Supporting Information

Song and Gunner 10.1073/pnas.1411119111

SI Methods

MCCE Analysis of HR and BR. All calculations use MCCE2.4, which includes local optimization of side-chain positions, pruning to delete energetically indistinguishable conformers and correcting the continuum electrostatics energy terms for errors in the dielectric boundary due to extra surface conformations (1). Monte Carlo sampling allows residues to sample different protonation states for acidic and basic side chains and conformers for all groups as well as ion binding-site occupancies to establish the Boltzmann distribution of these degrees of freedom at 25 °C (2, 3). Residue pK_a values are determined from the fraction group ionization in a series of Monte Carlo simulations at different pH values. Relative Cl^- affinities are determined by monitoring the probability of Cl^- binding as a function of the Cl^- chemical potential (3).

Preselected side-chain conformers are generated and optimized using the MCCE2.4 default analysis (1). Waters in the crystal structure are removed. The cavities are filled with implicit water with a dielectric constant of 80. The one chloride (Cl^-501) found in the X-ray crystal structure (1E12) is retained. Calculations add 179 additional chlorides to cavities using the IPECE subroutine (2). Chloride parameters are those used in previous MCCE calculations of ion binding to α-amylase (3). All chlorides have a conformation allowing them to leave the protein.

A 33-Å low-dielectric slab is added to simulate the low-dielectric environment of the membrane. The membrane position is determined with IPECE (2) using the trimeric biological unit obtained from the Macromolecular Structure Database at the European Bioinformatics Institute (pqs.ebi.ac.uk/pqs-doc.shtml) (4). The membrane orientation generated with the trimer is in better agreement with the crystal lipid positions than those found using the monomer. The pK_a calculations are carried out on a slab embedded monomer. The other two units of the HR trimer are assumed to have only a small effect on the calculated pK_a and chloride binding affinity because the active sites are separated by >25 Å.

The protein and the membrane slab are given a dielectric constant of 4, whereas the surrounding water has a dielectric constant of 80 with an implicit salt concentration of 150 mM. In the MCCE analysis, the first, default Cl^- conformer has no interaction with the protein to keep the cavities from being filled with ions that do not remain bound to the protein in the end of the simulation (3). The electrostatic pairwise interactions and reaction field (solvation) energies are calculated using the program DelPhi (5–7). The protein is given PARSE charges and radii (8). The same retinal Schiff base Density Functional Theory (DFT)-derived atomic charges (9) and an aqueous, reference pK_a of 7 (10) is applied to both *cis* and *trans* isomers. Four focusing runs (11) with 65^3 grids give a final resolution of 2.0 grids per Å.

Analysis of Cl^- Affinity in MCCE. MCCE has been developed to sample the pH and ion concentration dependence of ion binding to proteins (1, 3). The microstate energy can be written as follows (3):

$$\Delta G^x = \sum_{i=1}^{M} \delta_{x,i} \left\{ \left[2.3 m_i k_b T (\mathrm{pH} - \mathrm{pK}_{\mathrm{sol},i}) + n_i \left(\mu_{\mathrm{corr}} + kT \ln\left(\frac{\rho}{\rho'}\right)\right)\right] + \left(\Delta\Delta G_{\mathrm{rxn},i} + \Delta G^{\mathrm{CE}}_{\mathrm{bkbn},i} + \Delta G^{\mathrm{LJ}}_{\mathrm{bkbn},i} + \Delta G_{\mathrm{torsion},i} + \Delta\Delta G_{\mathrm{SAS},i}\right) + \sum_{j=i+1}^{M} \delta_{x,j} \left[\Delta G^{\mathrm{CE}}_{ij} + \Delta G^{\mathrm{LJ}}_{ij}\right] \right\}, \quad [\mathbf{S1}]$$

where M is the total number of conformers. $\delta_{x,i}$ is 1 if conformer i is present in the microstate or 0 otherwise. m_i is 1 for bases, −1 for acids, and 0 for neutral conformers. k_bT is 0.59 kcal/mol (0.43 pH units) at 298 K, the default temperature. The pH describes the ability of the solvent to donate protons. Group pK_a in aqueous solution ($pK_{a,sol,i}$) is the reference solution pK_a of groups involved in acid/base reactions. n_i is 1 if conformer i represents an unbound ion, which has a self-energy term $\mu_{\mathrm{corr}} + kT\ln(\rho/\rho')$ and zero interaction with the protein. ρ is the bulk ion density, $\rho' = 1/a^3$, where a is the spacing between Cl^- incorporated into the cavities. The finer spacing thus adds a larger number of Cl^- into cavities, increasing the effective ion density. Earlier studies of chloride binding in α-amylase, human serum albumin, and Omp32 (3) found that, although the relative pH and chloride concentration dependence calculated by MCCE agreed well with the experimental data, the absolute binding energy of chloride was difficult to determine. As a result of that study, an additional energy correction term, μ_{corr}, independent of pH and chloride concentration, is added to the bulk solution conformer. This value appears to be correlated with the solvent exposure of the ion binding site. The HR binding sites are deeply buried. The μ_{corr} of 7.1 kcal/mol penalizing Cl^- in solution, derived from the study of α-amylase, is used here. The default value for $kT\ln(\rho/\rho')$ of −6.15 kcal/mol, represents 0.050 M Cl^-. The default $\mu_{\mathrm{corr}} + kT\ln(\rho/\rho')$ yields a SB pK_a of 8.5, which corresponds to that found with 0.2 M Cl^- (3). Thus, the Cl^- binds approximately fourfold more tightly in these calculations than in experiment with a modest systematic error of only ∼0.8 kcal/mol. The concentration dependence of Cl^- binding is obtained by changing $kT\ln(\rho/\rho')$. This is analogous to changing the proton concentration by changing the pH in Eq. **S1**.

The second line of the equation describes the conformer self-energies, which are independent of the other conformers in the microstate. The third line gives the electrostatic and Lennard–Jones pairwise interactions, which depend on the conformers selected in the microstate.

1. Song Y, Mao J, Gunner MR (2009) MCCE2: Improving protein pK_a calculations with extensive side chain rotamer sampling. *J Comput Chem* 30(14):2231–2247.
2. Song Y, Mao J, Gunner MR (2003) Calculation of proton transfers in Bacteriorhodopsin bR and M intermediates. *Biochemistry* 42(33):9875–9888.
3. Song Y, Gunner MR (2009) Using multiconformation continuum electrostatics to compare chloride binding motifs in α-amylase, human serum albumin, and Omp32. *J Mol Biol* 387(4):840–856.
4. Henrick K, Thornton JM (1998) PQS: A protein quaternary structure file server. *Trends Biochem Sci* 23(9):358–361.
5. Nicholls A, Sharp KA, Honig B (1991) Protein folding and association: Insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins* 11(4):281–296.
6. Bharadwaj R, Windemuth A, Sridharan S, Honig B, Nicholls A (1995) The fast multipole boundary element method for molecular electrostatics: An optimal approach for large systems. *J Comput Chem* 16(7):898–913.
7. Rocchia W, Alexov E, Honig B (2001) Extending the applicability of the nonlinear Poisson-Boltzmann equation: Multiple dielectric constants and multivalent ions. *J Phys Chem B* 105(28):6507–6514.
8. Sitkoff D, Sharp KA, Honig B (1994) Accurate calculation of hydration free-energies using macroscopic solvent models. *J Phys Chem* 98(7):1978–1988.
9. Spassov VZ, Luecke H, Gerwert K, Bashford D (2001) pK_a calculations suggest storage of an excess proton in a hydrogen-bonded water network in bacteriorhodopsin. *J Mol Biol* 312(1):203–219.
10. Baasov T, Sheves M (1986) Alteration of pK_a of the bacteriorhodopsin protonated Schiff-base—a study with model compounds. *Biochemistry* 25(18):5249–5258.
11. Gilson MK, Honig B (1988) Calculation of the total electrostatic energy of a macromolecular system: Solvation energies, binding energies, and conformational analysis. *Proteins* 4(1):7–18.

```
HR   35   VNVALAGIAILVFVYMGRTIRPGRPRLIWGATLMIPLVSISSYLGLLSGLTVGMIEMPAG   73
BR   15   LGTALMGLGTLYFLVKGMGVSDPDAKKFYAITTLVPAIAFTMYLSMLLGYGLTMVP----   70

HR   95   HALAGEMVRSQWGRYLTWALSTPMILLALGLLADVDLGSLFTVIAADIGMCVTGLAAAMT   133
BR   71   --FGGEQNPIYWARYADWLFTTPLLLLDLALLVDADQGTILALVGADGIMIGTGLVGAL-   127

HR   155  TSALLFRWAFYAISCAFFVVVLSALVTDWAASASS--AGTAEIFDTLRVLTVVLWLGYPI   191
BR   128  TKVYSYRFVWWAISTAAMLYILYVLFFGFTSKAESMRPEVASTFKVLRNVTVVLWSAYPV   187

HR   213  VWAVGVEGLALVQSVGATSWAYSVLDVFAKYVFAFILLR   230
BR   188  VWLIGSEGAGIV-PLNIETLLFMVLDVSAKVGFGLILLR   225
```

Fig. S1. Sequence alignment between HR and BR. Buried ionizable residues in BR and their homologous residues in HR are shown in bold. Buried basic residues are in blue and acidic residues are in red.

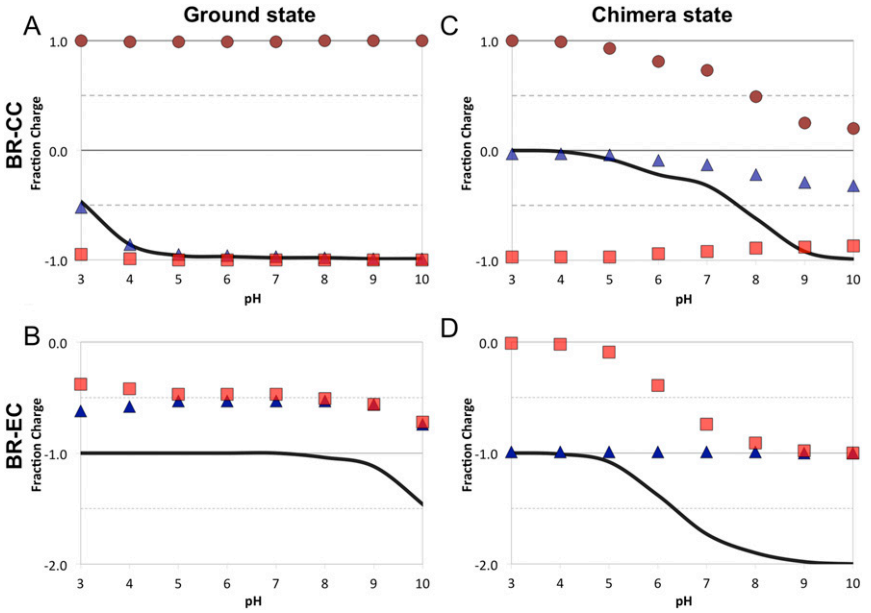

Fig. S2. Calculated pH dependence of the residue fractional charge for (*A* and *C*) central cluster Schiff base (brown circles), BR-D212 (red squares), and BR-D85 (blue triangles); and (*B* and *D*) extracellular cluster BR-E204 (red squares) and BR-E194 (blue triangles) in crystal structures captured in the (*A* and *B*) ground state, and (*C* and *D*) ground/late M chimera structures. The solid line is the cluster net charge. Data for BR are obtained from Song and Gunner (1).

1. Song Y, Gunner MR, Identifying the key conformational changes coupled to proton pumping in bacteriorhodopsin. *J Phys Chem B*, in press.

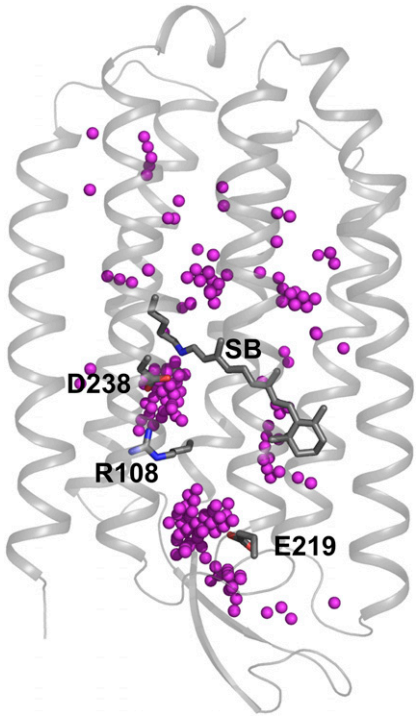

Fig. S3. HR structure with all potential Cl binding sites. Cl (magenta spheres) are filled in the cavity of HR using IPECE (1). Retinal, D238, R108, and E219 are shown.

1. Song Y, Mao J, Gunner MR (2003) Calculation of proton transfers in Bacteriorhodopsin bR and M intermediates. *Biochemistry* 42(33):9875–9888.