

Electrostatic Calculations of the pK_a Values of Ionizable Groups in Bacteriorhodopsin

Donald Bashford

The Scripps Research Institute
10666 North Torrey Pines Road
La Jolla, CA 92037, U.S.A.

and Klaus Gerwert

Max-Planck-Institut für Ernährungsphysiologie
Dortmund 4600, Germany

(Received 22 May 1991; accepted 11 November 1991)

The effects of solvation and charge–charge interactions on the pK_a of ionizable groups in bacteriorhodopsin have been studied using a macroscopic dielectric model with atom-level detail. The calculations are based on the atomic model for bacteriorhodopsin recently proposed by Henderson *et al.* Even if the structural data are not resolved at the atomic level, such calculations can indicate the quality of the model, outline some general aspects of electrostatic interactions in membrane proteins, and predict some features. The effects of structural uncertainties on the calculations have been investigated by conformational sampling. The results are in reasonable agreement with experimental measurements of several unusually large pK_a shifts (e.g. the experimental findings that Asp96 and Asp115 are protonated in the ground state over a wide pH range). In general, we find that the large unfavorable desolvation energies of forming charges in the protein interior must be compensated by strong favorable charge–charge interactions, with the result that the titrations of many ionizable groups are strongly coupled to each other. We find several instances of complex titration behavior due to strong electrostatic interactions between titrating sites, and suggest that such behavior may be common in proton transfer systems. We also propose that they can help to resolve structural ambiguities in the currently available density map. In particular, we find better agreement between theory and experiment when a structural ambiguity in the position of the Arg82 side-chain is resolved in favor of a position near the Schiff base.

Keywords: bacteriorhodopsin; electrostatic; titration; membrane; proton pump

1. Introduction

It is an important challenge in the life sciences to understand the function of membrane proteins at the molecular level. The light-driven proton pump bacteriorhodopsin is an attractive system in which to study such functions by a great variety of biophysical and molecular biological methods (Stoeckenius & Bogomolni, 1982; Oesterhelt & Tittor, 1989). Site-directed mutants of bacteriorhodopsin have been studied by FTIR† difference

spectroscopy to provide information about the protonation states of the four internal aspartic acid residues, 85, 96, 115 and 212, in the different intermediate states of the proton pumping cycle (Braiman *et al.*, 1988; Gerwert *et al.*, 1989, 1990). It was found that, on the proton release pathway, Asp85 accepts the Schiff base proton and, on the proton uptake pathway, Asp96 donates a proton to the Schiff base. Mutation of Asp85 and Asp96 leads to an inhibition of the proton release and the proton uptake, respectively (Butt *et al.*, 1989; Holz *et al.*, 1989). Bacteriorhodopsin has also been studied by electron microscopy techniques, which recently have provided diffraction data to about 3 Å (1 Å = 0.1 nm) parallel to the membrane plane and about 10 Å perpendicular to the membrane plane. The density maps from these data, in combination with many available experimental data obtained by molecular biological, spectroscopic and other structural methods, such as neutron diffraction,

This paper is dedicated to Dr Walter Stoeckenius on the occasion of his 70th birthday.

† Abbreviations used: FTIR, Fourier transform infrared spectroscopy; HE, Henderson *et al.* model; HM, energy minimized model; HEDnn, HM models with constrained dynamics; RU, HE model with Arg82 altered with harmonic and electrostatics constraints; RUDnn, RU models with constrained dynamics; r.m.s., root-mean-square; n.m.r., nuclear magnetic resonance.

have been used to construct an atomic model of bacteriorhodopsin (Henderson *et al.*, 1990).

Here, we use a continuum dielectric model and finite difference technique to calculate the pK_a values of the titrating groups of bacteriorhodopsin in its ground state. The calculations include pK_a shifts due to the electrostatic effects of burying titratable groups in a low dielectric medium, interactions with non-titrating charges, such as peptide dipoles, and interactions of titratable groups with each other. Recently, such calculations were performed for lysozyme, and it was found that about half of the calculated pK_a values were within one pK unit of the experimental value, and errors usually were overestimates of pK shifts rather than errors in the directions of the shifts (Bashford & Karplus, 1990). Because comparison of calculations for different crystal structures for lysozyme showed that the method is sensitive to the input structural data, and because the available structural data on bacteriorhodopsin are low resolution, we have used molecular dynamics in the present work to generate an ensemble of conformations on which calculations can be made to assess the effects of conformational uncertainty on the predicted pK_a values. The aim of the present calculations is not so much to predict very precise pK_a values of all titratable groups as to predict protonation states of key residues in the ground state and outline their titration behavior and their interactions.

We find that the calculated pK_a values in bacteriorhodopsin are mostly consistent with experimental findings on the protonation states of residues in the ground state of bacteriorhodopsin. Agreement with experiment is substantially improved by changing the proposed position of Arg82, for which no electron density was seen by Henderson *et al.* (1990), from the "down" position towards the extracellular side of the membrane to the "up" position towards the Schiff base.

2. Methods

(a) Electrostatics and titration calculations

The calculations presented here are based on the idea that the difference in the titration behavior of an ionizable group in a protein and in a model compound can be accounted for by calculating the difference in the electrostatic work of altering the group's charges from the unprotonated to the protonated state in the protein, and the work of making the same alteration in the model compound. We assume that the electrostatic potential is described by the macroscopic Poisson equation:

$$\nabla \epsilon(r) \nabla \phi(r) = -4\pi \rho(r), \quad (1)$$

where ϕ is the electrostatic potential due to the charge density, ρ ; and ϵ is the dielectric constant which takes on the value ϵ_m in the protein interior and ϵ_s in the solvent (Fig. 1). The Poisson equation is solved by the use of finite difference methods (Press *et al.*, 1986). We identify 3 types of electrostatic energy contribution to the differences in a group's titration in the model compound versus the protein: contributions arising from the group's interaction

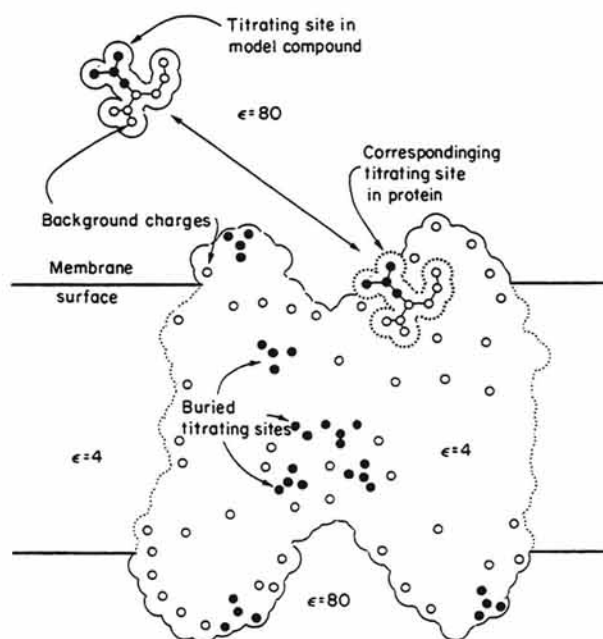


Figure 1. The macroscopic idealization used for the calculations. The protein and membrane are rigid low-dielectric objects with embedded, permanent background charges. Titration is the alteration of charges on a titrating site. The model compound is "carved out" from the parent protein. pK_a differences between the protein and model compound are due to differences in the electrostatic work required to make charge alterations in the protein titrating site and the corresponding model compound.

with the polarization that the group's charges induce in the protein and solvent dielectric environment (Born solvation energy or $\Delta\Delta G_{\text{Born}}$); the group's interaction with partial charges, such as the peptide dipoles (background charge interactions or $\Delta\Delta G_{\text{back}}$); and the group's interaction with other titrating groups in the protein. The interactions with other titrating groups complicate the discussion of titration because the charges of these groups are, themselves, pH dependent; so it is useful to concentrate first on a hypothetical quantity that Tanford has called the intrinsic pK or pK_{intr} (Tanford & Kirkwood, 1957). This is defined as the pK_a that a titrating group in the protein would have if all other groups were held in their neutral protonation states. If it is assumed that electrostatic interactions are the major determinant of the perturbation of titration behavior, then one can calculate the difference between the model compound pK_a and the protein pK_{intr} by calculating $\Delta\Delta G_{\text{Born}}$ and $\Delta\Delta G_{\text{back}}$.

Born solvation energy terms arise from the interaction of a set of charges with their own reaction field; that is, the field due to the polarization that the charges themselves induce in the surrounding medium. They will differ between the protein and the model compounds because of the differing exposure of the titrating groups to the high dielectric solvent. The method of calculating the solvation terms, $\Delta\Delta G_{\text{Born}}$, can be conveniently described in terms of Green functions, $\Phi(r, r')$, defined as the potential produced at point r by a unit charge placed at point r' , where the potential is determined by the Poisson equation. Thus, the field produced by a partial charge, q , at r' is $q\Phi(r, r')$. For the case of the protein with dielectric constant, ϵ_m , in

the ϵ_w solvent, when r and r' both lie in the ϵ_m region, the Green functions will have the form:

$$\Phi_{\text{protein}}(r, r') = \frac{1}{\epsilon_m |r - r'|} + \Phi_{\text{protein}}^*(r, r'), \quad (2)$$

in which the Coulombic $1/r$ term is modified by a Φ_{protein}^* term arising from the presence of the external medium, ϵ_w . The increment of work, δW , required to add increments of charge, δq_i , to each site i in the presence of partial charges, q_j , already on the sites is:

$$\delta W = \sum_{i,j} \delta q_i q_j \Phi_{\text{protein}}(r_i, r_j). \quad (3)$$

For the recharging process, integrating over the q from 0 to their full values, Q , gives:

$$W = \frac{1}{2} \sum_{i,j} Q_i Q_j \Phi_{\text{protein}}(r_i, r_j). \quad (4)$$

Substituting eqn (2) for Φ_{protein} , this becomes:

$$W = \frac{1}{2} \sum_{i,j} \frac{Q_i Q_j}{\epsilon_m |r_i - r_j|} + Q_i Q_j \Phi_{\text{protein}}^*(r_i, r_j). \quad (5)$$

This formula is not a solvation energy as it stands, since it contains Coulombic interactions between internal charge pairs ij , and Coulombic singularities in the ii , or self-energy terms. Solvation energy differences are obtained by subtracting the results of a 2nd calculation in which the same set of charges with the same internal geometry are placed in a model compound. The Coulomb terms cancel in the subtraction, since they are independent of the external dielectric constant and the geometry of the dielectric boundary. Only the Φ^* terms remain:

$$\Delta G_{\text{Born}} = \frac{1}{2} \sum_{i,j} Q_i Q_j [\Phi_{\text{protein}}^*(r_i, r_j) - \Phi_{\text{model}}^*(r_i, r_j)]. \quad (6)$$

In practice, one does not have the analytical form of the Green functions for the complex shape of the proteins and model compounds, so the following procedure for computing the W terms is used: the full set of charges, Q_i , is placed on the molecule and the potential, ϕ , due to these charges is found by finite difference methods, and the quantity:

$$\frac{1}{2} \sum_i Q_i \phi(r_i) \quad (7)$$

is calculated. This is equivalent to eqn (4), since ϕ_s can be expressed in terms of Green functions:

$$\phi(r_i) = \sum_j Q_j \Phi(r_i, r_j). \quad (8)$$

Eqn (7) contains Coulombic charge-charge interaction terms, and Coulomb self-energy terms analogous to the first term on the right-hand side of eqn (5). However, the finite difference equations do not give rise to infinite Coulomb self-energies. Instead, there arises a Coulomb-like artifact of the grid that depends on the charge, the grid spacing and the way in which the physical charge is distributed over the grid points. As before, the Coulombic contributions, including the grid artifacts, must be cancelled by subtraction from a model compound calculation. In addition, since we are interested in pK_a shifts, another subtraction must be made to calculate the work

of going from the unprotonated to the protonated state charges. The expression used to calculate the solvation contribution to pK_a shifts is then:

$$\Delta \Delta G_{\text{Born}} = \frac{1}{2} \sum_i Q_i^{(p)} [\phi_{\text{protein}}^{(p)}(r_i) - \phi_{\text{model}}^{(p)}(r_i)] - \frac{1}{2} \sum_i Q_i^{(u)} [\phi_{\text{protein}}^{(u)}(r_i) - \phi_{\text{model}}^{(u)}(r_i)], \quad (9)$$

where ϕ_{protein} and ϕ_{model} are potential fields calculated by the finite difference method in the protein and model compound, respectively, and the superscripts (p) and (u) indicate the protonated and unprotonated charge states, respectively. The fields are calculated with only the charges of the titrating site under consideration present. To ensure proper cancellation of Coulombic terms and grid artifacts, it is necessary to use the same atomic charges and co-ordinates and the same local arrangement of grid points in the model compound as in the protein (see below).

The 2nd type of contribution to pK_a differences, that due to interactions with non-titrating charges, is calculated using the values of ϕ_{protein} and ϕ_{model} at the locations of those charges:

$$\Delta \Delta G_{\text{back}} = \sum_j q_j [\phi_{\text{protein}}^{(p)}(r_j) - \phi_{\text{protein}}^{(u)}(r_j)] - \sum_j q_j [\phi_{\text{model}}^{(p)}(r_j) - \phi_{\text{model}}^{(u)}(r_j)], \quad (10)$$

where the q_j are the non-titrating partial charges located at r_j . The intrinsic pK_a of the site in the protein is then:

$$pK_{\text{intr}} = pK_{\text{model}} - \frac{1}{2.303 k_B T} [\Delta \Delta G_{\text{Born}} + \Delta \Delta G_{\text{back}}]. \quad (11)$$

In addition to the pK_{intr} of each titrating site in the protein, we require the interactions between titrating sites, or more precisely, the additional work, $W_{\mu\nu}$, required to add a proton to site μ due to the presence of a proton on site ν . This is:

$$W_{\mu\nu} = \sum_i [Q_{\mu,i}^{(p)} - Q_{\mu,i}^{(u)}] [\phi_{\text{protein},\nu}^{(p)}(r_i) - \phi_{\text{protein},\nu}^{(u)}(r_i)], \quad (12)$$

where $\phi_{\text{protein},\nu}$ is the potential field due to the charges on site ν , $Q_{\mu,i}$ is the charge on the i th atom of site μ , and the superscripts (p) and (u) refer to the protonated and unprotonated states, respectively.

Once the $pK_{\text{intr},\mu}$ for each site, μ , and the interaction energy, $W_{\mu,\nu}$, between each pair of sites μ and ν have been obtained, the titration curves for each site can be calculated by taking a Boltzmann weighted sum over all possible protonation states at each pH. The fraction of molecules having site i protonated is:

$$\theta_i = \frac{\sum_{\{x\}} x_i \exp \left(\sum_{\mu} x_{\mu} 2.303 (pK_{\text{intr},\mu} - \text{pH}) - \beta \sum_{\mu,\nu} (x_{\mu} + q_{\mu}^0)(x_{\nu} + q_{\nu}^0) W_{\mu\nu} \right)}{\sum_{\{x\}} \exp \left(\sum_{\mu} x_{\mu} 2.303 (pK_{\text{intr},\mu} - \text{pH}) - \beta \sum_{\mu,\nu} (x_{\mu} + q_{\mu}^0)(x_{\nu} + q_{\nu}^0) W_{\mu\nu} \right)}, \quad (13)$$

where $\mathbf{x}$ is an N -element protonation state vector whose elements, x_{μ} , are 0 or 1 according to whether site μ is

protonated or unprotonated, q_μ^0 is the charge of the unprotonated state of site μ , and $\{\mathbf{x}\}$ indicates summation over all 2^N possible protonation states (Bashford & Karplus, 1991). The $pK_{1/2}$ for a site is defined as the pH at which the site is half-protonated ($\theta = 0.5$).

For the internal dielectric constant, ϵ_m , we use the value 4.0 (Tanford & Roxby, 1972), and for the external dielectric constant, ϵ_w , we use the value 80.0. The boundary between the interior and the exterior is defined as the solvent contact and re-entrant surfaces (Richards, 1977) of a 1.4 Å spherical solvent probe rolling over the van der Waals' surface of the protein. Channels and cavities in the protein are taken to be water-filled and are assigned a dielectric constant of $\epsilon_w = 80.0$. The atomic radii for the protein's atoms and the charges for the protein's non-titrating atoms are taken from the polar hydrogen parameter set, CHARMM19, of the CHARMM computer program (Brooks *et al.*, 1983). The charge parameters used for the atoms of the titrating groups are shown in Table 1. CHARMM19 charges are used where available. For protonated carboxylate groups, charges on the oxygen atoms have been selected to represent protonation without explicitly introducing a new hydrogen atom. For deprotonated arginine, all charges are set to zero to avoid the complexities of tautomerism. Trial calculations (not shown) suggest that the calculations are less sensitive to the details of the charge model used for these side-chains than to conformational changes. For the retinal Schiff base, the charges of both the protonated and unprotonated states are taken from semi-empirical quantum calculations (Tavan *et al.*, 1985). The atomic radii for the retinal are taken from CHARMM19 ring carbon atom parameters. For the reasons described in the above discussion of the calculation of Born energy differences, the co-ordinates for the model compounds, the *N*-formyl *N*-methylamide derivatives of the amino acids, are derived from the protein co-ordinates. The co-ordinates of the atoms of the residue containing site i , along with the peptide C and O of the previous residue and the N, H and C α of the following residue, are taken as the co-ordinates of the model compound corresponding to site i . This selection of atoms gives overall neutrality for the peptide groups when the CHARMM19 (Brooks *et al.*, 1983) charge set is used. It is consistent with the idea implied in Fig. 1, where a model compound is "carved out" of the rigid dielectric of the protein.

The present calculations differ from previous ones (Bashford & Karplus, 1990) in that a membrane is now included. The membrane is represented as a slab having the same dielectric constant as the protein. The cytoplasmic surface is presumed to be even with the C α of Glu166, and the extracellular surface is presumed even with the C α of Glu74. This results in a membrane with a thickness of 45 Å (Henderson & Unwin, 1975; Henderson, 1975). In the finite difference procedures, grid points lying between these planes but more than 10 Å from the center of the channel are marked as "interior" points to be assigned a dielectric constant of 4.0, as are points interior to the protein. Although a lower value than 4.0 would have been more appropriate for the membrane, we were compelled to choose the same value for the membrane and the protein because our computer program allowed only 2 different dielectric constants, 1 for "interior" and 1 for solvent. Calculations without the membrane suggest that the choice of the membrane dielectric constant has only a minor effect. Charges due to the membrane head groups and their counterion atmospheres have not been included (see Discussion). A method of images solution for the potential, with a series of 20 image charges on each side of

Table 1
Charges used in the calculations

	Protonated state	Deprotonated state
Asp		
CB	0.0	-0.16
CG	0.59	0.36
OD-1	-0.16	-0.6
OD-2	-0.43	-0.6
Glu		
CG	0.0	-0.16
CD	0.59	0.36
OE-1	-0.16	-0.6
OE-2	-0.43	-0.6
Tyr		
OH	-0.65	-0.85
CZ	0.25	-0.85
HH	0.4	0.0
Arg		
CZ	0.50	0.0
CD	0.10	0.0
NE	-0.40	0.0
HE	0.30	0.0
NH-1	-0.45	0.0
HH-11	0.35	0.0
HH-12	0.35	0.0
NH-2	-0.45	0.0
HH-21	0.35	0.0
HH-22	0.35	0.0

each dielectric interface, has been used to initialize the outer boundary of the finite difference grid.

(b) Conformational models and generation of alternative conformations

To gain some insight into the effect of structural uncertainties on our calculations, we have used restrained molecular dynamics and minimization to generate a small ensemble of conformations whose range of variation is intended to reflect the range of conformational uncertainties; and we have made electrostatic titration calculations on the members of this ensemble to obtain a range of calculated pK values. Henderson *et al.* (1990) built their model co-ordinates from a density map having an effective Fourier resolution of 3.5 Å in directions parallel to the membrane and 10 Å in the perpendicular direction. To reflect the structural uncertainties that arise at such resolutions, they assigned atoms to 3 categories of accuracy (R. Henderson, personal communication). The category of highest accuracy includes helix backbone atoms, the retinal chromophore and a number of side-chains that could be seen clearly in the map; several other side-chains of intermediate quality were placed in the middle category; and some side-chains that were not visible in the map were placed in the category of lowest accuracy. As detailed below, weaker or stronger dynamics constraints were used to generate greater or lesser variations in the co-ordinates of atoms according to their level of accuracy in the model structure. Model co-ordinates for the links between helices were not available at the time of these calculations, so the links were omitted (see Discussion).

The overall scheme of the calculations to generate various co-ordinate sets is shown in Fig. 2. Minimization and molecular dynamics calculations were done using the X-PLOR computer program (Brünger, 1988) with the CHARMM19x parameter set for proteins with explicit

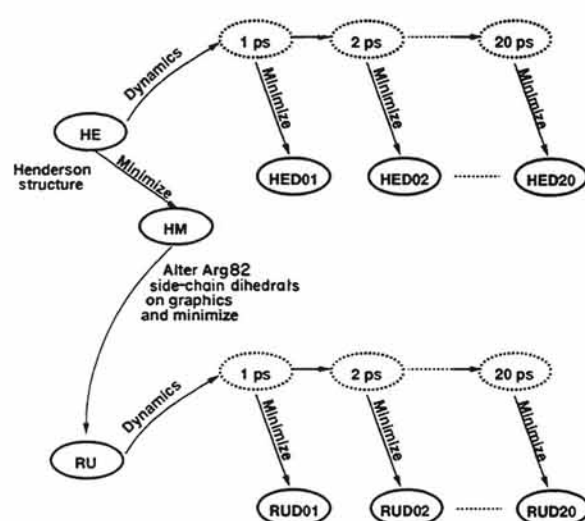


Figure 2. Scheme of calculations for generating atomic structural models for calculations (see the text).

polar hydrogen atoms. The charges for the retinal chromophore are as described above. For reasons given in Results, the charges of Asp96 and Asp115 were altered to make their side-chains neutral. Co-ordinates for polar hydrogen atoms were built using the HBUILD (Brünger & Karplus, 1988) facility of X-PLOR. The Henderson *et al.* model co-ordinates with the HBUILD polar hydrogen atoms is referred to here as the "HE" co-ordinate set. Initial minimizations without electrostatic terms were used to relieve conformational strain in the chromophore and side-chains of the HE co-ordinate set. They were carried out in the following stages: first, the β -ionone ring, then the polyene chain, and then the lysine side-chain were minimized in isolation by 100 steps of conjugate gradient minimization (Powell, 1977). Finally, 200 minimization steps were performed on all side-chains. Next, harmonic constraints were imposed with their minima centered about the HE co-ordinates and force constants of 0, 0.110 or 0.411 kcal/(mol Å²) (1 cal = 4.184 J) for atoms in the low, intermediate and high accuracy categories, respectively. With these constraints applied, the structure was minimized for another 200 steps without electrostatics and 200 steps with electrostatics. We shall refer to the resulting structure as the HM co-ordinate set. Starting from HM and using the harmonic constraints described above, 20 ps of molecular dynamics at 1000 K were performed using a 1000 K Maxwellian distribution of initial velocities, velocity rescaling to maintain temperature, and a 1 fs time step. Co-ordinates were written to files at 1 ps intervals and subsequently put through 400 steps of minimization. We shall refer to these co-ordinate sets as HEDnn, where "nn" ranges from 1 to 20 and indicates the picosecond of dynamics from which the co-ordinates were derived.

For reasons given in Results, an ensemble of structures having Arg82 pointing towards the Schiff base rather than towards the extracellular membrane was also prepared. Interactive molecular graphics were used to alter the Arg82 side-chain dihedral angles of the Henderson *et al.* (1990) model so as to fit the side-chain into a cavity in the protein interior. Minimization was performed first on the Arg82 side-chain, then on the all atoms with harmonic constraints included, but no electrostatics (there are no constraints on the Arg82 side-chain,

since it was in the low confidence category) and finally on all atoms with electrostatics included. We shall refer to the resulting co-ordinate set as RU. The side-chain dihedrals for Arg82 were:

$$\chi_1 = -76.5, \chi_2 = -65.8, \chi_3 = 81.6, \chi_4 = -126.4.$$

The RU and HE structures are shown in Fig. 3. Constrained dynamics were used to generate an ensemble of 20 Arg-up structures, referred to as RUDnn, by the method described above.

3. Results

(a) Solvent accessibility

The solvent accessibility algorithm used to define the solvent and interior regions shows large solvent-filled invaginations and cavities into the protein from each side of the membrane with the chromophore and associated side-chains in between (Fig. 3(b)). In the lower (extracellular side) channel Asp85 and Asp212 are partially exposed to solvent and block the direct access of solvent to the lysine-retinal Schiff base of the chromophore. Placing Arg82 in an upward configuration substantially blocks solvent access to Asp85 and Asp212 (Fig. 3(c)). This is consistent with infrared measurements of the carbonyl vibrations in the M state of the photocycle showing that Asp85 is in a hydrophobic environment (Gerwert *et al.*, 1990). However, no such measurements of its environment in the ground state are available. In the upper (cytoplasmic side) channel, Asp115 and Asp96 are only partially exposed to solvent, consistent with infrared spectral data at 1740/cm and 1742/cm, showing that these residues are in a partly hydrophobic environment (Braiman *et al.*, 1988; Gerwert *et al.*, 1989, 1990). The lysine-retinal Schiff base is not directly accessible to solvent.

(b) Conformational variation

The root-mean-square (r.m.s.) difference between the Henderson *et al.* (1990) co-ordinates (HE) and the energy minimized co-ordinates (HM) was 0.442 Å. All bonds in the HM structure were within 0.05 Å of their equilibrium length in the potential function and all but ten bond angles were within 10° of equilibrium with the largest deviation being 13.6°. The total van der Waals' energy was -1083 kcal/mol, indicating that there were very few unfavorable steric overlaps. Thus, the Henderson *et al.* model appears to be reasonable in terms of the CHARMM19 energy function used here (Brooks *et al.*, 1983). The r.m.s. difference between the HE structure and the energy minimized structure with Arg82 "up" (RU) was 0.410 Å for atoms not part of Arg82 itself. This is similar to the HM *versus* HE r.m.s. deviation, indicating that the repositioning of Arg82 could be accommodated with relatively little structural perturbation. After minimization, all bonds in the RU structure were within 0.05 Å of their equilibrium length in the potential function, and all but nine bond angles

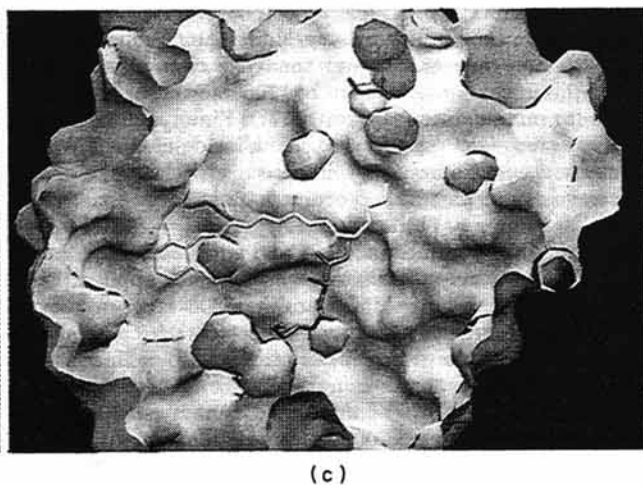
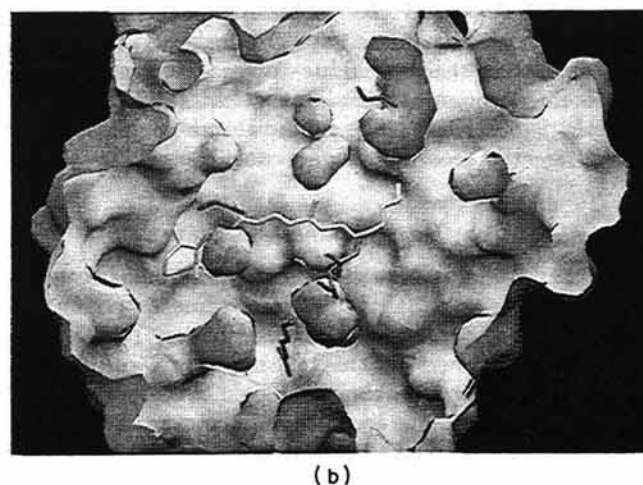
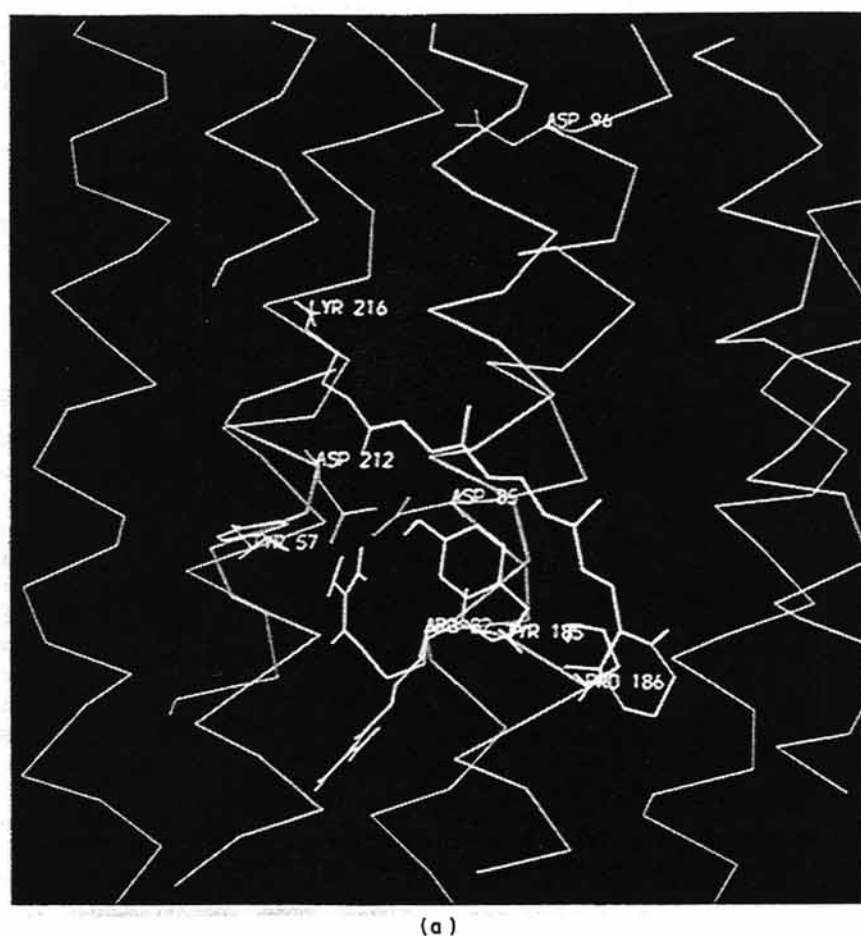


Figure 3. (a) The Henderson structure (HE) with the Arg82 side-chain from the "Arg up" (RU) structure superimposed. (b) and (c) A cut-away view through the solvent contact surface of (b) the Henderson structure and (c) the RU structure. The lysine-retinal Schiff base (purple) is in the center. Aspartate and arginine side-chains are shown in red, with Asp96 near the top, Asp115 above and to the right of the β -ionone ring, and Asp85, Asp212 and Arg82 below the Schiff base.

were within 10° of equilibrium with the largest deviation being 14.4° . The total van der Waals' energy was -1082 kcal/mol.

As expected, the ensemble of structures generated by high temperature dynamics and minimization had a much greater range of conformational variation. Figure 4(a) shows the r.m.s. deviations of the

HEDnn structures from the Henderson model; and Figure 4(b) shows the r.m.s. deviations of the RUDnn structures from the Henderson model. As above, the structures were generally free from adverse bond geometries and steric overlaps. Although Arg82 was not constrained in either dynamics simulation, it remained in an upward orien-

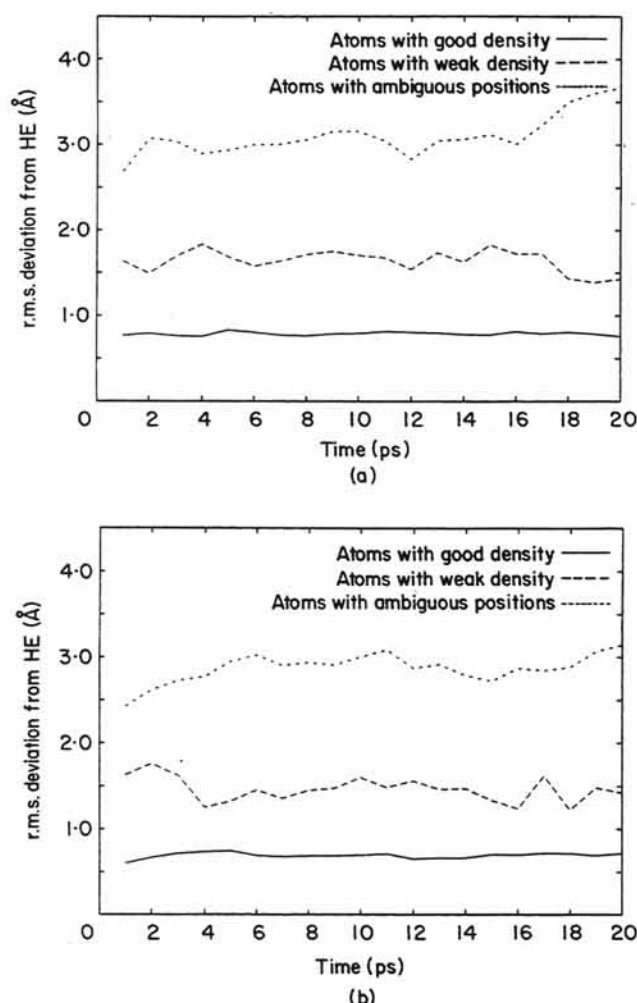


Figure 4. Conformational variation of the molecular dynamics structures with (a) Arg82 in the downward orientation (HED) and (b) Arg82 in the upward orientation (RUD). Structures have been generated each picosecond. Classification of atoms into those with "good density", "weak density" and those whose position is ambiguous is according to R. Henderson *et al.* (personal communication).

tation throughout the RUDnn ensemble and in a downward orientation throughout the HEDnn ensemble.

(c) Preliminary calculations and methodological adjustments

The results of preliminary calculations motivated some methodological adjustments. It is known from FTIR experiments that Asp96 and Asp115 have $pK_{1/2}$ values greater than 9.0 (Engelhard *et al.*, 1985) and electrostatic titration calculations on the HE, HM and RU structures correctly predict high $pK_{1/2}$ values for these side-chains. But when dynamics and minimization are performed using the standard CHARMM19 (Brooks *et al.*, 1983) parameter set in which aspartate residues are charged, Asp96 tends to form salt bridges with nearby lysine side-chains and subsequent electrostatic titration

calculations incorrectly predict low pK values for this residue due to the positive charges nearby. Such salt bridge formations are a common artifact of *in vacuo* molecular dynamics calculations caused by neglect of the charged side-chain's attraction to the solvent. The pK value of Asp115 is also lowered because of charge-induced structural changes. To eliminate these artifacts, we have neutralized the charge sets on these two aspartate side-chains in the dynamics calculations.

Preliminary electrostatic titration calculations included a number of residues on the membrane-exposed surface of the protein which, in the final calculations presented here, are treated as fixed, non-titrating charges. Tyr26, Tyr43, Tyr147 and Tyr150 are membrane exposed and located approximately halfway between the membrane surfaces. Preliminary calculations gave $pK_{1/2}$ values greater than 20 for these side-chains, indicating that they are always protonated. This is to be expected, since the apolar membrane environment will favor the neutral state. We therefore regarded these side-chains as non-titrating and neutral in subsequent calculations. Lys30, Lys40, Lys41 and Lys172 and Arg175 and Arg225 are membrane exposed and are located near the cytoplasmic membrane surface. Preliminary calculations on the HE structure gave $pK_{1/2}$ values of 7.2 for Arg225 and less than 1.0 for the others, implying that they are never protonated. However, we feel it is more likely that the calculations are wrong, and that these residues are protonated because they are positioned near the membrane surface, where the protonated forms could make strong favorable interactions with negative charges on the membrane head groups. (These head groups are not explicitly included in either the Henderson *et al.* model structure or in our calculations.) Therefore we regarded these protein side-chains as non-titrating and positively charged in subsequent calculations. Preliminary calculations on the HE and HM structures showed that the reverse assumption, that these residues were deprotonated, would raise the calculated $pK_{1/2}$ of Asp96 by as much as five pK units, but other titrating residues would be affected by one pK unit or less. The results would still have the same agreement with the experimentally determined protonation state in the neutral pH range (see below).

(d) pK_{intr} calculations

As shown in Table 2, the pK_{intr} values for most titrating groups are strongly shifted relative to the model compound pK_a values so as to favor the neutral protonation state. This occurs because most titratable groups are poorly solvated, as indicated by the large positive Born terms for Asp, Glu and Tyr, and large negative Born terms for Arg and the Schiff base. The terms due to the background of non-titrating protein charges tend to have signs opposite to the Born terms, but are not large enough to compensate for them. This means that if these residues are to become charged at moderate

Table 2
Calculation of pK_{intr} values

Site	pK_{mod}	Born terms†			Background terms†			pK_{intr}		
		HE	HM	RU	HE	HM	RU	HE	HM	RU
Asp85	4.0	8.4	9.3	10.6	-1.4	-3.0	-2.2	11.0	10.3	12.4
Asp96	4.0	6.3	7.4	7.2	-2.4	-2.7	-3.0	7.9	8.6	8.3
Asp115	4.0	9.7	9.9	9.7	-3.7	-2.0	-2.3	9.9	11.8	11.3
Asp212	4.0	8.4	9.0	10.5	-1.7	-4.7	-5.4	10.7	8.3	9.1
Glu9	4.0	1.6	3.5	2.0	-1.0	-2.0	-2.0	4.6	5.5	4.0
Glu204	4.0	2.9	3.4	2.3	-0.5	-0.7	-1.2	6.3	6.7	5.1
Tyr57	9.6	9.8	9.4	10.5	-2.1	-3.4	-3.2	17.3	15.6	16.9
Tyr79	9.6	1.6	1.2	1.1	0.4	0.7	0.6	11.6	11.5	11.3
Tyr83	9.6	7.4	8.0	7.7	-2.4	-1.6	-1.6	14.7	16.0	15.7
Tyr185	9.6	11.4	11.7	11.9	-0.7	0.1	-0.2	20.3	21.5	21.3
Arg82	12.0	-11.7	-11.8	-14.9	0.0	2.2	-0.6	0.4	2.4	-3.5
Schiff base‡	7.0	-9.8	-9.4	-9.3	-0.6	-0.7	-1.1	-3.4(F.P.)	-3.0(F.P.)	-3.4(-0.8)

† The Born and background terms are expressed as pK shifts.

‡ The last line shows calculated pK_{intr} values for the Schiff base, but in subsequent calculations the Schiff base is either forced to remain protonated (F.P.) or its pK_{intr} is set to -0.8, as explained in the text.

pH there must be interactions between titratable groups that stabilize charged states. The sensitivity of the calculations to conformational change is shown by the differences between the HE and HM calculations, of the order of 1.0 to 2.0 pK units, that are caused by the 0.442 Å r.m.s. change in co-ordinates due to minimization.

When the simple electrostatic model used here to calculate pK_{intr} for the amino acid side-chains is applied to the Schiff base, we generally find that lack of solvation shifts the pK_{intr} down by 9 to 11 pK units from its model compound pK_a value of 7. But semi-empirical quantum calculations (Tavan *et al.*, 1985) have shown that steric effects and redistribution of charges through the π -orbital system of the polyene chain can have strong effects on the pK_a of the Schiff base, and the present electrostatic method is clearly inadequate to model such effects. Therefore, we have tried to obtain a

$pK_{1/2}$ value consistent with experiment by adjusting the Schiff base pK_{intr} . For the RU structure this could be done by setting pK_{intr} to -0.8 rather than the value of -3.44 that would be calculated from electrostatic effects alone. The result is a two-stage titration of the Schiff base with the first pK_a near 12 and the second pK_a much higher (see below). The first stage pK_a is reasonably consistent with the experimental result that the Schiff base deprotonates with a $pK_a > 12$, and undergoes an irreversible change, probably denaturation, near pH 13 (Druckmann *et al.*, 1982; Doukas *et al.*, 1981; Lewis *et al.*, 1978). The pK_{intr} value of -0.8 for the Schiff base was used in all the calculations having Arg82 in the "up" position (i.e. the RU and RUDnn structures). For the HE structure it was not possible to select a pK_{intr} value that gave a $pK_{1/2}$ near 12 for the Schiff base; therefore we forced the Schiff base to remain protonated at all pH values.

Table 3
Calculated $pK_{1/2}$ values and experimental data for comparison

Site	Calculated $pK_{1/2}$					Experimental†
	Arg82 to extracellular			Arg82 to Schiff base		
	HE	HM	HED	RU	RUD	
Asp85	18.2	15.0	6.6±0.5	-1.6	-3.0±1.3	≤2.5
Asp96	7.4	8.1	10.6±1.3	8.2	11.6±1.2	≥9.5
Asp115	8.4	10.2	14.1±0.7	16.3	15.4±8.6	≥9.5
Asp212	3.3	-2.7	-5.4±0.8	-10.1	-16.4±2.0	≤2.5
Glu9	4.5	3.4	1.0±7.3	4.1	0.2±3.0	N.A.
Glu204	4.9	1.4	-5.5±8.1	6.9	5.3±1.3	N.A.
Tyr57	45.6	38.5	46.2±5.9	44.9	35.4±18.2	≥12
Tyr79	17.8	16.8	15.4±0.8	18.4	17.0±2.3	≥12
Tyr83	25.4	27.0	26.6±7.3	25.9	25.3±1.2	≥12
Tyr185	37.9	44.1	39.2±1.9	52.0	45.8±0.5	≥12
Arg82	9.9	20.0	25.2±10.7	30.6	55.1±30.1	N.A.
Schiff base	Fixed in protonated state			12.0‡	13.1±13.6	≥12

† Experimental data from Engelhard *et al.* (1985), Gerwert *et al.* (1987), Subramaniam *et al.* (1990), Herzfeld *et al.* (1990), Druckmann *et al.* (1982), Lewis *et al.* (1978) and Doukas *et al.* (1981). N.A., not available.

‡ Adjusted to 12 by manipulation of pK_{intr} value (see the text).

(e) $pK_{1/2}$ calculations and individual site titration curves

The most striking general feature of the calculated $pK_{1/2}$ values (Table 3) is their large deviation from the corresponding model compound pK_a values; many $pK_{1/2}$ values are below zero or well above 14. Since the calculations were made at fixed conformations, extremely low (high) values should be interpreted to mean that the site cannot be protonated (deprotonated) without significant conformational change. The large shifts are caused partly by the shifts in pK_{intr} values due to apolar burial, as noted above, and partly by strong electrostatic couplings between titrating sites in the protein channel. These sites are close together and are not strongly screened by water.

Plots of most of the calculated titration curves of individual sites have the familiar sigmoidal shape and can be usefully summarized by their $pK_{1/2}$ value (see Fig. 5(a)). However, for some sites the curves are more complex. Figure 5(b) shows the two-step titration of the Schiff base and Asp115 of the RU structure. Below pH 11, both residues are protonated, but at pH 13 there has been a net loss of one proton and the remaining proton is shared between the two sites, with the Schiff base having the proton approximately 40% of the time and Asp115 having it the other 60% of the time. Near pH 17 there is another titration in which this shared proton is lost. (Of course, at very high pH, denaturation may occur (Druckmann *et al.*, 1982): a possibility not provided for here. These calculations at extreme pH should be taken to indicate tendencies only.) The calculated $pK_{1/2}$ values for the Schiff base and Asp115 are 12.0 and 16.3, respectively, but these values are clearly misleading, since each residue titrates in two steps. Proton sharing occurs in these calculations because the residues involved have strong mutual electrostatic interactions. Without such interactions the sites would titrate independently.

An even more striking effect is seen in the titration curves of the Schiff base, Asp115 and Asp96 of the RUD13 structure (Fig. 5(c)). Between pH 8 and pH 11 the protonation of the Schiff base falls, then between pH 11 and pH 14 its protonation increases with increasing pH, and finally the proton is lost completely near pH 19. The increase in protonation with increasing pH would be consistent with the law of mass action only if another titrating group or groups were losing their protons more rapidly so that there is no net gain of protons for the molecule as a whole, but rather a net internal transfer of protons from one site to another. In this case, a net loss of approximately one proton from the Schiff base and Asp115 occurs between pH 8 and pH 11, after which the Schiff base and Asp115 share the remaining proton at a ratio of approximately 20% to 80%. Between pH 11 and pH 14 Asp96 loses its proton and because the electrostatic interaction (W_{uv} term) between Asp96 and the Schiff base is larger than the interaction between Asp96

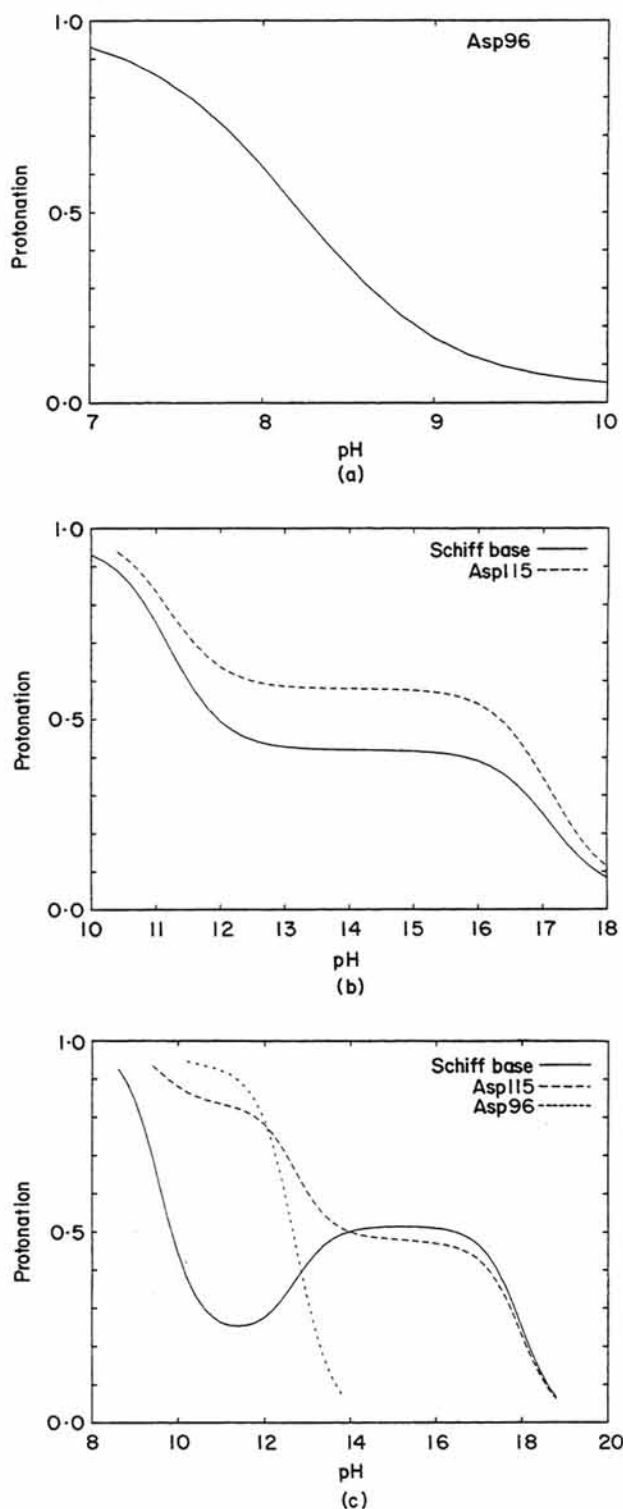


Figure 5. Titration curves of individual sites (pH versus protonation). (a) Asp96 in the RU structure. (b) Schiff base and Asp115 in the RU structure. (c) Schiff base, Asp115 and Asp96 in the RUD13 structure.

and Asp115, the negative charge of the deprotonated Asp96 tends to pull the shared proton onto the Schiff base and away from Asp115. Finally, at pH 18 the shared proton is lost. The possibility of increasing protonation with increasing pH in

molecules with electrostatically coupled titrating sites is not just a theoretical curiosity; it has been observed experimentally in a nuclear magnetic resonance (n.m.r.) study of the titration of diethylenetriaminepenta-acetic acid (DPTA; Sudmeier & Reilly, 1964).

Comparison of the HE and HM columns in Table 3 shows that relatively small conformational changes, in this case the 0.44 Å r.m.s. deviation that occurs upon minimization, can have very large effects on the computed $pK_{1/2}$. Still larger effects are seen in the HED column, which summarizes the results of the calculations done on a set of dynamics-generated conformations. For the residues having non-zero net charges in the minimization and dynamics procedures (Asp82, Asp212, the glutamate residues and Arg82), minimization and dynamics tend to produce $pK_{1/2}$ shifts reflecting conformations that favor the charged state. Comparison of the RU and RUD columns shows the same trend. For six out of 11 sites in the HED calculations and eight out of 12 sites in the RUD calculations, the standard deviations in $pK_{1/2}$ values is within 3.0 pK units. Larger deviations in other sites usually occur either because these sites participate in two-step titrations, as explained above, so that their calculated $pK_{1/2}$ can occur near the first or second titration step depending on whether the proton sharing section of the curve lies above or below 0.5 (Asp115 and the Schiff base in the RUD calculations, for example), or because one or two conformations out of the sample of 20 have $pK_{1/2}$ values drastically different from the others (Tyr57 and Arg82, for example). The sensitivity to conformational changes shown here is much greater than that for solvent-exposed side-chains in a water-soluble protein (Bashford & Karplus, 1990).

The calculated $pK_{1/2}$ values for structures with Arg82 in the downwards orientations are in agreement with much of what is known from experiment, although there are some important differences. Asp96 and Asp115 are predicted to have $pK_{1/2}$ values of 7.4 and 8.4, respectively, in the calculations on the HE structure and even higher values in calculations on minimized (HM) and dynamics (HEDnn) structures. This is in general agreement with experimental findings that Asp96 and Asp115 have abnormally high pK_a values (Engelhard *et al.*, 1985), but the $pK_{1/2}$ values calculated from the HE structure are not as high as the experimental findings. All tyrosine residues in the lower channel have $pK_{1/2}$ values above 14, indicating that they cannot deprotonate without significant structural changes. This is consistent with n.m.r. data showing the absence of tyrosinate up to pH 12 (Herzfeld *et al.*, 1990). Asp212 has a $pK_{1/2}$ of 3.3 or lower, while Asp85 is predicted to be protonated at neutral pH in the HE, and HM structures and has $pK_{1/2}$ values near 6.6 in HED structures. Both Asp85 and Asp212 have high pK_{intr} values due to a lack of favorable Born solvation energy for the deprotonated state. The deprotonation of Asp212 in the calculations occurs because of interactions with the

positively charged protonated Schiff base. Asp85 remains neutral because no other positive charge is available to stabilize its deprotonated state. However, from experiments it is known that Asp85 and Asp212 are both deprotonated at neutral pH (Gerwert *et al.*, 1989). The protonation of one of these residues may occur during the purple to blue transition near pH 2.5 (Gerwert *et al.*, 1987; Subramaniam *et al.*, 1990). Therefore, the calculations in the structures with Arg82 in the downwards orientation agree with experiment for Asp212 but not for Asp85.

Henderson *et al.* (1990) did not find any resolved density for the Arg82 side-chain and noted that it could plausibly be placed, "either upward near Asp85 or downwards away from Asp85, or could move between the two positions". In their model structure they chose the downwards (extracellular) position, but we reasoned that the upward position might provide the positive charge needed to stabilize the deprotonated state of Asp85. The calculations on the RU (Arg82 up) conformations confirmed this by predicting both Asp85 and Asp212 to be deprotonated at neutral pH, in agreement with experiment. The $pK_{1/2}$ values for Asp96 and Asp115 are in better agreement with experiment for RU than for HE, but this may be due as much to minimization effects as to the change in the Arg82 position. The $pK_{1/2}$ value of 16.3 given for Asp115 in the RU structure is misleading, since this is one of the residues that titrates in two steps (see above and Fig. 5(b)). In addition, for the RU structure it is possible to find an adjustment to the pK_{intr} value of the Schiff base that results in partial deprotonation, but for the HE structure this is not possible (see above). Trial calculations were carried out with a more detailed charge model of the deprotonated state of the Arg side-chain, and it was found that the results did not differ significantly from those with the no-charge model used here, since the important effects of Arg82 on its neighbors, Asp85 and Asp212, arise from its protonated state, and Arg82 itself is predicted to have such a high pK as to be non-titrating. On the whole, the calculations done with Arg82 in the upward orientation give substantially better agreement with experiment than those with Arg82 down.

4. Discussion

(a) The conformation of Arg82

As noted above, the calculations with Arg82 oriented upward are successful in predicting the correct protonation states of the important residues, Asp85, Asp212, Asp96 and Asp115 at neutral pH. Calculations with Arg82 downwards incorrectly predict Asp85 to be protonated. We therefore propose that, in the absence of resolved density to the contrary, the model for the ground state of bacteriorhodopsin should have Arg82 in the upwards orientation. This places Arg82 in close contact with Asp85 and Asp212, which are highly

conserved residues near the Schiff base, and raises the question of whether Arg82 is part of the proton transfer pathway. Asp85 receives the proton from the Schiff base in the L to M transition and remains protonated in M; but another proton is released to the extracellular side in the L to M transition. Therefore, another proton acceptor must be involved in the proton release pathway. Arg82 is a good candidate; mutation of this residue inhibits the proton release (Butt *et al.*, 1989; Otto *et al.*, 1990). The model is also consistent with the suggestion that the Schiff base has a complex counterion (de Groot *et al.*, 1989), although the Henderson *et al.* (1990) structure does not appear to include space for a water molecule that mediates hydrogen bonding between the protonated Schiff base and the carboxylate groups.

(b) Buried charges must occur in pairs

The need to place Arg82 in the interior in close proximity to the two negatively charged aspartate residues and the positively charged protonated Schiff base appears to be an example of the need to create charges in opposite pairs in an apolar medium (Honig *et al.*, 1979; Honig & Hubbell, 1984). A rough quantitative argument for such a general principle can be given as follows. Transfer of a charge q (or $-q$) of radius r from water into a medium of dielectric constant ϵ entails a loss of Born solvation energy of approximately $q^2/(\epsilon 2r)$. The loss can be partly compensated for by placement of two opposite charges, q and $-q$, at a distance of approximately $2r$. This suggests that, where buried charged residues are involved, care must be taken in the design and interpretation of site-directed mutagenesis experiments. It may prove more useful to alter ion-counterion pairs together than to attempt to introduce or remove single charges.

(c) Membrane proteins versus water soluble proteins

This situation is in sharp contrast to the Linderström-Lang smeared charge concept for water soluble, globular proteins, in which electrostatic energy is a function of the total charge on the protein and interactions between opposite pairs are neglected (Linderström-Lang, 1924). The difference arises because of the "inside-out" structure of membrane proteins; the outside surface is mostly hydrophobic, while the charged, hydrophilic residues are on the inside. In the low dielectric environment of the protein interior, desolvation of the charged states leads to large difference between pK_{model} and pK_{intr} and strong interactions between titrating sites. The end result is that $pK_{1/2}$ values tend to have larger deviations from their stereotypical values in model compounds and unfolded proteins, and there are strong correlations between the ionization states of many titrating sites. This implies two tendencies: that internal charge groups will occur in ion-counterion pairs whose titration

may be complex due to site-site coupling; and that ionizable groups lacking counterions will be found in their neutral forms over a wide pH range.

(d) Complex titration behavior in proton transfer systems

There is reason to believe that the kind of complex titration behavior seen in Figure 5(b) or (c) may be a common feature in transmembrane proton pumps. If a proton is to be transferred from one site to another without using up too much free energy, the pK_a values of the sites should be similar. The sites are likely to have strong electrostatic interactions with each other, since close proximity is needed for proton transfer and the membrane/protein environment in which the groups are likely to be buried has a low dielectric constant. Strong electrostatic coupling between sites with similar pK_a values are exactly the conditions required to produce the type of proton sharing seen in Figure 5(b)†. This has implications for the interpretation of pH-dependent experimental data; there is not necessarily a simple one-to-one correspondence between observed pK_a values and the titration of individual groups.

(e) Limitations

The accuracy to be expected from the present calculations is limited by the accuracy of the structural data, certain aspects of the computational methods used and uncertainty about the validity of the continuum approximation.

The current Henderson *et al.* (1990) model does not include the N- and C-terminal fragments of the protein or the connecting loops between helices. The partially conserved Arg227 near the top of the G helix is also missing from the model. This probably causes the Born solvation energy of the charged states of some sites to be too negative because they appear more exposed to solvent than they really are. On the other hand, charged or polar groups on the missing loops might contribute interactions that either stabilize or destabilize the charged states. Fortunately, one expects that functionally important residues deep inside the channels, such as Asp85, Asp212, Asp96, Asp115 and the Schiff base, will be less affected by the missing loops than residues near the lips of the channels, such as Asp38, Glu9 and Glu166.

Our calculations on ensembles of structures generated by constrained molecular dynamics suggests that structural uncertainty due to low resolution data generally produces uncertainty in the calculated $pK_{1/2}$ values of the order of several pK units. Even when these uncertainties are taken into

† These are also the conditions under which the mean field approximation to eqn (13), introduced by Tanford & Roxby (1972), is most likely to fail. The exact equation or a more robust approximation method (Bashford & Karplus, 1991) should be used instead.

account, the calculations on the structures with Arg82 pointed toward the Schiff base still produce a protonation state consistent with experimental data for the neutral pH range, while the calculations with Arg82 in the extracellular orientation fail to predict the deprotonation of Asp85.

Although we have used dynamics to generate an ensemble of structures, we have not accounted for the possibility that protonation may induce conformational change; instead, the calculations are based on a model of titration as the addition and removal of charges on a rigid object. For example, an aspartate residue might be predicted to have a very high $pK_{1/2}$ because its negatively charged, deprotonated state is not well solvated in the model co-ordinates, while in fact, the side-chain is able to move into a more solvent-exposed conformation upon deprotonation, which results in a lower $pK_{1/2}$. In this respect, the calculation will tend to be biased towards the protonation state of the sample from which the structural model is derived. If dynamics or minimization calculations are used to produce structural models, the subsequent $pK_{1/2}$ calculations will tend to be biased toward the protonation state assumed in the selection of side-chain charges for dynamics or minimization. Since the bacteriorhodopsin specimens for the model were prepared at pH 5.2 (Henderson *et al.*, 1990), we expect that our calculations will be most useful for prediction of the protonation state at mildly acid to neutral conditions, so we have given the most weight to such predictions in the interpretations of our results. We believe that errors due to neglect of conformational changes during titration should not be too severe, since there is no experimental evidence for large pH-dependent conformational changes in the pH range 4 to 10.

The real membrane has a layer of negatively charged phosphate head groups and a neutralizing layer of counterions from the solution on each side of the membrane. There may also be counterions that bind to specific sites on bacteriorhodopsin, but the locations of such sites are unknown. We have not included these in the present calculations because of a lack of data on the atomic-level structure of the membrane and counterions and because of the computational difficulties of including ionic strength effects in the presence of a membrane. The inclusion of ionic strength effects (Poisson-Boltzmann as opposed to Poisson equation) in the presence of a membrane is beyond the capability of our computer programs. To include the negatively charged phosphate head groups without the counterion layer would be likely to reduce the accuracy of the calculations, since it would introduce large unbalanced charges, therefore we have included neither. This lack of detail may have an effect on the results for Asp96, which appears to be sensitive to charges near the membrane surface. In addition, bacteriorhodopsin in purple membrane tends to cluster in regular two-dimensional lattices; but our calculations are for an isolated molecule. However, bacteriorhodopsin monomers function as

independent pumping units; no effects from trimerization or lattice formation have been observed (Dencher & Heyn, 1979).

Although continuum approximations such as those used here have been used successfully for water-soluble globular proteins (Sternberg *et al.*, 1987; Gilson & Honig, 1987; Bashford & Karplus, 1988, 1990), there is a serious question as to whether they are valid when applied to the water in a transmembrane channel and whether the dielectric constant of channel water should be the same as that of bulk water. If the dielectric constant were significantly lower, as one might expect from restrictions in water mobility imposed by the channel, individual terms in the calculation of shifts in pK values would be larger due to less favorable Born solvation of charges in the channel and reduced screening of charge interactions, and the predicted shifts would be even more extreme unless there was a cancellation of terms. It may be useful to explore this question by means of microscopic molecular dynamics or Monte Carlo simulations in which water molecules are treated explicitly. In interpreting such simulations it should be remembered that the relevant "dielectric property" is the response of the solvent to the large changes in electric field caused by protonation of sites rather than the response to vanishingly small fields to which an analysis of dipole fluctuations would pertain.

5. Summary and Conclusions

The FTIR and n.m.r. data on the protonation states of internal residues and availability of a structural model for bacteriorhodopsin provide an opportunity to apply the recently developed continuum electrostatic techniques to titration calculations on the proposed model. This allows an independent check of the structural model, a test of the theoretical techniques, and a better understanding of the role of electrostatics in the mechanism.

The calculations are mostly consistent with experimental data. In particular, the unusual protonation states of the upper channel aspartic acid residues, 96 and 115, are reproduced, and Asp212 in the lower channel is correctly predicted to be deprotonated. No tyrosine residues are deprotonated, consistent with the n.m.r. finding that no tyrosinate residues are present (Herzfeld *et al.*, 1990). However, in order to reproduce the experimental result that Asp85 is deprotonated, it is necessary to place Arg82, in an orientation towards the Schiff base rather than towards the extracellular membrane. According to Henderson *et al.* (1990), no density is seen for this side-chain, so it can be placed in either position. Thus, these theoretical calculations have led to an experimentally testable prediction. These results, together with our finding that minimization of the structure with the CHARMM19 empirical potential function (Brooks *et al.*, 1983) produces only small structural changes and indicates no

significant steric overlaps, are evidence of the general correctness of the model of Henderson *et al.* (1990).

The results provide examples of several general principles that may hold for bacteriorhodopsin and other transmembrane proton pumps. One principle is that it is most energetically favorable to form charges in a hydrophobic region in opposite pairs. With Arg82 pointing down out of the lower channel, the Schiff base was able to have only one negative counterion, Asp212; but, with Arg82 pointing up, an extra positive charge is provided allowing Asp85 to be deprotonated as well. Another principle is that titrating sites in proton pumps are likely to exhibit complex collective titration behavior, since they contain strongly interacting sites titrating simultaneously.

More complete and accurate structural information and computational procedures that account for conformational change upon protonation, and treat the membrane's electric double layer more realistically, should make more reliable calculations of this kind possible in the future.

During the preparation of this manuscript, we learned that Sampogna and Honig had independently performed similar calculations on bacteriorhodopsin, and also had obtained the result that general agreement with the experimentally determined protonation states is best when Arg82 is in an upward orientation. Their calculations did not include a membrane, which suggests that these general results are not very sensitive to the membrane or the choice of the membrane dielectric constant.

The computer programs and data files used to do the electrostatic calculations and titration curves for structures HE and RU are available by anonymous FTP from InterNet address, scripps.edu [192.42.82.27], or from Dr Bashford (bashford@scripps.edu).

The authors thank Richard Henderson for providing bacteriorhodopsin atomic model co-ordinates and Charles Perrin for pointing out the Sudmeier & Reilly paper on the titration behavior of DPTA. D.B. acknowledges support from the Research Institute of Scripps Clinic and National Institutes of Health (grant GM45607-01). K.G. acknowledges a Heisenberg fellowship of the DFG and the hospitality of Scripps during an extended visit.

References

- Bashford, D. & Karplus, M. (1988). Electrostatic effects of charge perturbations introduced by metal oxidation in proteins. *J. Mol. Biol.* **203**, 507–510.
- Bashford, D. & Karplus, M. (1990). pK_a 's of ionizable groups in proteins: atomic detail from a continuum electrostatic model. *Biochemistry*, **29**, 10219–10225.
- Bashford, D. & Karplus, M. (1991). Multiple-site titration curves of proteins: an analysis of exact and approximate methods for their calculation. *J. Phys. Chem.* **95**, 9556–9561.
- Braiman, M. S., Mogi, T., Marti, T., Stern, L. J., Khorana, H. G. & Rothschild, K. J. (1988). Vibrational spectroscopy of bacteriorhodopsin mutants: light-driven proton transport involves protonation changes of aspartic acid residues 85, 96, and 212. *Biochemistry*, **27**, 8516–8520.
- Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S. & Karplus, M. (1983). CHARMM: a program for macromolecular energy, minimization, and dynamics calculations. *J. Comp. Chem.* **4**, 187–217.
- Brünger, A. T. (1988). *X-PLOR (Version 1.5) Manual*. The Howard Hughes Medical Institute and Department of Molecular Biophysics and Biochemistry, Yale University, New Haven.
- Brünger, A. T. & Karplus, M. (1988). Polar hydrogen positions in proteins: empirical energy placement and neutron diffraction comparison. *Proteins*, **4**, 148–156.
- Butt, H. J., Fendler, K., Bamberg, E., Tittor, J. & Oesterhelt, D. (1989). Aspartic acids 96 and 85 play a central role in the function of bacteriorhodopsin as a proton pump. *EMBO J.* **8**, 1657–1663.
- de Groot, H. J. M., Harbison, G. S., Herzfeld, J. & Griffin, R. G. (1989). Nuclear magnetic resonance study of the Schiff base in bacteriorhodopsin: counterion effects on the ^{15}N shift anisotropy. *Biochemistry*, **28**, 3346–3353.
- Dencher, N. A. & Heyn, M. P. (1979). Bacteriorhodopsin monomers pump protons. *FEBS Letters*, **108**, 307–310.
- Doukas, A. G., Pande, A., Suzuki, T., Callender, R. H., Honig, B. & Ottolenghi, M. (1981). On the mechanism of hydrogen-deuterium exchange in bacteriorhodopsin. *Biophys. J.* **33**, 275–280.
- Druckmann, S., Ottolenghi, M., Pande, A., Pande, J. & Callender, R. H. (1982). Acid-base equilibrium of the Schiff base in bacteriorhodopsin. *Biochemistry*, **21**, 4953–4959.
- Engelhard, M., Gerwert, K., Hess, B., Kreutz, W. & Siebert, F. (1985). Light-driven protonation of changes of internal aspartic acids of bacteriorhodopsin: an investigation by static and time-resolved infrared difference spectroscopy using $[4-^{13}\text{C}]$ aspartic acid labeled purple membrane. *Biochemistry*, **24**, 400–407.
- Gerwert, K., Ganter, U. M., Siebert, F. & Hess, B. (1987). Only water-exposed carboxyl groups are protonated during the transition to the cation-free bacteriorhodopsin. *FEBS Letters*, **213**, 39–44.
- Gerwert, K., Hess, B., Soppa, J. & Oesterhelt, D. (1989). Role of aspartate-96 in proton translocation by bacteriorhodopsin. *Proc. Nat. Acad. Sci., U.S.A.* **86**, 4943–4947.
- Gerwert, K., Souvignier, G. & Hess, B. (1990). Simultaneous monitoring of light-induced changes in protein side-group protonation, chromophore isomerization, and backbone motion of bacteriorhodopsin by time-resolved Fourier-transform infrared spectroscopy. *Proc. Nat. Acad. Sci., U.S.A.* **87**, 9774–9778.
- Gilson, M. K. & Honig, B. H. (1987). Calculation of electrostatic potentials in an enzyme active site. *Nature (London)*, **330**, 84–86.
- Henderson, R. (1975). The structure of the purple membrane from *Halobacterium halobium*: analysis of the X-ray diffraction pattern. *J. Mol. Biol.* **93**, 123–138.
- Henderson, R. & Unwin, P. N. T. (1975). Three-dimensional model of purple membrane obtained by electron microscopy. *Nature (London)*, **257**, 28–32.
- Henderson, R., Baldwin, J. M., Ceska, T. A., Zemlin, F., Beckmann, E. & Downing, K. H. (1990). Model for the structure of bacteriorhodopsin based on high-resolution electron cryo-microscopy. *J. Mol. Biol.* **213**, 899–929.
- Herzfeld, J., Das Gupta, S. K., Farrar, M. R., Harbison,

- G. S., McDermott, A. E., Pelletier, S. L., Raleigh, D. P., Smith, S. O., Winkel, C., Lugtenburg, J. & Griffin, R. G. (1990). Solid-state ^{13}C NMR study of tyrosine protonation in dark-adapted bacteriorhodopsin. *Biochemistry*, **29**, 5567–5574.
- Holz, M., Drachev, L. A., Mogi, T., Otto, H., Kaulen, A. D., Heyn, M. P., Skulachev, V. P. & Khorana, H. G. (1989). Replacement of aspartic acid-96 by asparagine in bacteriorhodopsin slows both the decay of the M intermediate and the associated proton movement. *Proc. Nat. Acad. Sci., U.S.A.* **86**, 2167–2171.
- Honig, B. H. & Hubbell, W. L. (1984). Stability of "salt bridges" in membrane proteins. *Proc. Nat. Acad. Sci., U.S.A.* **81**, 5412–5416.
- Honig, B. H., Ebrey, T., Callender, R. H., Dinur, U. & Ottolenghi, M. (1979). Photoisomerization, energy storage, and charge separation: a model for light energy transduction in visual pigments and bacteriorhodopsin. *Proc. Nat. Acad. Sci., U.S.A.* **76**, 2503–2507.
- Lewis, A., Marcus, M. A., Ehrenberg, B. & Crespi, H. (1978). Experimental evidence for secondary protein-chromophore interactions at the Schiff base linkage in bacteriorhodopsin: molecular mechanism for proton pumping. *Proc. Nat. Acad. Sci., U.S.A.* **75**, 4642–4646.
- Linderström-Lang, K. (1924). On the ionisation of proteins. *C.r. Lab. Carlsberg*, **15**, 1–29.
- Oesterhelt, D. & Tittor, J. (1989). Two pumps, one principle: light driven ion transport in halobacteria. *Trends Biochem. Sci.* **14**, 57–61.
- Otto, H., Marti, T., Holz, M., Mogi, T., Stern, L. J., Engel, F., Khorana, H. G. & Heyn, M. P. (1990). Substitution of amino acids Asp-85, Asp-212, and Arg-82 in bacteriorhodopsin affects the proton release phase of the pump and the pK of the Schiff base. *Proc. Nat. Acad. Sci., U.S.A.* **87**, 1018–1022.
- Powell, M. J. D. (1977). Restart procedures for the conjugate gradient method. *Math. Program.* **12**, 241–254.
- Press, W. H., Flannery, B. P., Teukolsky, S. A. & Vetterling, W. T. (1986). *Numerical Recipes. The Art of Scientific Computing*, Cambridge University Press, Cambridge.
- Richards, F. M. (1977). Areas, volumes, packing, and protein structure. *Annu. Rev. Biophys. Bioeng.* **6**, 151–176.
- Sternberg, M. J. E., Hays, F. R. F., Russell, A. J., Thomas, P. G. & Fersht, A. R. (1987). Prediction of electrostatic effects of engineering of protein charges. *Nature (London)*, **330**, 86–88.
- Stoeckenius, W. & Bogomolni, R. A. (1982). Bacteriorhodopsin and related pigments of halobacteria. *Annu. Rev. Biochem.* **51**, 587–616.
- Subramaniam, S., Marti, T. & Khorana, H. G. (1990). Protonation state of Asp (Glu)-85 regulates the purple-to-blue transition in bacteriorhodopsin mutants Arg-82→Ala and Asp-85→Glu: the blue form is inactive in proton translocation. *Proc. Nat. Acad. Sci., U.S.A.* **87**, 1013–1017.
- Sudmeier, J. L. & Reilly, C. N. (1964). Nuclear magnetic resonance studies of protonation of polyamine and aminocarboxylate compounds in aqueous solution. *Anal. Chem.* **36**, 1699–1706.
- Tanford, C. & Kirkwood, J. G. (1957). Theory of protein titration curves. I. General equations for impenetrable spheres. *J. Amer. Chem. Soc.* **79**, 5333–5339.
- Tanford, C. & Roxby, R. (1972). Interpretation of protein titration curves. Application to lysozyme. *Biochemistry*, **11**, 2192–2198.
- Tavan, P., Schulten, K. & Oesterhelt, D. (1985). The effect of protonation and electrical interactions on the stereochemistry of retinal Schiff bases. *Biophys. J.* **47**, 415–430.

Edited by F. Cohen