ANNUAL REVIEWS

Annual Review of Biophysics

Hybrid Quantum Mechanical/Molecular Mechanical Methods For Studying Energy Transduction in Biomolecular Machines

T. Kubař,[1] M. Elstner,[1,2] and Q. Cui[3,4,5]

[1]Institute of Physical Chemistry, Karlsruhe Institute of Technology, Karlsruhe, Germany; email: tomas.kubar@kit.edu

[2]Institute of Biological Interfaces (IBG-2), Karlsruhe Institute of Technology, Karlsruhe, Germany; email: marcus.elstner@kit.edu

[3]Department of Chemistry, Boston University, Boston, Massachusetts, USA; email: qiangcui@bu.edu

[4]Department of Physics, Boston University, Boston, Massachusetts, USA

[5]Department of Biomedical Engineering, Boston University, Boston, Massachusetts, USA

Annu. Rev. Biophys. 2023. 52:525–51

First published as a Review in Advance on February 15, 2023

The *Annual Review of Biophysics* is online at biophys.annualreviews.org

https://doi.org/10.1146/annurev-biophys-111622-091140

Keywords

quantum mechanical/molecular mechanical, QM/MM, energy transduction, biomolecular machines, proton transfers, electron transfers, nucleotide hydrolysis

Abstract

Hybrid quantum mechanical/molecular mechanical (QM/MM) methods have become indispensable tools for the study of biomolecules. In this article, we briefly review the basic methodological details of QM/MM approaches and discuss their applications to various energy transduction problems in biomolecular machines, such as long-range proton transports, fast electron transfers, and mechanochemical coupling. We highlight the particular importance for these applications of balancing computational efficiency and accuracy. Using several recent examples, we illustrate the value and limitations of QM/MM methodologies for both ground and excited states, as well as strategies for calibrating them in specific applications. We conclude with brief comments on several areas that can benefit from further efforts to make QM/MM analyses more quantitative and applicable to increasingly complex biological problems.

Contents

1. INTRODUCTION

One of the hallmarks of living systems is the efficient transduction of energy among different forms via biomolecular machines (100, 119). Remarkable examples include the conversion of light energy (or O_2 activation) into proton motive force by bacteriorhodopsin (bR) (cytochrome c oxidase), utilization of the proton gradient to synthesize ATP by the F_0F_1-ATPase and to transport small molecules across the membrane by various transporters, and the driving of the motion of biomolecular motors by the chemical energy of ATP hydrolysis. The definition of efficiency for these biomolecular machines is somewhat subtle (12), but generally, the values are substantially higher compared to those of artificial machines. Therefore, uncovering the physical and chemical principles that govern the working mechanism of naturally evolved systems is not only of fundamental significance, but also valuable for providing guidance to the development of novel molecular machines at the nanoscale (2, 130).

Since the ultimate driving force for most biomolecular machines involves chemical reactions, such as ATP hydrolysis; coupled proton–electron transfers; or electronic excitation, theoretical and computational analysis of bioenergy transduction necessarily involves treating chemical reactions and dissecting their coupling to other processes, such as protein conformational transitions or translocation of small molecules or ions. Therefore, compared to the problem of enzyme catalysis, which also involves chemical reactions in biomolecules, the topic of bioenergy transduction is arguably even more challenging, as it requires methodologies that are able to bridge broader length and timescales. While striking the balance between computational accuracy and efficiency is relevant to most biophysical problems, its central role in the meaningful analysis of bioenergy transduction has driven the development and integration of multiscale computational methodologies, including semiempirical quantum mechanical/molecular mechanical (QM/MM) methods (15); ab initio QM/MM approaches (11); and, more recently, machine learning (ML) techniques.

To illustrate the value and limitations of these methodologies, as well as strategies for properly calibrating them in specific systems, we discuss several recent applications in the area of bioenergy transduction. Due to limited space, we mainly focus on studies in our own labs, although

complementary efforts by others are also mentioned. The discussions focus on the steps that are most tightly coupled with the chemical reactions; discussion of methodologies that tackle other steps, such as large-scale conformational transitions, can be found in recent reviews by others (134, 142), including general conceptual issues related to the efficiency of energy transduction (2, 97, 125). We end with a brief outlook section that comments on future developments that will enable the analysis of increasingly complex bioenergy transduction problems.

2. THEORY AND METHODS

The basic computational framework to study biological processes driven by chemical reactions is the hybrid QM/MM approach (30, 148) pioneered by Warshel, Levitt, and Karplus, who were awarded the Nobel Prize in Chemistry for their efforts. At the conceptual level, the approach is intuitive: The reactive region of the system is treated at a quantum mechanical (QM) level, while the environment is described with a more simplified molecular mechanical (MM) model. Through the contributions of many researchers, the QM/MM approach has become an indispensable tool for the analysis of condensed phase problems (16, 34, 55), especially biomolecules (122). Recent developments and applications, especially in chemical and enzymatic systems, have been summarized in excellent review articles (11, 83); some of the remaining challenges have also been discussed (20). In this section, we briefly touch upon several generally relevant methodological issues.

2.1. Quantum Mechanical/Molecular Mechanical Methods for Ground Electronic States

In the most standard scheme applied to biomolecules (**Figure 1**), the QM/MM energy is given in the additive form,

$$E^{\mathrm{Tot}} = \langle \Psi | \hat{H}^{\mathrm{QM}} + \hat{H}^{\mathrm{QM/MM}}_{\mathrm{elec}} | \Psi \rangle + E^{\mathrm{QM/MM}}_{\mathrm{vdW}} + E^{\mathrm{QM/MM}}_{\mathrm{bonded}} + E^{\mathrm{MM}}, \qquad 1.$$

which indicates that the QM and MM atoms interact through electrostatic, van der Waals, and bonded terms; in most implementations, only the QM/MM electrostatic interaction (see below) is

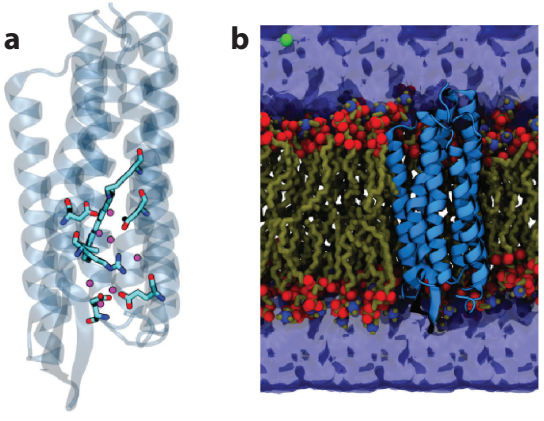


Figure 1

Illustration of a typical QM/MM set-up using the analysis of the O→bR transition in bacteriorhodopsin as an example. (*a*) The QM region, which includes key amino acids and water molecules, is highlighted in the licorice form, while the rest of the MM protein environment is shown in a transparent schematic. (*b*) The entire protein is then embedded in a solvated lipid bilayer; the lipids, water molecules, and salt ions are also treated at the MM level. Figure adapted with permission from Reference 87; copyright 2021 National Academy of Sciences. Abbreviations: MM, molecular mechanical; QM, quantum mechanical.

included in the self-consistent determination of the QM region wavefunction, Ψ (or the electron density), while the van der Waals and bonded terms are treated at the classical force field level. When the QM/MM partitioning is conducted across a covalent bond (e.g., between Cα and Cβ atoms of an amino acid), link atoms (22, 111), frozen orbitals (35), or pseudo potentials (161) are required to properly treat the boundary QM atoms; care needs to be exercised to avoid artificial polarization of the QM atoms (1, 63), and it is generally inadvisable to partition across highly polar covalent bonds.

The proper QM level depends on the problem of interest. While ab initio or density functional theory (DFT) is generally more reliable than semiempirical QM methods, they are computationally expensive. Even with modern hardware, ab initio– or DFT-based QM/MM simulations are typically limited to tens to hundreds of picoseconds of sampling (141), which are usually too short for a reliable computation of equilibrium (e.g., free energy) or dynamical properties. Therefore, carefully calibrated semiempirical QM methods remain an attractive choice, especially for bioenergetics problems; in recent years, density functional tight binding (DFTB) models (4, 15, 39) have emerged as promising choices in many applications.

An issue that has attracted much debate in recent years is the appropriate size of the QM region (23, 24, 31, 58, 69, 95); with the development of efficient algorithms and implementations on modern hardware, it has become possible to conduct QM/MM simulations with hundreds or thousands of QM atoms (141), or even with the entire system treated at the DFT (149) or DFTB (101) level. The computational cost associated with such large QM regions, however, limits the amount of conformational sampling. Therefore, depending on the properties of interest, it is important to choose the QM region to best balance computational cost and sampling efficiency. For example, for the analysis of reaction mechanism and free energy profiles, adequate sampling is essential. Sampling is expected to be particularly important for systems involved in bioenergy transduction, as the driving chemical reactions often involve the migration of charged species (e.g., protons, electrons, metal ions) over a long distance; thus, there are significant protein and solvent responses that need to be captured with extensive sampling. For properties that are particularly sensitive to electronic structure, such as nuclear magnetic resonance chemical shifts and hyperfine coupling constants, it is possible that large QM regions are required (24, 31, 121).

In the simplest model of QM/MM interactions, the electrostatic Hamiltonian ($\hat{H}_{\text{elec}}^{\text{QM/MM}}$) involves Coulombic interactions between the QM atoms (nuclei and electrons) and fixed MM partial charges. At short QM/MM distances, however, the point charge models for the MM atoms may lead to overpolarization of the QM atoms. Thus, a more physical model involves blurring the MM charges as spherical Gaussians (22, 54).

It is increasingly being realized that it is valuable to explicitly include electronic polarization at the MM level. In recent years, substantial progress has been made in the systematic parameterization and validation of polarizable MM force field models for biomolecules (56), including, for example, the CHARMM-Drude model (75), AMEOBA (59), and SIBFA (99). Improvements in computational efficiency and implementation on graphics processing units have made it possible to conduct extensive molecular dynamics (MD) simulations for realistic biomolecules, including the ribosome (70). The inclusion of electronic polarization is particularly relevant to a reliable treatment of bioenergy transduction, since many systems involve buried charges or ion pairs; as shown in a recent analysis (25), electronic polarization is critical to the description of the stabilization, conformational dynamics, and hydration levels of these buried charges or dipoles. In this context, while charge scaling has been advocated as an empirical approach for approximating the effect of electronic polarization (28, 76), such a phenomenological model may not be able to capture both energetic and dynamic properties correctly (18); thus, including the MM polarization explicitly in QM/MM simulations is preferable (9, 81).

2.2. Quantum Mechanical/Molecular Mechanical Methods for Excited Electronic States

Many bioenergy transduction processes are initiated by the absorption of photons; thus, the description of electronically excited states is required. Often, ground state structures are QM/MM geometry optimized, and the excited state can then be treated with a broad set of QM methods that include, for example, time-dependent DFT (TD-DFT) or post-Hartree-Fock methods. The typical errors for TD-DFT are in the range of 0.2–0.4 eV for singly excited states, strongly depending on the functional applied. Pure GGAs (generalized gradient approximations) tend to have smaller errors, and the size of the error increases with the amount of exact exchange. For example, an error of 0.26 eV was reported for B3LYP for a set of medium-sized molecules of biological relevance (105); since this behavior is often systematic, a simple shift in energies can help facilitate a comparison with experimental data. However, the limitations of common functionals and the adiabatic linear-response approximations are well known, especially for charge-transfer and doubly excited states, which are not uncommon in biological systems (27, 143). For some problems, range-separated functionals [long-range corrected DFT (LC-DFT)] provide major improvements, although they do not completely resolve the issues (8, 57).

The absorption spectrum is better computed by sampling of the ground state potential energy surface and subsequent vertical excitation energy calculations, in which nuclear quantum effects can be neglected. In this case, methods with significantly reduced computational cost are needed, since convergence of the spectrum requires the calculation of hundreds or even thousands of snapshots. Semiempirical methods are therefore suitable for these types of applications; however, the computational challenges also put strong constraints on the choice of methods, as has been discussed in detail elsewhere for retinal proteins (146). The Hartree-Fock-based OM2 and OM3 methods within a multireference configuration interaction (MRCI) implementation proved to be quite reliable, while the DFT-based linear response TD-DFTB method was not able to treat these types of systems. They show the same failures as the GGA-based TD-DFT methods. LC-DFT functionals have been shown to ameliorate the problems, and LC-TD-DFTB has been shown to be quite accurate for the calculation of absorption and fluorescence properties (65, 126). However, the typical TD-DFT problems are not yet completely solved, as discussed in detail elsewhere for the examples of retinal proteins and chlorophylls (8).

Since the ground and excited states usually feature rather different electronic distributions, polarization of the MM environment was found to make a nonnegligible contribution to the excitation energy (9, 81, 144). It is worth noting that in some cases, force fields have intrinsic limitations, as is the case for strong hydrogen-bonding interactions with the chromophore (see examples below), necessitating large QM regions in QM/MM simulations.

2.3. Sampling Transitions and Dynamics Relevant to Chemistry in Biomolecules

For most problems in bioenergy transduction, the key quantity of interest is the underlying free energy profile, which can be used to compute rate constants with well-accepted theoretical models such as the transition state theory (102). The rate constants can then be fed into kinetic network models, such as dominant kinetic pathway(s) and the overall timescale and efficiency of energy transduction, to gain further mechanistic insights (132). In this context, it is important to recall that rate constants depend exponentially on the free energy barrier; thus, small errors in the computed barrier can lead to large errors in kinetics, and effective computational techniques for adjusting computed rate constants based on experimental observables are highly valuable in this regard (110, 132). In principle, the methodologies commonly used for classical simulations, such as metadynamics (139) and finite-temperature string (29) methods, could readily be adopted for

the computation of free energy profiles. The key difference lies in the choice of collective variables (CVs), which might take unique forms for chemical processes. For example, charge centroids based on either structural (62), charge (43; D. Maag, J. Böser, B. Hourahine, M. Elstner, H.A. Witek and T. Kubař, unpublished manuscript), or electron (77) density have been used to construct CVs for studying long-range proton transfers.

Since QM/MM computations are generally more expensive than classical simulations, it is worthwhile to consider multilevel strategies. For example, semiempirical QM methods are usually more reliable for structural properties than for energetics. Thus, one effective protocol (159) is to first sample the reaction pathway with, for example, DFTB/MM string simulations; with configurations from the optimized minimal free energy path as the initial guess, string calculations with a higher-level QM/MM potential are expected to converge more rapidly for more accurate energetics. Alternatively, configurations from semiempirical QM/MM simulations can be used in conjunction with ML techniques to improve energetics (discussed further below).

A particularly interesting topic concerns the direct simulation of sequential electron transfers (ETs) in proteins, for which two of us developed the fragment orbital tight binding density functional theory (FO-DFTB) (66, 67). In this quantum-chemical calculation, the quantum region is divided into several fragments, with the general aim of reducing the computational cost, as well as allowing for an easy and efficient parallelization. The fragments are defined in such a way that any conjugated π-electron systems are kept intact, so the electronic structure of the isolated fragments can be computed straightforwardly. For biological ETs, the fragments may be side chains of aromatic amino acids, peptide bond moieties, or nucleic acid bases. Notably, while the electronic structure of a single fragment is calculated, the other fragments, as well as the entire condensed phase environment, are treated as point charges, so the fragment is polarized properly.

Next, the Hamiltonian (and overlap) matrices are set up on the basis of the computed fragment orbitals, and they are further used to propagate the electron density corresponding to the excess charge (electron or hole) by way of integrating a time-dependent Schrödinger equation; one of the many available nonadiabatic propagation schemes can be used, and the most common choices are the trajectory surface hopping and the mean-field Ehrenfest methods. The diagonal elements (site energies representing the ionization potentials for hole transfer or the electron affinities for excess ET), as well as the off-diagonal elements (electronic couplings), were benchmarked, and high accuracy was observed, although scaling is required in some cases (42, 68). Effectively, the entire molecular system, with the exception of an excess electron or hole, is treated using a classical MM force field. Note that this also includes the ET-active fragments at any moment that they do not carry any excess charge. For the purpose of analysis, the charge occupation (squared expansion coefficient) of each fragment is recorded along the trajectories being performed. This value ranges between 0 (neutral fragment) and 1 (fragment completely occupied by the hole or excess electron being transferred). After averaging over the ensemble of simulations, the time-dependent occupations can be fitted using a kinetic model (86) to obtain rate constants of the forward and backward ET events.

2.4. Computation of Experimental Observables

It is critical to compute experimental observables so that computational results can be validated and predictions tested. With QM/MM methods, a broad range of observables can in principle be computed (19), ranging from various spectra through free energy relations to kinetic isotope effects (36). As mentioned above, computed properties such as rate constants, equilibrium constants, and spectral densities can be fed into kinetic network models [either classical or quantum (102)] for probing macroscopic features relevant to the energy transduction process. The convergence of

these observables depends on the system and sometimes requires extensive sampling. For example, kinetic isotope effects can be computed with path-integral techniques (91), which may require the equivalent of multiple nanoseconds of sampling (24, 115), which is possible only with semiempirical methods. The computation of infrared (IR) spectra (45, 106), especially multidimensional spectra (3), also requires extensive sampling.

For semiempirical methods, the trade-off between computational efficiency and accuracy is a matter of concern; however, well-calibrated semiempirical methods have been shown to be very useful in interpreting and predicting spectral properties of complex systems. For example, the mid-IR C=O double bond vibration is highly sensitive to the hydrogen bonding environment, showing red-shifts up to 50 cm^{-1} in strongly hydrogen-bonded systems relative to the gas phase, and a reparametrized semiempirical method has been shown to capture these shifts reliably (150) and to be able to resolve details of water-bridged hydrogen bonds between amino acids, which were experimentally difficult to determine. This helped to differentiate between various intermediate structures in the bR photocycle at different temperatures (153, 154).

Another frequently computed property for energy transfer (e.g., light-harvesting) systems is the spectral density, which describes the coupling of the nuclear degrees of freedom to the electronic structure. In principle, it can be computed from a time series of excitation energies along a classical MD trajectory. However, since classical force fields often do not capture the geometry of the chromophore reliably, spectral density computed using MD trajectories sampled with classical force fields often suffers from the geometry mismatch problem (89). Accurate spectral densities compared to experimental data were obtained using a semiempirical method specifically reparameterized for an accurate description of vibrational frequencies (89, 90).

2.5. Integration with Machine Learning

In recent years, ML has become an increasingly powerful tool in computational analysis of molecular systems (103). In the context of QM/MM simulations and bioenergy transduction, ML techniques are uniquely valuable in several areas (**Figure 2**). First, ML techniques can be used to enhance the efficiency of free energy sampling, especially when multiple collective variables (160) are potentially important to describing the coupling between the reaction and environmental degrees of freedom (6). Second, ML methods can be used to improve the accuracy of semiempirical QM/MM simulations via iterative Δ-learning, in which configurations from low-level QM/MM (e.g., DFTB/MM) trajectories are used to learn the differences (Δ) between low-level and high-level QM/MM energies and forces; the differences are then used to either directly estimate correction of the free energy surface or resample the potential energy surface for improved free energies (44, 114, 123). These multilevel free energy simulation methods have to date been applied only to relatively simple solution reactions (5), as learning the Δ reliably for realistic systems with a large QM region is, in fact, not straightforward and requires further developments. Finally, ML can also be used to learn other properties, such as electronic coupling elements (64) or vertical excitation energies (14) as functions of molecular configurations; training such ML models can substantially reduce the cost of ET and spectra calculations.

3. APPLICATION AND CASE STUDIES

In this section, we discuss several recent applications to illustrate the value and limitations of QM/MM based simulations, including calibration of the methods in realistic applications for meaningful mechanistic insights. Due to the sampling requirement of these applications, we focus largely on DFTB/MM studies, although complementary efforts using other QM/MM methods are also mentioned for comparison.

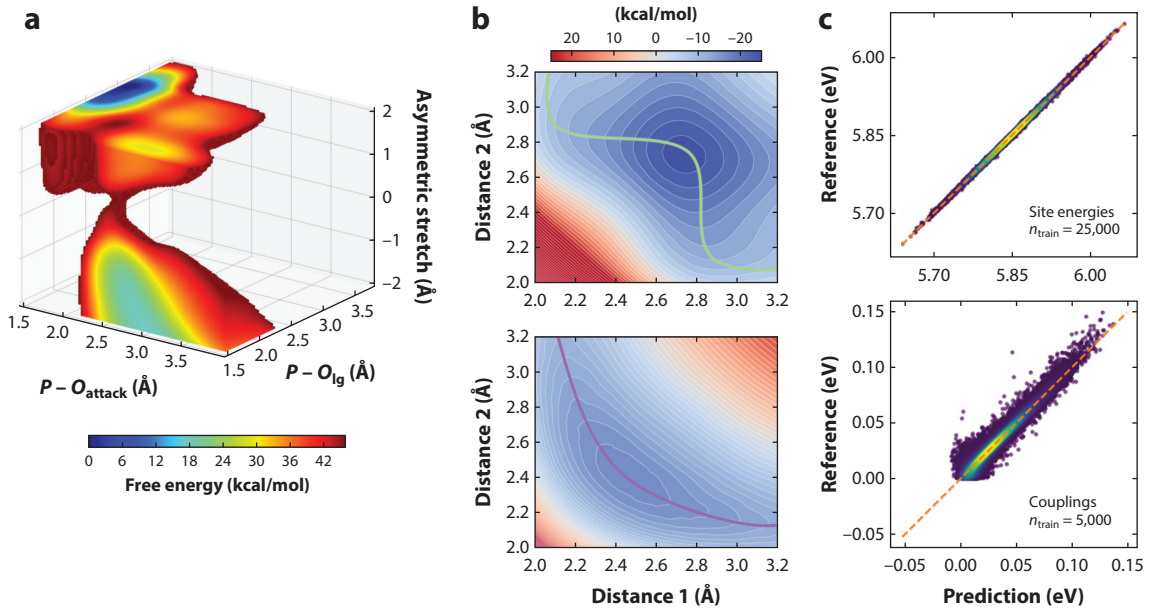


Figure 2

ML techniques can be integrated with QM/MM simulations in diverse ways. (*a*) Reinforcement learning and other techniques can be used to accelerate QM/MM free energy simulations. The example shown is the three-dimensional free energy surface for the hydrolysis of methyl phosphate in solution at the DFTB3/MM level (20). Panel adapted with permission from Reference 20; copyright 2021 American Chemical Society. (*b*) Neural networks using symmetry function representation of molecular structures are used in a Δ-ML scheme to correct the energies yielded by DFTB3. In this case, the qualitatively wrong DFTB3 energy landscape of the thiol–disulfide shuffling reaction (*top*) is quantitatively corrected to agree with the ab initio reference (*bottom*) (44). Panel adapted with permission from Reference 44; copyright 2022 American Chemical Society. (*c*) A kernel ridge regression combined with the Coulomb matrix representation was used to model the Hamiltonians describing ET processes in organic semiconductors. Shown is the accuracy of prediction versus reference for ET site energies (*top*) and couplings (*bottom*) for geometries taken from an MD simulation of crystalline anthracene (64). Panel adapted with permission from Reference 64; copyright 2021 American Chemical Society. Abbreviations: DFTB3, third-order density functional tight binding; ET, electron transfer; MD, molecular dynamics; ML, machine learning; MM, molecular mechanical; QM, quantum mechanical.

3.1. Long-Range Proton Transport in Proteins

Since the proton motive force plays key roles in bioenergetics, long-range proton transports coupled with either photon absorption or ETs are richly featured in bioenergy transductions. Representative examples include bR, Complex I, Complex IV (cytochrome c oxidase), and the F_0F_1-ATPase.

3.1.1. Key mechanistic questions. The key questions of common interest include: (*a*) What amino acids or cofactors are involved as the proton donor, acceptor, and mediating groups? This is difficult to answer based on experiments alone because the positions of hydrogen atoms are usually not visible in crystal or electron microscopy structures; mutation experiments and spectroscopic data are potentially informative, although a molecular-level interpretation is often not as straightforward as it may appear. An example that we touch on below concerns the proton storage group (PRG) in bR, which has attracted much attention over the past decades and up until very recently (10). (*b*) Is there a single, dominant proton transfer pathway or many pathways that are involved with similar fluxes? The typical approach for identifying potential proton transfer pathway(s) is to focus on water wires that mediate proton transfers via the canonical Grotthuss mechanism.

As discussed in an increasing number of studies (13, 78), however, water wires in static crystal structures in the absence of the excess proton may not represent the proton transfer pathway, and very often, multiple pathways may contribute. (*c*) Finally, arguably the most puzzling question is what controls the gating of proton transfers (49); the timing and directionality of long-range proton transfers lie at the heart of the efficiency of bioenergy transduction, and understanding the underlying molecular mechanism requires going beyond structural models to evaluate how thermodynamics and kinetics of proton transfers are modulated by other events, such as reduction of nearby cofactors (47, 60) and change in the local hydration level (78, 98, 107, 127).

3.1.2. Model validation. Proton transfer energetics depend critically on the pK_a values of the donor, acceptor, and mediating group(s). Therefore, microscopic pK_a calculations are essential validations for both the enzyme model (e.g., titration state of key residues) and the computational (QM/MM) methodology (112). The prediction of reliable microscopic pK_a values relies on an accurate treatment of both the proton affinity of the relevant titratable group and its interaction with the protein environment while in different protonation states. Moreover, the responses of the protein and internal water molecules to the change of titration state also need to be properly captured (162); these generally include both dipolar reorientations and electronic polarization, which require extensive conformational sampling and treatment of the electronic polarizability of the environment, respectively. In a study of cytochrome c oxidase, for example, DFTB/MM-based thermodynamic integration simulations (41, 46) were used to probe the microscopic pK_a value of the critical Glu286 residue, which was buried in a relatively hydrophobic internal cavity and thus featured a rather shifted experimental pK_a value of approximately 9.7. DFTB/MM free energy simulations with different enzyme models and approximate treatments of electronic polarization (46) found that reproducing the experimental pK_a value required penetration of water molecules into the cavity, which was coupled with the protonation of a propionate group of heme a_3. Therefore, in addition to serving as validation, microscopic pK_a calculations can potentially provide important mechanistic insights. Along this line, it is worth noting that microscopic pK_a calculations have also been used to calibrate reactive force field models (Multistate Empirical Valence Bond) for studying proton transfers in solution and proteins (74).

3.1.3. Example: the O to bR transition in bacteriorhodopsin. One prototypical example of proton pumps is bR, whose photocycle features several proton transfer reactions coupled to conformational transitions. Among the last few remaining challenges in the study of these complex processes were the atomic structure of the O state, especially with regard to the hydration of the active site, and the mechanism of the conversion to the ground state of bR, featuring a long-range proton transfer. Application of parallelized DFTB3/MM metadynamics protocols, as well as other extended sampling methods, made it possible to answer these questions with multidimensional free energy surfaces (87) (**Figure 3**).

Maag et al. (87) obtained the thermodynamics and kinetics of the proton transfer and found that they agreed well with experimental estimates, which underscored the credibility of the methodology used. The reaction energy of −3.6 kcal/mol was consistent with the measured pK_a differences between D85 and the PRG. Concerning the kinetics, the obtained energy barrier translated into a timescale of approximately 0.6 ms, which is in the range of experimental estimates.

This study (87) provided insight into the microscopic proton transfer mechanism and the key structural patterns, especially those related to charged amino acid side chains and water molecules, coupled to the actual proton transfer. The side chain of R82 switches between two possible distinct orientations to stabilize specific negatively charged amino acids: D85/D212 in the ground-state bR and PRG in the O state. Linked to this process is the control of the hydration level of the

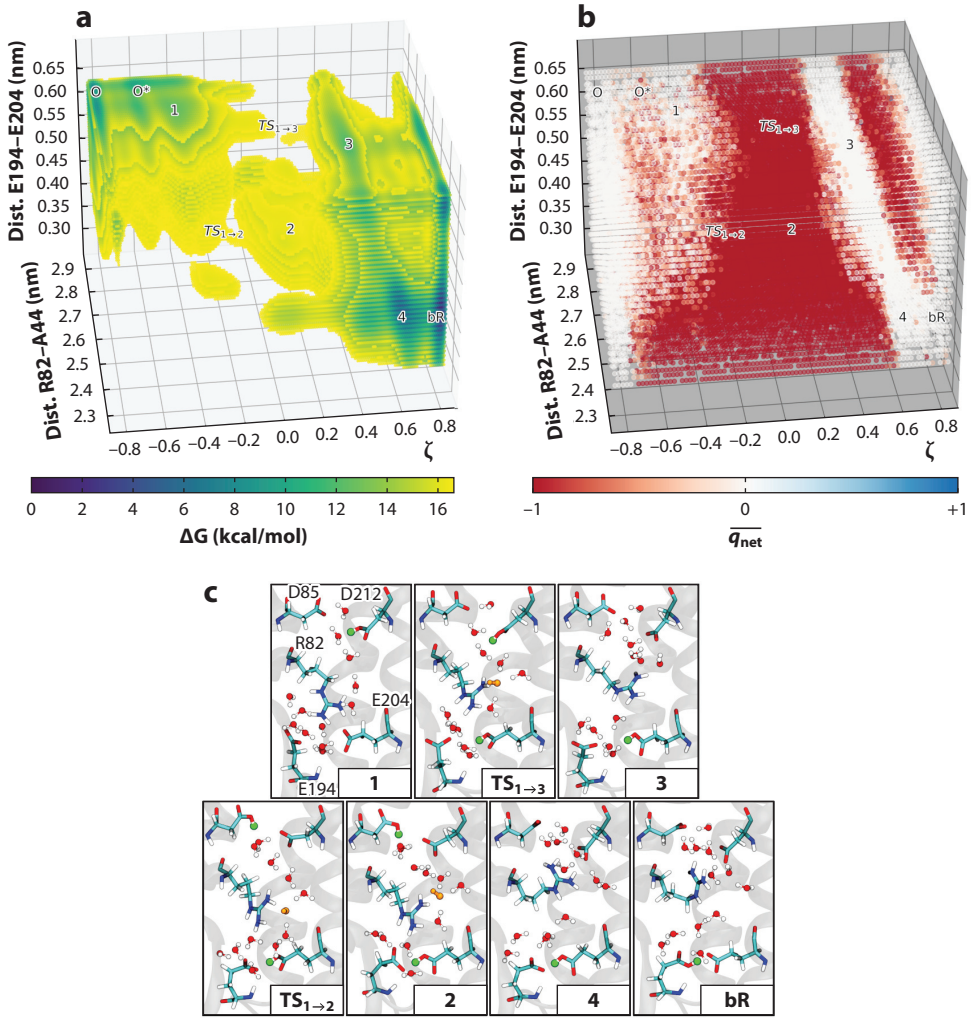


Figure 3

DFTB3/MM free energy simulations reveal a proton hole mechanism for the O→bR transition in bacteriorhodopsin (87), which involves a proton transfer from D85 to the PRG (E194, E204), mediated by D212, cavity water, and R82. (*a*) The 3D free energy surface from ∼100 ns multiwalker metadynamics simulations. The relevant states are labeled. (*b*) Net charge of QM water molecules. White indicates $\overline{q_{net}} = 0$, neutral water; red indicates $\overline{q_{net}} = -1$, OH^-; blue indicates $\overline{q_{net}} = +1$, H_3O^+. (*c*) Representative structures corresponding to the states identified in the free energy surface. Green indicates a proton bound to an Asp or Glu; orange indicates charged water species (in this case, always OH^-). Figure adapted with permission from Reference 87; copyright 2021 National Academy of Sciences. Abbreviations: MM, molecular mechanical; PRG, proton storage group; QM, quantum mechanical.

active site by means of the motion of the R82 side chain and opening and closing of the PRG. The ground-state bR features a low internal hydration level due to the closed PRG, and there is no continuous water wire because of the intervening R82 side chain. The latter feature helps prevent any wasteful back transfer of protons (49). In contrast, the O state has an elevated level of internal hydration, made possible by the open PRG; a continuous water wire is formed between D85 and the PRG, allowed by the reoriented side chain of R82. From a simple electrostatic perspective,

the positive charge of R82 may also aid in the generation and transfer of the hydroxide ion, which was observed in DFTB3/MM free energy simulations (87). The observation of the proton hole mechanism corroborates the chloride pumping ability of the D85S mutant of bR (118) and the similar mechanisms of bR and halorhodopsin, which feature conserved patterns of electrostatic interactions between the transferred anion and the protein's amino acid side chains (128).

3.1.4. Vibrational spectra. Two examples of vibrational spectra concern vibrational bands in bR from Fourier transform infrared spectroscopy studies by Gerwert and coworkers (33, 38). The first example focused on the PRG (see **Figure 3*c***), whose identity has been debated for many years. Numerous crystal structures, including the very recent high-resolution crystal structures for multiple kinetic states (10), revealed a pair of glutamate residues (Glu 194, Glu204) whose side chains are extremely close to each other, with carboxylate oxygens merely separated by approximately 2.4 Å. Such unusual geometry, together with mutation data, suggested that the pair of glutamates are involved in storing the proton. The FT-IR study of Gerwert and coworkers observed a diffuse band of approximately 2,000 cm^{-1}, which is commonly observed in protonated water clusters. Since several water molecules were observed near the pair of glutamates, it was initially suggested that the proton was in fact stored on these water molecules in the form of a delocalized proton. This interesting hypothesis stimulated a set of computational studies using different QM/MM methodologies and sizes of the QM region; nuclear quantum effects were also evaluated with ring-polymer MD (45). While it remains challenging to exactly match the experimental line shape, the latest DFT/MM simulations (135) continue to support the model (106) in which the excess proton is delocalized between the pair of glutamates, although nearby water molecules are clearly important to provide additional stabilization.

Another unique IR signature in bR was observed for several water molecules trapped between the protonated retinal Schiff base and deprotonated Asp96 in the N kinetic state (33). The vibrational peaks were observed around 2,540–2,775 cm^{-1}, which is substantially red-shifted compared to neutral water. DFTB3/MM simulations (152) were able to capture the key experimental observation (see **Figure 5*a*** below) and highlighted how the unique environment of the water cluster led to its significant polarization and therefore dramatic red-shift of the collective O-H vibration. An interesting observation was that the computed line shape was sensitive to the number of water molecules in the cavity, since including a larger number of water molecules perturbs the hydrogen bonding network and thus the vibrational coupling between the strongly polarized water molecules. Therefore, by systematically comparing computed and measured vibrational spectra, one would potentially be able to characterize the structure and composition of water molecules in protein internal cavities, which is difficult to accomplish otherwise, especially for transient kinetic states.

3.2. Fast Electron Transfer

Of paramount importance in bioenergetics are ET processes as the primary tool of biological energy transduction. ET taking place between an electron donor and an acceptor in a complex molecular environment may be described with the Marcus theory (102). For instance, Wu & Van Voorhis (156, 157) developed a scheme, based on constrained DFT, that generates the diabatic potential energy surfaces needed in the Marcus theory to express the reaction and reorganization energies. The application of constrained DFT was crucial as it turned out to effectively avoid the overdelocalization problem of DFT (155).

3.2.1. General remarks. Most biological ETs of interest, however, involve one or several electron relays, and a condition for the applicability of the Marcus theory is that any reorganization

processes have finished completely before the subsequent ET event. Thus, the Marcus model works for slow hopping ET but not for fast ET, where slow means that the individual events are infrequent rather than that they would take a long time to complete. This was described by Troisi (137) as a speed limit for hopping transfer. It appears that numerous ET pathways in bioenergetics exceed this speed limit, and the temporal scales of the reorganization and of the (relatively fast, or frequent) ET itself overlap. Similarly, Matyushov and colleagues (73) reasoned that the energy barriers to ET in proteins (as a representative of a complex molecular system) are not real equilibrium quantities; instead, they depend on the specific temporal scale of the ET reaction. Such cases thus need to be described by more flexible methodologies that do not imply any separation of temporal scales.

Biomolecular ET has been a frequent subject of computational studies, and the state of the art in 2015 was reviewed by Blumberger (7). Fast ET has proven challenging to describe, but it has still been investigated using a few different approaches, usually employing more advanced versions of the traditional ET models. For example, the light-driven ET in multiheme cytochrome STC from *Shewanella oneidensis* in aqueous solution occurs on a scale of several nanoseconds, and it was investigated by a combination of experimental and computational techniques (140). This work used a nonergodicity correction by Matyushov (93), which removes the contribution of slow motions from the total reorganization energy. In addition, the dynamics of ultrafast ET in flavodoxin protein was shown to depend on its coupling to environmental fluctuations (85). This coupling reduces the reaction free energy, as well as the reorganization energy. Most recently, the mechanism of the highly efficient ET supported by polymerized cytochrome OmcS in *Geobacter sulfurreducens* was discovered (21). The process takes place over micrometers, with elementary hopping on a subnanosecond scale—thus the notion of a protein wire—and accelerates upon cooling. On the other side of the spectrum of computational methods is the study of subpicosecond ET in a rhenium complex coupled to the azurin protein from *Pseudomonas aeruginosa* (88). Computationally expensive combinations of surface hopping and hybrid-DFT-based calculations of excited states were feasible given the short temporal scale to cover, making it possible to reveal the mechanism of the reaction.

By comparison, the FO-DFTB approach relies on the simultaneous propagation of the coupled electronic and nuclear degrees of freedom in a semiclassical fashion. This method does not involve any assumption of the relative rates of the ET reaction and the relaxation (reorganization) processes, making it possible to resolve any relevant temporal scales in an unbiased manner. The efficiency of the method makes it possible to explicitly describe ET processes on a nanosecond scale, as illustrated by the examples below.

3.2.2. Examples: photolyases and cryptochromes. Our labs applied the FO-DFTB/MM methodology to investigate multistep ETs in the proteins of the photolyase cryptochrome family (PCF), which are involved in DNA repair and in various signaling pathways. Woiczikowski et al. (151) investigated the part of the photoactivation process in *Escherichia coli* DNA photolyase, which constitutes an ET along three conserved Trp side chains (ET from the protein surface to the FAD cofactor). The ET was simulated on the naturally occurring temporal scales without introducing any system-specific parameters, and a kinetic analysis yielded rates in excellent agreement with experimental data. A detailed picture of the ET process emerged, and it became clear that the second ET step may occur before the structural relaxation following the first ET step has been completed. This illustrates the flexibility of the nonadiabatic simulation approach, which is applicable even in this case, to which the Marcus theory is not applicable, because of the nonequilibrium reaction conditions due to the overlapping temporal scales of the processes involved (ET itself, delocalization of charge, relaxation of protein, reorganization of solvent). As for the features

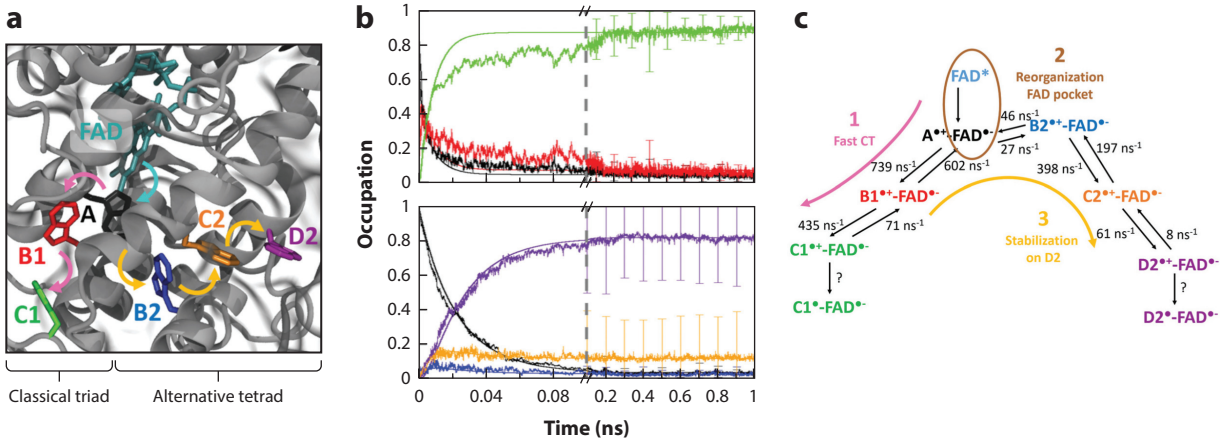


Figure 4

Multiscale nonadiabatic semiclassical simulations of electron transfer (ET) in complex and biomolecular systems involve a simultaneous propagation of the transferring electron and of the protein dynamics. Ths process allows them to provide atomic-level insight into the coupling of these processes. The example shown considers two different ET pathways in a photolyase protein PhrA (53). (*a*) The active site of PhrA, showing the amino acid side chain participating in the branching electron transfer pathways. (*b*) The temporal course of the averaged occupation of the individual participating amino acid side chains by the transferring electron hole along the two different pathways. (*c*) The kinetic scheme inferred from the temporal dependences of the occupations reveals a possible mechanism of the kinetic and thermodynamic control of the ET pathways in PhrA. All panels share the same color code to designate the different amino acid side chains considered as electron relays. Figure adapted with permission from Reference 53; copyright 2019 Royal Society of Chemistry.

of the ET in that protein and perhaps in PCF in general, it was shown that it is the polarization of the solvent at the surface of the protein that makes the process exergonic and thus unidirectional.

Holub et al. (53) then investigated the interesting case of a class III cyclobutane pyrimidine dimer photolyase from *Agrobacterium tumefaciens*, which features not just one but two different, branching ET pathways containing a total of six Trp side chains, which were both found to play a role in the photoreduction of FAD. They discovered an intriguing thermodynamic and kinetic competition between the two pathways: One pathway supports a faster ET, while the other pathway stabilizes the final product better. This balance manifested itself in Holub et al.'s simulations by a process in which the electron hole first transferred along the fast Trp triad, before a sequence of backward ET steps led to the transfer of the electron hole to the second pathway. The whole process occurred on a temporal scale of a nanosecond and was accompanied by complex repolarization of the environment (parts of the protein as well as, not least importantly, water molecules), emphasizing the need for a multiscale computational description (**Figure 4**).

3.3. Electronic Excitation

The computational determination of protein excited states poses a significant challenge. This is due to the large size of the molecular systems to be treated, on the one hand, and the requirements for both high accuracy and computational efficiency, on the other hand.

3.3.1. The challenge of excited states in biomolecules. The size of the chromophores requires the application of approximate methods like TD-DFT or more approximate post-Hartree-Fock methods like CC2. In some cases, the determination of an optimized structure using QM/MM geometry minimization techniques may be sufficient when average and optimized structures agree sufficiently well. This is the case for many retinal proteins like bR or ppR, as shown by

explicit calculations (52). In this case, the active site is characterized by a very stable and strongly hydrogen-bonded structure. We note that force fields using fixed point charges, however, may have difficulties in describing strongly hydrogen-bonded complexes and therefore that equilibrating the protein structure using such force fields may lead to conformations that are not suitable for further QM/MM investigations; this problem was described recently for the ChR2 protein (50), as well as for a glucose binding protein (104). Due to the long sampling times that are necessary, the system then requires using either a polarizable force field (104) or a semiempirical QM/MM approach, which has been shown to retain the active site structure well (50). In the latter case, however, a minimum energy structure may not be adequate, and absorption spectra have to be computed by sampling excited states along QM/MM MD trajectories (8).

All approximate quantum methods trade off accuracy with computational efficiency, which is particularly challenging for electronically excited states, as mentioned in Section 2.2. We note, however, that the effects of approximations are already visible in the determination of the ground state structures of conjugated electronic systems (8, 146), where the bond alternation (the difference in the bond lengths of neighboring single and double bonds) is sensitive to the method applied. This is important, since the ground state structure also determines the excited states properties; thus, a careful choice of methods for both ground and excited state calculations is needed. Similarly, torsional angles and planarity of the chromophore can be dependent on the approximation, which may also affect excited state properties.

The problems of TD-DFT for excited states are now well known, as recently discussed for the case of chlorophylls (113). It is important to mention the overestimation of excited state energies, which leads to a blue-shift in the computed spectra and is pronounced for hybrid and long-range corrected functionals. This systematic error can be effectively corrected by comparison to experimental data. More problematic is the inability of TD-DFT (or single-reference methods in general) to describe doubly excited states, and the problems with describing charge transfer states using this method (27); these problems were observed for retinal proteins early on (143, 146). For retinal, this can result in incorrect estimates of dependence of excitation energies on the chromophore structure, but also on the influence of the protein electrostatic field. The development of LC-TD-DFT methods could partially resolve many of the TD-DFT problems. However, LC-TD-DFT methods are far from being perfect for color-tuning studies in retinal proteins (see below) and still underestimate the response to the protein electrostatic field. Similar effects have been reported for chlorophylls (8) and can be expected to hold for other color pigments as well.

3.3.2. Color tuning of excited states in biomolecules. Many photoactive biomolecules, like rhodopsins, green fluorescent proteins, or blue light–using FAD proteins, contain one chromophore bound to a protein matrix. In contrast, light-harvesting systems like the Fenna-Matthews-Olson complex (FMO) or light-harvesting complex II (LH2) contain several (chlorophyll) chromophores. The maximum absorption wavelength is determined by several factors: (*a*) the geometrical and electronic features of the chromophore; (*b*) the steric, electrostatic, and hydrogen bonding interactions with the environment; and (*c*) a possible coupling between the chromophores within the protein complexes.

Due to their diverse molecular and electronic structure, the various chromophores have very different absorption maxima in the gas phase, but also very different responses to steric and electrostatic interactions with their environments. A steric interaction with nearby protein groups can twist the often planar structure, thereby disturbing the conjugation of the delocalized electronic system. Through hydrogen bonding, as in the case of the retinal Schiff base in retinal proteins, the localization of charge on the chromophore can be strongly affected, leading to large color shifts.

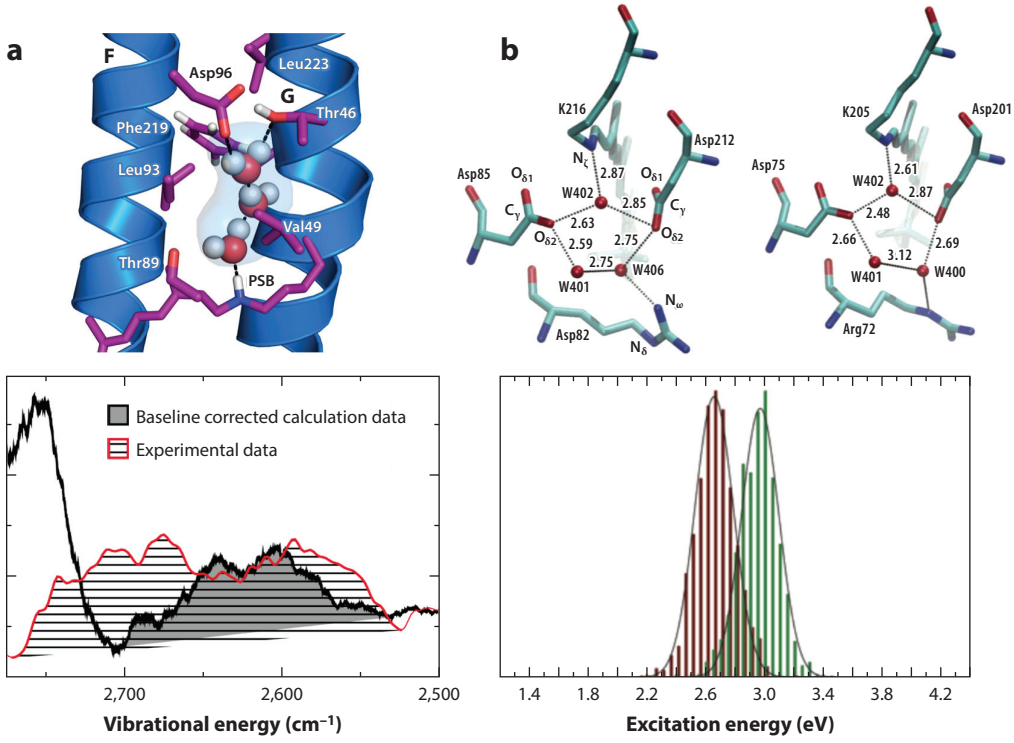


Figure 5

QM/MM calculations help explain shifts in vibrational and electronic spectra in retinal proteins. (*a*) A water cluster in the N state of bR has a unique local environment that leads to significant red shifts of the computed (DFTB3/MM) vibrational spectra, in agreement with experimental observations (152). Panel adapted with permission from Reference 152; copyright 2015 American Institute of Physics. (*b*) Difference in the hydrogen-bonding networks in the active sites of bR and ppR lead to significant shifts in the computed S0-S1 transitions: The histograms for bR (*red*) and ppR (*green*) are based on OM2/MRCI calculations sampled along ground state classical MD trajectories (52). Panel adapted with permission from Reference 52; copyright 2006 American Chemical Society. Abbreviations: bR, bacteriorhodopsin; MD, molecular mechanics; MM, molecular mechanical; MRCI, multireference configuration interaction; OM2, orthogonalization-corrected method 2; ppR, phoborhodopsin; QM, quantum mechanical.

Electrostatic interactions with the environment can lead to a further color shift, when ground and excited states have different dipole moments. Polarizable residues, in particular, in the immediate vicinity of the chromophore, can lead to a polarization shift. Retinal proteins, for example, absorb in a range of 300–700 nm, which illustrates their immense tunability due to their very sensitive response to external electrostatic potentials (8, 52); most other chromophores have much less tunability. For example, Hoffmann et al. (52) investigated the color shift of 70 nm between bR and phoborhodopsin (ppR) and found that the shift was partly due to the different hydrogen bonding networks around the retinal and partly due to the different electrostatic fields from the protein environment (**Figure 5**). In multichromophoric systems like LH2, close-lying chromophores lead to a coupling of excited states, i.e., a delocalization of the excited states over several chromophores, which results in a red shift of approximately 50 nm in LH2 (8, 71, 80, 124).

The impact of polarization on retinal excited state properties was emphasized by Warshel and coworkers (147) early on and has since been investigated in detail using QM and polarizable force field methods (144, 145). Including polarization allows one to compute accurate absolute excitation energies when a reliable and well-calibrated ab initio method like SORCI is used for the QM

region. This has allowed researchers to determine the protonation states of protein side chains (32) and discriminate between different structural models proposed for early intermediates of the bR photocycle (153, 154). Several other implementations of QM/MM with polarizable force fields have been reported to date (9, 81). For chlorophylls (82) in LH2, interestingly, direct electrostatic interactions seem to cancel the polarization response; therefore, it has been suggested that one should instead neglect the direct electrostatic interaction when not including polarization (17).

3.4. Mechanochemical Coupling

Many bioenergy transduction processes in cells are driven by the binding and hydrolysis of NTP (e.g., ATP or GTP). While in many cases large-scale conformational transitions have been established to be coupled directly to the binding of NTP and/or dissociation of hydrolysis products, it is an intrinsically interesting question how the hydrolysis of NTP is coupled with the conformational state of the biomolecule. Indeed, if the coupling were weak, then futile nucleotide hydrolysis would occur, leading to reduced thermodynamic efficiency of the energy transduction (51).

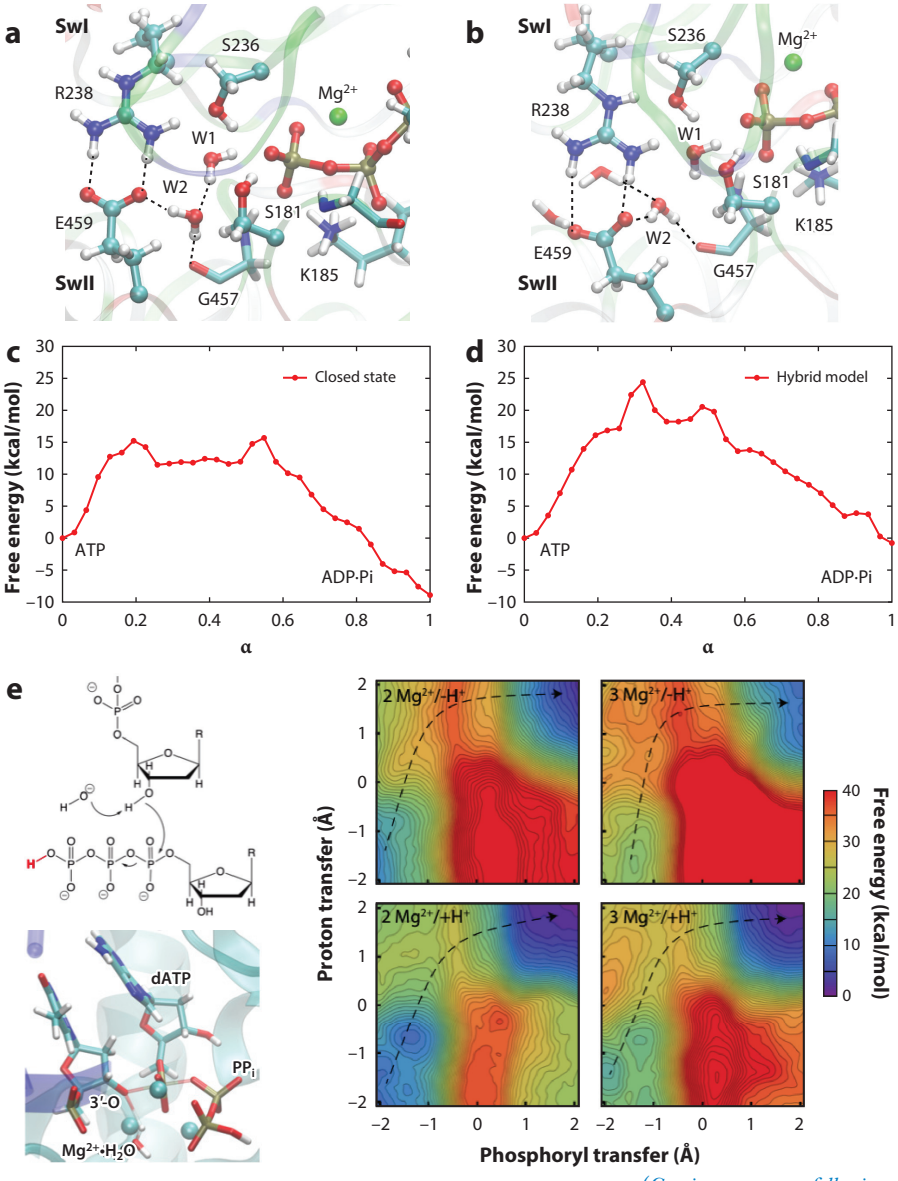
3.4.1. Key mechanistic questions. In addition to the more visible conformational transitions, other more subtle changes may make notable contributions to the chemical step. For example, an emerging theme in the study of nucleic acid enzymes such as DNA polymerases and RNase H is that transient cation recruiting into the active site may play a major role in activating the chemical step (120, 158). For example, in the DNA polymerase η, time-resolved crystallography has captured the presence of the third Mg^{2+} ion during the nucleotide addition process (37). Whether the third Mg^{2+} ion is essential to lowering the activation barrier of nucleotide addition or whether its main role is to stabilize the product has been debated. Nevertheless, the importance of such transiently recruited metal ions (i.e., not highly populated in the ground state structure) to the chemical activity has been well documented in several nucleic acid enzyme systems (92, 109). After all, it is common to invoke change of protonation state of active site residues for efficient catalysis in enzymes, despite the fact that the concentration of protons is often substantially lower than those for common metal ions under physiological conditions.

In short, the key mechanistic questions commonly encountered in biomolecular machines using NTP as fuels include: (*a*) In what conformational state does the chemical step (e.g., ATP hydrolysis) occur? (*b*) What are the roles of the key regulatory events, which are likely local in nature, such as closure of the nucleotide binding pocket, reduction of the level of local hydration, and recruitment of additional metal ion(s)? (*c*) How are these local regulatory transitions coupled to the more global conformational transitions so as to ensure a tight mechanochemical coupling and therefore a high thermodynamic efficiency of the bioenergy transduction process?

3.4.2. Model validation. Since nucleotide hydrolysis is the driving reaction for many biomolecular machines, it is important to calibrate the QM/MM methodology for phosphoryl transfer chemistry (117). This has been a challenging task due to the highly charged nature of phosphates and the existence of multiple competing mechanisms (61, 72), which place major demands on both computational accuracy and sampling efficiency. In addition to energetic properties, kinetic isotope effects are valuable for confirming that the nature of the transition state is adequately captured (72, 115).

3.4.3. Competing pathways and mechanochemical coupling. To illustrate the value of QM/MM simulations to the analysis of chemical steps in biomolecular machines, we briefly discuss two recent examples (**Figure 6**). The first is a classical biomolecular motor, myosin II, in which the hydrolysis of ATP requires closure of the active site, which in turn is coupled with the

remarkable rotation of the converter domain more than 40 Å away (133). Since our labs' QM/MM analysis of myosin (84) has been summarized in several recent articles (20, 117), we only highlight two key points in this review. First, even with the crystal structure that features a fully closed active site (prepowerstroke state), DFTB3/MM simulations were able to identify several reaction pathways for the hydrolysis of ATP that have rather similar rate-limiting free energy barriers. These pathways differ in terms of the mechanism through which the lytic water deprotonates to generate the strong nucleophile (OH^-) and the routes that the ionized proton takes to reach the hydrolysis product, the inorganic phosphate and ADP. This observation highlights that, even in enzyme active sites that have been optimized by evolution, competing pathways exist; thus, it is



(*Caption appears on following page*)

Figure 6 (*Figure appears on preceding page*)

QM/MM studies of reaction mechanisms in the biomolecular motors (*a–d*) myosin (84) and (*e*) DNA polymerase η (116). (*a*) The prepowerstroke state based on the crystal structure of the myosin motor domain complexed with ADP·vanadate (PDB ID 1VOM). (*b*) A computational hybrid model in which the active-site protein residues are restrained to adopt the structure of the prepowerstroke state, while the rest adopts the postrigor crystal structure (PDB ID 1FMW). (*c*,*d*) Free energy profiles computed for the two structural models based on DFTB3/MM string calculations. Note that, although the active-site structures are very similar in the two models, the hydrolysis free energy profiles are significantly different. Panels *a–d* adapted with permission from Reference 117; copyright 2018 Elsevier. (*e*) Mechanism, transition state structure, and free energy surfaces for a mechanism with a Mg^{2+}-coordinated hydroxide as the base under four different conditions: 2 Mg^{2+} and deprotonated leaving group, 3 Mg^{2+} and deprotonated leaving group, 2 Mg^{2+} and protonated leaving group, and 3 Mg^{2+} and protonated leaving group. Panel adapted with permission from Reference 116; copyright 2019 National Academy of Sciences. Abbreviations: PDB, Protein Data Bank; QM/MM, quantum mechanical/molecular mechanical.

essential to develop (almost) automated methodologies for the exploration of reaction pathways with little prior human bias. Along this line, the computational efficiency of the DFTB3/MM approach was essential to the successful integration between metadynamics and finite temperature string methods for the exploration of alternative reaction mechanisms (84).

Another unique piece of insight was obtained by studying the hydrolysis energetics in an artificial hybrid construct that featured a closed active site in a postrigor state of the myosin motor domain; this model was established to explicitly probe the coupling between ATP hydrolysis and distal conformational transitions. As shown in **Figure 6*a–d***, although the hybrid and prepowerstroke active sites feature essentially the same set of first coordination shell interactions for the lytical water and the γ-phosphate, the computed free energy profiles differ significantly in terms of both the rate-limiting barrier height and the overall exergonicity; both quantities are less favorable by approximately 9 kcal/mol for the hybrid structural model. This result clearly underscores the notion that reaction free energy profiles in enzyme active sites are not solely determined by first-shell interactions. In particular, in the absence of conformational transitions beyond the active site, the reactive moieties and catalytic motifs (e.g., Switch II) in the hybrid model are surrounded by a cluster of water molecules, which are not observed in the prepowerstroke state. As a result, the active site is less preorganized in the hybrid model, leading to less favorable reaction energetics. This observation is reminiscent of the results from a recent computational analysis of loop closure in triose phosphate isomerase (79); it was found that the proton transfer energetics remained largely similar to the bulk values as long as a few water molecules were trapped in the active site, even with the lid loop adopting essentially the fully closed configuration. Evidently, establishing a well-preorganized active site for efficient catalysis requires conformational rearrangements beyond the nearest neighbors of the catalytic motifs, which together with cooperative conformational transitions form the basis of mechanochemical coupling in biomolecular machines.

As a second example, we briefly discuss DNA polymerase η, especially the role of the third Mg^{2+} ion in nucleotide addition. While energy transduction is not the primary function of DNA and RNA polymerases, they are no doubt sophisticated biomolecular machines in that their functional cycles involve complex and coordinated structural transitions at different spatial and temporal scales; the goal of many mechanistic studies is to establish the driving force and functional significance of these structural transitions (138). For DNA polymerase η, a much-debated issue in recent years concerns the role of the third Mg^{2+} ion identified in time-resolved crystallography studies (37); a closely related question concerns the timing of the deprotonation of the 3′-OH in the terminal base, for which multiple mechanisms have been suggested (40). To shed light on these issues, Roston et al. (116) conducted extensive free energy simulations using DFTB3/MM following

model calibration, including microscopic pK_a calculations for both model compounds and enzyme active sites. By systematically comparing free energy profiles for 10 plausible mechanistic models, they found that the lowest activation barrier occurred for a reaction where a Mg^{2+}-coordinated water deprotonates the nucleophilic 3′-OH in concert with the phosphoryl transfer step. The presence of a third Mg^{2+} in the active site was observed to indeed lower the activation barrier for the water-as-base mechanism, as did protonation of the pyrophosphate leaving group. This mechanistic model, which does not invoke any specific protein residue as the catalytic base to activate 3′-OH, is consistent with a recent study (48) that systematically removed potential hydrogen-bonding partners of the 3′-OH; it was observed that no single or combined perturbations eliminated the catalytic activity; i.e., neither the proton acceptor nor the departure path of the nucleophile deprotonation is fixed. These observations supported a model in which the 3′-OH deprotonation does not require a specific general base and is readily activated by the three Mg^{2+} ions with flexible proton exit routes. DFT/MM simulations have also been employed to probe various mechanistic issues in related systems (6, 26, 40, 94, 108, 129, 131, 136), sometimes reaching similar conclusions as DFTB/MM studies, such as the key features of ATP/GTP hydrolysis transition states in biomolecular motors (94, 108, 131, 136) and the role of the third Mg^{2+} ion in DNA polymerase η (129). In other cases, different conclusions were drawn, such as in a study of the mechanism of 3′-OH deprotonation in DNA polymerase (40). With substantial differences in the timescale of sampling and general computational accuracy between DFT/MM and DFTB/MM simulations, it remains difficult to resolve discrepancies based solely on computational results; as emphasized in many studies, it is imperative to view the computational results in the context of available experimental data, and only by combining information obtained from experiments with that obtained from simulation can we obtain microscopic answers to questions of chemical reactivity in complex settings.

4. CONCLUSIONS AND OUTLOOKS

In this review, we discuss QM/MM methods and their applications to various bioenergy transduction processes, such as long-range proton transport, fast ETs, and nucleotide hydrolysis. These applications require balancing computational efficiency and accuracy for the system of interest, making semiempirical QM/MM models particularly valuable, although calibration and comparison to higher-level QM/MM methods is indispensable. Looking forward, there are several areas that can benefit from further efforts to make such QM/MM analyses more quantitative and applicable to increasingly complex problems.

First, further improvements in both semiempirical and ab initio QM methods, especially for transition metal ions and noncovalent interactions, are of major significance, since many systems in bioenergetics involve metal cofactors and require treating a large number of atoms at the QM level for the description of long-range proton or electron transfers or mechanochemical coupling; further integration with ML techniques is likely required to reach quantitative accuracy for these challenging problems.

Second, for MM methods, it is important to further enhance not only the accuracy (e.g., an explicit consideration of electronic polarization), but also the complexity that better represents the working environment of the biomolecular machine(s) under investigation. Examples of the latter include representation of the transmembrane potential, pH gradient, and molecular crowding. In other words, the general aim is to develop truly multiscale computational models that enable the analysis of energy transduction under realistic cellular conditions.

Finally, while providing a better understanding of experimental observations remains a major goal for computational studies, we believe that blind predictions will play an important role in critically evaluating different QM/MM methodologies, similar to such efforts in the areas of

protein structure prediction (CASP, CAPRI) and ligand binding (SAMPLX). As QM/MM calculations become increasingly affordable and standardized, along with advances in high-throughput experiments (96) that are able to efficiently generate a large amount of data, the time might become ripe for initiating such efforts in the QM/MM community.

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

We thank our collaborators over the years for making the developments and applications of QM/MM methods possible. We also thank various agencies for generous grant support, especially the National Institutes of Health through grants R01-GM106443 and R35-GM141930 to Q.C., the German Science Foundation (DFG) through projects GRK 2450 and INST 40/575-1 FUGG (JUSTUS 2 cluster) to T.K. and M.E. and grants EL 206/18-1 to M.E. and KU 3677/2-1 to T.K., and the state of Baden-Württemberg through bwHPC to T.K. and M.E.

LITERATURE CITED

1. Antes I, Thiel W. 1999. Adjusted connection atoms for combined quantum mechanical and molecular mechanical methods. *J. Phys. Chem. A* 103:9290–95
2. Astumian RD, Mukherjee S, Warshel A. 2016. The physics and physical chemistry of molecular machines. *Phys. Chem. Chem. Phys.* 17:1719–41
3. Baiz CR, Blasiak B, Bredenbeck J, Cho M, Choi JH, et al. 2020. Vibrational spectroscopic map, vibrational spectroscopy, and intermolecular interaction. *Chem. Rev.* 120:7152–218
4. Bannwarth C, Caldeweyher E, Ehlert S, Hansen A, Pracht P, et al. 2020. Extended tight-binding quantum chemistry methods. *WIREs Comput. Mol. Sci.* 11:e01493
5. Barros EP, Ries B, Boselt L, Champion C, Riniker S. 2022. Recent developments in multiscale free energy simulations. *Curr. Opin. Struct. Biol.* 72:55–62
6. Berta D, Buigues PJ, Badaoui M, Rosta E. 2020. Cations in motion: QM/MM studies of the dynamic and electrostatic roles of H^+ and Mg^{2+} ions in enzyme reactions. *Curr. Opin. Struct. Biol.* 61:198–206
7. Blumberger J. 2015. Recent advances in the theory and molecular simulation of biological electron transfer reactions. *Chem. Rev.* 115(20):11191–238
8. Bold BM, Sokolov M, Maity S, Wanko M, Dohmen PM, et al. 2020. Benchmark and performance of long-range corrected time-dependent density functional tight binding (LC-TD-DFTB) on rhodopsins and light-harvesting complexes. *Phys. Chem. Chem. Phys.* 22:10500–18
9. Bondanza M, Nottoli M, Cupellini L, Lipparini F, Mennucci B. 2020. Polarizable embedding QM/MM: the future gold standard for complex (bio)systems? *Phys. Chem. Chem. Phys.* 22:14433–48
10. Borshchevskiy V, Kovalev K, Round E, Efremov R, Astashkin R, et al. 2022. True-atomic-resolution insights into the structure and functional role of linear chains and low-barrier hydrogen bonds in proteins. *Nat. Struct. Mol. Biol.* 29:440–50
11. Brunk E, Rothlisberger U. 2015. Mixed quantum mechanical/molecular mechanical molecular dynamics simulations of biological systems in ground and electronically excited states. *Chem. Rev.* 115:6217–63
12. Bustamante C, Keller D, Oster G. 2001. The physics of molecular motors. *Acc. Chem. Res.* 34:412–20
13. Cai XH, Haider K, Lu JX, Radik S, Son CY, et al. 2018. Network analysis of a proposed exit pathway for protons to the P-side of cytochrome c oxidase. *Biochem. Biophys. Acta Bioenerg.* 1859:997–1005
14. Chen MS, Zuehlsdorff TJ, Morawietz T, Isborn CM, Markland TE. 2020. Exploiting machine learning to efficiently predict multidimensional optical spectra in complex environments. *J. Phys. Chem. Lett.* 11:7559–68

15. Christensen AS, Kubař T, Cui Q, Elstner M. 2016. Semi-empirical quantum mechanical methods for non-covalent interactions for chemical and biochemical applications. *Chem. Rev.* 116:5301–37
16. Chung LW, Sameera WMC, Ramozzi R, Page AJ, Hatanaka M, et al. 2015. The ONIOM method and its applications. *Chem. Rev.* 115:5678–769
17. Cignoni E, Cupellini L, Menucci B. 2022. A fast method for electronic couplings in embedded multichromophoric systems. *J. Phys. Condens. Matter* 34:304004
18. Cui K, Yethiraj A, Schmidt JR. 2019. Influence of charge scaling on the solvation properties of ionic liquid solutions. *J. Phys. Chem. B* 123:9222–29
19. Cui Q. 2016. Quantum mechanical methods in biochemistry and biophysics. *J. Chem. Phys.* 145:140901
20. Cui Q, Pal T, Xie L. 2021. Biomolecular QM/MM simulations: What are some of the "burning issues"? *J. Phys. Chem. B* 125:689–702
21. Dahl PJ, Yi SM, Gu Y, Acharya A, Shipps C, et al. 2022. A 300-fold conductivity increase in microbial cytochrome nanowires due to temperature-induced restructuring of hydrogen bonding networks. *Sci. Adv.* 8(19):eabm7193
22. Das D, Eurenius KP, Billings EM, Sherwood P, Chatfield DC, et al. 2002. Optimization of quantum mechanical molecular mechanical partitioning schemes: Gaussian delocalization of molecular mechanical charges and the double link atom method. *J. Chem. Phys.* 117:10534–47
23. Das S, Nam K, Major DT. 2018. Rapid convergence of energy and free energy profiles with quantum mechanical size in quantum mechanical–molecular mechanical simulations of proton transfer in DNA. *J. Chem. Theory Comput.* 14(3):1695–705
24. Demapan D, Kussman J, Ochsenfeld C, Cui Q. 2022. Factors that determine the variation of equilibrium and kinetic properties of QM/MM enzyme simulations: QM region, conformation and boundary condition. *J. Chem. Theory Comput.* 18:2530–42
25. Deng J, Cui Q. 2022. Electronic polarization is essential for the stabilization and dynamics of buried ion pairs in staphylococcal nuclease mutant. *J. Am. Chem. Soc.* 144:4594–610
26. Donati E, Genna V, De Vivo M. 2020. Recruiting mechanism and functional role of a third metal ion in the enzymatic activity of 5' structure-specific nucleases. *J. Am. Chem. Soc.* 142:2823–34
27. Dreuw A, Head-Gordon M. 2005. Single-reference ab initio methods for the calculation of excited states of large molecules. *Chem. Rev.* 105(11):4009–37
28. Duboué-Dijon E, Javanainen M, Delcroix P, Jungwirth P, Martinez-Seara H. 2020. A practical guide to biologically relevant molecular simulations with charge scaling for electronic polarization. *J. Chem. Phys.* 153:050901
29. E W, Vanden-Eijnden E. 2010. Transition-path theory and path-finding algorithms for the study of rare events. *Annu. Rev. Phys. Chem.* 61:391–420
30. Field MJ, Bash PA, Karplus M. 1990. A combined quantum-mechanical and molecular mechanical potential for molecular-dynamics simulations. *J. Comput. Chem.* 11(6):700–33
31. Flaig D, Beer M, Ochsenfeld C. 2012. Convergence of electronic structure with the size of the QM region: example of QM/MM NMR shieldings. *J. Chem. Theory Comput.* 8(7):2260–71
32. Frähmcke J, Wanko M, Phatak P, Mroginski A, Elstner M. 2010. The protonation state of Glu181 in rhodopsin revisited: Interpretation of experimental data on the basis of QM/MM calculations. *J. Phys. Chem. B* 114:11338–52
33. Freier E, Wolf S, Gerwert K. 2011. Proton transfer via a transient linear water-molecule chain in a membrane protein. *PNAS* 108:11435–39
34. Gao J. 1996. Methods and applications of combined quantum mechanical and molecular mechanical potentials. In *Reviews in Computational Chemistry*, Vol. VII, ed. KB Lipkowitz, DB Boyd, pp. 119–85. New York: Wiley
35. Gao JL, Amara P, Alhambra C, Field MJ. 1998. A generalized hybrid orbital (GHO) method for the treatment of boundary atoms in combined QM/MM calculations. *J. Phys. Chem. A* 102:4714–21
36. Gao JL, Truhlar DG. 2002. Quantum mechanical methods for enzyme kinetics. *Annu. Rev. Phys. Chem.* 53:467–505
37. Gao Y, Yang W. 2016. Capture of a third Mg^{2+} is essential for catalyzing DNA synthesis. *Science* 352:1334–37

38. Garczarek F, Gerwert K. 2006. Functional waters in intraprotein proton transfer monitored by FTIR difference spectroscopy. *Nature* 439:109–12
39. Gaus M, Cui Q, Elstner M. 2014. Density functional tight binding (DFTB): application to organic and biological molecules. *WIREs Comput. Mol. Sci.* 4:49–61
40. Genna V, Vidossich P, Ippoliti E, Carloni P, De Vivo M. 2016. A self-activated mechanism for nucleic acid polymerization catalyzed by DNA/RNA polymerases. *J. Am. Chem. Soc.* 138:14592–98
41. Ghosh N, Prat-Resina X, Gunner M, Cui Q. 2009. Microscopic pKa analysis of Glu286 in cytochrome c oxidase (*Rhodobacter sphaeroides*): toward a calibrated molecular model. *Biochemistry* 48:2468–85
42. Gillet N, Berstis L, Wu XJ, Gajdos F, Heck A, et al. 2016. Electronic coupling calculations for bridge-mediated charge transfer using constrained density functional theory (CDFT) and effective Hamiltonian approaches at the density functional theory (DFT) and fragment-orbital density functional tight binding (FODFTB) level. *J. Chem. Theory Comput.* 12:4793–805
43. Gillet N, Elstner M, Kubař T. 2018. Coupled-perturbed DFTB-QM/MM metadynamics: application to proton-coupled electron transfer. *J. Chem. Phys.* 149(7):072328
44. Gómez-Flores CL, Maag D, Kansari M, Vuong VQ, Irle S, et al. 2022. Accurate free energies for complex condensed-phase reactions using an artificial neural network corrected DFTB/MM methodology. *J. Chem. Theory Comput.* 18(2):1213–26
45. Goyal P, Ghosh N, Phatak P, Clemens M, Gaus M, et al. 2011. Proton storage site in bacteriorhodopsin: new insights from QM/MM simulations of microscopic pK(a) and infrared spectra. *J. Am. Chem. Soc.* 133:14981–97
46. Goyal P, Lu J, Yang S, Gunner MR, Cui Q. 2013. Changing hydration level in an internal cavity modulates the proton affinity of a key glutamate in cytochrome c oxidase. *PNAS* 110:18886–91
47. Goyal P, Yang S, Cui Q. 2015. Microscopic basis for kinetic gating in cytochrome c oxidase: insights from QM/MM analysis. *Chem. Sci.* 6:826–41
48. Gregory MT, Gao Y, Cui Q, Yang W. 2021. Multiple deprotonation paths of the nucleophile 3'-OH in the DNA synthesis reaction. *PNAS* 118:e2103990118
49. Gunner MR, Amin M, Zhu X, Lu J. 2013. Molecular mechanisms for generating transmembrane proton gradients. *Biochim. Biophys. Acta Bioeng.* 1827:892–913
50. Guo Y, Beyle FE, Bold BM, Watanabe HC, Koslowski A, et al. 2016. Active site structure and absorption spectrum of channelrhodopsin-2 wild-type and C128T mutant. *Chem. Sci.* 7(6):3879–91
51. Hill TL. 1977. *Free Energy Transduction in Biology*. New York: Acad. Press
52. Hoffmann M, Wanko M, Strodel P, König PH, Frauenheim T, et al. 2006. Color tuning in rhodopsins: the mechanism for the spectral shift between bacteriorhodopsin and sensory rhodopsin II. *J. Am. Chem. Soc.* 128(33):10808–18
53. Holub D, Lamparter T, Elstner M, Gillet N. 2019. Biological relevance of charge transfer branching pathways in photolyases. *Phys. Chem. Chem. Phys.* 21(31):17072–81
54. Hou G, Zhu X, Elstner M, Cui Q. 2012. A modified QM/MM Hamiltonian with the self-consistent-charge density-functional-tight-binding theory for highly charged QM regions. *J. Chem. Theory Comp.* 8:4293–304
55. Hu H, Yang WT. 2008. Free energies of chemical reactions in solution and in enzymes with ab initio quantum mechanics/molecular mechanics methods. *Annu. Rev. Phys. Chem.* 59:573–601
56. Inakollu VSS, Geerke DP, Rowley CN, Yu H. 2020. Polarisable force fields: What do they add in biomolecular simulations? *Curr. Opin. Struct. Biol.* 61:182–90
57. Jacquemin D, Perpète EA, Scuseria GE, Ciofini I, Adamo C. 2008. TD-DFT performance for the visible absorption spectra of organic dyes: conventional versus long-range hybrids. *J. Chem. Theory Comput.* 4(1):123–35
58. Jindal G, Warshel A. 2016. Exploring the dependence of QM/MM calculations of enzyme catalysis on the size of the QM region. *J. Phys. Chem. B* 120(37):9913–21
59. Jing Z, Liu C, Cheng SY, Qi R, Walker BD, et al. 2019. Polarizable force fields for biomolecular simulations: recent advances and applications. *Annu. Rev. Biophys.* 48:371–94
60. Kaila VRI. 2021. Resolving chemical dynamics in biological energy conversion: long-range proton-coupled electron transfer in respiratory complex I. *Acc. Chem. Res.* 54:4462–73

61. Kamerlin SCL, Sharma PK, Prasad RB, Warshel A. 2013. Why nature really chose phosphate. *Q. Rev. Biophys.* 46:1–132
62. Koenig P, Ghosh N, Hoffman M, Elstner M, Tajkhorshid E, et al. 2006. Towards theoretical analysis of long-range proton transfer kinetics in biomolecular pumps. *J. Phys. Chem. A* 110:548–63
63. König PH, Hoffmann M, Frauenheim T, Cui Q. 2005. A critical evaluation of different QM/MM frontier treatments with SCC-DFTB as the QM method. *J. Phys. Chem. B* 109:9082–95
64. Krämer M, Dohmen PM, Xie WW, Holub D, Christensen AS, Elstner M. 2021. Charge and exciton transfer simulations using machine-learned Hamiltonians. *J. Chem. Theory Comput.* 16:4061–70
65. Kranz JJ, Elstner M, Aradi B, Frauenheim T, Lutsker V, et al. 2017. Time-dependent extension of the long-range corrected density functional based tight-binding method. *J. Chem. Theory Comput.* 13(4):1737–47
66. Kubař T, Elstner M. 2013. Efficient algorithms for the simulation of non-adiabatic electron transfer in complex molecular systems: application to DNA. *Phys. Chem. Chem. Phys.* 15:5794–813
67. Kubař T, Elstner M. 2013. A hybrid approach to simulation of electron transfer in complex molecular systems. *J. Royal Soc. Interface* 10:20130415
68. Kubař T, Woiczikowski PB, Cuniberti G, Elstner M. 2008. Efficient calculation of charge-transfer matrix elements for hole transfer in DNA. *J. Phys. Chem. B* 112:7937–47
69. Kulik HJ, Zhang J, Klinman JP, Martínez TJ. 2016. How large should the QM region be in QM/MM calculations? The case of catechol *O*-methyltransferase. *J. Phys. Chem. B* 120(44):11381–94
70. Lagardere L, Jolly LH, Lipparini F, Aviat F, Stamm B, et al. 2018. Tinker-HP: a massively parallel molecular dynamics package for multiscale simulations of large complex systems with advanced point dipole polarizable force fields. *Chem. Sci.* 9:956–72
71. Lahav Y, Noy D, Schapiro I. 2021. Spectral tuning of chlorophylls in proteins: electrostatics versus ring deformation. *Phys. Chem. Chem. Phys.* 23(11):6544–51
72. Lassila JK, Zalatan JG, Herschlag D. 2011. Biological phosphoryl transfer reactions: understanding mechanism and catalysis. *Annu. Rev. Biochem.* 80:669–702
73. LeBard DN, Martin DR, Lin S, Woodbury NW, Matyushov DV. 2013. Protein dynamics to optimize and control bacterial photosynthesis. *Chem. Sci.* 4:4127–36
74. Lee S, Liang RB, Voth GA, Swanson JMJ. 2016. Computationally efficient multiscale reactive molecular dynamics to describe amino acid deprotonation in proteins. *J. Chem. Theory Comput.* 12:879–91
75. Lemkul JA, Huang J, Roux B, MacKerell AD Jr. 2016. An empirical polarizable force field based on the classical Drude oscillator model: development history and recent applications. *Chem. Rev.* 116:4983–5013
76. Leontyev IV, Stuchebrukhov AA. 2011. Accounting for electronic polarization in non-polarizable force fields. *Phys. Chem. Chem. Phys.* 13:2613–26
77. Li CH, Voth GA. 2021. Using constrained density functional theory to track proton transfers and to sample their associated free energy surface. *J. Chem. Theory Comput.* 17:5759–65
78. Liang RB, Swanson JMJ, Peng YX, Wikström M, Voth GA. 2016. Multiscale simulations reveal key features of the proton-pumping mechanism in cytochrome c oxidase. *PNAS* 113:7420–25
79. Liao QH, Kulkarni Y, Sengupta U, Petrovic D, Mulholland AJ, et al. 2018. Loop motion in triosephosphate isomerase is not a simple open and shut case. *J. Am. Chem. Soc.* 140:15889–903
80. List NH, Curutchet C, Knecht S, Mennucci B, Kongsted J. 2013. Toward reliable prediction of the energy ladder in multichromophoric systems: a benchmark study on the FMO light-harvesting complex. *J. Chem. Theory Comput.* 9(11):4928–38
81. Loco D, Lagardere L, Adjoua O, Piquemal JP. 2021. Atomistic polarizable embeddings: energy, dynamics, spectroscopy, and reactivity. *Acc. Chem. Res.* 54:2812–22
82. Loco D, Polack E, Capresa S, Lagardere L, Lipparini F, et al. 2016. A QM/MM approach using the amoeba polarizable embedding: from ground state energies to electronic excitations. *J. Chem. Theory Comput.* 12:3654–61
83. Lonsdale R, Mulholland A. 2014. QM/MM modelling of drug-metabolizing enzymes. *Curr. Top. Med. Chem.* 14(11):1339–47
84. Lu X, Ovchinnikov V, Roston DR, Demapan D, Cui Q. 2017. Regulation and plasticity of catalysis in enzymes: insights from analysis of mechanochemical coupling in myosin. *Biochemistry* 56:1482–97

85. Lu Y, Kundu M, Zhong D. 2020. Effects of nonequilibrium fluctuations on ultrafast short-range electron transfer dynamics. *Nat. Commun.* 11:2822
86. Lüdemann G, Solov'yov I, Kubař T, Elstner M. 2015. Solvent driving force ensures fast formation of a persistent and well-separated radical pair in plant cryptochrome. *J. Am. Chem. Soc.* 137(3):1147–56
87. Maag D, Mast T, Elstner M, Cui Q, Kubař T. 2021. O to bR transition in bacteriorhodopsin occurs through a proton hole mechanism. *PNAS* 118:e2024803118
88. Mai S, Menger MFSJ, Marazzi M, Stolba DL, Monari A, González L. 2020. Competing ultrafast photoinduced electron transfer and intersystem crossing of $Re(CO)_3(Dmp)(His124)(Trp122)]^+$ in *Pseudomonas aeruginosa* azurin: a nonadiabatic dynamics study. *Theor. Chem. Acc.* 139(3):65
89. Maity S, Bold BM, Prajapati JD, Sokolov M, Kubař T, et al. 2020. DFTB/MM molecular dynamics simulations of the FMO light-harvesting complex. *J. Phys. Chem. Lett.* 11(20):8660–67
90. Maity S, Daskalakis V, Elstner M, Kleinekathöfer U. 2021. Multiscale QM/MM molecular dynamics simulations of the trimeric major light-harvesting complex II. *Phys. Chem. Chem. Phys.* 23(12):7407–17
91. Major DT, Gao J. 2007. An integrated path integral and free-energy perturbation-umbrella sampling method for computing kinetic isotope effects of chemical reactions in solution and in enzymes. *J. Chem. Theory Comput.* 3(3):949–60
92. Manigrasso J, De Vivo M, Palermo G. 2021. Controlled trafficking of multiple and diverse cations prompts nucleic acid hydrolysis. *ACS Catal.* 11:8786–97
93. Matyushov DV. 2013. Protein electron transfer: dynamics and statistics. *J. Chem. Phys.* 139(2):025102
94. McCullagh M, Saunders MG, Voth GA. 2014. Unraveling the mystery of ATP hydrolysis in actin filaments. *J. Am. Chem. Soc.* 136:13053–58
95. Mehmood R, Kulik HJ. 2020. Both configuration and QM region size matter: zinc stability in QM/MM models of DNA methyltransferase. *J. Chem. Theory Comput.* 16:3121–34
96. Mokhtari DA, Appel MJ, Fordyce PM, Herschlag D. 2021. High throughput and quantitative enzymology in the genomic era. *Curr. Opin. Struct. Biol.* 71:259–73
97. Mugnai ML, Hyeon C, Hinczewski M, Thirumalai D. 2020. Theoretical perspectives on biological machines. *Rev. Mod. Phys.* 92:025001
98. Muhlbauer ME, Saura P, Nuber F, Di Luca A, Friedrich T, Kaila VRI. 2020. Water-gated proton transfer dynamics in respiratory complex I. *J. Am. Chem. Soc.* 142:13718–28
99. Naseem-Khan S, Lagardere L, Narth C, Cisneros GA, Ren P, et al. 2022. Development of the quantum-inspired SIBFA many-body polarizable force field: enabling condensed-phase molecular dynamics simulations. *J. Chem. Theory Comput.* 18:3607–21
100. Nicholls DG, Ferguson SJ. 2002. *Bioenergetics*. New York: Acad. Press. 3rd ed.
101. Nishimura Y, Nakai H. 2019. DCDFTBMD: divide-and-conquer density functional tight-binding program for huge-system quantum mechanical molecular dynamics simulations. *J. Comput. Chem.* 40:1538–49
102. Nitzan A. 2014. *Chemical Dynamics in Condensed Phases: Relaxation, Transfer, and Reactions in Condensed Molecular Systems*. Oxford, UK: Oxford Univ. Press
103. Noe F, Tkatchenko A, Müller KR, Clementi C. 2020. Machine learning for molecular simulation. *Annu. Rev. Phys. Chem.* 71:361–90
104. Pang Z, Sokolov M, Kubař T, Elstner M. 2022. Unravelling the mechanism of glucose binding in a protein-based fluorescence probe: molecular dynamics simulation with a tailor-made charge model. *Phys. Chem. Chem. Phys.* 24(4):2441–53
105. Parac M, Grimme S. 2002. Comparison of multireference Møller-Plesset theory and time-dependent methods for the calculation of vertical excitation energies of molecules. *J. Phys. Chem. A* 106:6844–50
106. Phatak P, Ghosh N, Yu H, Cui Q, Elstner M. 2008. Amino acids with an intermolecular proton bond as the proton storage site in bacteriorhodopsin. *PANS* 105:19672–77
107. Pisliakov AV, Sharma PK, Chu ZT, Haranczyk M, Warshel A. 2008. Electrostatic basis for the unidirectionality of the primary proton transfer in cytochrome c oxidase. *PNAS* 105:7726–31
108. Priess M, Göddeke H, Groenhof G, Schäfer LV. 2018. Molecular mechanism of ATP hydrolysis in an ABC transporter. *ACS Central Sci.* 4:1334–43
109. Raper AT, Reed AJ, Suo ZC. 2018. Kinetic mechanism of DNA polymerases: contributions of conformational dynamics and a third divalent metal ion. *Chem. Rev.* 118:6000–25

110. Reinhardt CR, Konstantinovsky D, Soudackov AV, Hammes-Schiffer S. 2022. Kinetic model for reversible radical transfer in ribonucleotide reductase. *PNAS* 119:e2202022119
111. Reuter N, Dejaegere A, Maigret B, Karplus M. 2000. Frontier bonds in QM/MM methods: a comparison of different approaches. *J. Phys. Chem. A* 104:1720–35
112. Riccardi D, Schaefer P, Cui Q. 2005. pK_a calculations in solution and proteins with QM/MM free energy perturbation simulations. *J. Phys. Chem. B* 109:17715–33
113. Richter R, Foerster JM, Schelter I, Kümmel S. 2020. Self-interaction correction, electrostatic, and structural influences on time-dependent density functional theory excitations of bacteriochlorophylls from the light-harvesting complex 2. *J. Chem. Phys.* 153:2441–53
114. Rizzi A, Carloni P, Parrinello M. 2021. Targeted free energy perturbation revisited: accurate free energies from mapped reference potentials. *J. Phys. Chem. Lett.* 12:9449–54
115. Roston D, Cui Q. 2016. Substrate and transition state binding in alkaline phosphatase exhibited by computational analysis of isotope effects. *J. Am. Chem. Soc.* 138:11946–57
116. Roston D, Demapan D, Cui Q. 2019. Extensive free energy simulations identify water as the base in nucleotide addition by DNA polymerase. *PNAS* 116:25048–56
117. Roston D, Lu X, Fang D, Demapan D, Cui Q. 2018. Analysis of phosphoryl transfer enzymes with QM/MM free energy simulations. *Methods Enzymol.* 607:53–90
118. Rouhani S, Cartailler JP, Facciotti MT, Walian P, Needleman R, et al. 2001. Crystal structure of the D85S mutant of bacteriorhodopsin: model of an O-like photocycle intermediate. *J. Mol. Biol.* 313(3):615–28
119. Roux B, ed. 2011. *Molecular Machines*. Singapore: World Sci.
120. Samara NL, Yang W. 2018. Cation trafficking propels RNA hydrolysis. *Nat. Struct. Mol. Biol.* 25:715–21
121. Schulz CE, van Gastel M, Pantazis DA, Neese F. 2021. Converged structural and spectroscopic properties for refined QM/MM models of azurin. *Inorg. Chem.* 60:7399–412
122. Senn HM, Thiel W. 2009. QM/MM methods for biomolecular systems. *Angew. Chem. Int. Ed.* 48:1198–229
123. Shen L, Yang WT. 2018. Molecular dynamics simulations with quantum mechanics/molecular mechanics and adaptive neural networks. *J. Chem. Theory. Comput.* 14:1442–55
124. Sirohiwal A, Pantazis DA. 2021. Electrostatic profiling of photosynthetic pigments: implications for directed spectral tuning. *Phys. Chem. Chem. Phys.* 23(43):24677–84
125. Sivak DA, Crooks GE. 2012. Thermodynamic metrics and optimal paths. *Phys. Rev. Lett.* 108:190602
126. Sokolov M, Bold BM, Kranz JJ, Höfener S, Niehaus TA, Elstner M. 2021. Analytical time-dependent long-range corrected density functional tight binding (TD-LC-DFTB) gradients in DFTB+: implementation and benchmark for excited-state geometries and transition energies. *J. Chem. Theory Comput.* 17(4):2266–82
127. Son CY, Yethiraj A, Cui Q. 2017. Cavity hydration dynamics in cytochrome c oxidase and functional implications. *PNAS* 114:E8830–36
128. Song Y, Gunner MR. 2014. Halorhodopsin pumps Cl^- and bacteriorhodopsin pumps protons by a common mechanism that uses conserved electrostatic interactions. *PNAS* 111(46):16377–82
129. Stevens DR, Hammes-Schiffer S. 2018. Exploring the role of the third active site metal ion in DNA polymerase η with QM/MM free energy simulations. *J. Am. Chem. Soc.* 140:8965–69
130. Stoddart JF. 2017. Mechanically interlocked molecules (MIMs)—molecular shuttles, switches, and machines (Nobel Lecture). *Angew. Chem. Int. Ed.* 56:11094–125
131. Sun R, Sode O, Dama JF, Voth GA. 2017. Simulating protein mediated hydrolysis of ATP and other nucleoside triphosphates by combining QM/MM molecular dynamics with advances in metadynamics. *J. Chem. Theory Comput.* 13:2332–41
132. Swanson JMJ. 2022. Multiscale kinetic analysis of proteins. *Curr. Opin. Struct. Biol.* 72:169–75
133. Sweeney HL, Houdusse A. 2010. Structural and functional insights into the myosin motor mechanism. *Annu. Rev. Biophys.* 39:539–57
134. Thirumalai D, Hyeon C, Zhuravlev PI, Lorimer GH. 2019. Symmetry, rigidity, and allosteric signaling: from monomeric proteins to molecular machines. *Chem. Rev.* 119:6788–821
135. Tripathi R, Forbert H, Marx D. 2019. Settling the long-standing debate on the proton storage site of the prototype light-driven proton pump bacteriorhodopsin. *J. Phys. Chem. B* 123:9598–608

136. Tripathi R, Glaves R, Marx D. 2017. The GTPase hGBP1 converts GTP to GMP in two steps via proton shuttle mechanisms. *Chem. Sci.* 8:371–80
137. Troisi A. 2011. The speed limit for sequential charge hopping in molecular materials. *Org. Electron.* 12:1988–91
138. Unarta IC, Zhu LZ, Tse CKM, Cheung PPH, Yu J, Huang XH. 2018. Molecular mechanisms of RNA polymerase II transcription elongation elucidated by kinetic network models. *Curr. Opin. Struct. Biol.* 49:54–62
139. Valsson O, Tiwary P, Parrinello M. 2016. Enhancing important fluctuations: rare events and metadynamics from a conceptual viewpoint. *Annu. Rev. Phys. Chem.* 67:159–84
140. van Wonderen JH, Hall CR, Jiang X, Adamczyk K, Carof A, et al. 2019. Ultrafast light-driven electron transfer in a Ru(II)tris(bipyridine)-labeled multiheme cytochrome. *J. Am. Chem. Soc.* 141(38):15190–200
141. Vennelakanti V, Nazemi A, Mehmood R, Steeves AH, Kulik HJ. 2022. Harder, better, faster, stronger: large-scale QM and QM/MM for predictive modeling in enzymes and proteins. *Curr. Opin. Struct. Biol.* 72:9–17
142. Wang W, Cao SQ, Zhu LZ, Huang XH. 2018. Constructing Markov state models to elucidate the functional conformational changes of complex biomolecules. *WIREs Comput. Mol. Sci.* 8:e1343
143. Wanko M, Garavelli M, Bernardi F, Niehaus T, Frauenheim T, Elstner M. 2004. A global investigation of excited state surfaces within time-dependent density-functional response theory. *J. Chem. Phys.* 120:1674–92
144. Wanko M, Hoffmann M, Frähmcke J, Frauenheim T, Elstner M. 2008. Effect of polarization on the opsin shift in rhodopsins. 2. Empirical polarization models for proteins. *J. Phys. Chem. B* 112:11468–78
145. Wanko M, Hoffmann M, Frauenheim T, Elstner M. 2008. Effect of polarization on the opsin shift in rhodopsins. 1. A combined QM/QM/MM model for bacteriorhodopsin and pharaonis sensory rhodopsin II. *J. Phys. Chem. B* 112:11462–67
146. Wanko M, Hoffmann M, Strodel P, Koslowski A, Thiel W, et al. 2005. Calculating absorption shifts for retinal proteins: computational challenges. *J. Phys. Chem. B* 109:3606–15
147. Warshel A, Chu ZT. 2001. Nature of the surface crossing process in bacteriorhodopsin: computer simulations of the quantum dynamics of the primary photochemical event. *J. Phys. Chem. B* 105:9857–71
148. Warshel A, Levitt M. 1976. Theoretical studies of enzymic reactions—dielectric, electrostatic and steric stabilization of carbonium-ion in reaction of lysozyme. *J. Mol. Biol.* 103(2):227–49
149. Weber C, Cole DJ, O'Regan DD, Payne MC. 2014. Renormalization of myoglobin-ligand binding energetics by quantum many-body effects. *PNAS* 111:5790–95
150. Welke K, Watanabe HC, Wolter T, Gaus M, Elstner M. 2013. QM/MM simulations of vibrational spectra of bacteriorhodopsin and channelrhodopsin-2. *Phys. Chem. Chem. Phys.* 15(18):6651–59
151. Woiczikowski PB, Steinbrecher T, Kubař T, Elstner M. 2011. Nonadiabatic QM/MM simulations of fast charge transfer in *Escherichia coli* DNA photolyase. *J. Phys. Chem. B* 115(32):9846–63
152. Wolf S, Freier E, Cui Q, Gerwert K. 2015. Infrared spectral marker bands characterizing a transient water wire inside a hydrophobic membrane protein. *J. Chem. Phys.* 141:22D524
153. Wolter T, Elstner M, Fischer S, Smith JC, Bondar AN. 2015. Mechanism by which untwisting of retinal leads to productive bacteriorhodopsin photocycle states. *J. Phys. Chem. B* 119(6):2229–40
154. Wolter T, Welke K, Phatak P, Bondar AN, Elstner M. 2013. Excitation energies of a water-bridged twisted retinal structure in the bacteriorhodopsin proton pump: a theoretical investigation. *Phys. Chem. Chem. Phys.* 15(30):12582–90
155. Wu Q, Van Voorhis T. 2005. Direct optimization method to study constrained systems within density-functional theory. *Phys. Rev. A* 72:024502
156. Wu Q, Van Voorhis T. 2006. Constrained density functional theory and its application in long-range electron transfer. *J. Chem. Theory Comput.* 2(3):765–74
157. Wu Q, Van Voorhis T. 2006. Direct calculation of electron transfer parameters through constrained density functional theory. *J. Phys. Chem. A* 110(29):9212–18
158. Wu WJ, Yang W, Tsai MD. 2017. How DNA polymerases catalyze replication and repair with contrasting fidelity. *Nat. Rev. Chem.* 1:0068
159. Yagi K, Ito S, Sugita Y. 2021. Exploring the minimum-energy pathways and free-energy profiles of enzymatic reactions with QM/MM calculations. *J. Phys. Chem. B* 125:4701–13

160. Zhang LF, Wang H, E W. 2018. Reinforced dynamics for enhanced sampling in large atomic and molecular systems. *J. Chem. Phys.* 148:124113
161. Zhang Y. 2006. Pseudobond ab initio QM/MM approach and its applications to enzyme reactions. *Theor. Chem. Acc.* 116:43–50
162. Zheng YQ, Cui Q. 2017. Microscopic mechanisms that govern the titration response and pK_a values of buried residues in staphylococcal nuclease mutants. *Proteins* 85:268–81