JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

ARTICLE

pubs.acs.org/JACS

Proton Storage Site in Bacteriorhodopsin: New Insights from Quantum Mechanics/Molecular Mechanics Simulations of Microscopic pK_a and Infrared Spectra

Puja Goyal,[†,‡] Nilanjan Ghosh,[†,‡] Prasad Phatak,[§,‡] Maike Clemens,[§] Michael Gaus,[⊥] Marcus Elstner,[*,§,⊥] and Qiang Cui[*,†]

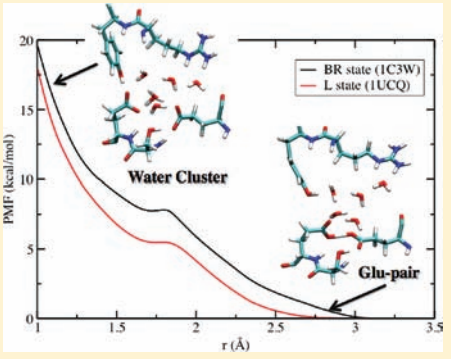
[†]Department of Chemistry and Theoretical Chemistry Institute, University of Wisconsin, Madison, 1101 University Avenue, Madison, Wisconsin 53706, United States

[§]Department of Physical and Theoretical Chemistry, Technische Universität Braunschweig, Hans-Sommer-Strasse 10, D-38106 Braunschweig, Germany

[⊥]Institute of Physical Chemistry, Karlsruhe Institute of Technology, Kaiserstrasse 12, 76131 Karlsruhe, Germany

Supporting Information

ABSTRACT: Identifying the group that acts as the proton storage/loading site is a challenging but important problem for understanding the mechanism of proton pumping in biomolecular proton pumps, such as bacteriorhodopsin (bR) and cytochrome *c* oxidase. Recent experimental studies of bR propelled the idea that the proton storage/release group (PRG) in bR is not an amino acid but a water cluster embedded in the protein. We argue that this idea is at odds with our knowledge of protein electrostatics, since invoking the water cluster as the PRG would require the protein to raise the pK_a of a hydronium by almost 11 pK_a units, which is difficult considering known cases of pK_a shifts in proteins. Our recent quantum mechanics/molecular mechanics (QM/MM) simulations suggested an alternative "intermolecular proton bond" model in which the stored proton is shared between two conserved Glu residues (194 and 204). Here we show that this model leads to microscopic pK_a values consistent with available experimental data and the functional requirement of a PRG. Extensive QM/MM simulations also show that, independent of a number of technical issues, such as the influence of QM region size, starting X-ray structure, and nuclear quantum effects, the "intermolecular proton bond" model is qualitatively consistent with available spectroscopic data. Potential of mean force calculations show explicitly that the stored proton strongly prefers the pair of Glu residues over the water cluster. The results and analyses help highlight the importance of considering protein electrostatics and provide arguments for why the "intermolecular proton bond" model is likely applicable to the PRG in biomolecular proton pumps in general.



1. INTRODUCTION

Proton pumps are central players in bioenergetics[1] and convert light or chemical (e.g., oxygen reduction) energy into the electrochemical concentration gradient of protons across a membrane. Although three-dimensional structures for several proton pumps have become available with crystallography, many key questions regarding their working mechanism remain unanswered due to the difficulty of directly observing protons in these complex macromolecules. For example, all proton pumps require a group that transiently stores a proton, which is released at a later stage of the catalytic cycle; such a group is referred to in the literature as a proton storage/release group (PRG) or a proton loading site (PLS). The precise location and chemical identity of the PRG/PLS are often difficult to resolve. In the arguably most intricate proton pump, cytochrome *c* oxidase,[2−4] for example, several proposals have been put forward for the identity of the PLS, which include an amino acid side chain,[5] a propionate group in a heme cofactor,[4] and some potentially strongly coupled titratable groups.[3,6]

Bacteriorhodopsin (bR), a well-studied light-driven proton pump,[7] is another system for which the identity of the PRG has been hotly debated over the years. Under physiological conditions, following the first proton transfer from the Schiff base to Asp85, a proton is released to the extracellular side during the L-to-M transition. This released proton does not originate from Asp85[8] but from a residue initially termed as XH, which in recent literature[9] has been referred to as the PRG. The initial suggestions for the PRG were the highly conserved Glu204 and/or Glu194, based on FTIR studies of Glu204Gln and Glu204Asp mutants[10,11] and pK_a analysis for Asp85;[12] indeed, a short distance (less than 3 Å, Figure 1) between the Glu204 and Glu194 side chains for both the bR ground- and L-state crystal

Received: February 28, 2011
Published: July 15, 2011

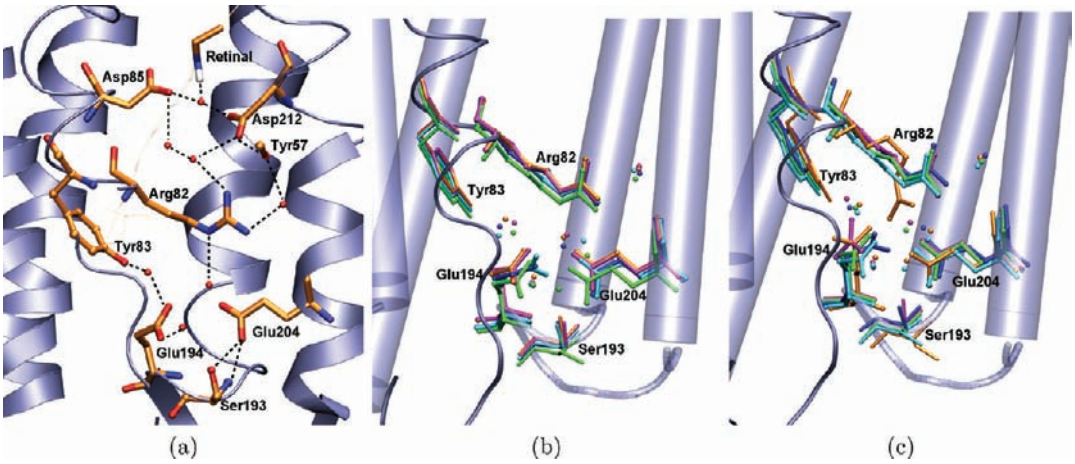


Figure 1. Structural features of the proposed proton release group region[10,11,15,16] in X-ray structures of wild-type bacteriorhodopsin. (a) Spatial location of key residues and water molecules in the PRG region based on a ground-state bR structure (PDB code: 1C3W). The retinal chromophore is shown in the transparent line form; water molecules resolved in the X-ray structure are shown as spheres. (b) Comparison of various X-ray structures for bR in the L state for the PRG region: 1E0P in orange, 1O0A in green, 1UCQ in blue, 1VJM in cyan, and 2NTW in magenta. (c) Comparison of various X-ray structures for bR in the ground state for the PRG region: 1C3W in orange, 1QHJ in green, 1KGB in blue, 1IW6 in cyan, and 1C8R in magenta. The structures are generated using VMD (Humphrey, W.; Dalke, A.; Schulten, K. *J. Mol. Graphics* **1996**, *14*, 33−38).

structures is consistent with the idea of trapping a proton between these two residues.[13,14] More recent analysis of the IR spectra for wild-type (WT) bR and a collection of mutants,[15,16] however, led to the refutal of Glu194/204 as possible candidates for the PRG because the expected spectral shifts in carboxylate groups upon Glu deprotonation were not observed for the L-to-M transition in WT bR. Instead, the authors noted the decay of a continuum band near 2000 cm^{-1} during the formation of the M state, and the continuum spectral feature was reminiscent of the IR signature of protonated water clusters.[17] Accordingly, it was proposed that the proton is stored on a water cluster trapped in a region surrounded by Glu194, Glu204, and Arg82.[15,16]

Although the idea of invoking a water cluster as the PRG was novel and received indirect support from several previous computational studies,[18,19] we note that the underlying energetic basis is in dire contrast with the general knowledge about protein electrostatics. For a motif to serve as an effective PRG in a proton pump, it needs to maintain a fairly high pK_a (>7) in the relevant functional state. The pK_a of a hydronium in solution is ∼−1.7, thus the water cluster model requires the protein to raise the pK_a of an Eigen or Zundel ion by as much as 9 pK_a units (or 11.4 pK_a units using the measured pK_a of 9.7 for the PRG in bR[9,20,21]), which is very difficult given known examples of pK_a shifts in proteins.[22,23] With a single acidic residue such as Glu or Asp, which has a standard pK_a ∼4.0, shifting the pK_a to the measured value of 9.7 for the PRG is very demanding, considering the fact that the PRG needs to be in a fairly polar environment so that the proton can be picked up/released in the relevant phase of the functional cycle; for example, even putting Glu/Asp into the hydrophobic core of staphylococcal nuclease typically raises the pK_a by less than 4 pK_a units![23]

These considerations led us to investigate the PRG issue using extensive quantum mechanics/molecular mechanics (QM/MM) simulations[24] with an efficient approximate density functional theory (SCC-DFTB[25]) or the popular B3LYP method as the QM level. Both SCC-DFTB/MM simulations[26,27] at the nanosecond scale and B3LYP/MM simulations at the picosecond scale pointed to a consistent molecular picture in which the PRG consists of both Glu194/204 residues with the stored proton delocalized between the corresponding side chains. Although this "intermolecular proton bond" [28] model of the PRG led to IR signatures and structural features consistent with available experimental observations (see discussions in ref 24), the discussion above highlights that it is crucial to directly address the energetic basis of the model by explicitly computing the microscopic pK_a of the putative PRG, which is an important task for the current work. Along this line, we note that the previous continuum electrostatics calculations that invoked a titratable Zundel ion ($H_5O_2^+$) between the two Glu residues (Glu 194, Glu 204) did not reproduce the experimental pK_a for the PRG. Instead of the value of 9.7,[9,20,21] the calculations predicted that the excess proton is held up to a pH of 15 or more,[18] which was rationalized by the lack of side-chain sampling during continuum electrostatics calculations. Therefore, a microscopic pK_a calculation using explicit free energy simulation is required. Moreover, since the two Glu residues and nearby water molecules are strongly coupled, a QM/MM potential is also necessary in the microscopic pK_a calculation, which has become possible in recent years.[27,29,30]

In addition to the issue of microscopic pK_a, additional features of the PRG have been brought to light by new experimental studies. First, time-resolved FTIR analysis[31] showed that, at pH 5, when proton release is expected to be slowed, the protonation of Glu194/204 is visible (∼1706−1720 cm^{-1}) during the rise of the M state; since no significant spectral features in this region are observed at pH 7, it was concluded that neither Glu194 nor 204 is protonated in the ground state. Moreover, the ground-state crystal structures of several mutants (Glu194Asp, Glu204Asp) were solved, which showed different positions for water molecules and Arg82 as compared to the WT. These structures were interpreted to reflect different stages of the ground-state to M-like transition that ultimately leads to the release of the stored proton. In another interesting study, Kandori and co-workers observed that the characteristic continuum band at ∼1800−2000 cm^{-1} disappeared at low temperature (∼230 K),[32] which was explained by the conjecture that proton release at low temperature is inhibited in the M state (thus, the continuum feature cancels out between the M- and ground-state spectra). Whether this is the only interpretation needs to be explored at the molecular level.

These recent experimental findings have been interpreted mostly in terms of a protonated water cluster between the Glu194/204 pair and Arg82, which was thought to form the PRG.[16] The model with the shared proton between the Glu194/204 pair, on the other hand, has been shown to reproduce most spectroscopic data and is also in good agreement with X-ray structural data concerning the distance between the Glu's.[24] However, recent classical molecular dynamics (MD) simulations with a classical hydronium ion (H_3O^+, i.e., not dissociable) stuck between Glu194 and Glu204 showed that the resulting average distance between Glu194 and Glu204 is only slightly (by 0.7 Å) longer than that in the ground-state crystal structure, suggesting that the protonated water cluster model is not in severe contrast to X-ray crystallography data.

Therefore, both models are, to a varying degree, in agreement with the spectroscopic and structural data (though it is unclear whether a hydronium stuck between the pair of Glu's leads to a continuum band), and other factors have to be investigated to shed additional light. In this work, we show that (i) the energy difference for the two models is quite large, favoring the shared proton model, which explains why the proton leaves the water cluster very quickly during QM/MM simulations; (ii) the large pK_a value of the PRG can be successfully computed using the shared proton model, while this is questionable for the protonated water cluster model; and (iii) the shared proton model appears to be qualitatively consistent with the observed temperature and pH dependence of the continuum band. As emphasized above, distinguishing the protonated water cluster and "intermolecular proton bond" models is not a simple semantic issue, and the two models, despite their seemingly small structural differences, have significantly different energetic properties.

In the following, we first summarize computational methods and models; we take this opportunity to also address a number of technical issues, such as the influence of starting crystal structure for simulation for a given state, the size of the QM region, and nuclear quantum effects. In section 3, we present the results of simulations and discuss their significance in the context of PRG identity. Finally, we draw a few conclusions in section 4.

2. COMPUTATIONAL METHODS

2.1. Gas-Phase Models. In this work we use the approximate DFT method SCC-DFTB[25] and one of its recent extensions.[33,34] SCC-DFTB is derived from DFT by a second-order expansion of the DFT total energy, and it has been shown that this expansion has to be extended to third-order when molecules with localized charges are involved.[34,35] This concerns in particular proton affinities of acids, as studied here. To enable a direct comparison of the second- and third-order SCC-DFTB method[25,27,33] used in this work to high-level *ab initio* results, we study several simple gas-phase models for the PRG. In particular, we aim to compare proton affinities of these models with the excess proton either on water or shared between the Glu side chains; it is important that SCC-DFTB captures the correct trend as compared to high-level DFT and *ab initio* calculations. The smaller model (**M1**) contains the side chains of Glu194/204, three active-site water molecules, and an excess proton (for the protonated state); the larger model (**M2R**) contains in addition the side chain of Arg82; and the largest model (**MLR**) further includes water molecules and other small molecules (e.g., methanol) to mimic the hydrogen-bonding environment of the Glu pair in bR. For the protonated state of each model, the excess proton is either delocalized between the Glu side chains or constrained to be on a water molecule that bridges the two Glu side chains. Some atoms in the model Glu or Arg are fixed in space so that the relative orientations of the model "side chains" mimic that in the crystal structure (based on 1UCQ[36] for the protonated state and on 2NTW[37] for the deprotonated state). To avoid proton transfers in the gas-phase environment (e.g., from "Arg82" to "Glu204"), the N−H bond distances in the model Arg82 and O−H bonds in the hydronium are constrained to their values optimized for isolated species. Other details of the SCC-DFTB and *ab initio* calculations are summarized in the footnote of Table 1; optimized structures are included in the Supporting Information.

Table 1. Proton Affinity (kcal/mol) of Active-Site Cluster Models from Different Levels of Theory[a]

			SCC-DFTB[d]	
model[b]	B3LYP BS1[c]	MP2 BS2[c]	second-order	DFTB3-diag
M1:Glu_pair	422.6 (−0.8)	423.4	440.6 (17.2)	431.5 (8.1)
M1:hydronium	411.1 (−0.7)	411.8	422.6 (10.8)	415.2 (3.4)
M2R:Glu_pair	356.1 (2.1)	354.0	366.3 (12.3)	355.6 (1.6)
M2R:hydronium	336.3 (0.4)	335.9	345.5 (9.6)	335.5 (−0.4)
MLR:Glu_pair	335.9 (5.6)	330.3	347.5 (17.2)	338.5 (8.2)
MLR:hydronium	323.3 (1.3)	322.0	332.8 (10.8)	326.0 (4.0)

[a] Values without parentheses are calculated proton affinities (zero-point energy and thermal corrections not included); those with parentheses are deviations from the MP2 result. [b] The smaller model (**M1**) contains the side chains of Glu194/204, three active-site water molecules, and an excess proton (for the protonated state); the larger model (**M2R**) contains in addition the side chain of Arg82; and the largest model (**MLR**) further includes water molecules and other small molecules (e.g., methanol) to mimic the hydrogen-bonding environment of the Glu pair in bR. For each model, the label "Glu_pair" indicates that the excess proton is delocalized between the Glu side chains in the protonated state, while "hydronium" indicates that the excess proton is constrained to be on a water molecule that bridges the two Glu side chain (for optimized structures, see Supporting Information). The structures have been fully optimized at the respective level of theory, except for the following: for **M1**, MP2 uses B3LYP/BS1 optimized structures; for **M2R**/**MLR**, both MP2 and B3LYP use B3LYP/6-31+G(d,p) optimized structures, and MP2 energies use BS1. [c] BS1 is 6-311++G(2d,2p), and BS2 is the "G3Large" basis set used in G3 calculations (see, for example, http://chemistry.anl.gov/compmat/g3theory.htm). [d] Second-order indicates the standard SCC-DFTB approach;[25] DFTB3-diag indicates the third-order extension of SCC-DFTB,[34] including the on-site contribution from the fitted Hubbard derivative and the modification in γ_{XH} ("set 5" in ref 33).

2.2. Stochastic Boundary Setup and Initial Protein Structure. Equilibrium simulations of the PRG region with a stochastic boundary condition and explicit solvent are carried out using the CHARMM program (c32a2 version).[38] Starting from the crystal structure (see below), hydrogen atoms are added with HBUILD.[39] All basic and acidic amino acids are kept in their physiological protonation state, except Asp96 and Asp115, which are known to be protonated in the ground and L states.[9] The protein atoms are described with the all-atom CHARMM force field for proteins,[40] and the water molecules are described with the TIP3P model.[41] The all-*trans*-retinal with protonated Schiff base is used for the ground-state simulations, while a 13-*cis* conformation is used for the L-state simulations; they are treated using force field parameters taken from refs 42 and 43. Water molecules are added to the system following the standard protocol of superimposing the system with a water sphere of 25 Å radius (see section 2.3 for additional discussions). In all cases, the system is partitioned into a 22 Å inner region centered at the N_ε atom of Arg82, while the rest of the protein is treated as the outer region within the framework of the generalized solvent boundary potential (GSBP) approach.[44] Newton's

equations of motion are solved for the MD region (within 18 Å), and Langevin equations of motion are solved for the buffer region (18−22 Å) with a temperature bath of 300 K.[45] All water molecules in the inner region are subjected to a weak GEO-type restraining potential to keep them inside the inner sphere with the MMFP module of CHARMM. The GEO restraining potential is in the form of a quartic polynomial on each oxygen atom in water: $k\Delta^2(\Delta^2 - V_P)$, where $\Delta = r - r_{off}$; k is the restraining quartic force constant (0.5 kcal/(mol·Å^4)), r is the distance of the oxygen from the center of the simulation sphere, r_{off} is the cutoff distance (22.0 − 1.5 = 20.5 Å) below which the GEO restraint is set to zero, and V_P is an offset value taken to be 2.25 Å^2. These parameters lead to a restraining potential on water that smoothly turns on at 20.5 Å, reaches a well at 21.5 Å with a depth of −0.625 kcal/mol, and then quickly rises to be repulsive beyond 22.0 Å. All protein atoms in the buffer region are harmonically restrained with force constants determined directly from the *B*-factors in the PDB file.[45] Langevin atoms are updated heuristically during the simulation to consistently treat protein groups and water molecules that may switch regions during the simulation. Nonbonded interactions within the inner sphere are treated with an extended electrostatics model, in which groups beyond 12 Å interact as multipoles.[46] The entire system is heated gradually to 300 K and equilibrated for ∼150 ps prior to the production simulations.

To account for the electrostatics between the inner- and outer-region atoms and the effect of bulk solvation, GSBP is used.[44] The static field due to outer-region atoms, ϕ_s^o, and the reaction field matrix, **M**, are evaluated using Poisson−Boltzmann (PB) calculations using a focusing scheme that places a 56 Å cube of fine grid (0.4 Å) into a larger 132 Å cube of coarse grid (1.2 Å). The inner-region charge density is expressed using the first 20th-order spherical harmonics with a total of 400 basis functions. The membrane environment is treated implicitly using a dielectric model with dielectric constant of 2.0 and membrane thickness set to 35 Å. The optimized radii reported by Roux and Nina[47,48] are adopted to define the solvent−solute dielectric boundary. Dielectric constants of 1.0 and 80.0 are used for protein and solvent, respectively.

As starting structures for the simulations, the X-ray structure with PDB code 1C3W[49] (1.55 Å resolution) is used for the ground-state simulations, while 1UCQ[36] (2.4 Å resolution) and 2NTW[37] (1.53 Å resolution) are used for the L-state simulations. In our recent study[24] of the L state, the 1UCQ structure was selected because it has nearly identical structural features for the region of interest compared to other X-ray structures that have been unambiguously assigned to the L kinetic state. The most distinct feature in the 2NTW structure is the much expanded distance between Glu194 and Glu204 (3.8 vs 2.6 Å in 1UCQ, see Figure 1). However, the authors themselves stated that "some of the changes resemble those in M2, although smaller in magnitude";[37] in other words, the 2NTW structure exhibits structural signatures of L as well as plausibly early M state. Nevertheless, the most stringent test of the "intermolecular proton bond" model would be to start the simulation using the 2NTW structure with the protonation pattern for Asp85 and retinal consistent with the L state and observe whether the same proton-mediated configuration of the Glu194/204 pair is formed; this is an important motivation of the current study.

It is worth stressing that all the crystal structures employed for the ground- and L-state simulations in this study have a cytoplasmic orientation of Arg82. In a separate work,[50] we have shown that the Arg82 side chain changes its orientation from cytoplasmic to extracellular upon deprotonation of the Schiff base and concurrent protonation of Asp85 in the early M state. This destroys the pentagonal hydrogen-bonded network above Arg82, with two water molecules moving toward Asp212 and Tyr57, and with Asp85 and Asp212 getting closer to each other. The extracellular rotation of Arg82 also causes Glu194 to reorient toward Tyr83, thus increasing the separation between Glu194 and Glu204, a key reason for the drop of the PRG's pK_a as illustrated below. Arg82 movement also seems to allow additional water molecules from the bulk to enter the PRG region. For additional details, please refer to ref 50.

2.3. Water Molecules in the Active Site. The number of waters in the PRG region of bR is a contentious issue. Although it is generally agreed that at least three waters are present in the PRG region based on the ground- and L-state X-ray structures, the QM/MM-CPMD studies by Marx et al.[19] added a fourth water to the ground-state (1C3W) structure based on the observation from a different X-ray structure (1QHJ[51]). In all of our WT simulations initiated with three water molecules in the PRG region,[24] two "bulk" water molecules were consistently observed to enter this region during the nanosecond simulations, suggesting that the PRG region is flexible enough to accommodate additional waters.

To better understand the stability of extra waters, Grand Canonical Monte Carlo (GCMC) simulations[52] are carried out for both the L- and ground-state bR structures. All water molecules within 10 Å of Arg82 are subjected to GCMC equilibration while keeping the protein fixed. An important technical issue related to GCMC simulations is the choice of the excess chemical potential for water (μ_{ex}), since it is not clear what is the most appropriate value of μ_{ex} with a specific protein force field. Following the work of Roux et al. on the KcsA channel,[52] we chose a value of $\mu_{ex} = -5.8$ kcal/mol, although a range of μ_{ex} values from −6.5 to −5.4 kcal/mol have been reported for the TIP3P water model.[41] All five waters in the PRG region are found to be stable during GCMC equilibration. In total, GCMC simulations predict eight water molecules within the 10 Å region around Arg82 in the L state, six of which have been resolved in the X-ray structure (for both 2NTW and 1UCQ); for the ground state, nine water molecules are found to be stable during GCMC simulations within the 10 Å region around Arg82 in the 1C3W structure, seven of which are resolved in the X-ray structure. The difference between 1C3W and 2NTW/1UCQ is that an extra water molecule between the carboxylic groups of Asp85 and Asp212 is resolved in 1C3W. Collectively, these GCMC simulations suggest that five waters can reside in the PRG region of bR. For the purpose of our study, it was observed that initiating SCC-DFTB/MM simulations with both three and five waters as part of the QM region leads to similar IR spectra in the region of interest, suggesting that the delocalization of the stored proton and the corresponding IR spectra are not significantly perturbed by the number of waters in the PRG.

2.4. QM/MM Setup. In our recent study,[24] results from SCC-DFTB[25]/MM[26,27,53] and B3LYP/6-31+G(d,p)/MM were compared, and very similar trends were observed. In this work, we focus on SCC-DFTB/MM simulations since long (nanosecond scale) simulations are required to observe potential structural transitions in the key Glu pair starting from the 2NTW structure; extensive sampling is also important for converging IR spectra calculations. As in our previous study, the third-order extension[33] of SCC-DFTB[25] plus the modification of the γ function for atom pairs involving hydrogen, which were found to be important for improving proton affinity and hydrogen-bonding interactions, respectively, are used as the QM level for both the ground- and L-state simulations. As shown by benchmark calculations using B3LYP and MP2 as references for active-site models (for details, see Supporting Information and Table 1), SCC-DFTB is appropriate for the current problem.

Simulations with different regions treated as QM are compared to explore the sensitivity of the result (e.g., distribution of the excess proton and IR spectra) to the size of the QM region. Unless explicitly stated, the simulations use the same QM region as in our recent study,[24] which includes the side chains of both Glu194 and Glu204 along with five water molecules in the PRG region. A substantially larger QM region (∼170 atoms) is also tested, which further expands to include all side chains within hydrogen-bonding distance of the pair of Glu residues and the five water molecules (Arg82, Tyr83, Asp85, Ser193, Asp212, Lys216, and the retinal plus another four water molecules). Simulations with the smaller

QM region are carried out for 2 ns, and at least three independent trajectories are carried out for each setup. For the larger QM region setup, which is much more expensive, simulations are carried out for 400 ps. To explore the temperature dependence, simulations are carried out at 300, 230, and 170 K with the smaller QM region and the L-state structure; given the similar results observed for the ground- and L-state simulations (see below), we expect similar behaviors for the ground state.

The QM/MM boundary (between the Cα and Cβ of the QM amino acids) is treated using link-atoms with the divided frontier charge scheme;[54] previous systematic benchmark calculations[54] suggest that this scheme is satisfactory, especially when the MM atom at the QM/MM frontier has very small charges, which is the case here (Cα has a charge of 0.07).

2.5. Microscopic pK_a Calculations. The pK_a calculations for the Glu194/204 pair are carried out using a thermodynamic integration approach within the dual topology single coordinate (DTSC) scheme.[30,55] As discussed in previous work, the dominant contribution to the free energy of deprotonation relative to a well-defined solution reference (acetic acid in the current work) is from the electrostatic free energy change ($\Delta G_{\mathrm{E\cdot PRG(H/D)}}$) associated with converting the acidic proton to a dummy atom (D) that interacts with the environment only through van der Waals and bonded terms, i.e., the transformation from E·PRG−H to E·PRG^{-1} −D, where E represents the protein. The corresponding free energy derivative is given by

$$\frac{\partial \Delta G_{\mathrm{E\cdot PRG(H/D)}}}{\partial \lambda} = \langle U_{\mathrm{elec}}^{\mathrm{E\cdot PRG^{-1}-D}}(\mathbf{X}_{\mathrm{E\cdot PRG(H/D)}}) - U_{\mathrm{elec}}^{\mathrm{E\cdot PRG-H}}(\mathbf{X}_{\mathrm{E\cdot PRG(H/D)}})\rangle_{\lambda} \quad (1)$$

which represents the QM/MM energy difference averaged for a specific coupling parameter λ using the same set of coordinates ($\mathbf{X}_{\mathrm{E\cdot PRGH}}$) for both protonation states, E·PRG−H and E·PRG^{-1}−D. The total electrostatic free energy contribution ($\Delta G_{\mathrm{E\cdot PRG(H/D)}}$) is determined by integrating the converged free energy derivatives ($\partial G/\partial \lambda$) over λ from 0 to 1. Specifically, five λ windows are used (0, 0.25, 0.5, 0.75, and 1.0). Since $\Delta G_{\mathrm{E\cdot PRG(H/D)}}$ is electrostatic in nature, the free energy derivatives depend very linearly on λ, despite the structural variations observed in the PRG region as λ varies (see discussions in section 3). To carefully monitor the statistical and sampling errors and reproducibility of the microscopic pK_a simulations, multiple independent trajectories have been carried out. For each set of pK_a simulations, the average free energy derivative value for a particular λ is determined using a block averaging scheme.[56] The amount of simulation identified as the equilibrating phase, the number/size of blocks for the equilibrated data, the statistical average, and the error for different λ values are summarized in Table S1 in the Supporting Information. With the similar DTSC-TI- and QM/MM-GSBP-based protocols, the microscopic pK_a value for Asp85 in bR has been successfully calculated in a recent study[57] as an important validation.

The QM region for the pK_a calculations is the same as that for the spectra calculations; previous studies on several systems[30,57−59] indicate that a QM/MM description for the interaction between a titratable group and its environment provides a good approximation, especially for pK_a shift calculations. An important modification here to the original DTSC-TI scheme[55] concerns the treatment of the bonded term for the acidic proton. In previous studies,[30,55] the bond between the acidic proton and the heavy atom is constrained using SHAKE.[60] In the current case, however, to allow the delocalization of the acidic proton between the Glu pair, a soft harmonic term (with a force constant of 25 kcal/(mol·Å^2) and an equilibrium distance of 0.96 Å) is applied between the acidic proton and $O_{\varepsilon 2}$ of Glu204. Our previous simulations found that the proton has a higher distribution near Glu204 than Glu194 (also see below). Moreover, the potential of mean force (PMF) results below show that the free energy of localizing the proton to one of the Glu residues is only slightly larger than 1.0 kcal/mol at 300 K. Therefore, the computed pK_a is not expected to be very sensitive to the precise value of the harmonic potential.

To further improve the quantitative accuracy of the computed pK_a values, we have carried out additional single-point B3LYP/6-311+G(d,p)/MM calculations using snapshots from the trajectories collected at the SCC-DFTB/MM level; 500 structures are used for the two end-states (λ = 0, 1), and the correction over the SCC-DFTB/MM free energy results is computed on the basis of a second-order cumulant expansion.[30,57] These single-point corrections account for two sources of systematic errors associated with SCC-DFTB-based calculations: proton affinity and QM/MM interactions. As shown in previous studies of titratable groups in the interior of a protein,[30,57] the magnitude of B3LYP/MM correction can be substantial when the titratable group has a very different local environment compared to the bulk solution, or when the titratable group is chemically distinct from the solution reference. Otherwise, such as for the microscopic pK_a calculation for Asp85 in bR, the magnitude of correction is fairly small, on the order of 1 kcal/mol.[57] In the current case, we expect that B3LYP/MM mainly corrects for the relative proton affinity errors between an acetic acid and a protonated Glu pair at the SCC-DFTB level.

In addition to the ground- and L-state simulations described above, we have also carried out microscopic pK_a calculations for a model of the "early M state". This structure was generated in our recent study of the movement of Arg82 in bR by changing the protonation states of the Schiff base and Asp85 to neutral in the K-state structure 1M0K.[61] With this structure, 20 MD simulations were performed, and six of them showed an extracellular rotation of Arg82. This reorientation was coupled to the breakage of the hydrogen-bonded network at the extracellular side of the Schiff base and happened within 300 ps after the breakage was observed. The Arg82 movement itself took about 50 ps to complete. Snapshots taken from one set of simulation where Arg82 rotated were used to set up the pK_a calculations. In principle, other crystal structures (e.g., those for the M or L states) can also be used as the template. Here we used the K state as the template to be sure that the conformation of the protein best resembles the conformation immediately after the first proton transfer. For additional discussions, see ref 50.

2.6. IR Spectra from Classical Molecular Dynamics and Ring Polymer Molecular Dynamics Calculations. For the various bR systems, the IR spectrum is computed by the Fourier transform of the classical dipole autocorrelation function (FT-DAC),[62,63]

$$I_{\mathrm{cl}} = \frac{1}{2\pi}\int_{-\infty}^{\infty} \mathrm{d}t\, \exp[-i\omega t]\langle \vec{\mu}(0)\cdot\vec{\mu}(t)\rangle \quad (2)$$

Within this framework, the absorption coefficient is given by

$$\alpha_{\mathrm{qc}} = \left[\frac{4\pi^2\omega}{3V\hbar c n(\omega)}\right](1-\exp(-\beta\hbar\omega))\mathrm{Q}_{\mathrm{qc}}(\omega)I_{\mathrm{cl}}(\omega) \quad (3)$$

where V is the sample volume, c the speed of light, and $n(\omega) \approx 1$ the refractive index of the medium. To partially account for the quantum effects due to nuclear motion (e.g., the delocalized proton between Glu194 and Glu204), an *ad hoc* harmonic quantum correction factor (Q_{qc})[64,65] is used,

$$\mathrm{Q}_{\mathrm{qc}} = \frac{\beta\hbar\omega}{[1-\exp(-\beta\hbar\omega)]} \quad (4)$$

Among the various proposed quantum correction factors,[65−67] the harmonic correction factor has been shown[64] to be effective for improving the IR intensities (but not the frequencies) of floppy molecules with large anharmonicities.

With this correction, the theoretical absorption coefficient simplifies to

$$\alpha_{\mathrm{qc}}(\omega) = \left[\frac{4\pi^2\omega}{3V\hbar c n(\omega)}\right]\beta\hbar\omega I_{\mathrm{cl}}(\omega) \quad (5)$$

To construct the IR spectra, SCC-DFTB/MM simulations are carried out on the nanosecond scale using a time step of 0.5 fs to properly

sample the high-frequency vibrational modes in the system. The choice is based on careful benchmark studies on protonated water clusters[68] where a 0.5 fs time step was found to be sufficient to produce reliable IR results for the ~2000 cm^{-1} region. No SHAKE constraint is applied to any bond. The coordinates and Mulliken charges of QM atoms are saved at each step, from which the dipole moment and dipole autocorrelation function are calculated. The IR spectrum for the QM region is computed by the Fourier transform of the dipole autocorrelation function collected from SCC-DFTB/MM trajectories with a Blackmann filter to minimize noise; multiple independent simulations are carried out for statistical significance. In Fourier transform analysis, 8192 points in the time correlation function are included, which results in a frequency resolution of about 8 cm^{-1} in the power spectrum.

An important technical issue for the calculation of the dipole autocorrelation function concerns the choice of the coordinate origin because the QM region has a net charge of −1; it is well known that the dipole moment is origin-dependent for systems with a net charge. In the context of IR spectra calculations, we expect that the origin dependence is small since the spectral region of interest is at fairly high frequency (~2000 cm^{-1}). This is because a coordinate transformation would lead to a constant shift in the dipole moment ($-Q\Delta\mathbf{R}$, where Q is the total charge and $\Delta\mathbf{R}$ is the displacement vector for the coordinate transformation); upon Fourier transform of the dipole autocorrelation function, the effect on the calculated IR spectra is small, except for the spectral region that corresponds to the frequency of the coordinate origin fluctuation. We show this explicitly in the Supporting Information by comparing the IR spectra from SCC-DFTB/MM simulations using three different coordinate origins: the center of charge (COC), the center of mass (COM) for the QM atoms, and the geometric center (GO) of the entire system; they give very consistent results.

Since the delocalization of the stored proton between Glu194 and Glu204 is important to the continuum band in the IR spectra, it is of interest to examine the influence of nuclear quantum effects. To this end, we have carried out ring polymer molecular dynamics (RPMD)[69,70] on a model system that contains two propionates bonded with a delocalized proton. Detailed discussions of the computational setup and results are included in the Supporting Information. The key point is that including nuclear quantum effects has a relatively minor impact on the continuum band at both temperatures of interest (300 and 170 K). Therefore, although it would be straightforward to carry out QM/MM RPMD for bR, we do not find it compelling to do so.

2.7. Potential of Mean Force (PMF) Calculations. Although one can compute the microscopic pK_a for a protonated water cluster model to compare with the result for the Glu pair model, it is more direct to compute the PMF for the proton transfer from the protonated water cluster to the Glu pair. We do not include nuclear quantum effects in such PMF calculations; although a quantum PMF can be computed using, for example, path-integral simulations,[71] the effect of proton quantization is partially (fortuitously) reflected because proton-transfer barriers are underestimated by popular DFT methods,[72] including SCC-DFTB.[26] The PMF calculations are done for both the ground- and L-state bR, and the state with the proton on the water cluster mimics that studied in ref 31, i.e., with a hydronium bridging two Glu side chains. The reaction coordinate, r, is taken to be the distance of the stored proton from the oxygen atom of the hydronium, and it is sampled in the range of 1.0−3.1 Å employing nine windows and force constants ranging from 50 to 300 kcal/mol·Å^2, ensuring sufficient overlap between the windows. Each window consists of 200 ps of equilibration and 900 ps of data collection. The statistics of r are converted into a PMF using the WHAM algorithm.[73] To further confirm that the results do not depend sensitively on the choice of the QM region, another set of calculations is carried out (for both ground and L states) with a larger QM region which further includes the side chains of Tyr83 and Ser193; this choice is made since these two side chains directly interact with the side chains of Glu194/204.

To characterize how temperature impacts proton delocalization between Glu194 and Glu204, PMF calculations are also performed for the L state at different temperatures; since calculations in this work show that the ground and L states are very similar in the PRG properties, the delocalization PMF calculations are not repeated for the ground state. The reaction coordinate in these PMF calculations is the asymmetric stretch coordinate (δ) involving the shared (stored) proton and side-chain oxygen on Glu194/204; i.e., $\delta = r_{O_{E194}H} - r_{O_{E204}H}$, where $r_{O_{E194/E204}H}$ is the shortest distance between the shared proton and Glu194/204. Umbrella sampling[74] simulations are used to sample δ in the range from −0.8 to 0.8 Å with an increment of 0.2 Å and a force constant of 50 kcal/(mol·Å^2). For each temperature, nine windows are used, with each window consisting of 200 ps of equilibration and 600 ps of data collection.

As explained below, we find that the proton PMF is influenced by the orientation of the Ser193 side chain, which can either hydrogen bond to Glu204 or the main chain of Pro77. The preference to these orientations appears to be temperature dependent. Therefore, PMF is also calculated for the orientation of Ser193 side chain at different temperatures, using the reaction coordinate, $\delta' = r_{O_{P77}H_{S193}} - r_{C_{E204}H_{S193}}$, where O_{P77} refers to the backbone carbonyl oxygen of Pro77, H_{S193} the hydroxyl hydrogen of Ser193 side chain and C_{E204} the carboxylate carbon of Glu204 side chain. δ' is sampled in the range from −4.0 to 2.4 Å with an increment of 0.4 Å and a force constant of 50 kcal/(mol·Å^2). For each temperature, 17 windows are used, with each window consisting of 150 ps of equilibration and production runs that range from 600 ps to 1.2 ns (longer simulations for lower temperatures).

3. RESULTS AND DISCUSSION

In this section, we first briefly discuss a pair of gas-phase models for the PRG to establish the accuracy of SCC-DFTB for the problem of interest, and then we discuss the impact of several technical issues on the results for the protein simulations (structure and IR spectra of the PRG), which include starting crystal structure and the size of the QM region; discussions on the nuclear quantum effects are included in the Supporting Information. Next, we discuss microscopic pK_a calculations for the PRG and the proton-transfer PMF between the water cluster and the Glu pair. Finally, we present simulation results for the temperature dependence and compare to recent experimental studies.

3.1. Gas-Phase Models. As shown in Table 1, the proton affinity (PA) is substantially higher for the Glu pair than the water cluster; without Arg (i.e., **M1**), the difference in PA is ~11 kcal/mol at the MP2 level, it increases to ~18 kcal/mol with **M2R**, and the value is ~8 kcal/mol for the largest model, **MLR**. At the SCC-DFTB levels, the absolute PA values depend on the parametrization, although the qualitative trend is properly captured with both parametrizations tested here. For example, with the standard second-order SCC-DFTB,[25] the absolute PAs are too large, especially for the Glu pair with **M1**; at the qualitative level, on the other hand, the trend is consistent with MP2 and B3LYP calculations. When third-order terms are included, as discussed in previous work,[27,33,35] the PAs are much improved; for model **M2R**, even the absolute errors are at the same scale as B3LYP as compared to MP2. The errors in absolute PAs appear larger in **MLR**, although we note that for large cluster models, small differences in both the strength and geometry of non-covalent interactions between the many components of the model tend to accumulate and complicate the comparison between different methods. Nevertheless, we note that the SCC-DFTB results with the third-order extension are close to the B3LYP values, which also have considerable difference as compared to MP2. Therefore, missing dispersion interactions in SCC-DFTB

and popular functionals may make a notable contribution for large QM models; indeed, calculations using functionals with improved descriptions of dispersion (e.g., M06[75] and B97D[76]) give values closer to MP2 (data not included). However, results for the gas-phase models clearly indicate that the SCC-DFTB approach we use here is adequate for distinguishing the relative stability of different PRG models; most importantly, as shown below by the PMF calculations, the relative stability of different PRG models in the protein environment is much larger than the systematic errors in SCC-DFTB revealed by the gas-phase model calculations.

Another point worth mentioning is that, although proton delocalization between the Glu side chains is essential to the continuum spectral feature,[24] it is not crucial to the PA. For example, localizing the proton onto one of the Glu residues by constraining the O−H distance to be 0.98 Å in the geometry optimization for the protonated state of the Glu pair leads to a change of only ∼1 kcal/mol in the calculated PA. This is also consistent with the relatively flat PMF profile for the proton motion between the two Glu side chains at 300 K (see below).

3.2. Impact of Starting Crystal Structure and Size of the QM Region. *3.2.1. Effects of Starting Crystal Structure.* In our recent work,[24] where the 1UCQ structure was used to initiate the QM/MM MD simulations, we observed in all cases (regardless of the initial location of the stored proton or level of QM) that the stored proton moved out of the water cluster and became delocalized between Glu194 and Glu204 oxygens within ∼10 ps of initiating the simulations. One could argue that this rapid proton delocalization is intrinsic to the 1UCQ crystal structure, in which the Glu pair was found to be spatially very close (see Figure 1). Therefore, we have carried out SCC-DFTB/MM simulations using a more recent structure (2NTW[37]) reported at 1.53 Å resolution. Even though the 2NTW structure is controversial (see section 2.2), the choice of the structure is motivated by the fact that it has the best resolution among all available L-state structures. In the reported 2NTW X-ray structure, the shortest distance between Glu194 and Glu204 is 3.8 Å, and Glu194 is within hydrogen-bonding distance (2.7 Å) of Tyr83. If delocalization of the stored proton between Glu194 and Glu204 can still occur, then it clearly suggests that the intermolecular proton bond is, in fact, an energetically stable arrangement.

In five independent sets of simulations of ∼2 ns each, the 2NTW-based SCC-DFTB/MM simulations exhibit features similar to those of the 1UCQ-based trajectories: the stored proton is unstable on the water cluster and immediately hops onto one of the glutamates. Following the initial hop, the side chain of Glu194 reorients itself toward Glu204 upon breaking its hydrogen-bonding interaction with Tyr83, which allows the proton to become shared between Glu194 and Glu204. As shown in Figure 2, reorientation of the Glu194 side chain and eventual formation of the intermolecular proton bond take about 200 ps, which is difficult to achieve with *ab initio*-based QM/MM simulations; this again highlights the unique value of the SCC-DFTB/MM approach. A snapshot of the PRG region is shown in Figure 3b; similar to the 1UCQ simulations (Figure 3a), two bulk waters are observed to diffuse into the PRG region during the nanosecond-scale simulations. The histograms for key distances in the PRG region further show that the 2NTW-based simulations are consistent with those based on 1UCQ (compare Figure 4 panels a and b); the only major difference is that the distance between Glu204 and the stored proton has a broader

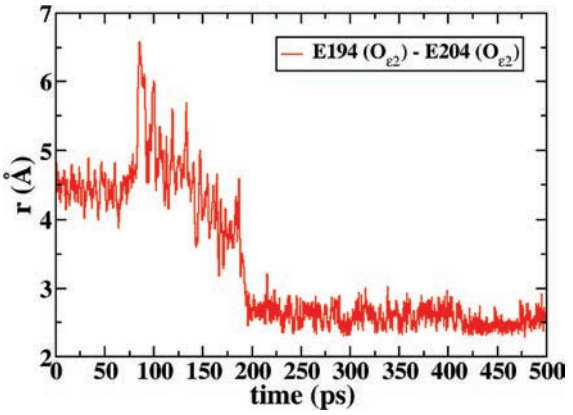


Figure 2. Plot showing the change in Glu194($O_{\varepsilon 2}$)−Glu204($O_{\varepsilon 2}$) distance as a function of time for the 2NTW structure. The side chain of Glu194 reorients itself toward Glu204 upon breaking its hydrogen-bonding interaction with Tyr83, which allows Glu194 and Glu204 to come close to each other, ultimately resulting in the sharing of the stored proton.

distribution with the 2NTW-based trajectories, which is not unexpected considering the larger initial distance between Glu204 and Glu194.

Apart from the simulations for the L-state, we have also carried out simulations for the ground state of bR using the 1C3W structure to directly compare our simulation results with the CPMD-QM/MM study by Marx et al.,[19] where the 1C3W structure was used as well. Similar to the L-state simulations, the stored proton moves out of the water cluster and becomes delocalized between Glu194 and Glu204 oxygens as shown in Figure 3c. The distance histograms in Figure 4c show very similar behaviors compared to the L-state simulations.

Another point of interest concerns the degree of spatial delocalization of the stored proton and water molecules near PRG at finite temperature. As shown in Figure 5a using the 1UCQ simulation at 300 K as an example (other simulations give qualitatively very similar results), the stored proton is delocalized between the Glu194/204 pair, and there is little overlap between its distribution and the water distribution, confirming that the stored proton is not stable on the water cluster. Concerning the position of water molecules near PRG, it is clear that they sample quite a number of locations at finite temperature, especially the one near Tyr83 and one of the water near Arg82. In Figure 5b, we overlap the water distribution from MD simulations (based on the last 1.5 ns of one 2 ns long trajectory with coordinates saved every 1000 MD steps), the average protein configuration, and several crystal structures. In terms of the protein configuration, it seems that the simulation using the smaller QM region gives a somewhat different orientation for the guanidinium part of the Arg82 side chain, which is farther away from the Glu194/204 pair compared to the crystal structures for the ground-state WT (yellow) and mutant (orange/red) structures; we note, however, that the Arg82 side chain actually spans a fairly broad set of configurations in the simulation, including those resolved in the crystal structure (see Supporting Information and discussions below). As to the water distributions, the simulations are largely consistent with the crystallographic results, although as noted earlier, two additional water molecules enter the PRG region during the simulation; it is possible that only three water molecules have been resolved in the crystal structure because of the

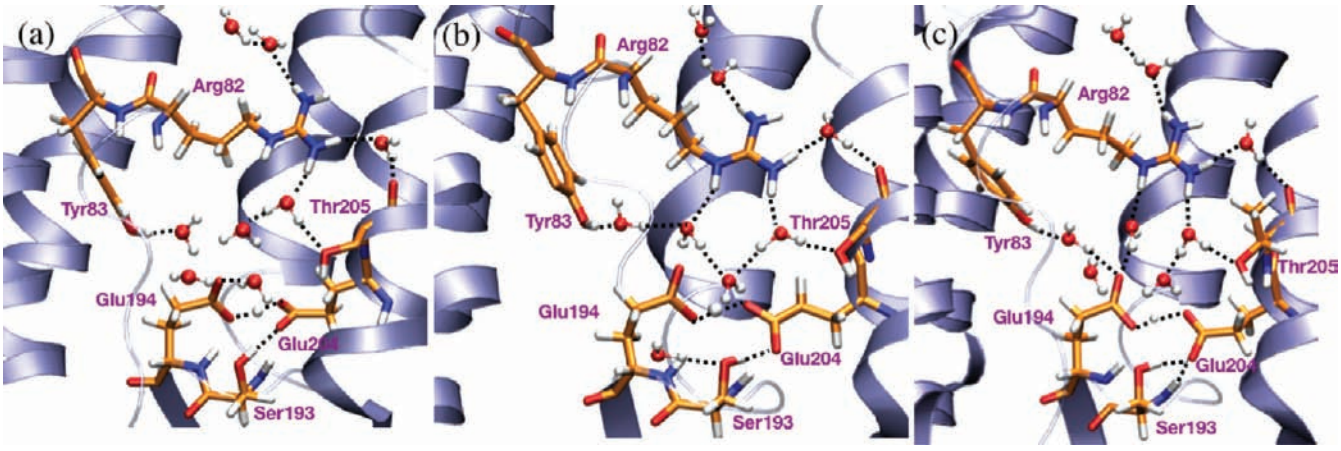


Figure 3. Snapshots of the PRG region from SCCDFTB/MM simulation, showing that the stored proton is shared between Glu194 and Glu204 side chains. (a) L-state 1UCQ structure, (b) L-state 2NTW structure, and (c) ground-state 1C3W structure. Although the simulations are initiated with three water molecules in the active site based on the X-ray structure, two water molecules are consistently observed to move to the region during the nanosecond time scale simulation (also see Supporting Information).

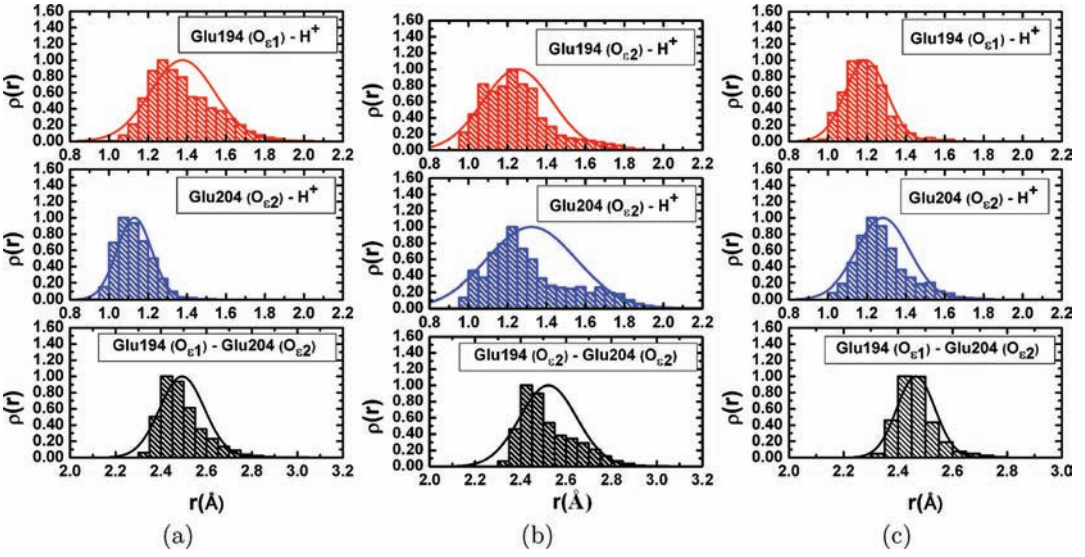


Figure 4. Histograms for key distances involving the stored proton and Glu194/204 side chains from SCCDFTB/MM simulations for (a) L-state 1UCQ, (b) L-state 2NTW, and (c) ground-state 1C3W. The distance variation clearly demonstrates that the stored proton is almost equally shared between the two Glu residues.

relatively high mobility of water molecules, which leads to weak diffraction signals. Furthermore, protein crystals are grown under low humidity and high salt conditions. Therefore, it is not uncommon that the water molecules resolved by crystallography provide only a lower limit to protein hydration (also see additional discussions in the Supporting Information).

Consistent with the similar structural features, the calculated IR spectra also resemble each other. As shown in Figure 6, all three sets of simulations (1UCQ, 2NTW, and 1C3W) lead to reasonably similar IR spectra, especially for the 1800−2000 cm^{-1} region of interest here. Both 1UCQ and 1C3W simulations tend to give a broader continuum band that has non-negligible intensity between 2000 and 2500 cm^{-1} (shown as an inset for the ground-state spectrum in Figure 6c), while the continuum band from the 2NTW model does not carry significant intensity beyond 2000 cm^{-1}. This is likely because the 2NTW simulations contain a significant fraction of the trajectory in which the two Glu residues are far apart (see Figure 2).

In short, regardless of the initial crystal structure used for either the ground or L state of bR, the QM/MM simulations show a consistent picture in which the stored proton does not reside on the water cluster in the PRG region but is delocalized between the highly conserved pair of Glu residues. This proton delocalization leads to the broad diffuse band in the 1800−2000 cm^{-1} region, in qualitative agreement with experiments.[15,16]

3.2.2. Effects of the QM Region Size. Since experimental studies using different mutants have indicated that the hydrogen-bonding network in the PRG region is essential in stabilizing the stored proton,[15] it is a legitimate concern that the results of QM/MM simulations may depend on the choice of the QM region. Therefore, we have carried out simulations with a much larger QM region that includes all residues within hydrogen-bonding interaction

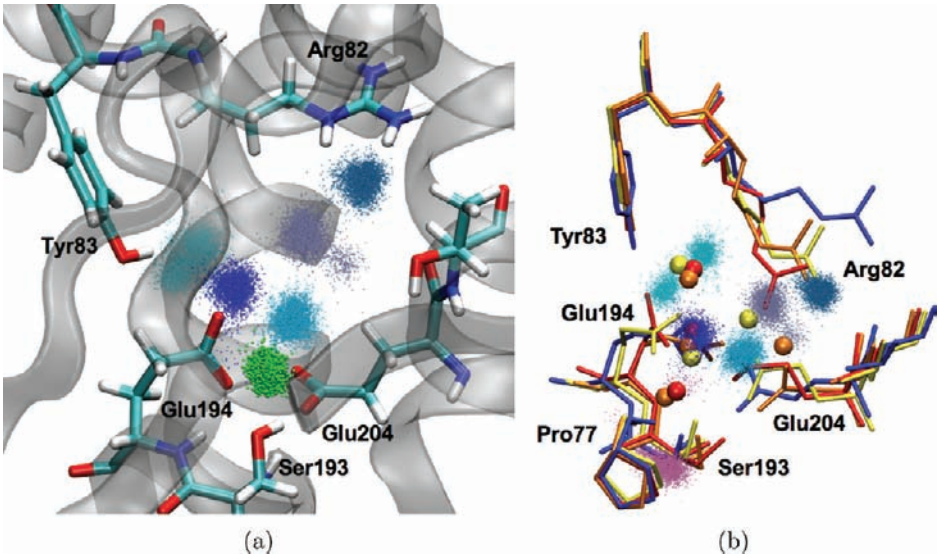


Figure 5. Distribution of water and the stored proton from 300 K L-state simulations. (a) Distribution of the stored proton (in green) overlapped with those of water molecules in the active site. (b) Comparison of water distribution from the simulation to positions (indicated by large spheres) identified in several ground-state crystal structures (yellow; WT, 1C3W; orange, Glu204Asp mutant, 2WJK; red, Glu194Asp mutant, 2WJL). Also shown are key side chains from these ground-state crystal structures and the average structure from the L-state simulation (in dark blue); note that the guanidinium group of Arg82 samples a fairly broad set of configurations during the simulation (see Supporting Information).

with the PRG, which contains as many as 170 atoms (see section 2.4). In Figure 7a, we show the overlap of the average structure from such large QM-region simulations with that from our previous 1UCQ simulations, which employed the smaller QM region (Glu pair plus five nearby water molecules), as well as with the crystal structure (1UCQ). Clearly, the QM/MM simulations with different QM regions give very consistent active-site structural features, such as the orientation of Tyr83 and Arg82, as well as the distribution of water molecules in the PRG region; these structural features are also consistent with those seen in the crystal structure. The agreement between the Arg82 side-chain configuration and the crystal structure is even better with the larger QM region simulations. Importantly, the large QM region simulations also lead to proton delocalization between the conserved Glu pair, although the O−H distance distribution for Glu194 is shifted to lower values compared to the smaller QM region results (compare Figures 4 and 7b). The calculated IR spectrum (Figure 7c) is qualitatively similar to the spectra calculated using the smaller QM region (Figure 6), although the continuum bands appear to be narrower, which might be considered consistent with the higher distributions of the excess proton being localized onto Glu194 in the large-QM simulations; on the other hand, we also note that the expensive large-QM calculations have much limited sampling (200 ps production run), which may have contributed to the difference from smaller QM region results.

The results so far have confirmed that the key features of the intermolecular proton bond model for the PRG[24] is not sensitive to several technical details, such as size of the QM region, the starting crystal structure, or the nuclear quantum effects (see Supporting Information). Therefore, in the following discussions, we will use 1UCQ as the starting crystal structure unless stated otherwise and the smaller QM region (Glu pair plus nearby water molecules), and we do not explicitly include nuclear quantum effects in the calculations.

3.3. Microscopic pK_a Calculations for the PRG. The experimental value of 9.7 for the pK_a of the PRG is not based on a direct measurement but inferred from the complex titration curve of Asp85 during the L-to-M transition.[9,20,21] Specifically, two transitions were observed, where the transition at low pH (pK_a ∼2.6) was interpreted to reflect the deprotonation of Asp85. The second transition at high pH (pK_a ∼9.5−9.7) was explained by suggesting that the deprotonation of the PRG increases the pK_a of Asp85, which further shifts the equilibrium in favor of the deprotonated Asp85.

We have carried out microscopic pK_a calculations for both the ground- and L-state bR, using 1C3W and 1UCQ as the starting crystal structure, respectively; for the L state, two independent sets of simulations have been carried out to verify the reproducibility of the results. As shown in Table 2, the three sets of calculations give rather similar results. Without the B3LYP/MM correction, the calculations lead to pK_a's of ∼15; with the B3LYP/MM corrections (see section 2.5), the microscopic pK_a values fall to ∼11, which should be considered to be in very encouraging agreement with the experimental estimate of 9.7. These rather high pK_a values are consistent with the expectation that a stored proton stabilizes the two otherwise negatively charged Glu residues in proximity. The fact that the ground and L states have similar calculated pK_a values is consistent with the observation that they share very similar local structures in the PRG region.

It is interesting to examine the structural changes in the PRG region during the titration simulations. As shown by the snapshots from different λ windows (Figure 8), the most apparent structural variations involve the key Glu residues. As λ increases, the charges on the two Glu side chains become increasingly negative, and the repulsion between them becomes stronger. As a result, the Glu residues start to repel each other, and the stored "proton" becomes more localized on Glu194. In the limit of λ = 1.0, the two Glu side chains become fully negatively charged; thus, Glu194 reorients to become stabilized by Tyr83 through

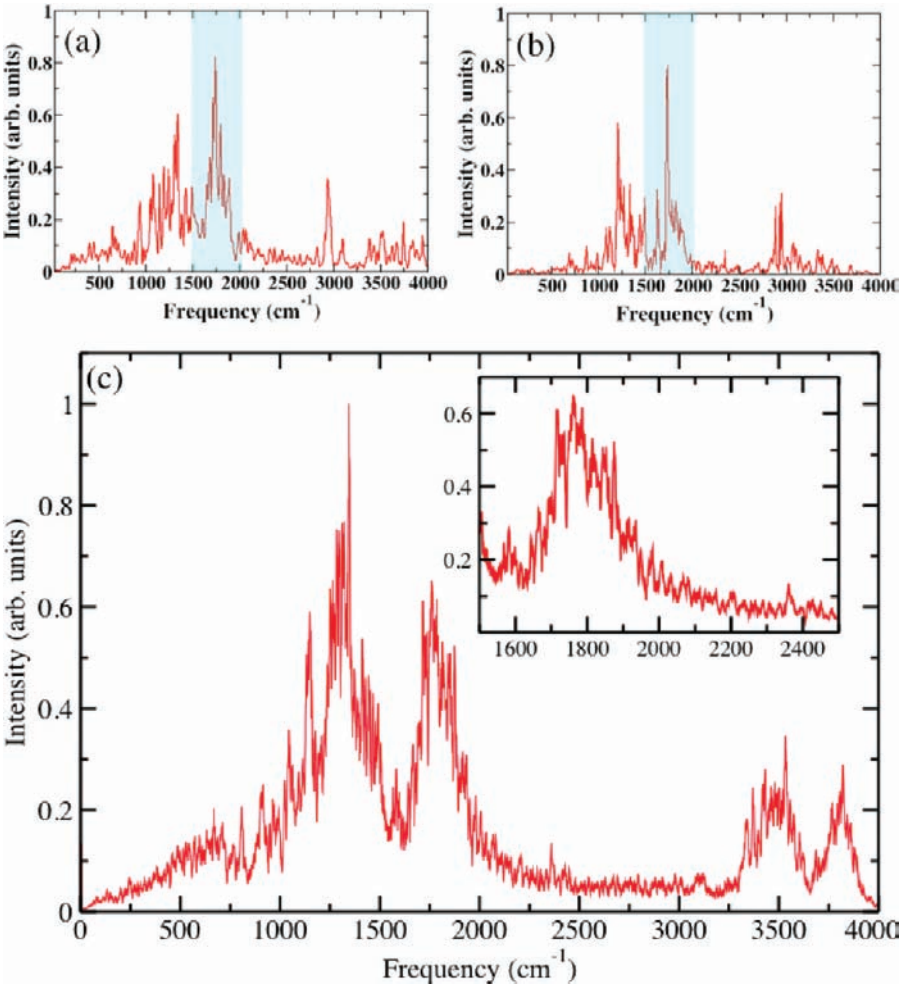


Figure 6. Computed IR spectra for the WT bacteriorhodopsin from SCC-DFTB/MM simulations. The IR spectra for the QM atoms, Glu194/204 side chains, the stored proton, and three/five water molecules treated with SCC-DFTB are shown. The spectral region of importance is highlighted in blue. (a) L-state (1UCQ), (b) L-state (2NTW), and (c) ground-state (1C3W). For the ground-state spectrum, the inset shows the region between 1500 and 2500 cm^{-1}.

hydrogen-bonding interactions. These somewhat expected but nontrivial structural variations highlight the value of microscopic pK_a calculations since it is nontrivial to select an effective dielectric constant for the PRG region in a continuum electrostatic model, explaining the difficulty for continuum electrostatic calculations[18] to reproduce the measured pK_a for the PRG.

As another independent validation of the microscopic approach, we have computed the pK_a for the PRG when Arg82 has isomerized to the configuration observed in the M state. Since the proton is released in the M state, the expectation is that the PRG's pK_a is lowered below 7, due largely to the electrostatic interaction with Arg82. This is supported by the calculations (see Table 2); with the B3LYP/MM corrections, the predicted pK_a in the "early M" configuration is 6.1.

In short, the microscopic pK_a calculations further illustrate that the intermolecular proton bond model for the PRG is consistent with not only experimental spectra but also energetic (pK_a) constraints.

3.4. Potential of Mean Force for Proton Transfer between the Water Cluster and the Glu Pair. The calculations so far indicate that, regardless of the initial location of the stored proton, it prefers to be delocalized between the pair of Glu residues rather than the water cluster, which suggests that the driving force for proton transfer from water to the Glu pair is likely significant. This is explicitly supported by the PMF results shown in Figure 9. In both the ground- (1C3W) and L-state (1UCQ) simulations, and independent of the QM region size, the configuration with the proton on the water is not even a local minimum on the PMF, and it is more than 15 kcal/mol higher than the low-energy segment of the PMF, which corresponds to a shared proton configuration. The plateau around $r \approx 2$ Å corresponds to configurations with larger Glu pair separations and also different hydrogen-bonding patterns, especially concerning Glu194. The most important point revealed by the PMF is that the energetic difference between the two PRG models is not small and much larger than the systematic errors in SCC-DFTB revealed in the gas-phase benchmark calculations.

3.5. Temperature Dependence of Proton Delocalization and IR Signature. As mentioned in the Introduction, an interesting experimental observation is that the continuum band disappeared as the temperature was lowered to ~230 K, which was interpreted as proton release being inhibited at low temperature, due perhaps to the limited flexibility of the protein;[32] bR is known to undergo a dynamic transition at ~230−250 K.[77,78]

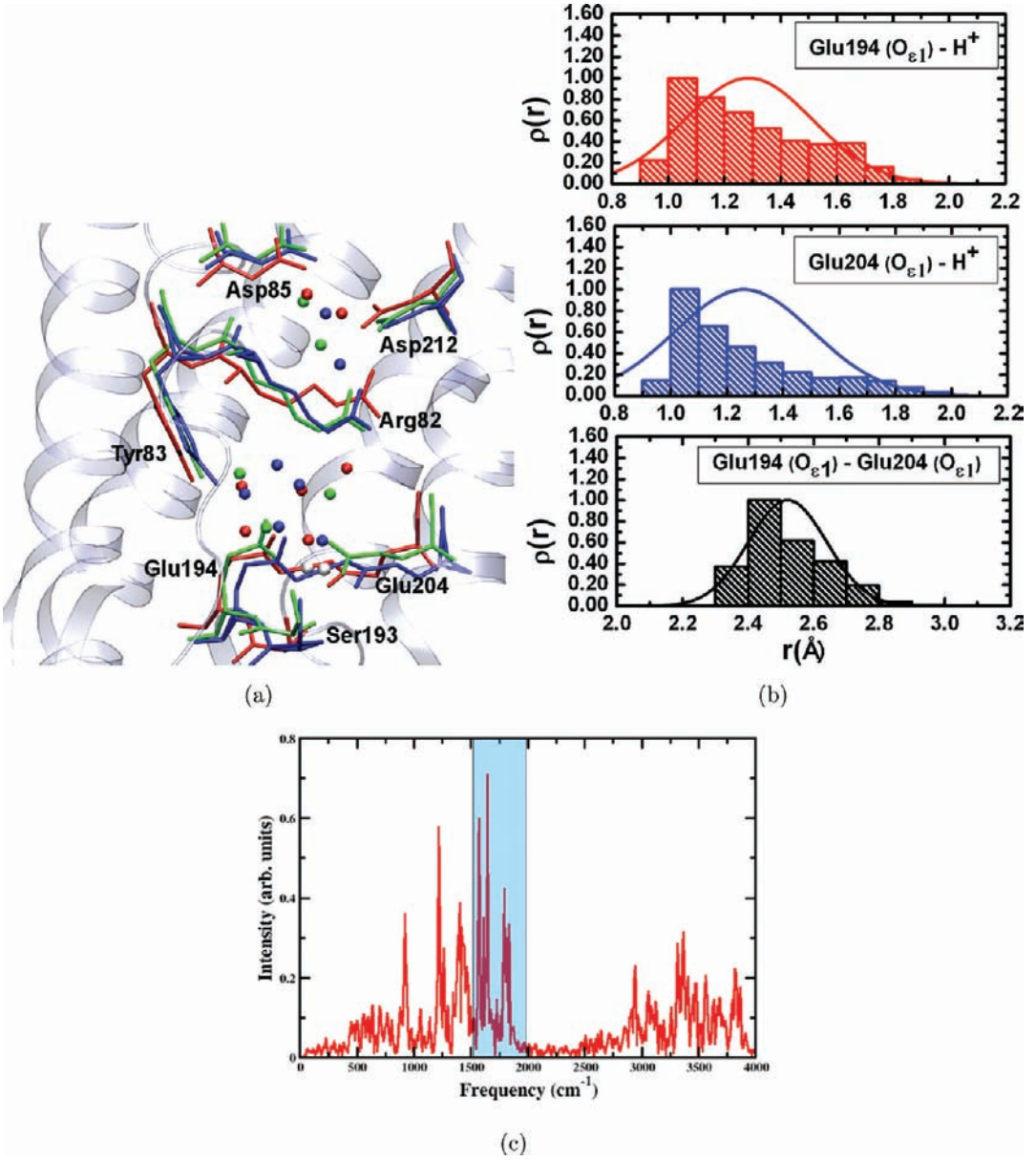


Figure 7. Results from the large-QM region simulations. (a) Comparison of the active-site geometries from the QM/MM simulations with different QM sizes to the L-state structure (PDB code 1UCQ). Color code: the small QM region (same as that in our previous study[24]) in red; the larger QM region (see section 2.4) in blue; the L state in green; and the stored proton from QM/MM simulations is represented as a white sphere. (b) Histograms for key distances involving the stored proton and Glu194/204 side chains. (c) Corresponding IR spectra out of 200 ps of production simulation (400 ps in total), which is substantially shorter than the smaller-QM region simulations and therefore not as well converged. The continuum band in the 1800−2000 cm^{-1} region is highlighted in blue.

Although this interpretation is plausible, especially considering that the M state of bR was shown to evolve to the N state upon increasing the temperature above 250 K,[32] we use simulations here to explore what effects temperature has on the PRG region and whether there is an alternative interpretation to the proposal in ref 32. In this context, we note that, given the limitations of nonpolarizable force field (e.g., the thermodynamically most stable phase of ice for the TIP3P model at its normal melting point is ice II instead of ice I_h; the normal melting point for TIP3P at which ice II melts is 210 K, whereas ice I_h melts at 146 K,[79,80] as compared to the experimental normal melting point of 273.15 K; however, also see ref 81) and the GSBP framework, we do not anticipate a quantitative mapping between the temperature dependence in simulation and in reality. The main goal is to explore, at a qualitative level, how reduction in protein/solvent mobility perturbs the local environment of the PRG and therefore its spectroscopic feature.

For the 230 K simulations, the results are similar to those at 300 K for both structural and spectral features (also see below). Among the independent trajectories we have carried out under 170 K, however, we note that two types of behaviors can be observed which lead to different spectroscopic features in the region of interest: the continuum band either has intensity similar to that in high-temperature simulations (Figure 10a) or is substantially quenched (Figure 10b, note that the continuum band did not fully disappear). Analysis of the trajectories indicates that the

key structural origin for the difference seems to be the orientation of Ser193 side chain, which is hydrogen-bonded with either Glu204 (Figure 10c) or the main-chain carbonyl of Pro77 (Figure 10d). In the former case, the negative charge of Glu204 is better stabilized by Ser193, and the stored proton is more equally shared between Glu204 and Glu194; with this hydrogen bond lost in the latter case, the stored proton is more localized on Glu204. Since we showed in previous work that the continuum feature arises due to the delocalized nature of the stored proton,[24] it makes sense that the continuum feature is much weaker when proton is largely localized on Glu204.

To better quantify the degree of proton delocalization, we have computed the PMFs for the proton between the Glu pair at different temperatures. As evident from Figure 11a, the PMFs at both 230 and 300 K show that the stored proton energetically favors being midway between the Glu residues. On the other hand, the free energy bias over being localized on one of the Glu residues (local minima near $\delta = -0.7$ and 0.7) is only on the order of 1.5−2 kcal/mol; the PMF also indicates that it is slightly more favorable to localize the proton on Glu204 than Glu194, which is consistent with the distance histograms in Figure 4a.

At 170 K, the proton PMF looks qualitatively different with the two Ser193 orientations (Figure 11a). When Ser193 is hydrogen-bonded with Glu204 (referred to as the "in" orientation), the PMF is qualitatively similar to those at higher temperatures, except that localization onto Glu194 becomes substantially less favorable. When Ser193 is flipped out to hydrogen-bond with the carbonyl of Pro77, the minimum of the proton PMF is shifted to a substantially larger δ value (∼0.5), which indicates that the stored proton is largely localized on Glu204 rather than delocalized between the Glu pairs, consistent with the spectroscopic features discussed above.

Natural questions at this point are whether and why alternative Ser193 orientations are observed only in low-temperature simulations. To this end, we have computed the PMF for Ser193 reorientation at different temperatures. As shown in Figure 11b, the PMF favors the "in" orientations at all three temperatures studies here, although the "out" orientation becomes more similar to the "in" configuration in stability as temperature decreases. For example, at 300 K, the "out" configuration is hardly a local minimum and ∼4.5 kcal/mol higher than the "in" configuration; this is consistent with both current QM/MM and previous MM simulations[31,82] with Ser193 engaged in a stable hydrogen-bonding interaction with Glu204, prompting the idea[31,82] that Ser193 may act as a gate to prevent premature proton release (though see discussions in section 4 below). At 170 K, the "out" configuration is less than 1 kcal/mol higher than the "in" configuration and therefore becomes fairly accessible; in fact, in one trajectory, both "in" and "out" configurations have been sampled, and, interestingly, the continuum band intensity also gets substantially quenched in this trajectory (see Supporting Information).

To better understand why the "out" configuration is somewhat better stabilized at the low temperature, we analyze the structural and solvation features of Ser193 and Pro77. The results suggest that the level of solvation for the backbone of Pro77 is notably

Table 2. Results for pK_a Simulations in bR[a]

	L-state, set 1	L state, set 2	ground state	"early M" state[c]
$\partial\Delta G/\partial\lambda\|_{0.0}$	215.0(1.0)[b]	216.8(1.5)	217.1(1.2)	215.0(1.0)
$\partial\Delta G/\partial\lambda\|_{0.25}$	176.3(1.2)	178.1 (0.8)	179.1(1.4)	175.3(1.1)
$\partial\Delta G/\partial\lambda\|_{0.5}$	142.5(1.1)	142.3(1.2)	143.1(1.2)	138.5(1.0)
$\partial\Delta G/\partial\lambda\|_{0.75}$	110.5(0.9)	112.3(1.0)	112.5(1.1)	108.5(0.9)
$\partial\Delta G/\partial\lambda\|_{1.0}$	83.3(1.2)	84.0(0.6)	81.2(1.2)	71.3(0.5)
ΔG	145.5(0.99)[d]	146.7(0.99)	146.0(0.99)	140.2(0.99)
pK_a	10.6	11.0	11.4	6.1

[a] The final pK_a is computed using the calculated pK_a shift relative to an acetic acid in solution (i.e., 0.73× the difference between ΔG shown here and the value calculated for an acetic acid in solution), whose experimental value is 4.8;[88] the values include correction using single-point B3LYP/6-311+G**/MM energy calculations at 500 SCC-DFTB/MM geometries at two end states (λ = 0,1) with a second-order cumulant expansion. [b] The free energy derivatives ($\partial\Delta G_{E\cdot PRG(H/D)}/\partial\lambda$) are given in kcal/mol; the values in parentheses are statistical errors (see Supporting Information about statistical analysis of the pK_a calculations). [c] A model for the "early M" state by equilibrating after modifying the protonation state of the retinal and Asp85 (see section 2.5). [d] Computed on the basis of the linear fit of the free energy derivatives vs λ and subsequent integration over λ; the value in parentheses is the R^2 for the linear fit.

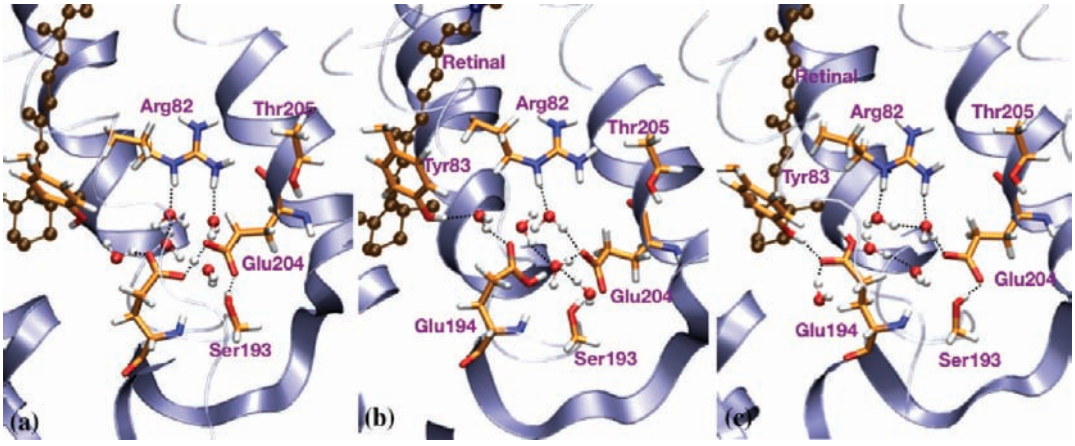


Figure 8. Snapshots from SCCDFTB/MM pK_a calculations for the L state (set 1, see Table 2) showing the structural features of the PRG region: (a) λ = 0.0, (b) 0.5, and (c) 1.0. As evident, the stored proton is shared between the Glu194/204 pair in the protonated state (λ = 0.0), while in the λ = 0.5 window the stored proton is preferentially localized on Glu194. In the deprotonated state (λ = 1.0) where the acidic proton has been converted into a dummy atom (with only van der Waals and bonded terms), the side chains of the two Glu residues move away from each other due to electrostatic repulsion.

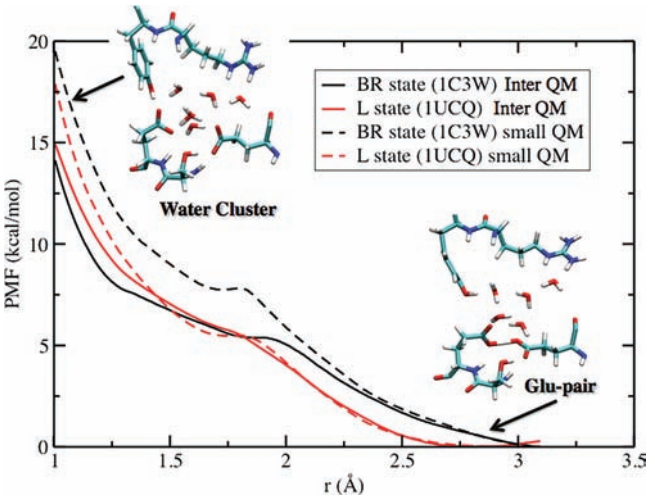


Figure 9. Potential of mean force for the proton transfer from the water cluster to the Glu pair in the PRG region of bR. The "small QM" set includes the side chains of Glu194/204 and nearby water molecules as the QM region (as that in our previous study[24]), and the "inter(mediate) QM" set further includes the side chains of Tyr83 and Ser193. The reaction coordinate (r) is the distance between the stored proton to the oxygen in a hydronium that bridges the two Glu side chains.[31] Representative snapshots from different umbrella sampling windows ($r \approx 1$ and 3.0 Å) are shown for the ground-state simulations.

lower in 170 K simulations, while those at 230 and 300 K are fairly similar (see Figure 11c). Indeed, among five independent trajectories (10 ns in total) at 170 K, a water penetrates to bridge Ser193 and Pro77 only twice for very short intervals. An indirect support of this is the observation from our simulation that, when an additional water is placed near Pro77 (e.g., by using a snapshot from 300 K simulation instead of the crystal structure to start the simulation at 170 K), the distance between Ser193 and Pro77 side chains becomes too large compared to the crystal structure collected at very low (~100 K) temperature; e.g., the value is between 4.5 and 5 Å, as compared to 3.6 Å in 1UCQ. Hence, since the presence of a water-bridging Pro77 and Ser193 is unfavored, because they are closer and less solvated than at high temperatures, it seems reasonable that the movement of Ser193 to an "out" configuration to hydrogen-bond with Pro77 is more favored at low temperature. Regarding a direct comparison of the temperature dependence of the continuum band from simulations and experiment, we note that previous neutron scattering experiments indicated that a dynamic transition of bR in purple membrane occurs around 230−250 K,[77,78] and Kandori et al. interpreted their 230 K experimental condition as below the dynamic transition temperature. Considering the limitations in the force field, it is perhaps more appropriate to take the simulation results at 170 K, as those characterize the behavior of bR at a temperature below the dynamic transition. If so, the current simulation results are qualitatively consistent with the notion that the protein and solvent dynamics are substantially perturbed at such conditions[83] that local solvation structure and therefore spectral features of the PRG are also altered; however, considering the limitations of the force field, this observation remains a speculation and needs to be tested experimentally. The current results, nevertheless, highlight an additional factor that may contribute to the temperature dependence of the continuum band.[32]

4. DISCUSSIONS AND CONCLUDING REMARKS

The problem of PRG in bR provides an unique opportunity to combine a large body of diverse experimental data and state-of-the-art computations to define the identity of an unconventional functional motif that likely exists in many proton pumping systems. As highlighted in the Introduction, although the protonated water cluster and "intermolecular proton bond" models for the PRG in bR look somewhat similar in structural terms, they are very different from an energetic perspective. The protonated water model requires the protein to raise the pK_a of a group by almost 11 pK_a units, which is at odds with our general knowledge of protein electrostatics.[22,23] By contrast, a high pK_a is obtained when a stored proton is shared between two acidic residues even in polar environments (as required by the function of PRG), such as in various "proton sponges".[84] By breaking apart the two acidic residues through rearranging the hydrogen-bonding network (e.g., moving Arg82 down in bR), the pK_a of the motif is readily shifted to the normal range of a single Glu/Asp, which is an ingenious way of triggering proton release from the PRG.

The extensive calculations reported in our previous study[24] and here, based on both SCC-DFTB/MM and B3LYP/MM simulations as well as multiple crystal structures for the ground and L states, strongly suggest that the proton is stored on the pair of conserved Glu residues and not on the water cluster, consistent with the above energetic argument. Most explicitly, the PMFs for the proton transfer from water cluster to the Glu pair show that, in both ground and L states, the stored proton strongly prefers the Glu pair over the water cluster. In other words, although nearby water and polar residues provide the proper electrostatic environment and therefore critically modulate the energetic properties of the Glu pair, the stored proton is explicitly delocalized only between the Glu side chains in the ground and L states. This "intermolecular proton bond" model gives a key spectroscopic signature (continuum band in the range 1800−2000 cm^{-1}) and temperature dependence that are in qualitative agreement with experiments from two research groups for the WT bR and several mutants (Glu204Asp, Glu194Asp, and Ser193Ala;[24] the interpretation for the Arg82Gln mutant is complicated by the fact that the mutant structure can be substantially different from the WT since a critical charge in the region is removed); we note, however, that the interpretation of the temperature dependence is not yet unambiguous, considering the limitations of the force fields. The microscopic pK_a calculated for the PRG is also consistent with the experimental value.

In recent IR studies,[15,31] it was emphasized that spectroscopic feature (1706−1720 cm^{-1}) for the protonated Glu194/204 in the ground-state bR (or, more precisely, the difference between ground- and M-state spectra) was observed not at pH 7 but at pH 5[31] during the L-to-M transition. The interpretation of the observation was that protonation of Glu194/204 occurs only during the L state, which becomes more visible at pH 5 due to the delay in proton release. Furthermore, on the basis of the movements of the Arg82 side chain and water in the PRG region as observed in two ground-state mutant structures relative to the WT, it was further argued that the "downshift" of the stored proton from the water cluster in the ground state to the Glu194/204 pair in the L state was due to a combination of electrostatic and steric effects associated with these movements.

How do our simulation results fit in with these experimental observations? Regarding the lack of spectroscopic signature of a protonated Glu194/204 in the ground-state bR under

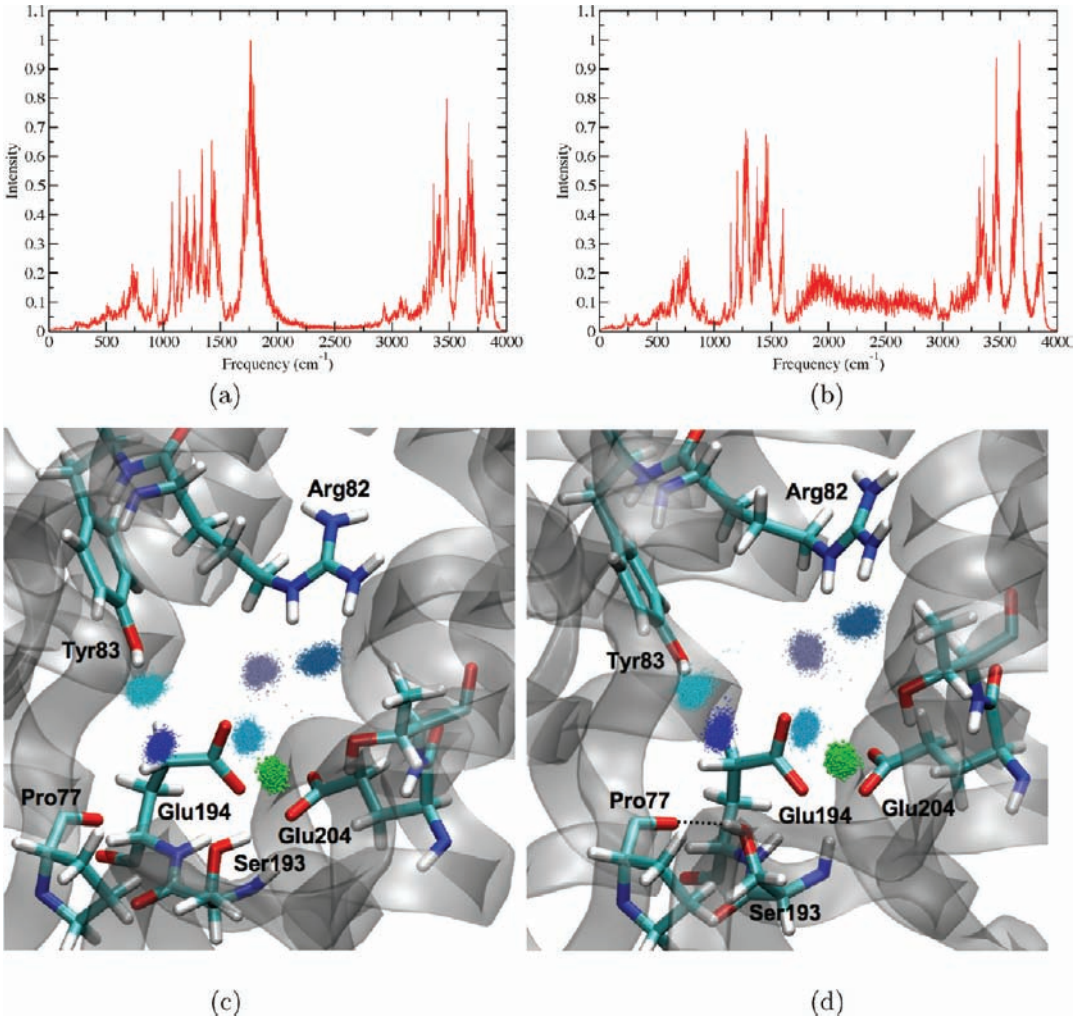


Figure 10. IR spectra and active-site structure/water distribution from simulations for the wild-type L state at 170 K. Panels on the left represent results for trajectories in which Ser193 is engaged in a hydrogen-bonding interaction with Glu204 (the "in" configuration of Ser193); panels on the right represent results for trajectories in which Ser193 is engaged in a hydrogen-bonding interaction with the backbone carbonyl of Pro77 (the "out" configuration of Ser193).

physiological pH, one possible explanation was provided in our previous study.[24] We showed that the strong hydrogen-bonding interaction between the two Glu side chains through the stored proton makes the carbonyl stretches significantly red-shifted, very much below the expected region (1706−1720 cm^{-1}); a larger model calculation further supports this (see Supporting Information). This explains why no band in 1706−1720 cm^{-1} was detected at pH 7 (also see below). Regarding the pH 5 observation, there is also an alternative interpretation. At pH 5, the proton release is delayed, which is in fact qualitatively consistent with our calculated pK_a for the PRG in an early M state (6.1 in Table 2) and the experimental observation that the band near 1706−1720 cm^{-1} arises as the M state accumulates.[31] Since the Glu pair is already broken in the (early) M state, the stored proton is localized onto Glu204, which has a characteristic carbonyl stretch in the 1706−1720 cm^{-1} region. In other words, the band observed at pH 5 corresponds to a singly protonated Glu204 rather than a pair of Glu's with a shared proton; this band was not observed at pH 7 for the M state because the proton is released in M at this pH.

Concerning the role of the Arg82 movement and "downshift" of the stored proton onto the Glu pair, although our pK_a calculations support the idea that Arg82 movement is crucial for reducing the pK_a of the PRG, our simulations also show that the proton delocalization between the Glu pair does not require any significant downshift of Arg82. In fact, relative to the WT bR crystal structures in the ground and L states, the Arg82 side chain is found to shift slightly upward in our simulations to interact with Asp212 (which was also observed in previous CP2K/MM simulations[85]), and proton delocalization between the Glu pair still occurs (see the Supporting Information for additional discussions).

These discussions highlight the importance of establishing the proper spectral assignments for a singly protonated Glu/Asp vs a carboxylate pair with a shared proton, for which thorough comparisons of experiments and high-level calculations for model systems[28] will be useful. Along this line, isotopically labeling the oxygen in Glu's will be helpful for identifying carbonyl bands for the Glu pair with a shared proton in bR (see additional discussions in the Supporting Information). Therefore, a future experiment that may help distinguish the two PRG models is to use ^{18}O-substituted Glu's to explore whether there are carbonyl bands that clearly correspond to a Glu pair with a shared proton. Another possibility, if it is experimentally feasible for a membrane

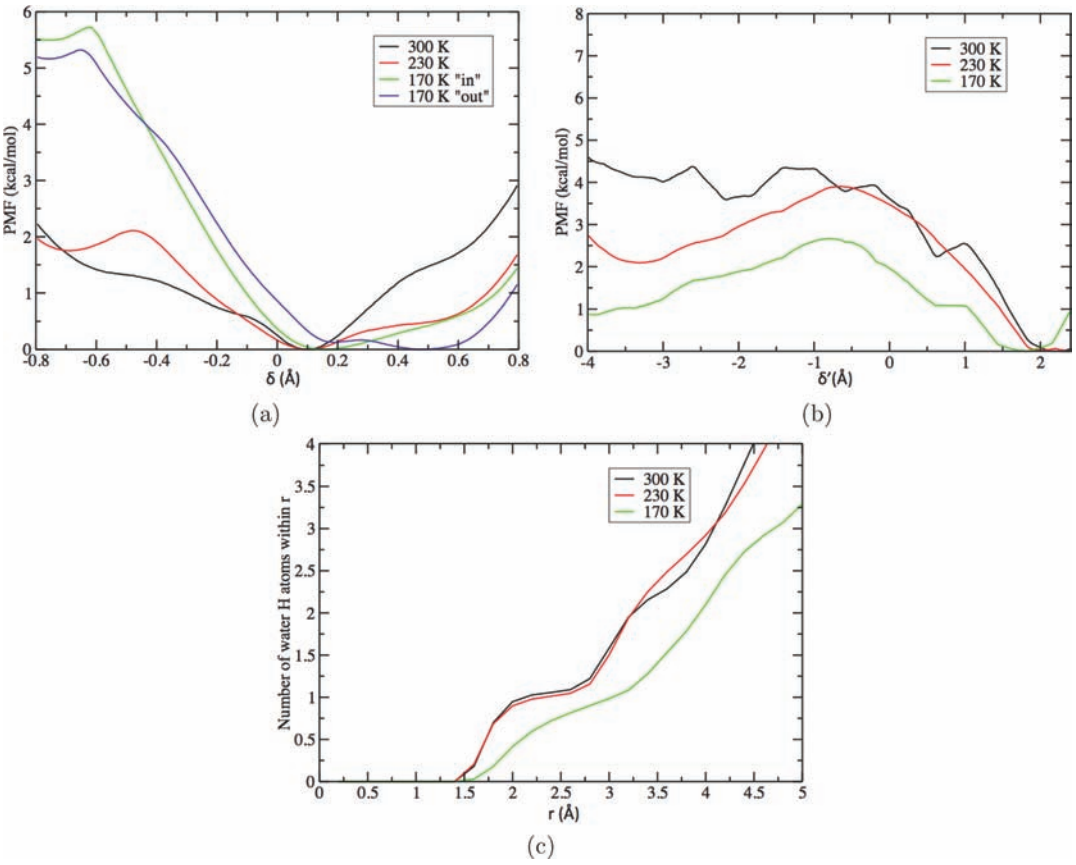


Figure 11. Potentials of mean force and solvation of Pro77 backbone from simulations at different temperatures. (a) PMF for the stored proton between Glu194 and Glu204 using the asymmetric stretch coordinate (in Å). (b) PMF for the orientation of Ser193 between the "in" (right side) and "out" (left side) configurations (also see Figure 10c,d); the reaction coordinate (in Å) is also an asymmetric stretch coordinate involving the Pro77 carbonyl and Ser193 side chain (see section 2.7). (c) Integrated number of H atoms in water around the carbonyl oxygen of Pro77.

protein, is to use ^{17}O NMR on ^{17}O-labeled Glu194/204 and compare to QM/MM calculations.[86] The shared proton is expected to cause a major perturbation of the chemical shift of the Glu oxygens, while the pattern of perturbation is expected to be different if the proton resides on the water molecules (see the Supporting Information for model calculations). These isotopic substitutions are expected to provide much more direct information than solvent ^{18}O effects, which are very small in magnitude[32] and therefore difficult to interpret or compute with any currently available methods.

Finally, we briefly comment on the "proton diode" model proposed in recent studies,[31,82] which argued that Ser193 can serve as a gate to prevent premature release of the stored proton and also ensure the vectorial nature of bR. Although this is clearly an interesting and visually attractive model, and our QM/MM PMF simulations also find that Ser193 forms a favorable hydrogen bond with Glu204 under physiological condition, the "proton diode" model appears to emphasize the role of water connectivity between the bulk and PRG over the role of protein electrostatics in regulating the pK_a of key groups. In our opinion, the latter is even more important because it correlates directly with the thermodynamic driving force for the proton release. In the ground state and L state, the pK_a of the PRG is much higher than 7; therefore, proton release is inhibited, even if the PRG is solvent accessible, as shown by L-state MD simulations here (Figure 5b) and in ref 87. In the M state, upon Arg82 isomerization, the pK_a of the PRG is significantly shifted to be below 7, as shown both experimentally and by our "early M" simulation here, which allows proton release. Therefore, the basic properties of bR can be readily explained by considering pK_a regulation of the PRG, which has a clearer thermodynamic basis directly related to proton transfers. For example, the observation that proton release is delayed in the Ser193Ala mutant[82] can be readily explained: due to the lack of hydrogen-bonding interaction between Ala193 and Glu204, the stabilization of the deprotonated state is more affected than the protonated state, thus the pK_a of the PRG is expected to be further raised in the Ser193Ala mutant, making it harder (thermodynamically) to release the proton than in the WT. Along this line, we note that even substituting Ser193 with a Cys will also have an effect similar to the Ser193Ala mutation; this is because a thiol group is much less effective than a hydroxyl group in terms of stabilizing a carboxylate through hydrogen-bonding interactions (see Supporting Information).

To sum up, our simulation and analyses suggest that the "intermolecular proton bond" model for PRG in bR is consistent with most available experimental data, although there remain interesting experimental observations that deserve further analysis, especially concerning the temperature dependence of the continuum band. From a technical point of view, we emphasize that it remains difficult to quantitatively compare the calculated and measured line shapes of continuum bands, which deserves

further developments and collaborative efforts between theory and experiment. The qualitative conclusion from this study, however, is robust by looking collectively at multiple types of observables. Considering the polar environment of PRG and the limited range of pK_a shifts observed for proteins, we argue that the "intermolecular proton bond" model is energetically more physical than a "protonated water cluster" model and likely involved in many proton pumping systems. For example, in cytochrome *c* oxidase, the identity of the proton loading site (the equivalent of the PRG in bR) remains elusive after many experimental and computational studies because there is not an obvious group in the relatively polar region that can reach an apparent pK_a of 9.[4] Examination of the crystal structure suggests that propionate A of heme a_3 is an interesting candidate because it is spatially close to a conserved Asp residue; whether they act as the loading site in a similar "intermolecular proton bond" fashion is a fascinating possibility waiting to be evaluated.

■ ASSOCIATED CONTENT

Supporting Information. Additional calculation results and discussions on the structure and normal modes of gas-phase models, dependence of the calculated IR spectra on the choice of origin for the dipole moment, ring-polymer MD simulations for the nuclear quantum effects on continuum band in a gas-phase model compound, behavior of Arg82 and Ser193 side chains, the possibility of having two protons in the PRG region, statistical analysis of sampling in pK_a simulations, and model NMR chemical shifts calculations; complete ref 40. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

marcus.elstner@kit.edu; cui@chem.wisc.edu

Author Contributions

‡These authors contributed equally.

■ ACKNOWLEDGMENT

The research has been supported by the National Institutes of Health (R01-GM084028 to Q.C.) and the Deutsche Forschungsgemeinschaft (Forschergruppe 490 to M.E.). Computational resources from the National Center for Supercomputing Applications at the University of Illinois and the Center of High Throughput Computing at UW-Madison are greatly appreciated; computations are also supported in part by the National Science Foundation through a major instrumentation grant (CHE-0840494). The authors thank Prof. Klaus Gerwert and Dr. S. Wolf for stimulating discussions.

■ REFERENCES

(1) Nicholls, D. G.; Ferguson, S. J. *Bioenergetics 3*; Academic Press: San Diego, CA, 2002.
(2) Hosler, J. P.; Ferguson-Miller, S.; Mills, D. A. *Annu. Rev. Biochem.* **2006**, *75*, 165–187.
(3) Brzezinski, P.; Gennis, R. B. *J. Bioenerg. Biomember* **2008**, *40*, 521–531.
(4) Kaila, V. R.; Verkhovsky, M. V.; Wikström, M. *Chem. Rev.* **2010**, *110*, 7062–7081.
(5) Quenneville, J.; Popovic, D. M.; Stuchebrukhov, A. A. *Biochim. Biophys. Acta* **2006**, *1757*, 1035–1046.
(6) Song, Y.; Michonova-Alexova, E.; Gunner, M., G. *Biochemistry* **2006**, *45*, 7959.
(7) Oesterhelt, D.; Stoeckenius, W. *Proc. Natl. Acad. Sci. U.S.A.* **1973**, *70*, 2853–2857.
(8) Gerwert, K.; Souvignier, G.; Hess, B. *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 9774–9778.
(9) Balashov, S. *Biochim. Biophys. Acta* **2000**, *1469*, 75–95.
(10) Brown, L. S.; Sasaki, J.; Kandori, H.; Maeda, A.; Needleman, R.; Lanyi, J. K. *J. Biol. Chem.* **1995**, *270*, 27122–27126.
(11) Balashov, S. P.; Imasheva, E. S.; Ebrey, T. G.; Chen, N.; Menick, D. R.; Crouch, R. K. *Biochemistry* **1997**, *36*, 8671–8676.
(12) Richter, H. T.; Brown, L. S.; Needleman, R.; Lanyi, J. K. *Biochemistry* **1996**, *35*, 4054–4062.
(13) Essen, L. O.; Siegert, R.; Lehmann, W. D.; Oesterhelt, D. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 11673–11678.
(14) Zscherp, C.; Schlesinger, R.; Tittor, J.; Oesterhelt, D.; Heberle, J. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 5498–5503.
(15) Garczarek, F.; Brown, L. S.; Lanyi, J. K.; Gerwert, K. *Proc. Acad. Natl. Sci. U.S.A.* **2005**, *102*, 3633.
(16) Garczarek, F.; Gerwert, K. *Nature* **2006**, *439*, 109–112.
(17) Headrick, J. M.; Diken, E. G.; Walters, R. S.; Hammer, N. I.; Christie, R. A.; Cui, J.; Myshakin, E. M.; Duncan, M. A.; Johnson, M. A.; Jordan, K. D. *Science* **2005**, *308*, 1765–1769.
(18) Spassov, V. Z.; Luecke, H.; Gerwert, K.; Bashford, D. *J. Mol. Biol.* **2001**, *312*, 203–219.
(19) Mathias, G.; Marx, D. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 6980–6985.
(20) Balashov, S. P.; Imasheva, E.; Govindjee, R.; Ebrey, T. G. *Biophys. J.* **1996**, *70*, 473–481.
(21) Balashov, S. P.; Govindjee, R.; Imasheva, E.; Misra, S.; Ebrey, T. G.; Feng, Y.; Crouch, R. K.; Menick, D. R. *Biochemistry* **1995**, *34*, 8820–8834.
(22) Schutz, C. N.; Warshel, A. *Proteins: Struct., Funct. Genet.* **2001**, *44*, 400–417.
(23) Isom, D. G.; Castaneda, C. A.; Velu, P. D.; B. Garcia-Moreno, E. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 16096–16100.
(24) Phatak, P.; Ghosh, N.; Yu, H.; Cui, Q.; Elstner, M. *Proc. Acad. Natl. Sci. U.S.A.* **2008**, *105*, 19672–19677.
(25) Elstner, M.; Porezag, D.; Jungnickel, G.; Elstner, J.; Haugk, M.; Frauenheim, Th.; Suhai, S.; Seifert, G. *Phys. Rev. B* **1998**, *58*, 7260–7268.
(26) Cui, Q.; Elstner, M.; Kaxiras, E.; Frauenheim, Th.; Karplus, M. *J. Phys. Chem. B* **2001**, *105*, 569–585.
(27) Riccardi, D.; Schaefer, P.; Yang, Y.; Yu, H.; Ghosh, N.; Prat-Resina, X.; König, P.; Li, G.; Xu, D.; Guo, H.; Elstner, M.; Cui, Qiang *J. Phys. Chem. B* **2006**, *110*, 6458–6469.
(28) Roscioli, J. R.; McCunn, L. R.; Johnson, M. A. *Science* **2007**, *316*, 249–254.
(29) Li, G.; Cui, Q. *J. Phys. Chem. B* **2003**, *107*, 14521–14528.
(30) Riccardi, D.; Schaefer, P.; Cui, Q. *J. Phys. Chem. B* **2005**, *109*, 17715–17733.
(31) Wolf, S.; Freier, E.; Potschies, M.; Hofmann, E.; Gerwert, K. *Angew. Chem., Int. Ed.* **2010**, *49*, 6889–6893.
(32) Lorenz-Fonfria, V. A.; Furutani, Y.; Kandori, H. *Biochemistry* **2008**, *47*, 4071–4081.
(33) Yang, Y.; Yu, H.; York, D.; Cui, Q.; Elstner, M. *J. Phys. Chem. A* **2007**, *111*, 10861–10873.
(34) Gaus, M.; Cui, Q.; Elstner, M. *J. Chem. Theor. Comput.* **2011**, *7*, 931–948.
(35) Elstner, M. *J. Phys. Chem. A* **2007**, *111*, 5614–5621.
(36) Kouyama, T.; Nishikawa, T.; Tokuhisa, T.; Okumura, H. *J. Mol. Biol.* **2004**, *335*, 531–546.
(37) Lanyi, J. K.; Schobert, B. *J. Mol. Biol.* **2007**, *365*, 1379–1392.
(38) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187–217.
(39) Brunger, A. T.; Karplus, M. *Proteins* **1988**, *4*, 148–156.
(40) MacKerell, A. D.; et al. *J. Phys. Chem. B* **1998**, *102*, 3586–3616.
(41) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926.

(42) Tajkhoshid, S.; Suhai, E. *J. Phys. Chem. B* **1999**, *103*, 5581–5590.
(43) Tajkhoshid, E.; Paizs, B.; Suhai, S. *J. Phys. Chem. B* **1999**, *103*, 4518–4527.
(44) Im, W.; Berneche, S.; Roux, B. *J. Chem. Phys.* **2001**, *114*, 2924–2937.
(45) Brooks, C. L., III; Karplus, M. *J. Mol. Biol.* **1989**, *208*, 159–181.
(46) Steinbach, P. J.; Brooks, B. R. *J. Comput. Chem.* **1994**, *15*, 667–683.
(47) Nina, M.; Beglov, D.; Roux, B. *J. Phys. Chem. B* **1997**, *101*, 5239–5248.
(48) Nina, M.; Im, W.; Roux, B. *Biophys. Chem.* **1999**, *78*, 89–96.
(49) Luecke, H.; Schobert, B.; Richter, H. T.; Cartailler, J. P.; Lanyi, J. K. *J. Mol. Biol.* **1999**, *291*, 899–911.
(50) Clemens, M.; Phatak, P.; Cui, Q.; Bondar, A. N.; Elstner, M. *J. Phys. Chem. B* **2011**, *115*, 7129–7135.
(51) Belrhali, H.; Nollert, P.; Royant, A.; Menzel, C.; Rosenbusch, J. P.; Landau, E. M.; Pebay-Peyroula, E. *Structure* **1999**, *7*, 909–917.
(52) Woo, H. J.; Dinner, A. R.; Roux, B. *J. Chem. Phys.* **2004**, *121*, 6392–6400.
(53) Elstner, M. *Theo. Chem. Acc.* **2006**, *116*, 316–325.
(54) König, P. H.; Hoffmann, M.; Frauenheim, T.; Cui, Q. *J. Phys. Chem. B* **2005**, *109*, 9082–9095.
(55) Li, G.; Zhang, X.; Cui, Q. *J. Phys. Chem. B* **2003**, *107*, 8643–8653.
(56) Frenkel, D.; Smit, B. *Understanding Molecular Simulations: from Algorithms to Applications*; Academic Press: San Diego, 1996.
(57) Ghosh, N.; Prat-Resina, X.; Gunner, M.; Cui, Q. *Biochemistry* **2009**, *48*, 2468–2485.
(58) Riccardi, D.; Cui, Q. *J. Phys. Chem. A* **2007**, *111*, 5703–5711.
(59) Ghosh, N.; Cui, Q. *J. Phys. Chem. B* **2008**, *112*, 8387–8397.
(60) Rychaert, J. P.; Ciccotti, G.; Berendsen, H. J. *J. Comput. Phys.* **1977**, *23*, 327–341.
(61) Schobert, B.; Cupp-Vickery, J.; Hornak, V.; Smith, S. O.; Lanyi, J. K. *J. Mol. Biol.* **2002**, *321*, 715–726.
(62) Gordon, R. G. *J. Chem. Phys.* **1965**, *43*, 1307.
(63) McQuarrie, D. A. *Statistical Mechanics*; Harper and Row: New York, 1976.
(64) Ramirez, R.; Lopez-Ciudad, T.; Kumar, P.; Marx, D. *J. Chem. Phys.* **2004**, *121*, 3973–3983.
(65) Lawrence, C. P.; Nakayama, A.; Makri, N.; Skinner, J. L. *J. Chem. Phys.* **2004**, *120*, 6621–6624.
(66) Berens, P. H.; White, S. R.; Wilson, K. R. *J. Chem. Phys.* **1981**, *75*, 515–529.
(67) Bader, J. S.; Berne, B. J. *J. Chem. Phys.* **1994**, *100*, 8359–8366.
(68) Yu, H.; Cui, Q. *J. Chem. Phys.* **2007**, *127*, 234504.
(69) Craig, I. R.; Manolopoulos, D. E. *J. Chem. Phys.* **2004**, *121*, 3368–3373.
(70) Witt, A.; Ivanov, S. D.; Shiga, M.; Forbert, H.; Marx, D. *J. Chem. Phys.* **2009**, *130*, 194510.
(71) Chandler, D.; Wolynes, P. G. *J. Chem. Phys.* **1981**, *74*, 4078–4095.
(72) Koch, W.; Holthausen, M. C. *A chemist's guide to density functional theory*; Wiley-VCH: Weinheim, 2002.
(73) Kumar, S.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A.; Rosenberg, J. M. *J. Comput. Chem.* **1992**, *13*, 1011–1021.
(74) Torrie, G.; Valleau, J. *J. Comput. Phys.* **1977**, *23*, 187–199.
(75) Zhao, Y.; Truhlar, D. G. *Theor. Chem. Acc.* **2008**, *120*, 215–241.
(76) Goerigk, L.; Grimme, S. *Phys. Chem. Chem. Phys.* **2011**, *13*, 6670–6688.
(77) Ferrand, M.; Dianoux, A. J.; Petry, W.; Zaccai, G. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 9668–9672.
(78) Lechnert, U.; Reat, V.; Weik, M.; Zaccai, G.; Pfister, C. *Biophys. J.* **1998**, *75*, 1945–1952.
(79) Vega, C.; Abascal, J. L. F.; Conde, M. M.; Aragones, J. L. *Faraday Discuss.* **2008**, *141*, 1–26.
(80) Vega, C.; Sanz, E.; Abascal, J. L. F. *J. Chem. Phys.* **2005**, *122*, 114507.
(81) Glass, D. C.; Krishnan, M.; Nutt, D. R.; Smith, J. C. *J. Chem. Theor. Comput.* **2010**, *6*, 1390–1400.
(82) Wolf, S.; Freier, E.; Gerwert, K. *Chem. Phys. Chem.* **2008**, *9*, 2772–2778.
(83) Vitkup, D.; Ringe, D.; Petsko, G. A.; Karplus, M. *Nat. Struct. Biol.* **2000**, *7*, 34–38.
(84) Perrin, C. L.; Nielson, J. B. *Annu. Rev. Phys. Chem.* **1997**, *48*, 511–544.
(85) Baer, M.; Mathias, G.; Kuo, I. F. W.; Tobias, D. J.; Mundy, C. J.; Marx, D. *ChemPhysChem* **2008**, *9*, 2703–2307.
(86) Cui, Q.; Karplus, M. *J. Phys. Chem. B* **2000**, *104*, 3721–3743.
(87) Chaumont, A.; Baer, M.; Mathias, G.; Marx, D. *ChemPhysChem* **2008**, *9*, 2751–2758.
(88) Lide, D. R., Ed. *CRC Handbook Chemistry and Physics*, 85 ed.; CRC Press: Boca Raton, FL,2005.