Calculation of Proton Transfers in Bacteriorhodopsin bR and M Intermediates†

Yifan Song, Junjun Mao, and M. R. Gunner*

Physics Department J-419, City College of New York, 138th Street and Convent Avenue, New York, New York 10031

Received March 25, 2003; Revised Manuscript Received June 22, 2003

ABSTRACT: Residue ionization states were calculated in nine crystal structures of bacteriorhodopsin trapped in bR, early M, and late M states by multiconformation continuum electrostatics. This combines continuum electrostatics and molecular mechanics, deriving equilibrium distributions of ionization states and polar residue and water positions. The three central cluster groups [retinal Schiff base (SB), Asp 85 and Asp 212] are ionized in bR structures while a proton has transferred from SB^+ to $Asp\ 85^-$ in late M structures matching experimental results. The proton shift in M is due to weaker SB^+-ionized acid and more favorable SB^0-ionized acid interactions following retinal isomerization. The proton release cluster (Glu 194 and Glu 204) binds one proton in bR, which is lost to water by pH 8 in late M. In bR the half-ionized state is stabilized by charge−dipole interactions while full ionization is disallowed by charge−charge repulsion between the closely spaced acids. In M the acids move apart, permitting full ionization. Arg 82 movement connects the proton shifts in the central and proton release clusters. Changes in total charge of the two clusters are coupled by direct long-range interactions. Separate calculations consider continuum or explicit water in internal cavities. The explicit waters and nearby polar residues can reorient to stabilize different charge distributions. Proton release to the low-pH, extracellular side of the protein occurs in these calculations where residue ionization remains at equilibrium with the medium. Thus, the key changes distinguishing the intermediates are indeed trapped in the structures.

Cells use a transmembrane proton gradient to fuel crucial metabolic processes such as ATP synthesis. In the protein bacteriorhodopsin in the archaeon *Halobacterium salinarum*, the energy of a photon absorbed by a retinal drives proton transfer across the cell membrane, generating the needed electrochemical gradient (*1*−*5*). Light absorption leads to retinal isomerization, which initiates protein structural changes. These changes somehow control internal proton transfers, ensuring that proton uptake is from the high pH and proton release to the low-pH side of the membrane-embedded protein.

Bacteriorhodopsin belongs to the class of retinylidene proteins that carry out transmembrane ion transport, photosignal reception, and transduction (*6*, *7*). They all have seven transmembrane α-helices, with a Lys buried in the protein covalently attached to a retinal forming a Schiff base (SB). The Schiff base is protonated in the ground state, and this proton is released from the SB after retinal photoisomerization, triggering conformational changes needed for ion transport (*7*).

The reaction cycle starts with retinal photoisomerization from the all-trans to 13-cis configuration (*4*, *5*). The intermediates are designated bR (ground state), K, L, early and late M, N, and O (Table 1). The submicrosecond changes in K and L are localized on the retinal SB and do not involve proton transfers and, so, will not be discussed here. The SB and nearby Asp 85 and Asp 212 constitute the central cluster. In bR the retinal is in its trans configuration, the SB is protonated, and each acid is ionized. By M a proton has been transferred from the SB to Asp 85 at the start of the extracellular proton pathway. The proton release cluster, comprised of Glu 194 and Glu 204 and perhaps a bound hydronium (*8*, *9*), is 12 Å from the central cluster. In bR this cluster binds a proton which is released into the extracellular space by late M, completing the proton release

Table 1: Ionization States of Functional Residues in Each Bacteriorhodopsin Intermediate[a]

	Intra-cellular (High pH)	Asp 96	SB	Asp 85+212	Glu 194+204	Intra-cellular (Low pH)	SB*
bR		0	+1	-2	-1		T
M_{early}		0	**0**	**-1**	-1		C
M_{late}		0	0	-1	**-2**	***H⁺ gained***	C
N		-1	+1	-1	-2	*H⁺ gained*	C
O	***H⁺ lost***	**0**	+1	-1	-2	*H⁺ gained*	T
bR	*H⁺ lost*	0	+1	**-2**	**-1**	*H⁺ gained*	T

[a] Bacteriorhodopsin reaction cycle showing five intermediates with four key protonation sites (*4*, *5*). Protons are bound from the intracellular and released to the extracellular sides of the membrane. In an energized membrane this moves protons from high- to low-pH regions. Charges on Asp 85 + Asp 212 and on Glu 194 + Glu 204 are summed. A group or cluster charge is in bold for the step where it has changed charge. SB* gives the configuration of the retinal Schiff base (T, all-trans; C, 13-cis). L and K, the short-lived intermediates that connect bR and M, have the same protonation states as bR and are not shown. The bR state is repeated at the top and bottom to emphasize the cyclic nature of the sequence. The calculations presented here consider the proton release portion of the reaction cycle from bR to late M.

† Supported by the National Institutes of Health, Grant RO1-GM64540.

* To whom correspondence should be addressed. Telephone: 212-650-5557. Fax: 212-650-6940. E-mail: gunner@sci.ccny.cuny.edu.

10.1021/bi034482d CCC: $25.00
Published on Web 08/02/2003

portion of the reaction cycle which will be discussed here. The next stage occurs in the cytoplasmic half-channel as the protonated Asp 96 reprotonates the SB, forming the N intermediate. Following reisomerization of the retinal, Asp 96^- takes up a proton from the cytoplasmic side, forming the O intermediate. Finally proton transfers from Asp 85 to the extracellular proton release cluster re-forms the bR ground state.

Using lipidic cubic phase crystallization (*10*), it has been possible to obtain a remarkable series of high-resolution bacteriorhodopsin structures trapped in various stages of the photocycle. Structures are available of bR (*11*−*13*), K (*14*, *15*), L (*14*, *16*), early (K. Takeda et al., unpublished; *13*, *17*) and late (*12*, *18*) M, and O-like (*19*) intermediates. Comparison of bR and M structures shows only small differences, with atomic RMS deviations of about 1 Å. The calculations presented here aim to see if the changes required for bR to M proton transfers are trapped in these intermediate structures.

Bacteriorhodopsin has previously been studied by various computational approaches. Continuum electrostatic studies estimated the pK_as of the key residues in bR (*20*, *21*). MD calculations have been used to generate structures for pK_a calculations (*22*, *23*) and to investigate the dynamics of the protein changes along the reaction cycle (*24*). Quantum mechanical calculations have focused on the retinal isomerization in the protein (*25*−*31*). The configuration of Arg 82 in the proton release channel (*32*, *33*), the proton release mechanism (*34*), and the location of the proton in the proton release group (*9*) have been investigated. Most previous studies have focused on the bR, K, and L intermediates.

Multiconformation continuum electrostatics (MCCE) was used to investigate the equilibrium proton distribution within crystal structures of bacteriorhodopsin trapped in bR, early M, and late M. MCCE combines continuum electrostatics and molecular mechanics to determine the equilibrium distribution of position and ionization state of side chains and ligands such as the SB and waters. Polar groups and waters can rearrange to stabilize proton shifts through the reaction cycle. The results show that the bacteriorhodopsin intermediate crystal structures do stabilize the expected ionization shifts in the central and proton release cluster bR and M states.

MATERIALS AND METHODS

Atomic Coordinates. Bacteriorhodopsin bR structures from the Protein Data Bank (*35*), 1C8R (*12*), 1C3W (*11*), 1KG9 (*17*), and 1M0M-1 (model 1 in 1M0M) (*13*), were analyzed. The early M structures 1DZE (Takeda et al.), 1KG8 (*17*), and 1M0M-2 (model 2 in 1M0M) (*13*), illuminated at ≈220 K until the purple to yellow transition has taken place and then frozen to 100 K, were used. The late M structures 1C8S (*12*) and 1F4Z (*18*) were analyzed. These crystals were frozen to 100 K and illuminated while the cyrosolvent stream was blocked for several seconds. 1C8S is a D96N mutant and 1F4Z an E204Q mutant, each of which slows the transition from M to N. There are only small differences between the atomic positions in the different coordinate files. The RMSD among the three structures, 1C8R (bR), 1DZE (early M), and 1C8S (late M), are 0.5−0.7 Å for backbone and ∼1 Å for backbone and side chains. Except for 1M0M-2 all M structures have the SB nitrogen pointing toward the cytoplasm and Arg 82 pointing toward the extracellular space. Mutated residues in the structure files were replaced by the wild-type residues in the analysis. The 1M0M-1 structure has a very nonstandard conformation for Asp 212, which was relaxed in the Swiss-PDB Viewer (*36*). No other minimization was carried out.

A full-length bacteriorhodopsin has 248 residues. No structure is complete. Missing residues in the N- and C-termini and the E−F loop (Appendix I in Supporting Information) could change the results by exposing the protein interior to solvent or by removing long-range electrostatic interactions. Comparison of the loss in reaction field (desolvation) energy in the smaller 1C8S and the other proteins shows that the deleted residues do not significantly change the solvent exposure of central or proton release cluster residues (see Appendixes III-A and IV-A in Supporting Information). The loss of pairwise interactions with missing residues also creates only minor errors (see Appendix II in Supporting Information for pairwise interactions of all side chains with cluster residues).

MCCE. Multiconformation continuum electrostatics (MCCE) (*37*−*39*) calculates the equilibrium distribution of the protein side chain, buried water, and ligand position and ionization states at a given pH. Backbone and nonpolar side chain positions are fixed. Polar and ionizable side chains and ligands have multiple predefined atomic positions and ionization states, called conformers. A rotamer library (*40*, *41*) provides additional side chain positions. The 44 polar and 28 ionizable residues yield 400−500 conformers in the different protein structures. The SB has 1 ionized and 1 neutral conformer, Asp 85 has 2 ionized and 5−9 neutral conformers, Asp 212 has 3 ionized and 14−18 neutral conformers, Glu 194 has 4−5 ionized and 19−33 neutral conformers, and Glu 204 has 2−3 ionized and 9−20 neutral conformers. Other residues have similar numbers of available conformers. Choosing one conformer for each residue and ligand yields a microstate. These are subjected to Metropolis sampling, yielding the Boltzmann distribution of conformers. Conformation and ionization degrees of freedom come to equilibrium in a single simulation.

The energy for each conformer in a microstate is the sum of five precalculated terms: (1) reference energy; (2) reaction field energy; (3) torsion energy; (4) pairwise electrostatic and Lennard-Jones interactions with fixed groups (backbone and nonpolar side chains); (5) pairwise electrostatic and Lennard-Jones interactions with all other conformers in that microstate. MCCE calculates only the shift in ionization free energy as a residue is moved from water into the protein (*42*, *43*). Protein microstates are created, given one conformer for each group including residues, cofactors, and buried waters. The energy of microstate x $(\Delta G_x)^1$ is

$$\Delta G_x = \sum_{i=1}^{M} \delta_x(i)\{\gamma(i)[2.3k_bT(\mathrm{pH} - \mathrm{p}K_{\mathrm{sol},i})] + (\Delta\Delta G_{\mathrm{rxn},i} + \Delta G_{\mathrm{torsion},i} + \Delta G_{\mathrm{pol},i} + \Delta G^{\mathrm{nonel}}_{\mathrm{fixed},i})\} + \sum_{i=1}^{M} \delta_x(i) \sum_{j=i+1}^{M} \delta_x(j)[\Delta G^{\mathrm{el}}_{ij} + \Delta G^{\mathrm{nonel}}_{ij}] \quad (1)$$

The first line accounts for the reference chemistry of the

group in solution (pK_a) and the solvent conditions (pH), the second line accounts for conformer self-energies and interactions with fixed portions of the protein, and the third line treats interactions between conformers in the considered microstate. k_bT is 25 meV (0.43 ΔpK_a units = 0.58 kcal/mol at 20 °C), M is the total number of conformers, and $\delta_x(i)$ is 1 for selected conformers and 0 for all others. $\gamma(i)$ is 1 for bases, −1 for acids, and 0 for neutral conformers (polar groups, waters, neutral acids, and bases). $pK_{sol,i}$ is the solution pK_a of group i, $\Delta\Delta G_{rxn,i}$ the difference between the conformer reaction field energy in solution and protein, $\Delta G_{torsion,i}$ the conformer torsion energy, $\Delta G_{pol,i}$ the conformer electrostatic interaction with the backbone dipoles, and ΔG_{fixed}^{nonel} the Lennard-Jones interaction with the backbone and side chains with no conformers. ΔG_{ij}^{el} and ΔG_{ij}^{nonel} are pairwise interactions between side chain conformers i and j. Monte Carlo sampling yields conformer occupancies in a Boltzmann distribution of states as a function of pH.

All electrostatic terms are calculated with the finite difference Poisson−Boltzmann procedure with the program Delphi (*44*). The protein dielectric constant is 4, while 80 is used for the solvent. The external salt concentration is 0.15 M. PARSE parameters provide atomic charges and radii (*45*). Torsion and Lennard-Jones parameters were previously reported (*37*, *39*). The SOFT function (*39*) was not used. See ref *39* for a complete description of MCCE.

The retinal in bacteriorhodopsin is covalently attached to the side chain nitrogen of Lys 216, forming a Schiff base (SB) which can be neutral or protonated. In bR the retinal is in the all-trans configuration while in M it is 13-cis. The same atomic charges (*9*) were used for both isomers. The SB pK_{sol} is set to 7 (*46*).

Orientation of the Neutral Schiff Base Dipole. The retinal moves from all-trans to 13-cis on the transition from bR to M. This changes the orientation of the SB nitrogen, so it points toward Asp 85 and Asp 212 in bR and away from them in M. An angle, δ, is defined for the SB C15−NZ−CE dipole as described in Appendix V in Supporting Information. When the SB is pointing down toward the two acids, δ is negative. A value between −90 and 90 indicates that the N is pointing toward Asp 212 rather than Asp 85.

IPECE (Implement Protein Environment for Continuum Electrostatics). Bacteriorhodopsin is an integral membrane protein. The protein is placed in a 30 Å low dielectric slab oriented to bury the fewest surface ionizable residues. To determine the optimal orientation and burial depth, the exposed surface areas (SA) of Asp OD1 and OD2, Glu OE1 and OE2, Arg NH1 and NH2, and Lys NZ were determined with SURFV (*47*) with all ligands and water molecules stripped off. For each position of the protein in the slab, burial of ionizable group is determined by $\sum(\gamma_i SA_i)$. SA is the exposure without the slab. γ is 1 for each terminal oxygen or nitrogen of Asp, Glu, and Arg and 2 for Lys in the slab and 0 if it is exposed. Random trial movements translating the protein parallel to the slab normal or rotating the protein around an axis along the slab minimize the score. If 20000 movements do not lower the score, the slab position is accepted. The protein is placed in a 0.55 Å lattice with all protein atoms at least 10 Å from its edge. Lattice points are assigned to protein, solvent, or intraprotein cavity. Cavities have a radius of >2.2 Å with no surface access through a 2.2 Å corridor. Dummy atoms are placed at every forth lattice point at z coordinates inside the slab and outside the protein and cavities. This creates a low dielectric region for DelPhi calculations.

Waters with more than 10% of their surface exposed without a slab are always deleted from the structure file in MCCE (*39*). For calculations with continuum waters in cavities all waters were removed. Added side chain rotomers can fill internal cavities. Calculations with continuum water were run twice. In the second run all buried conformers that were not selected by Monte Carlo sampling were deleted. This provides a more reproducible water cavity while maintaining important side chain positions.

Calculations with explicit waters were carried out with 1C8R, 1DZE, and 1C8S. All buried crystallographic waters were included. Extra water oxygens were added at every fifth IPECE lattice point (2.75 Å) to fill cavities. For the central and exit cluster extra water oxygens were added to make good hydrogen bonds to the protein. Standard MCCE algorithms added protons to the oxygens so waters can donate or accept hydrogen bonds to neighboring residues (*37*). Each water oxygen gets ≈10 conformers differing in their proton position. Each water also has a conformer with no interactions with the protein representing movement into the bulk. The atomic charge is −0.8 on oxygen and 0.4 on each hydrogen. The reaction field energy of an individual water in bulk water is −1.30 ΔpK_a unit.

Reduced Model. The 400−500 conformers in each structure generate an extremely large number of possible microstates. This makes it difficult to identify the factors that control ionization changes. A reduced model is used to simplify the complexity. Here analysis focuses on a single cluster composed of several residues where strong interactions yield interdependent ionization states. Conformational changes of noncluster residues, which are weakly coupled to cluster residue ionization states, are ignored. The noncluster residues interact with the cluster via a mean-field approximation given their occupancy in a full MCCE calculation. For the cluster, only the most populated ionized and neutral conformers in MCCE are considered. Therefore, for n residues in the cluster, there are 2^n ionization states. So for SB, Asp 85, and Asp 212 in the central cluster there are 8 microstates and for Glu 194 and Glu 204 in the proton release cluster there are 4 microstates. The primary advantage of the reduced model is that the number of microstates is now enumerable so all state energies can be compared. In the reduced model, the M conformers are split into m residues in the cluster and M' surrounding residues. The pairwise energy in a given cluster ionization microstate is

$$\{\sum_{i=1}^{m}\delta_n(i)\sum_{j=i+1}^{m}\delta_n(j)[\Delta G_{ij}]\} + \{\sum_{i=1}^{m}\delta_n(i)\{\sum_{j=1}^{M'}\rho(j)[\Delta G_{ij}]\}\} + \{\sum_{i=1}^{M'}\rho(i)\sum_{j=i+1}^{M'}\rho(j)[\Delta G_{ij}]\} \quad (2)$$

When calculating the full microstate energy in the reduced model, eq 2 replaces the third line in eq 1 which contains the conformer−conformer pairwise interactions in the full

[1] All free energies, ΔG, refer to standard conditions at pH 7 ($\Delta G°'$). 1 ΔpK_a unit = 1.36 kcal/mol, the energy needed to change a pK_a by 1 pH unit at 20 °C.

MCCE energy function. Here $\rho(j)$ is the Boltzmann probability of noncluster conformer j in the full MCCE calculations. The first term contains the interaction among cluster conformers, calculated explicitly; the second term, $\sum_{j=1}^{M'}\rho(j)[\Delta G_{ij}]$, has the mean-field interactions between cluster conformer i and all noncluster conformers; the third term has the interactions between noncluster conformers, which is the same for all cluster microstates.

In addition to determining the total microstate energy the energy of each residue in each microstate can be determined. Without intracluster interactions, the protein shifts the energy of a given conformer i by $\Delta\Delta G_{\mathrm{rxn},i} + \Delta G_{\mathrm{pol},i} + \sum_{j=1}^{M'}\rho(j)[\Delta G_{ij}]$; the first two terms are same as in full MCCE and the last term is the mean-field energy from noncluster residues. The energy difference between ionized (I) and neutral (N) conformers of residue k is

$$\Delta G'_k = \{(\Delta\Delta G_{\mathrm{rxn},k(I)} + \Delta G_{\mathrm{pol},k(I)} + \sum_{j=1}^{M'}\rho(j)\Delta G_{jk(I)}) - (\Delta\Delta G_{\mathrm{rxn},k(N)} + \Delta G_{\mathrm{pol},k(I)} + \sum_{j=1}^{M'}\rho(j)\Delta G_{jk(N)})\} \quad (3)$$

So without intracluster interaction, the pH where ionized and neutral conformers of residue k are of equal energy is

$$\mathrm{p}K'_k = \mathrm{p}K_{\mathrm{sol},k} - \gamma_k \Delta G'_k \quad (4)$$

where $\mathrm{p}K_{\mathrm{sol}}$ is the $\mathrm{p}K_a$ of the ionizable group in the solution and γ is 1 for bases and -1 for acids. Now the total energy of a given microstate in the reduced model is

$$\Delta G_n = \sum_{i=1}^{m}\delta_n(i)[\gamma_i(\mathrm{p}K'_i - \mathrm{pH})] + \sum_{i=1}^{m}\delta_n(i)\sum_{j=i+1}^{m}\delta_n(j)[\Delta G_{ij}] \quad (5)$$

where γ is 0 for all neutral conformers, 1 for ionized bases, and -1 for ionized acids.

RESULTS

The pH dependence of the ionization of all acidic and basic residues (Figure 1) was determined by MCCE in nine bacteriorhodopsin protein structures trapped in the bR, early M, and late M states. Nine ionizable residues are buried inside the protein: the Schiff base, Asp 85, Asp 96, Asp 115, Asp 212, Glu 194, Glu 204, Arg 82, and Arg 134. The Schiff base, Asp 85, and Asp 212 are in the center of the protein forming the central cluster. Glu 194 and Glu 204 form the proton release cluster on the extracellular surface. Residue ionization states within these two groups are strongly interdependent, shifting through the reaction cycle as will be described in detail below (Tables 1 and 2). All other buried residues maintain the same ionization state in all intermediates (Appendix I in Supporting Information). The buried Asp 96 and Asp 115 are neutral from pH 0 to pH 10. Arg 82 is located between the central and proton release clusters, and Arg 134 is near the extracellular surface. Both are deeply buried, losing more than 4.5 $\Delta\mathrm{p}K_a$ units (6.2 kcal/mol) of reaction field energy, which would lower their $\mathrm{p}K_a$s to 8.0 in the absence of other interactions with the protein (*48*). However, both remain fully ionized to pH 14. The negatively charged central and proton release clusters

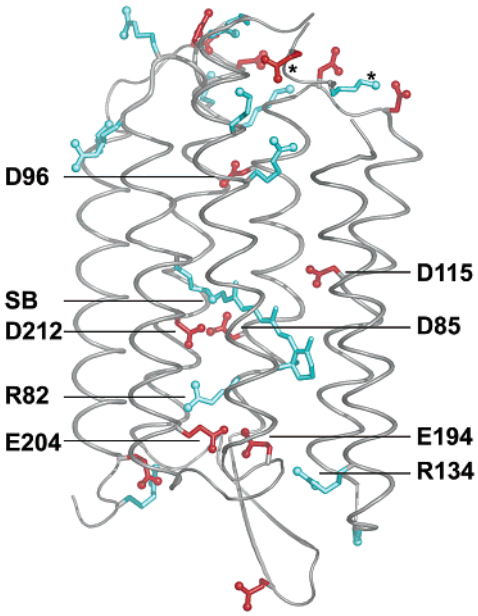

FIGURE 1: Ionizable residues in bacteriorhodopsin. Bacteriorhodopsin bR structure 1C8R. Acidic (red) and basic (blue) side chains with terminal nitrogens and oxygens are accented as spheres. All deeply buried residues are labeled. The retinal Schiff base (SB) is in the protein center. SB, D85, and D212 are the central cluster, and E194 and E204 are the proton release cluster. (*) Lys 159 and Glu 161 and the backbone connecting them are not resolved in 1C8R. These were added from the early M structure (1DZE).

stabilize ionization of Arg 82 by more than 12 $\Delta\mathrm{p}K_a$ units (16.3 kcal/mol). Hydrogen bonds with the backbone carbonyls of Glu 194, Ala 126, and Thr 128 favor ionization of Arg 134 by >9 $\Delta\mathrm{p}K_a$ units. There are 19 exposed Arg, Lys, Asp, and Glu on the cytoplasmatic surface and 4 on the extracellular side. With the exception of Lys 30 in 1M0M-1, Glu 74 in 1KG8 and 1C8S, and Lys 129 in 1DZE, they are all >90% ionized at neutral pH (Appendix I in Supporting Information). All Tyr in the protein are neutral at pH 7. The buried Tyr 57, Tyr 83, and Tyr 185 remain neutral to high pH in all structures.

Central Cluster Titration. The calculations to be discussed first have all waters deleted and cavities filled by DelPhi with $\epsilon = 80$. Calculations with explicit buried waters will be considered below. The retinal Schiff base (SB) and two acids, Asp 85 and Asp 212, are in the center of bacteriorhodopsin. In all three bR structures (1C8R, 1C3W, and 1KG9) the central cluster has a net charge of -1 (pH 5−10) (Figure 2) and the Schiff base is 83−97% ionized (pH 0−11) (Table 2). When the cluster charge is -1 (pH 5−10), any residual neutral SB is matched by protonated Asp 85 or Asp 212. The central cluster in 1C8R and 1C3W begins to lose a proton from SB^+ at pH >11 but not in 1KG9. The $\mathrm{p}K_a$ for the cluster moving from a charge of 0 to -1 is ≈3 (Table 2). 1M0M-1 and 1M0M-2 are resolved in crystals with 40% bR and 60% early M (*13*). While 1M0M-1 is identified with bR and 1M0M-2 with M, the results of each lie between bR and M. Thus, less SB is ionized in 1M0M-1 than in any pure bR structure while there is more SB^+ in 1M0M-2 than in any M structure (Table 2). The results from these structures will be treated separately.

In early M (1DZE and 1KG8), the cluster still has a charge of -1 but the SB is now 88−95% neutral (pH 6−11) (Figure 3). The cluster proton is shared by Asp 85 and Asp 212

Table 2: Fraction Cluster Residue Ionization at pH 7, pK_as for Total Cluster Ionization, and Schiff Base Orientation in Each Structure[a]

	PDB	resolution (Å)	PDB author	fractional ionization, pH 7					pH dependence of summed charge of both clusters		
				SB	Asp 85	Asp 212	Glu 194	Glu 204	lower pK_a	higher pK_a	δ
bR	1C8R	2.3	Luecke et al.	0.95	0.98	0.97	0.00	1.00	3.5	13.5*	−61.6
	1C3W	1.55	Luecke et al.	0.83	0.93	0.90	0.01	1.00	3.5	12.0*	−69.8
	1KG9	1.81	Facciotti et al.	0.99	1.00	0.99	0.17	0.83	2.5	11.1*	−65.0
	1M0M-1	1.43	Lanyi and Schobert	0.62	0.73	0.90	0.01	0.99	3.1	11.5*	−64.3
early M	1DZE	2.5	Takeda et al.	0.12	0.67	0.43	0.00	1.00	5.0	12.6*	75.9
	1KG8	2.0	Facciotti et al.	0.05	0.14	0.91	0.00	1.00	4.2	11.6	99.5
	1M0M-2	1.43	Lanyi and Schobert	0.32	0.39	0.92	0.00	1.00	4.8	10.2	−51.3
late M	1C8S	2.0	Luecke et al.	0.13	0.07	0.92	0.51	0.65	5.5	8.3	102.8
	1F4Z	1.8	Luecke et al.	0.00	0.00	1.00	0.79	0.21	3.3	12.4	122.6

[a] Lower and higher pK_a from data in Figures 2−4. *: pK_a of total ionization of two clusters plus Asp 115. δ: angle defining orientation of SB relative to Asp 85 and Asp 212 (Appendix IV in Supporting Information).

(Table 2). The neutral cluster stability has increased, moving its pK_a from 3 in bR to ≈4.5 in early M. At high pH (>11) the cluster begins to lose a proton. Thus, moving from bR to early M, the Schiff base changes from being almost fully ionized to mostly neutral. Asp 85 and Asp 212 also change from being fully ionized to sharing a proton. There is 9% protonated Asp 212 in 1KG8 and 57% in 1DZE. The presence of protonated Asp 212 is unexpected (*49*) and will be discussed in detail below.

In late M, the SB is ≈90% neutral from pH 7 to pH 14 (Figure 4) with the cluster proton on Asp 85, in agreement with experiment (*49*). The pK_a of the neutral cluster has shifted up to 5.5 in 1C8S while it remains near 3 in the E204Q mutant 1F4Z (Table 2).

Proton Release Cluster Titration. Glu 194 and Glu 204, located near the bacteriorhodopsin periplasmic surface, are implicated in binding the proton to be released to the extracellular space in late M (*5*, *50*, *51*). While a hydrated hydronium may be the bR proton reservoir (*8*, *9*), no hydronium was included in the model tested here. Since adding another binding site would increase the overall proton affinity, these calculations should provide lower limits for the cluster pK_as. The Glu 194 and Glu 204 ionization states are strongly coupled (Figures 2−4). In all bR and early M structures the cluster has a single proton and a charge of −1 from pH 0 to pH 10. Glu 204 prefers to be ionized and Glu 194 neutral. In bR there is little proton loss even at pH 14, while in early M ionized Glu 194 begins to be seen by pH 10. The pK_a of Glu 204 remains below pH 0. In the late M (1C8S) 50% of the clusters have a −2 charge by pH 8. Here, the titration is bimodal with some proton loss by pH 5. The late M E204Q mutant (1F4Z) has a Gln at position 204. This is replaced by an Glu in the calculation. However, while the central cluster ionization is appropriate for late M, the proton release cluster still binds a proton to pH 10. In addition, the protonated Glu 204 is more stable than in any other structure. Thus the neutral Glu 204, imposed by the mutation, is stabilized by changes that are captured in the crystal structure.

The Reduced Model of Cluster Ionization. The titrations calculated for bR, early M, and late M structures show shifts in cluster proton distributions (Figures 2−4). The aim is to understand how the subtle differences between the structures yield these changes. A reduced model was explored which extracts information from the full MCCE treatment. Each of the three central cluster residues have an ionized and neutral conformer yielding eight microstates: one with a −2 charge, three at −1, three neutral, and one at +1. There are four release cluster microstates: two at −1, one at 0, and one at −2. There are modest differences between cluster ionization in the reduced model and full Monte Carlo analysis especially in late M when residues outside of the cluster change their orientation when the cluster changes ionization state. These errors represent a breakdown of the reduced model mean-field energy approximation.

(A) Why the Fully Ionized Central Cluster Is Favored in the bR State. The three groups in the central cluster are deeply buried. The two acids each lose more than 8.5 ΔpK_a units of reaction field (solvation) energy (Appendix III-A in Supporting Information). The SB, with its delocalized charge (and larger Born radius), loses 3.4 ΔpK_a units, which would lower its pK_a from 7, the solution value (*46*), to 3.6 without other interactions. The rest of the protein favors ionization of each acid by 5−13 ΔpK_a units while destabilizing the ionized Schiff base by 3−5 ΔpK_a units [eq 3, Appendix III-A in Supporting Information (the appendix provides interaction and microstate energies for 1C8R in bR, 1DZE in early M, and 1C8S in late M)]. The general positive potential from the backbone (*52*) contributes 1−4 ΔpK_a units to each cluster residue. The buried Arg 82 and Arg 134 and nearby polar groups also raise the potential. With no intracluster interactions the acids would be ionized and the SB neutral (Appendix III-A in Supporting Information). In bR, since the acids favor ionization of the SB, and each acid is stabilized by the SB^+ enough to overcome their mutual repulsion, the microstate $SB^+85^-212^-$ is lowest in energy (Figure 5A, Appendix III-C in Supporting Information). The other two states with a charge of −1, with the SB neutral and either Asp 85 or Asp 212 ionized, are closest in energy. All microstates with a different net charge are much higher in energy. Thus, if both acids are not ionized, the SB will not be charged.

(B) Why the Proton Shifts from the SB to the Cluster Acids in the M States. In early and late M the fractionally ionized SB is reduced to 0−13% (Table 2) primarily because intracluster interactions stabilize full ionization less (Appendix III-B in Supporting Information). Retinal isomerization lengthens the distance between the basic SB proton and the acids (Appendix V in Supporting Information), reducing the favorable interactions found in bR by ≈0.8 ΔpK_a unit

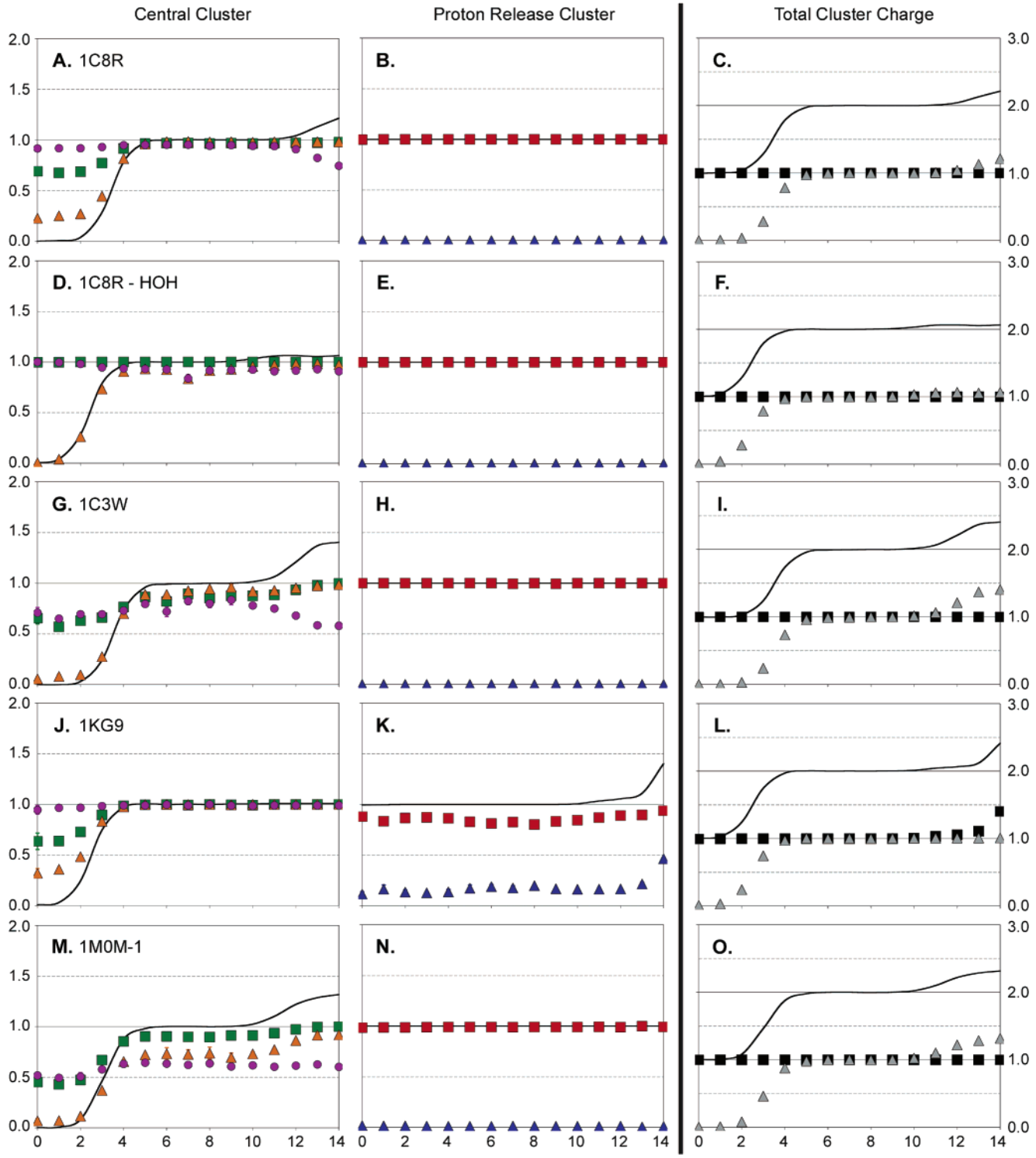

FIGURE 2: Calculated bR titrations. MCCE calculated bR ionization states of the (A, D, G, J, M) central and (B, E, H, K, N) release cluster. (C, F, I, L, O) Sum of charges on the two clusters in (A−F) 1C8R, (G−I) 1C3W, (J−L) 1KG9, and (M−O) 1M0M-1. Calculation with protein cavities filled by (A−C, G−O) continuum and (D−F) explicit water. (A, D, G, J, M) Central cluster: line = −1 × net charge; ionization (●) SB, (▲) Asp 85, and (■) Asp 212. (B, E, H, K, N) Release cluster: line = −1 × net charge; ionization (▲) Glu 194 and (■) Glu 204. (C, F, I, L, O) line = −1 × sum of total charges on the two clusters; −1 × cluster charge in the (▲) central and (■) release cluster. Values and their error bars represent the average and standard deviation of five rounds of Monte Carlo sampling.

(Appendix III-B in Supporting Information). Also, the C15−NZ−CE dipole (δ) is reoriented. When the SB nitrogen points toward the acids as it does in bR, the neutral SB lone pair is destabilized by either ionized acid. When the nitrogen rotates toward the cytoplasm, either ionized acid stabilizes the neutral SB by as much as 3 ΔpK_a units. The noncluster portions of the protein do stabilize acid ionization less and SB ionization more in M than in bR (changes of ≈1−1.5 ΔpK_a units). In particular, Arg 82 moves toward the extracellular release cluster, destabilizing the ionized acids while favoring SB^+ (Appendix II in Supporting Information). The SB and acids also move into a region with smaller

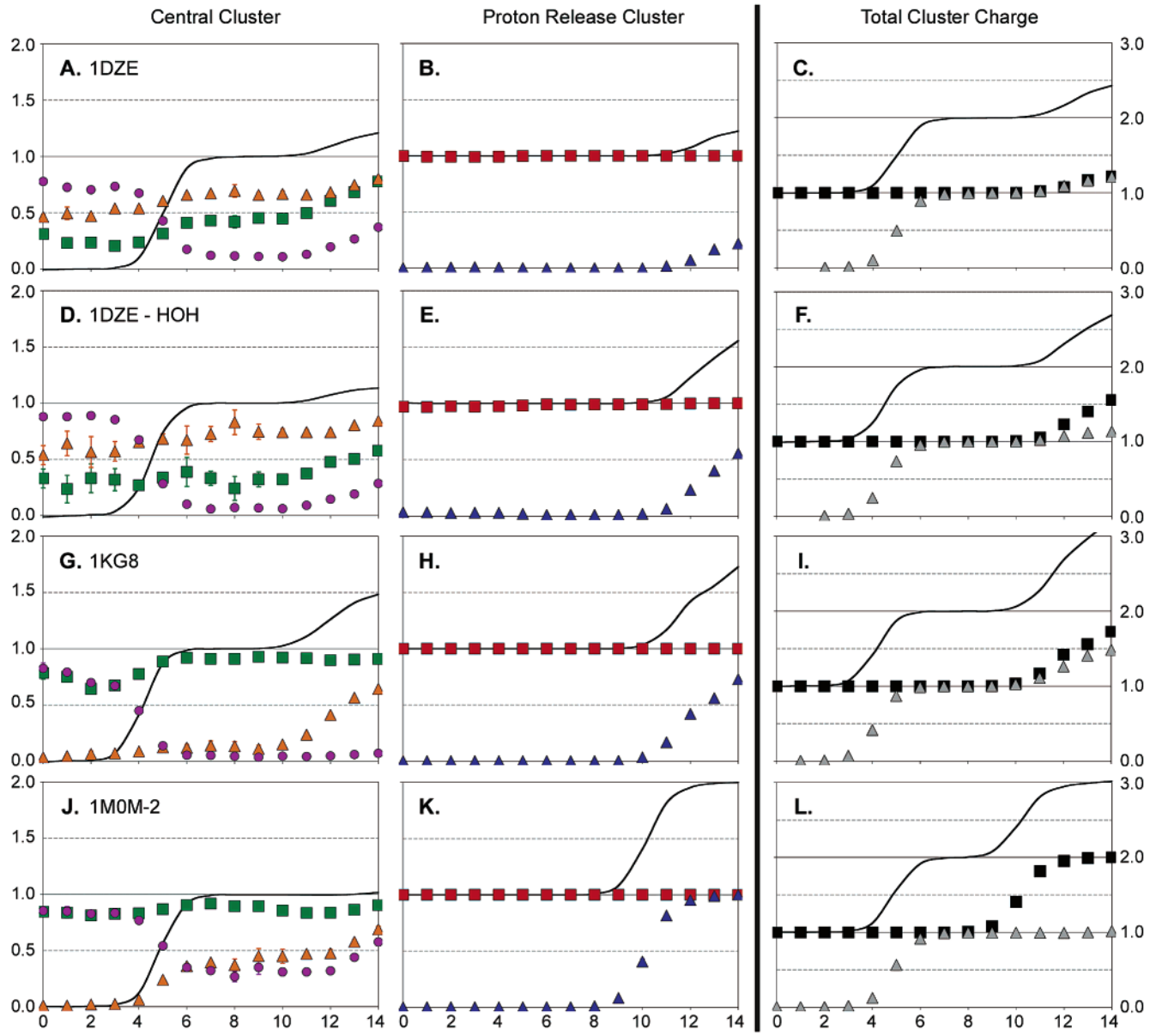

FIGURE 3: Calculated early M titrations. MCCE calculated ionization states in early M of the (A, D, G, J) central and (B, E, H, K) release cluster. (C, F, I, L) Sum of total charges on the two clusters in (A−F) 1DZE, (G−I) 1KG8, and (J−L) 1M0M-2. Calculation with protein cavities filled by (A−C, G−L) continuum and (D−F) explicit water. Lines and symbols are the same as in Figure 2.

positive backbone potential (Table 3-A). Thus, changes in the noncluster portion of the protein shift the balance between protonated SB and acid little since they stabilize SB^+ more and either ionized acid less. However intracluster interactions destabilize SB^+ and stabilize SB^0, shifting the proton off the Schiff base.

(C) What Determines the Distribution of Protonated Asp 85 and Asp 212. While experiments have shown no indication of a protonated Asp 212 in M (*53*, *54*), the calculations show that $SB^0 85^- 212^0$ and $SB^0 85^0 212^-$ are close in energy especially in early M. Four factors control the balance of ionized Asp 85 and Asp 212 in different structures: Thr 89, Tyr 57, Tyr 185, and the dipole orientation (δ) of the neutral SB C15−NZ−CE. Smaller $|\delta|$ (absolute value of δ) and the Thr 89 hydroxyl favors Asp 85^-, while the two Tyr stabilize Asp 212^- (Table 2, Appendix V in Supporting Information). The early M structure 1DZE has the most protonated Asp 212. It has a short distance between Asp 85 and Thr 89 as well as a small $|\delta|$. In late M the proton is bound to Asp 85 in all structures (Figure 4, Table 2) because $|\delta|$ is larger so SB^0 stabilizes Asp 212^- more (Table 2, Appendix V in Supporting Information). In late M the shifted F helix moving Tyr 185 and the reoriented Thr 89 (Appendix II in Supporting Information) also stabilize Asp 212^-.

(D) Why the Proton Leaves the Release Cluster in Late M. Although Glu 194 and Glu 204 are closer to the surface than Asp 85 and Asp 212, they also loose >8 ΔpK_a units of reaction field energy (Appendix IV-A in Supporting Information). As in the central cluster, the rest of the protein stabilizes negative charges on the release cluster. The backbone dipoles stabilize each ionized Glu by more than 2 ΔpK_a units (Appendix IV-A in Supporting Information). Arg 82 and Arg 134 also stabilize negative charges, as do the nearby Tyr 83 and Ser 193. The result is that in the absence of intracluster interactions both acids would have very low pK_as and the microstate $194^- 204^-$ would be fully occupied.

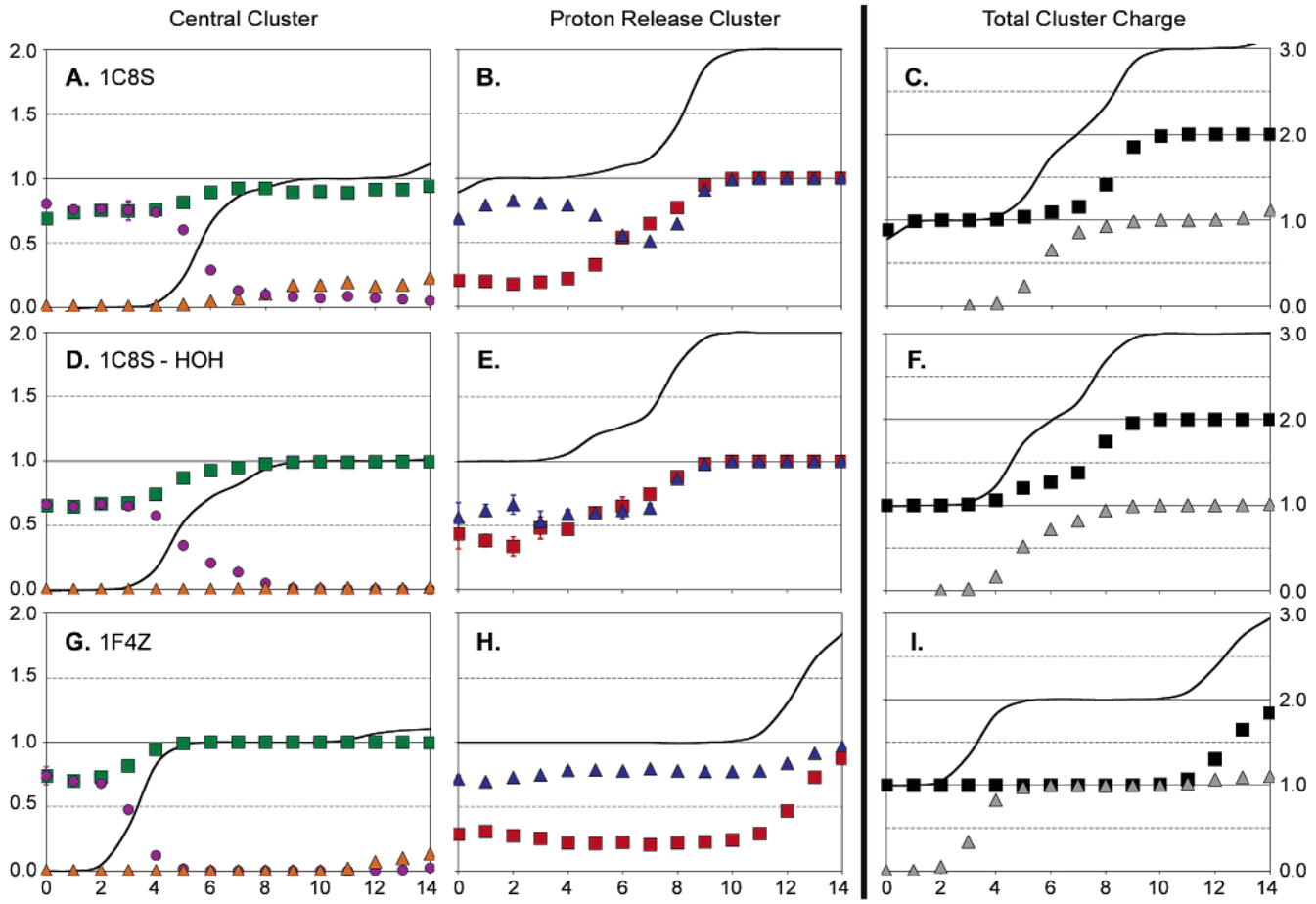

FIGURE 4: Calculated late M titrations. MCCE calculated late M ionization states of the (A, D, G) central and (B, E, H) proton release cluster and (C, F, I) sum of total charges on the two clusters in (A−F) 1C8S and (G−I) 1F4Z. Calculation with protein cavities filled by (A−C, G−I) continuum and (D−F) explicit water. Lines and symbols are the same as in Figure 2.

It is the mutual repulsion of the charges on the two acids that destabilizes the fully ionized microstate in each intermediate (Appendix IV-B in Supporting Information).

In bR, $194^{-}204^{-}$ is disallowed by interacid repulsions of 10−15 ΔpK_a units in different structures. In addition, the cluster charge−dipole interactions favor the singly ionized state by 2−3 ΔpK_a units. The backbone amide interactions with Glu 204^{-} are ≈2.5 ΔpK_a units more favorable than with Glu 194^{-}, helping to decide between the two singly ionized microstates. Since the fully neutral state loses interactions with the positive protein potential, it can be above the fully ionized microstate in energy.

In early M, Glu 204 has twisted slightly around its CG−CD bond, and Glu 194 moves away from the extracellular surface; Arg 82 has moved down toward the release cluster stabilizing ionization of Glu 194 and to a lesser extent Glu 204, and Tyr 83 moves along with the adjacent Arg 82, providing Glu 194 with a stronger hydrogen bond (Figure 6B, Appendix II in Supporting Information). However, the interacid repulsion is still 6−8 ΔpK_a units, while there is still 2−3 ΔpK_a units of favorable charge−dipole interaction stabilizing the half-ionized cluster. The fully ionized state is therefore still not accessible at pH 7 but is now sufficiently low in energy that it is seen at high pH.

In late M, Glu 194 now points directly up toward the cavity near Arg 82 (Figure 6B), while the OE1−CD−OE2 plane of Glu 204 is perpendicular to the membrane normal. At pH 7, the fully ionized state is only 1.3 ΔpK_a units higher in energy than $194^{0}204^{-}$ so now the fully ionized release cluster is seen at neutral pH. The most important reason is that the two acids have moved apart. This reduces the interacid repulsion to be ≈6 ΔpK_a units and the favorable charge−dipole interactions to be less than 1 ΔpK_a unit. There are also small changes in the interaction with Arg 82 and the backbone (Appendix IV in Supporting Information). In the E204Q mutant 1F4Z structure, the release cluster remains protonated. Arg has moved down, but Glu 194 has not moved up so it is still too close to Glu 204 for both to be ionized. As the protein is trapped with a neutral Glu 204, the preferred state has Glu 194^{-} rather than Glu 204^{-} because Arg 82 is positioned to stabilize Glu 194^{-}, while in the other late M structure, with a wild-type release cluster, both acids have similar interactions with the base.

Coupling between the Central and Release Clusters. The M state is characterized by a shift of the proton from the SB^{+} to Asp 85^{-}, maintaining a central cluster charge of −1. There is no direct electrostatic coupling between this change in charge distribution and release cluster deprotonation. Thus, the early M structures and E204Q late M structure have a protonated release cluster and a neutral SB. Arg 82 does facilitate coupling between the clusters. The $SB^{0}85^{0}212^{-}$ central cluster has a less favorable interaction with the Arg than does $SB^{+}85^{-}212^{-}$. This allows the Arg to move toward the release cluster, stabilizing a −2 release cluster charge. However, Arg movement alone does not ensure proton release as seen in the early M and E204Q structures where

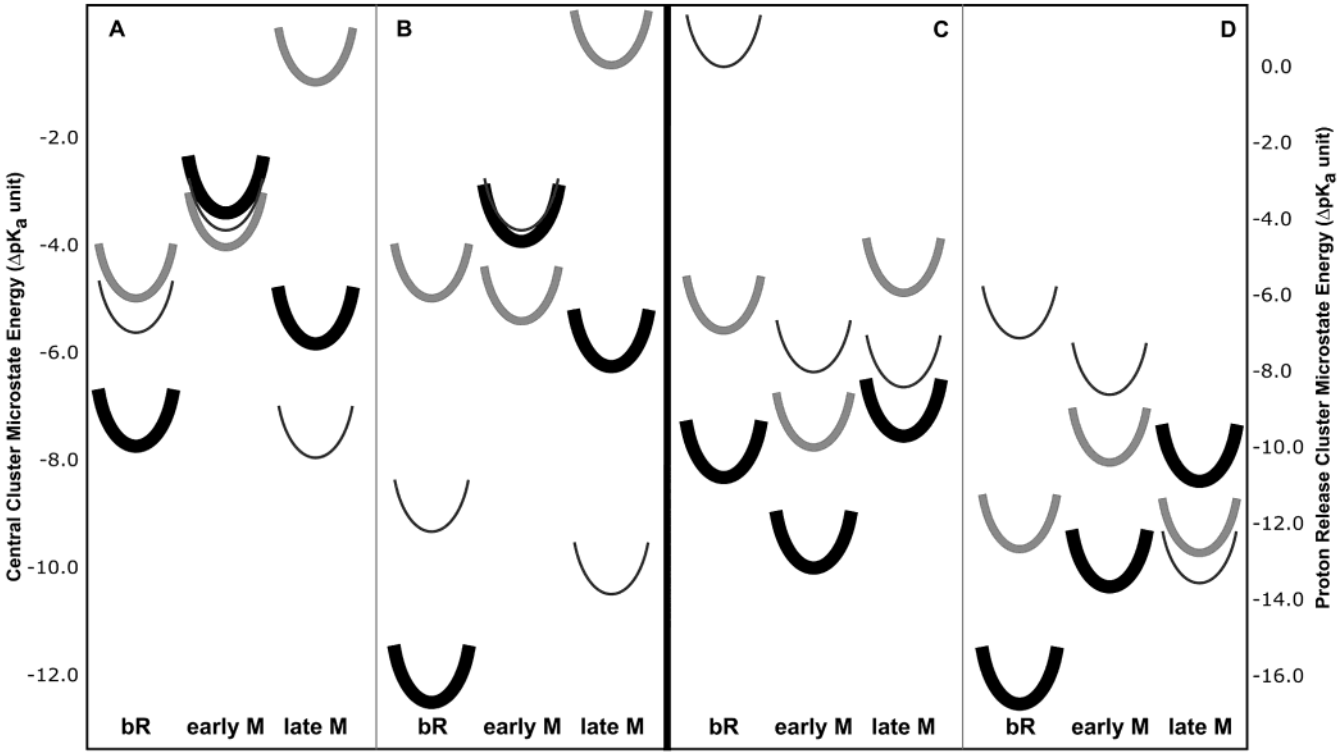

FIGURE 5: Reduced model energy levels of the three central cluster microstates with a charge of −1 and the low-energy proton release cluster microstates. (A, B) Central cluster microstate energies of (thick) $SB^{+}85^{-}212^{-}$, (medium) $SB^{0}85^{-}212^{0}$, and (thin) $SB^{0}85^{0}212^{-}$. (C, D) Proton release cluster of (thick) $E194^{0}E204^{-}$, (medium) $E194^{-}E204^{0}$, and (thin) $E194^{-}E204^{-}$. The pH is 7. (A, C) Continuum and (B, D) explicit water calculations. Each microstate is represented as an arbitrarily shaped harmonic energy well. The minimum energy is derived from the reduced model (Appendix III in Supporting Information).

the two release cluster acids are still too close together so they keep a proton. In late M Glu 194 and Glu 204 move apart, lowering the pK_a for proton release.

While the two clusters are separated by 12 Å (between 85:212 and 194:204 carboxylates), the interaction between their net charges is 2 ΔpK_a units (2.7 kcal/mol), providing direct coupling between their total ionization. Some of the differences between the titrations calculated with different structures trapped in the same state reflect this. For example, in bR structures 1C8R and 1C3W the central cluster begins to lose a proton, moving to a charge of −2 at pH >10 while the release cluster retains its −1 charge to pH >14. In 1KG9 the results are reversed as the exit cluster begins to be deprotonated and the central cluster keeps its −1 charge. The total charge on the two clusters is quite similar in all structures (Figure 2). Thus, these pK_as in the two clusters are close. Small structural differences change the ordering of the cluster pK_as. Their interaction then ensures that ionization of one cluster suppresses the proton release from the other.

The 1C8S (late M) titration highlights the coupling of the cluster net charges (Figure 4). The central cluster 0 to −1 and proton release cluster −1 to −2 titrations are now bimodal with pK_as at 4.5 and 7.5 as the two clusters move from a net charge of −1 (pH <4) to −3 (pH >8). The break in the global titration at pH 6 occurs when the total charge is −2, with ≈74% of the central cluster and 26% of the proton release cluster deprotonated. The cluster pK_as in M differ from those in bR because Arg 82 motion now stabilizes the central cluster 0 charge state. The release cluster −2 state is also stabilized by the Arg motion as well as by separation of its two acids.

Addition of Explicit Waters to Cavities inside the Protein. (A) Calculations with Explicit Waters. MCCE uses a low, continuum dielectric constant for the protein of 4 but allows polar group reorientation which stabilizes the equilibrium charge distribution. This provides an effective dielectric greater than 4 and yields information about the atomic details of the response. When there are cavities in a protein, DelPhi fills them with high dielectric ($\epsilon = 80$). Nearby charges are now stabilized by the continuum reaction field energy which models an averaged water orientation. Calculations were initially carried out without explicit waters because MCCE calculations are sensitive to the preselection of starting conformers. If a conformer capable of a favorable hydrogen bond is not included, charge state energies may be wrong. This is a particular problem with waters which have translational and rotational degrees of freedom which need to be carefully considered when assigning initial positions.

MCCE titrations were calculated in bR (1C8R), early M (1DZE), and late M (1C8S) intermediates with explicit waters in cavities. There are 13 internal water oxygen positions in the bR, 10 in the early M, and 10 in late M structure. Additional water oxygens were added as described in Materials and Methods. Each oxygen has protons placed to optimize local hydrogen-bonding opportunities, resulting in 896, 279, and 313 water conformers for the 66, 36, and 34 oxygen positions tested in bR, early M, and late M. In Monte Carlo sampling waters have a conformer that has left the protein. This is chosen if the interaction of an individual water with bulk solvent (−1.3 ΔpK_a unit) is more favorable than those found in this binding site (*38*). Cavities were overfilled, including oxygens separated by small translations which will compete during Monte Carlo sampling. There are

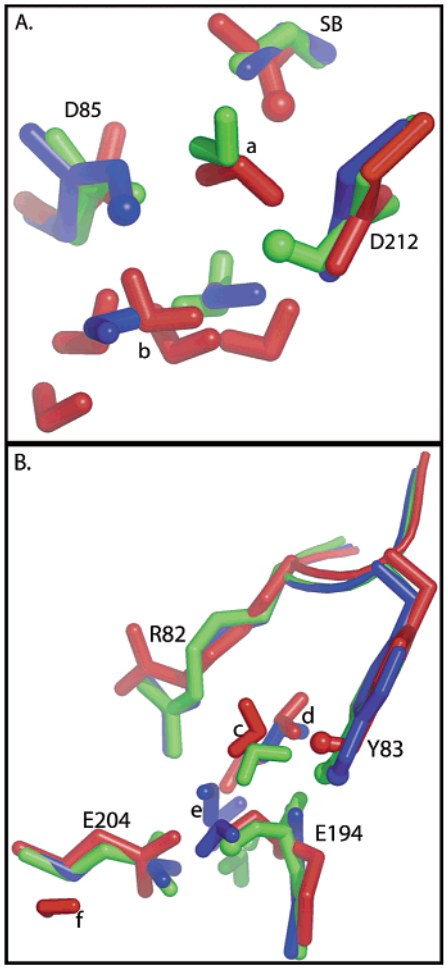

FIGURE 6: Water molecules in the (A) central and (B) proton release cluster. Waters reorient to stabilize ionization states in (red) bR, (green) early M, and (blue) late M states. Waters with occupancy of >40% in the Monte Carlo sampling are shown. Thinner line in (B): backbone connecting R82 and Y83. The active, transferred proton on cluster residues and the Y83 hydroxyl proton are small spheres. The crystallographic waters often have unfavorable Lennard-Jones interactions with the protein and so are not retained in the Monte Carlo sampling. Waters were added to provide alternate positions and to fill cavities. The numbers refer to the nearest crystallographic waters. (A) Central cluster: (a) Water 402 between SB, Asp 85, and Asp 212 is 97% occupied in bR and early M, stabilizing the proton on SB in bR and on D212 in early M. The cavity is not big enough for water in late M. (b) Water cluster 401/406 (bR, late M) and 603 (early M) lies to the extracellular side of D85 and D212. Waters were added to fill the cavity. The average cavity water occupancy is 1.7−1.9 in bR, 0.61 in early M, and 1.9 in late M. The cavity is smaller in early M. This cluster can rearrange to stabilize ionization of either acid. In late M the waters form a hydrogen bond chain from D85 through W401/406 to $D212^{-}$. (B) Proton release cluster (c) added waters between E194 and Y83 in bR and early M. One accepts a hydrogen bond from Y83 in bR. In early M, Y83 moves downward along with Arg 82 to hydrogen bond to E194, reducing the Y83−water interaction. (d) Water cluster 403/405 and 404 (bR only) with additional waters added to fill the cavity lies between R82, E194, and E204 in bR and late M. Occupied conformers accept a hydrogen bond from Arg 82 and donate one to Glu 204. (e) Waters 608 (early M) and 404 (late M) and additional waters lie between Glu 194 and Glu 204. The average water occupancy in the cavity is 1 in early and 3 in late M. This cavity is closed in bR. One water is a hydrogen bond acceptor from Arg 82 and donor to Glu 204; another is a hydrogen bond donor to both Glu 194 and Glu 204. (f) Water 944 is added to the extracellular side of the cluster in bR. It makes a hydrogen bond to E204.

on average 17.5 water molecules bound in bR, 12.8 in early M, and 7.8 in late M.

(B) Titration with Explicit Water. For both clusters the calculated titrations are similar with continuum or explicit waters in cavities (Figures 2−4). The total central cluster charge is independent of the water model at low and intermediate pH. This cluster retains its single proton somewhat better at high pH with explicit waters. There are some changes in the residue ionization states. At low pH in bR the proton is exclusively bound by Asp 85, in better agreement with the experimental estimates of the pK_as of the two Asps (*53*, *55*, *56*). Above pH 9.5 Asp 85 ionization increases from 92% to 98%, as the cluster as a whole begins to deprotonate. This may capture the small secondary wave in the Asp 85 titrations near pH 9.5 found experimentally (*4*, *57*). This deprotonation is coupled to proton loss from Asp 115. At the plateau of this wave of titration the central cluster has an average charge of −1.07 and Asp 115 of −0.93. In early M explicit waters modify the distribution of protonated Asp 212 and Asp 85 slightly at all pHs. In late M at higher pH explicit waters stabilize $SB^{0}85^{0}212^{-}$ so no occupied microstates have ionized SB or Asp 85.

There are modest changes in release cluster titrations when the cavities are filled with explicit rather than continuum water (Figures 2−4). There are no changes in the bR state. In early M the cluster is more fully deprotonated at high pH. In late M the fully deprotonated cluster is more stable at lower pH as explicit water provides additional stabilization of the $194^{-}204^{-}$ state. This lower pK_a is closer to the experimentally determined value of 5.8 (*51*). Long-range coupling, discussed above, yields the bimodal titration curves.

(C) Reduced Model with Explicit Water. When explicit waters fill cavities residues lose favorable continuum reaction field energy but gain favorable pairwise interactions with water (Appendixes III and IV in Supporting Information). The waters rearrange to stabilize the low-energy ionization states. In general, the most occupied microstate found in the continuum cavity analysis is stabilized most, as seen in the reduction of minor ionization states in the titrations (Figures 2−4).

Waters in the cavities defined by 402 and 401/406 in the bR structure 1C8R are the only ones that influence the central cluster (Figure 6). In bR waters support ionization of all three cluster groups (Appendix III in Supporting Information), although long-range interactions with the buried Arg 82 and Arg 134 destabilize SB^{+} more when interior cavities are filled with explicit waters with $\epsilon = 4$ rather than continuum water with $\epsilon = 80$. The water cluster is closer to Asp 85 than Asp 212 and so stabilizes Asp 85^{-} by about 2 ΔpK_a units and Asp 212^{-} by ≈1 ΔpK_a unit. An added water near 402 is a hydrogen bond donor to both Asp 85 and Asp 212, stabilizing their ionization by 1.5−1.7 ΔpK_a units. This lowers the energy of microstates with ionized Asp 85 and Asp 212. In early M, two waters stabilize Asp 85 ionization by 2.7 ΔpK_a units without influencing Asp 212^{-}. Asp 85 is now ≈5% more ionized and Asp 212 10% more protonated than with continuum water. In late M a chain from the protonated Asp 85, through two waters in the 401/406 cluster to Asp 212^{-} stabilizes the microstate with Asp 85 neutral and Asp 212 ionized by 3.4 ΔpK_a units.

The switch from the averaged water response encapsulated in $\epsilon = 80$ to explicit, movable waters changes the release

cluster behavior in modest ways. In bR, a water added near 403 stabilizes both Glu 194$^-$ and Glu 204$^-$ by 1.5−2 ΔpK_a units, but this is still insufficient to overcome the interacid repulsion. In early M, water stabilizes ionization of both residues by a small amount, increasing cluster ionization at high pH. The water changes microstate energies less in early M than in bR (Table 5), but since the state energies are closer together, it is easier to modify the site titrations. In late M, two waters stabilize ionization of Glu 194 and Glu 204 by $\approx$2 ΔpK_a units. Two others are oriented to stabilize 194$^-$204$^-$ at high pH by 3−4 pK_a units while they reorient to favor the 194$^-$204^0 microstate at low pH.

DISCUSSION

MCCE was used to calculate the Boltzmann distribution of residue side chain position and ionization state in bacteriorhodopsin structures trapped in the bR (1C8R, 1C3W, and 1KG9), early M (1DZE and 1KG8), and late M (1C8S and 1F4Z) intermediates as well as in structures which mix bR and M states (1M0M models 1 and 2) (Appendix I in Supporting Information). All of these structures are very similar (<1 Å RMSD). However, protonation states of the central cluster [Asp 85, Asp 212, and retinal Schiff base (SB)] and the extracellular proton release cluster (Glu 194 and Glu 204) change as expected from numerous earlier experiments (*5*, *50*, *51*, *58*, *59*) in the appropriate structure (Figures 2−4).

The SB is predominantly ionized in bR and neutral in M. In early M both Asp 85 and Asp 212 are partially protonated. By late M the proton held by the SB in bR has been transferred to Asp 85. The proton shift in the central cluster from bR to M is primarily due to the intracluster changes when the SB reorients on retinal isomerization. This moves the dipole to face the intracellular space, favoring the neutral SB and ionized cluster acid while diminishing the ion-pair stabilization between SB$^+$ and the acids. The movement of Arg 82 toward the extracellular space and smaller changes of other residues destabilize ionization of the acids and stabilize, to a lesser extent, ionization of the SB. However, when both acids are not ionized, the SB cannot remain protonated. While the SB$^+$ proton moves out of the central cluster pocket with Thr 89, Tyr 185, and several waters into a less polar environment on retinal isomerization (*5*), this does not influence SB ionization significantly. Isomerization does change long-range interactions with the backbone dipoles and Arg 82 that tend to destabilize SB$^+$ in the trans bR retinal (Appendix II-A in Supporting Information).

The release cluster ionization also shifts, moving from bR through M states. A proton is bound to Glu 194 in bR and early M even at high pH. Significant proton release is not seen until late M where $\approx$50% of the clusters are deprotonated by pH 8. Movements of Arg 82 and Tyr 83 stabilize the doubly ionized state in both early and late M. However, the acids are still too close together to reduce the charge−charge repulsion in the fully ionized state until late M. Also, as in the central cluster, intracluster charge−dipole interactions are important. They stabilize the cluster with a single proton in bR and early M states and are diminished in late M.

Intracluster Shift Rather than Individual Site pK_a Shift. Calculations and experiments show significant ionization changes in the central cluster moving from bR to M states. The SB has a pK_a of >12 in bR and shifts down 6 pH units in M. Asp 85 is predominantly ionized by pH 2 in bR and mostly neutral to pHs above 14 in late M, a remarkable change. If these were isolated residues where a proton must be bound from or lost to solution, a 12 pH unit pK_a shift would indicate a change in proton affinity relative to that of water of 16.3 kcal/mol (12 ΔpK_a units). Given the small differences between bR and M structures, large changes in free energy seem unlikely and are in fact unnecessary. Asp 85 binds a proton from the protonated SB, not solution moving from bR to M (Figure 3). In bR SB$^+$85$^-$212$^-$ is 2 ΔpK_a units above SB^{0}85^{0}212$^-$. With shifts between the 2 states of $\approx$4 ΔpK_a units the order of stability is reversed in late M. Thus, relatively modest changes in the relative energy of competing microstates with different distributions of bound protons are sufficient to yield close to stoichiometric shifts of intracluster protonation.

The Impact of Dipolar Rearrangement. In each intermediate, polar side chains move to stabilize the lowest energy ionization state. The changes in microstate energies with explicit waters in cavities illustrate how polar groups stabilize low-energy microstates while destabilizing minority species (Figure 5). For example, there is a hydrogen-bonding network near Asp 85 and Asp 212 consisting of Thr 89, Tyr 57, Tyr 185, and at least three waters. In late M, when these are equilibrated around the low-energy state (SB^{0}85^{0}212$^-$), they destabilize SB$^+$85$^-$212$^-$ largely by raising the energy of Asp 85$^-$. However, if a late M structure is equilibrated around SB$^+$85$^-$212$^-$, the nearby polar groups now stabilize the imposed Asp 85$^-$ by as much as 4 ΔpK_a units. Thus, MCCE captures the local polarizability of the protein, where specific polar residues reorient to stabilize appropriate ionization states. Since these residues can stabilize different ionization states, they do not in general provide a strong driving force for ionization changes. However, within the heterogeneous protein environment, the polar groups are placed so that they can add a modest bias to the microstate energy differences.

The reorientation of polar groups creates interior regions with high effective dielectric constants. However, this does not lead to significant screening of long-range interactions as would be found in a true medium with a uniform high dielectric constant. Thus, MCCE finds that the central and proton release clusters are coupled through direct long-range interactions while each is in a locally polarizable environment formed by buried waters and polar side chains.

Comparison of Calculated and Experimentally Determined pK_as. Bacteriorhodopsin is active from pH 2 to pH 12.4 (*60*, *61*). In bR below pH 2, the cluster has a net charge of 0 with one acid and the SB protonated. At high pH the protein denatures (*60*, *61*). The pK_a of 2.6 has been assigned to Asp 85 from the rate of thermal isomerization of the SB in the dark (*53*, *55*, *57*) and a pK_a <1 assigned to Asp 212 (*53*, *56*). These are in good agreement with the low-pH cluster titration (Table 2). However, the calculations, which can see ionization of minority species more easily than experiment, suggest that there is a mixture of protonated Asp 85 and Asp 212 at low pH (Figure 2).

There is a minority titration of Asp 85 near pH 9.7 in bR calculations with 1C8R and 1C3W as a cluster with a net charge of −2 is formed. This may correspond to a transition in the thermal isomerization of the SB assigned to a second

wave of Asp 85 deprotonation (*57*). In 1KG9, the −2 state of the release cluster is seen above pH 9, close to experimental release cluster pK_a of 9.2 (*51, 62*). The high-pH deprotonation of the central and release clusters and the buried Asp 115 are interdependent. Asp 115^- has a lower pK_a than the central cluster, so it is the majority species. In 1C8R and 1C3W the central cluster has a lower pK_a than the proton release cluster, so the central cluster and Asp 115 lose one proton while the proton release cluster retains its proton. In 1KG9, the −2 state of the proton release cluster is more stable. Now proton loss from Asp 115 perturbs the release cluster titration.

The calculated early to late M transition releases a proton from the extracellular side of the protein via the proton release cluster, consistent with experiments (*51*). A pK_a at 5.8 for the steady-state rate of transition through the photocycle was assigned to proton loss from the release cluster in the M state (*51*). A late M pK_a of 8.2 was assigned to deprotonation of the central cluster in a D96N mutant (*63*). These pK_as are found in the calculations of the total charge in the wild-type protein (Figure 4); however, here 75% of the exit cluster and 25% of the release cluster have the higher pK_a. This mixing of ionization states shows both that the two clusters interact significantly (2−3 ΔpK_a units) and that in the absence of these interactions the two clusters have pK_as differing by <1 pH unit. Very small changes in energy of either cluster would shift the order of titration. The coupling between the two clusters could play a role in the formation of the N state since the doubly charged release cluster will stabilize the proton transfer from Asp 96, neutralizing the central cluster (Table 1).

Protonation of Asp 212. The calculations suggest that protonated Asp 212 is a transient, minority species in early M. While Asp 212 is needed for bR function (*64*), the proton acceptor in M has been shown by FTIR measurements to be Asp 85 (*53, 54*). Thus, it was not expected that $SB^0 85^- 212^0$ and $SB^0 85^0 212^-$ would be so close in energy in early M. Thr 89, Tyr 57, Tyr 185, the dipole of the neutral SB, and waters near the central cluster all influence the balance between Asp 85^- and Asp 212^- when the SB is neutral. Transient neutralization of Asp 212 could play a functional role, since Asp 212^- interacts more strongly with Arg 82 than does Asp 85^-. Thus, a neutral Asp 212 could lower the barrier to motion of the Arg which couples the proton shift in the central cluster to the release cluster rearrangement (*65*−*67*).

Kinetic or Thermodynamic Control of Proton Pumping in Bacteriorhodopsin. The key to proton pumping is that protons are released to one side of the protein and bound when needed from the other. This could be accomplished by kinetic barriers to proton transfer that change the side of the protein that is accessible to the solvent at different stages of the reaction cycle. For example, if kinetic barriers controlled the direction of proton release, the MCCE equilibrium calculations might show that Asp 96 rather than the extracellular cluster could release a proton in M, which is not found. Rather the MCCE calculations show that the small changes in the structures trapped in different intermediates are sufficient to change the equilibrium protonation states appropriately. The calculations do suggest that central and release cluster deprotonations are in competition in late M, with the proton release cluster pK_a being higher than the central cluster. Here kinetic constraints could favor the required proton loss at the extracellular exit. This would then shift the pK_a of the central cluster up, so it holds onto its proton.

Prior Computational Studies of Ionization States in Bacteriorhodopsin. Residue pK_as have been previously investigated by continuum electrostatic calculations (*9, 20, 21, 68*). Early calculations that studied structures obtained by electron microscopy (*69*) highlighted the importance of Arg 82 in stabilizing the fully ionized state of the central cluster in bR (*20, 21*). Structures prepared for continuum electrostatic analysis by placing an appropriately rotated retinal SB in a bR structure and relaxing the protein by molecular dynamics show that there is little difference in the SB pK_a in bR and K, but the proton affinity is smaller in the L state, preparing the SB for the proton transfers found in M (*68*). More recent continuum electrostatic studies of the bR structure, 1C3W (*11*), using the program MEAD (*21, 70*) provided values for the pH dependence of cluster ionization states close to those found here. This shows that the results from the analysis are robust, being relatively independent of the differences in parameters and approximations used in MEAD and MCCE.

Most early calculations focused on one structure from one intermediate. In contrast, the current work applies one methodology and parameter set to nine structures from three intermediates. The side chain and water relaxation found in MCCE help the structure models of the same intermediate yield similar behavior, while protonation states shift in the different intermediates as expected. However, MCCE omits analysis of several important variables. MCCE fixes the protein backbone, reducing the conformation space that can be sampled in the calculation. These calculations do not include the possibility of a protonated water in the exit cluster as described below (*8, 9*). In addition, the partial charges for the trans retinal Schiff base from Spassov et al. (*9*) are used for both all-trans and 13-cis configurations. Since the long-range potential is most positive near the Schiff base nitrogen, changes in retinal conjugation could influence the proton distribution in the central cluster. A delocalized charge on the SB (*9, 71*) reduces the long-range repulsion from the backbone and Arg 82 that destabilize SB^+ as well as the intracluster interactions, which stabilize SB^+ in bR. Delocalization also diminishes the differences in intracluster interactions on retinal isomerization.

The Composition of the Release Cluster. Glu 194 and Glu 204 change ionization state as the protein changes from bR to M in the results presented here (Figures 2−4). These acids have been previously implicated as the proton release cluster (*72*−*74*), without a firm identification of which remains protonated in bR (*51*). However, FTIR studies do not see a Glu undergoing protonation changes during the photocycle, suggesting that a protonated water species such as H_3O^+ or $H_5O_2^+$ binds the proton during the bR state and releases it in M (*8*). Here Glu 194 and Glu 204 would both be ionized in bR and M states. Continuum electrostatic calculations of bR have explored the stability of $H_5O_2^+$ at a position near water 403 (*9*) (Figure 4). At physiological pH the ionization state with Glu 194^-, Glu 204^-, and $H_5O_2^+$ was found to be stable. It is not yet known how this cluster will behave in the M states. Ionized water molecules were not included in the MCCE calculations reported here. Adding another proton

binding group would increase the proton affinity of the release cluster. Thus, the calculations presented here can be viewed as providing a lower limit for the pK_as in this cluster.

SUPPORTING INFORMATION AVAILABLE

Appendix I, MCCE calculated fractional ionization of each residue at pH 7; Appendix II, mean-field interactions between each residue and the central and proton release cluster residues in calculations with explicit waters; Appendix III, energy terms contributing to cluster microstate energies in the reduced model of the central cluster; Appendix IV, energy terms contributing to cluster microstate energies in the reduced model of the proton release cluster; Appendix V, position of the Schiff base relative to Asp 85 and Asp 212. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Lanyi, J. K. (1993) *Biochim. Biophys. Acta 1183*, 241−261.
2. Lanyi, J. K. (1997) *J. Biol. Chem. 272*, 31209−31212.
3. Haupts, U., Tittor, J., and Oesterhelt, D. (1999) *Annu. Rev. Biophys. Biomol. Struct. 28*, 367−399.
4. Balashov, S. P. (2000) *Biochim. Biophys. Acta 1460*, 75−94.
5. Luecke, H. (2000) *Biochim. Biophys. Acta 1460*, 133−156.
6. Oesterhelt, D. (1998) *Curr. Opin. Struct. Biol. 8*, 489−500.
7. Spudich, J. L., Yang, C.-S., Jung, K.-H., and Spudich, E. N. (2000) *Annu. Rev. Cell Dev. Biol. 16*, 365−393.
8. Rammelsberg, R., Huhn, G., Lubben, M., and Gerwert, K. (1998) *Biochemistry 37*, 5001−5009.
9. Spassov, V. Z., Luecke, H., Gerwert, K., and Bashford, D. (2001) *J. Mol. Biol. 312*, 203−219.
10. Rummel, G., Hardmeyer, A., Widmer, C., Chiu, M. L., Nollert, P., Locher, K. P., Pedruzzi, I. I., Landau, E. M., and Rosenbusch, J. P. (1998) *J. Struct. Biol. 121*, 82−91.
11. Luecke, H., Schobert, B., Richter, H. T., Caartailler, J.-P., and Lanyi, J. K. (1999) *J. Mol. Biol. 291*, 899−911.
12. Luecke, H., Schobert, B., Richter, H. T., Cartailler, J. P., and Lanyi, J. K. (1999) *Science 286*, 255−261.
13. Lanyi, J., and Schobert, B. (2002) *J. Mol. Biol. 321*, 727−737.
14. Schobert, B., Cupp-Vickery, J., Hornak, V., Smith, S., and Lanyi, J. (2002) *J. Mol. Biol. 321*, 715−726.
15. Edman, K., Nollert, P., Royant, A., Belrhali, H., Pebay-Peyroula, E., Hajdu, J., Neutze, R., and Landau, E. M. (1999) *Nature 401*, 822−826.
16. Royant, A., Edman, K., Ursby, T., Pebay-Peyroula, E., Landau, E. M., and Neutze, R. (2000) *Nature 406*, 645−648.
17. Facciotti, M. T., Rouhani, S., Burkard, F. T., Betancourt, F. M., Downing, K. H., Rose, R. B., McDermott, G., and Glaeser, R. M. (2001) *Biophys. J. 81*, 3442−3455.
18. Luecke, H., Schobert, B., Cartailler, J. P., Richter, H. T., Rosengarth, A., Needleman, R., and Lanyi, J. K. (2000) *J. Mol. Biol. 300*, 1237−1255.
19. Rouhani, S., Cartailler, J. P., Facciotti, M. T., Walian, P., Needleman, R., Lanyi, J. K., Glaeser, R. M., and Luecke, H. (2001) *J. Mol. Biol. 313*, 615−628.
20. Sampogna, R. V., and Honig, B. (1994) *Biophys. J. 66*, 1341−1352.
21. Bashford, D., and Gerwert, K. (1992) *J. Mol. Biol. 224*, 473−486.
22. Engels, M., Gerwert, K., and Bashford, D. (1995) *Biophys. Chem. 56*, 95−104.
23. Sandberg, L., and Edholm, O. (1997) *Biophys. Chem. 65*, 189−204.
24. Baudry, J., Tajkhorshid, E., Molnar, F., Phillips, J., and Schulten, K. (2001) *J. Phys. Chem. B 105*, 905−918.
25. Tajkhorshid, E., Baudry, J., Schulten, K., and Suhai, S. (2000) *Biophys. J. 78*, 683−693.
26. Crouzy, S., Baudry, J., Smith, J. C., and Roux, B. (1999) *J. Comput. Chem. 20*, 1644−1658.
27. Nonella, M. (2000) *J. Phys. Chem. B 104*, 11379−11388.
28. Warshel, A., and Chu, Z. T. (2001) *J. Phys. Chem. B 105*, 9857−9871.
29. Baudry, J., Crouzy, S., Roux, B., and Smith, J. C. (1999) *Biophys. J. 76*, 1909−1917.
30. Humphrey, W., Lu, H., Logunov, I., Werner, H. J., and Schulten, K. (1998) *Biophys. J. 75*, 1689−1699.
31. Zhang, G. P., Zong, X. F., and George, T. F. (1999) *J. Chem. Phys. 110*, 9765−9768.
32. Scharnagl, C., and Fischer, S. F. (1996) *Chem. Phys. 212*, 231−246.
33. Nagel, J., Edholm, O., Berger, O., and Jahnig, F. (1997) *Biochemistry 36*, 2875−2883.
34. Sampogna, R. V., and Honig, B. (1996) *Biophys. J. 71*, 1165−1171.
35. Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N., and Bourne, P. E. (2000) *Nucleic Acids Res. 28*, 235−242.
36. Guex, N., and Peitsch, M. C. (1997) *Electrophoresis 18*, 2714−2723.
37. Alexov, E. G., and Gunner, M. R. (1997) *Biophys. J. 72*, 2075−2093.
38. Alexov, E., and Gunner, M. (1999) *Biochemistry 38*, 8253−8270.
39. Georgescu, R. E., Alexov, E. G., and Gunner, M. R. (2002) *Biophys. J. 83*, 1731−1748.
40. Dunbrack, R. L., and Karplus, M. (1994) *Nat. Struct. Biol. 1*, 334−340.
41. Dunbrack, R. L., and Cohen, F. E. (1997) *Protein Sci. 6*, 1661−1681.
42. Gilson, M. K., and Honig, B. (1988) *Proteins 3*, 32−52.
43. Simonson, T. (2001) *Curr. Opin. Struct. Biol. 11*, 243−252.
44. Nicholls, A., and Honig, B. (1991) *J. Comput. Chem. 12*, 435−445.
45. Sitkoff, D., Sharp, K. A., and Honig, B. (1994) *J. Phys. Chem. 98*, 1978−1988.
46. Baasov, T., and Sheves, M. (1986) *Biochemistry 25*, 5249−5258.
47. Sridharan, S., Nicholls, A., and Honig, B. (1992) *Biophys. J. 61*, A174.
48. Gunner, M. R., and Alexov, E. (2000) *Biochim. Biophys. Acta 1458*, 63−87.
49. Fahmy, K., Weidlich, O., Engelhard, M., Tittor, J., Oesterhelt, D., and Siebert, F. (1992) *Photochem. Photobiol. 56*, 1073−1083.
50. Lewis, A., Spoonhower, J., Bogomolni, R. A., Lozier, R. H., and Stoeckenius, W. (1974) *Proc. Natl. Acad. Sci. U.S.A. 71*, 4462−4466.
51. Zimanyi, L., Varo, G., Chang, M., Ni, B., Needleman, R., and Lanyi, J. K. (1992) *Biochemistry 31*, 8535−8543.
52. Gunner, M. R., Saleh, M., Cross, E., ud-Doula, A., and Wise, M. (2000) *Biophys. J. 78*, 1126−1144.
53. Metz, G., Siebert, F., and Engelhard, M. (1992) *FEBS Lett. 303*, 237−241.
54. Fahmy, K., Weidlich, O., Engelhard, M., Sigrist, H., and Siebert, F. (1993) *Biochemistry 32*, 5862−5869.
55. Subramniam, S., Marti, T., and Khorana, H. G. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 1013−1017.
56. Balashov, S. P., Govindjee, R., Imasheva, E. S., Misra, S., Ebrey, T. G., Feng, Y., Crouch, R. K., and Menick, D. R. (1995) *Biochemistry 34*, 8820−8834.
57. Balashov, S. P., Imasheva, E. S., Govindjee, R., and Ebrey, T. G. (1996) *Biophys. J. 70*, 473−481.
58. Brown, L. S., Bonet, L., Needleman, R., and Lanyi, J. K. (1993) *Biophys. J. 65*, 124−130.
59. Subramaniam, S., Marti, T., and Khorana, H. G. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 1013−1017.
60. Druckmann, S., Ottolenghi, M., Pande, A., Pande, J., and Callender, R. H. (1982) *Biochemistry 21*, 4953−4959.
61. Balashov, S. P., Govindjee, R., and Ebrey, T. G. (1991) *Biophys. J. 60*, 475−490.
62. Kono, M., Misra, S., and Ebrey, T. G. (1993) *FEBS Lett. 331*, 31−34.
63. Brown, L. S., and Lanyi, J. K. (1996) *Proc. Natl. Acad. Sci. U.S.A. 93*, 1731−1734.
64. Needleman, R., Chang, M., Ni, B., Varo, G., Fornes, J., White, S. H., and Lanyi, J. K. (1991) *J. Biol. Chem. 266*, 11478−11484.
65. Richter, H. T., Brown, L. S., Needleman, R., and Lanyi, J. K. (1996) *Biochemistry 35*, 4054−4062.
66. Luecke, H., Schobert, B., Richter, H. T., Cartailler, J. P., and Lanyi, J. K. (1999) *Science 286*, 255−261.
67. Brown, L. S., Dioumaev, A. K., Needleman, R., and Lanyi, J. K. (1998) *Biochemistry 37*, 3982−3993.

68. Engles, M., Gerwert, K., and Bashford, D. (1995) *Biophys. Chem. 56*, 95−104.
69. Henderson, R., Baldwin, J. M., Ceska, T. A., Zemlin, F., Beckmann, E., and Downing, K. H. (1990) *J. Mol. Biol. 213*, 899−929.
70. Bashford, D., and Karplus, M. (1990) *Biochemistry 29*, 10219−10225.
71. Hayashi, S., Tajkhorshid, E., and Schulten, K. (2002) *Biophys. J. 83*, 1281−1297.
72. Balashov, S. P., Imasheva, E. S., Ebrey, T. G., Chen, N., Menick, D. R., and Crouch, R. K. (1997) *Biochemistry 36*, 8671−8676.
73. Dioumaev, A. K., Richter, H. T., Brown, L. S., Tanio, M., Tuzi, S., Saito, H., Kimura, Y., Needleman, R., and Lanyi, J. K. (1998) *Biochemistry 37*, 2496−2506.
74. Essen, L., Siegert, R., Lehmann, W. D., and Oesterhelt, D. (1998) *Proc. Natl. Acad. Sci. U.S.A. 95*, 11673−11678.

BI034482D