **439**, 109–112 (2006).
- ¹⁹⁸O. P. Ernst, D. T. Lodowski, M. Elstner, P. Hegemann, L. S. Brown, and H. Kandori, *Chem. Rev.* **114**, 126–163 (2014).
- ¹⁹⁹E. Freier, S. Wolf, and K. Gerwert, *Proc. Natl. Acad. Sci. U. S. A.* **108**, 11435–11439 (2011).
- ²⁰⁰Q. Cui and M. Elstner, *Phys. Chem. Chem. Phys.* **16**, 14368–14377 (2014).
- ²⁰¹O. A. Vydrov, J. Heyd, A. V. Krukau, and G. E. Scuseria, *J. Chem. Phys.* **125**, 074106 (2006).
- ²⁰²H. J. Kulik, *J. Chem. Phys.* **142**, 240901 (2016).
- ²⁰³V. Lutsker, B. Aradi, and T. A. Niehaus, *J. Chem. Phys.* **143**, 184107 (2015).
- ²⁰⁴H. J. Kulik, N. Seelam, B. D. Mar, and T. J. Martinez, *J. Phys. Chem. A* **120**, 5939–5949 (2016).
- ²⁰⁵S. Wolf, E. Freier, and K. Gerwert, *Biophys. J.* **107**, 174–184 (2014).
- ²⁰⁶A. Fersht, *Structure and Mechanism in Protein Science: A Guide to Enzyme Catalysis and Protein Folding* (W. H. Freeman and Company, 1999).
- ²⁰⁷D. Roston, D. Demapan, and Q. Cui, *J. Am. Chem. Soc.* **138**, 7386–7394 (2016).
- ²⁰⁸D. Roston and Q. Cui, *J. Am. Chem. Soc.* **138**, 11946–11957 (2016).
- ²⁰⁹D. T. Major and J. L. Gao, *J. Chem. Theory Comput.* **3**, 949–960 (2007).
- ²¹⁰L. Wang, M. Ceriotti, and T. E. Markland, *J. Chem. Phys.* **141**, 104502 (2014).
- ²¹¹J. L. Gao and D. G. Truhlar, *Annu. Rev. Phys. Chem.* **53**, 467–505 (2002).
- ²¹²S. Hammes-Schiffer, *Acc. Chem. Res.* **34**, 273–281 (2001).
- ²¹³R. Blomberg, H. Kries, D. M. Pinkas, P. R. E. Mittl, M. G. Grutter, H. K. Privett, S. L. Mayo, and D. Hilvert, *Nature* **503**, 418 (2013).
- ²¹⁴J. V. Rodrigues, S. Bershtein, A. Li, E. R. Lozovsky, D. L. Hartl, and E. I. Shakhnovich, *Proc. Natl. Acad. Sci. U. S. A.* **113**, E1470–E1478 (2016).
- ²¹⁵M. P. Jacobson, C. Kalyanaraman, S. W. Zhao, and B. X. Tian, *Trends Biochem. Sci.* **39**, 363–371 (2014).