doi:10.1006/jmbi.2001.4902 available online at http://www.idealibrary.com on IDEAL® *J. Mol. Biol.* (2001) **312**, 203–219

JMB

pK_a Calculations Suggest Storage of an Excess Proton in a Hydrogen-bonded Water Network in Bacteriorhodopsin

Velin Z. Spassov[1,4], Hartmut Luecke[2], Klaus Gerwert[3]* and Donald Bashford[1]*

[1]*Department of Molecular Biology, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla CA 92037, USA*

[2]*Department of Molecular Biology and Biochemistry University of California, Irvine 3205 BioSci II, Irvine CA 92697-3900, USA*

[3]*Ruhr-Uni Bochum, Lehrstuhl für Biophysik, Universitätsstr. 150, Gebäude ND 04 Nord D-44780, Bochum, Germany*

[4]*Institute of Biophysics Bulgarian Academy of Sciences 1113 Sofia, Bulgaria*

**Corresponding authors*

Calculations of protonation states and pK_a values for the ionizable groups in the resting state of bacteriorhodopsin have been carried out using the recently available 1.55 Å resolution X-ray crystallographic structure. The calculations are in reasonable agreement with the available experimental data for groups on or near the ion transport chain (the retinal Schiff base; Asp85, 96, 115, 212, and Arg82). In contrast to earlier studies using lower-resolution structural data, this agreement is achieved without manipulations of the crystallographically determined heavy-atom positions or *ad hoc* adjustments of the intrinsic pK_a of the Schiff base. Thus, the theoretical methods used provide increased reliability as the input structural data are improved. Only minor effects on the agreement with experiment are found with respect to methodological variations, such as single *versus* multi-conformational treatment of hydrogen atom placements, or retaining the crystallographically determined internal water molecules *versus* treating them as high-dielectric cavities. The long-standing question of the identity of the group that releases a proton to the extracellular side of the membrane during the L-to-M transition of the photocycle is addressed by including as pH-titratable sites not only Glu204 and Glu194, residues near the extracellular side that have been proposed as the release group, but also an $H_5O_2^+$ molecule in a nearby cavity. The latter represents the recently proposed storage of the release proton in an hydrogen-bonded water network. In all calculations where this possibility is included, the proton is stored in the $H_5O_2^+$ rather than on either of the glutamic acids, thus establishing the plausibility on theoretical grounds of the storage of the release proton in bacteriorhodopsin in a hydrogen-bonded water network. The methods used here may also be applicable to other proteins that may store a proton in this way, such as the photosynthetic reaction center and cytochrome *c* oxidase.

Keywords: bacteriorhodopsin; electrostatic model; H-bonded network; Zundel proton; proton release

Introduction

This paper reports calculations, using a semi-macroscopic electrostatic model, of the states of protonation of various ionizable groups in the ground state of the light-driven, trans-membrane ion pump, bacteriorhodopsin. In particular, we have modeled various alternative sites for the storage of a proton that is released to the extracellular side of the membrane early in the photocycle, including the possibility that this proton resides in a hydrogen-bonded water network inside the protein rather than on any particular protein side-chain.

Present address: V. Z. Spassov, MSI Inc., 9685 Scranton Road, San Diego, CA 92121, USA.

Abbreviations used: bR, bacteriorhodopsin; FTIR, Fourier transform infrared (spectroscopy).

E-mail addresses of the corresponding authors: bashford@scripps.edu

0022-2836/01/010203–17 $35.00/0

Bacteriorhodopsin (bR) is a transmembrane proton-transport protein for which there exists a wealth of structural and mechanistic information from a variety of sources, including visible light and FTIR spectroscopy, electron microscopy, electron diffraction, X-ray crystallography, and NMR spectroscopy; and it has become a prototype for other transmembrane proteins whose function includes ion transport, such as the photosynthetic reaction center and cytochrome *c* oxidase.[1–4] The retinal chromophore is bound to a lysine side-chain of bR by a Schiff base linkage, which in the ground state of bR, is protonated. Excitation by light causes a rapid isomerization of the chromophore triggering a series of protonation state changes involving the Schiff base and several other groups in bR, as well as structural transitions, that ultimately result in the release of a proton to the extracellular side of the membrane, and a proton uptake from the cytoplasmic side.[5–8] The cycle involves intermediates that are designated, J, K, L, M, N, and O, in order of their appearance. Of particular interest here is the release of a proton from a group, X, to the extracellular surface of the protein during the L-to-M transition, coupled to the movement of a proton from the Schiff base to Asp85.

The identity of the group X, upon which the release proton is stored in the ground state, remained unknown even after the identities of a number of other groups in the proton transport chain were established. As more complete and accurate structural models became available, it appeared that a close pair of glutamic acid residues near the extracellular surface, 194 and 204, were well positioned to play the role of group X. Site-directed mutagenesis of these residues clearly showed that they were involved in proton release, and early FTIR experiments appeared to suggest that one or the other these groups was itself group X.[9–13] But recently, time-resolved nanosecond step scan FTIR experiments by Rammelsberg *et al.*[13] exclude protonation changes of R82, E204, E194 or other side-chains in this region during the transition to the M state, ruling out the identification of these groups as the release group. Instead, these experiments showed that an IR continuum absorption band present in the ground state disappears during the transition to M, when the proton is released. Spectroscopic studies on a number of model systems have shown that such continuum bands are characteristic of an excess proton within an hydrogen-bonded network;[14] and in *ab initio* calculations of an excess proton in water, the fluctuation of the excess proton between oxygen atoms in species such as $H_5O_2^+$ and higher-order complexes are predicted to give rise to continuum IR bands.[15] It was therefore proposed that the disappearance of the continuum band in the L-to-M transition is related to the deprotonation of the proton release group, and that group X was therefore an excess proton within a hydrogen-bonded network involving at least two internal water molecules.[13] Recently, FTIR experiments on deuterated samples have provided additional support for this proposal.[16] An $H_5O_2^+$ molecule in a protein cavity is the minimal model for such a protein release group.

Although the FTIR results provide strong evidence for the chemical character of the release group, they cannot resolve its specific location. Specific water positions within several cavities in the extracellular half of the protein are provided by a recent bR structure determined at 1.55 Å resolution by X-ray crystallography,[17] but the X-ray data do not give direct evidence of protonation states. Furthermore, consideration of typical pK_a values make it seemingly implausible that a proton would reside preferentially on a species such as $H_5O_2^+$ rather than on carboxylic acid groups. We therefore take an approach integrating the available experimental data with theoretical calculations aimed at finding a specific model for the storage of the release proton, and evaluating its plausibility by energetic calculations using semi-macroscopic electrostatic models that have proven successful in the past for studies of protonation states and pK_a values in proteins, including bacteriorhodopsin.[18–23] This might also provide a paradigm for studies of other proton transporting membrane proteins, such as the photosynthetic reaction center and cytochrome *c* oxidase, in which IR continuum bands have also been observed.

In addition, the availability of high-resolution structural data, including internal water molecules, makes it possible to re-visit the problem of calculating the titration properties of other functionally important groups in bR, to see whether improved structural data result in improved predictions, as one would hope, and to compare models treating internal water molecules explicitly to models treating water-filled cavities as cavities of high-dielectric continuum. Because of the high quality of the structural data, we have taken a conservative approach to structural modeling, generally preferring to leave heavy-atom positions unchanged. An extension of the usual theoretical models used in semi-macroscopic pK_a calculations was necessary to incorporate the idea of $H_5O_2^+$ as an ionizable site within a protein.

Results

Since a major aim of this work is to explore the ground-state location of the extracellular release proton, the present calculations were designed to allow for each of the following alternative locations that have been proposed on experimental grounds:[11–13] Glu204, Glu194 or a hydrogen-bonded water network, so that their plausibility on theoretical grounds can be evaluated. Calculations were performed using structural models that differed as to whether crystallographic water molecules were explicitly included or whether they were removed, so that water-filled cavities were

represented as cavities of bulk-water-like dielectric; whether a single conformer was used or whether a multi-conformer calculation with respect to hydrogen-atom placements was done; and whether the possibility of an $H_5O_2^+$ molecule in a protein cavity was included. In all cases, glutamic acid residues 194 and 204 were included as possible proton storage sites. Other aims of the work are to determine whether pK_a values and ionization states for other groups involved in the photocycle can be predicted when a high resolution (1.55 Å) X-ray crystallographic structure[17] is used rather than the lower-resolution structures used in the first calculations of this kind for bR,[18] and whether this agreement can be achieved without the *ad hoc* structure manipulations, or the fixing of the Schiff base protonation that were previously required. For this reason, the residues considered as ionizable are not confined to the proton-release region, but include ionizable groups throughout the molecule, including the Schiff base. In most calculations here, no changes are made to the crystallographically determined heavy-atom positions, otherwise changes are small and confined to a few atoms.

Calculations involving only the Schiff base and standard side-chains as ionizable groups were carried out using the now-standard combination of multi-site titration methods with a model assuming macroscopic electrostatics with atomic detail (MEAD model) for the effects of the protein environment on ionization energetics.[18,24] Extentions to this methodology to allow the possibility of $H_5O_2^+$ as a proton storage site, and multiple hydrogen atom placements, are described in Theory, Methods and Structural Models. That section also provides an outline of the standard theory, including definitions of several quantities used in presenting the results, pK_{half} (the closest analog to pK_a in a coupled, multi-site system), pK_{intr}, and the ''Born'' and ''background'' terms.

In some cases the calculated pK_{half} values are below 0 or above 15, and some care is needed in the interpretation of the results. Since the calculations are based on either a single conformer, or a model with conformational flexibility limited to hydrogen atoms, extreme pK_{half} values should be understood as predictions that it is not possible to protonate (in the p$K_{\mathrm{half}} < 0$ case) or deprotonate (in the p$K_{\mathrm{half}} > 15$ case) the site within the normal pH range while keeping the structure unchanged. This implies that any actual titration of these groups would be energetically coupled to structural changes not accounted for in the model. In particular, a number of sites are known to remain in particular protonation states throughout the pH range in which bacteriorhodopsin is stable, so that these sites should be coupled to the acid or alkaline denaturation of the protein. It is to be expected that extreme pK_{half} values will be calculated for such sites.[18]

Electrostatic potentials in the channels

Internal cavities large enough to be occupied by one or more water molecules were located by a simple algorithm that checked possible solvent-sized probe positions on a cubic lattice for overlap with protein heavy-atoms, and by graphical examination to distinguish internal pockets from the external surface. It was confirmed that all internal crystallographic water molecules fall into these cavities, except for one water molecule making a very short H-bond with Asp115. The boundaries of these cavities are displayed in Figure 1 where they are also colored according to electrostatic potential.

Immediately below (to the extracellular side of) the Schiff base is a water accessible cavity, and below that is a second, larger cavity. We denote these cavities as EC_u and EC_d, respectively. They are separated by the guanidinium group of Arg82 which forms both a steric and (for proton transport) an electrostatic barrier between the two cavities. EC_d is near the extracellular membrane surface, in direct contact with glutamic acid residues 194 and 204, which have been implicated in the proton release. If residues 194 and 204 are both deprotonated, EC_d contains a strong negative electrostatic potential. However, if one of these residues is protonated, the cavity's potential becomes mildly positive. Therefore, EC_d is an attractive candidate for the possible location of a proton stored in a hydrogen-bonded water network provided Glu194 and Glu204 are deprotonated.

The change in color of the EC_d cavity in Figure 1(a) from deep red at the bottom to pink at the top reflects the fact that the electrostatic potential is substantially more negative in the bottom part of the pocket, near the negative charges of the glutamate residues, than at the top, near Arg82. This fact is of particular significance in the $H_5O_2^+$ calculations described subsequently.

Models without explicit water

The simplest model for protonation state calculations, which we denote here as S, is a single conformer with no explicit water molecules. The internal cavities, which in the crystal structure contain water molecules, here are filled with high-dielectric ($\varepsilon = 80$) continuum. The hydrogen atom positions were those generated by the HBUILD and the minimization procedure described in Theory, Methods and Structural Models. The single-conformer version of the protein titration theory was used. In contrast to previous studies[18,20–22] based on older structural models, there was no *ad hoc* adjustment of the Schiff-base pK_{mod}, or re-positioning of the Arg82 side-chain or alteration of any other protein heavy-atom positions from the crystallographic structure.

The calculated pK_{half} values, as well as the Born and background terms contributing to the intrinsic pK, are shown in Table 1. The ΔpK_{Born} terms reflect the energetic penalty of burying the charged

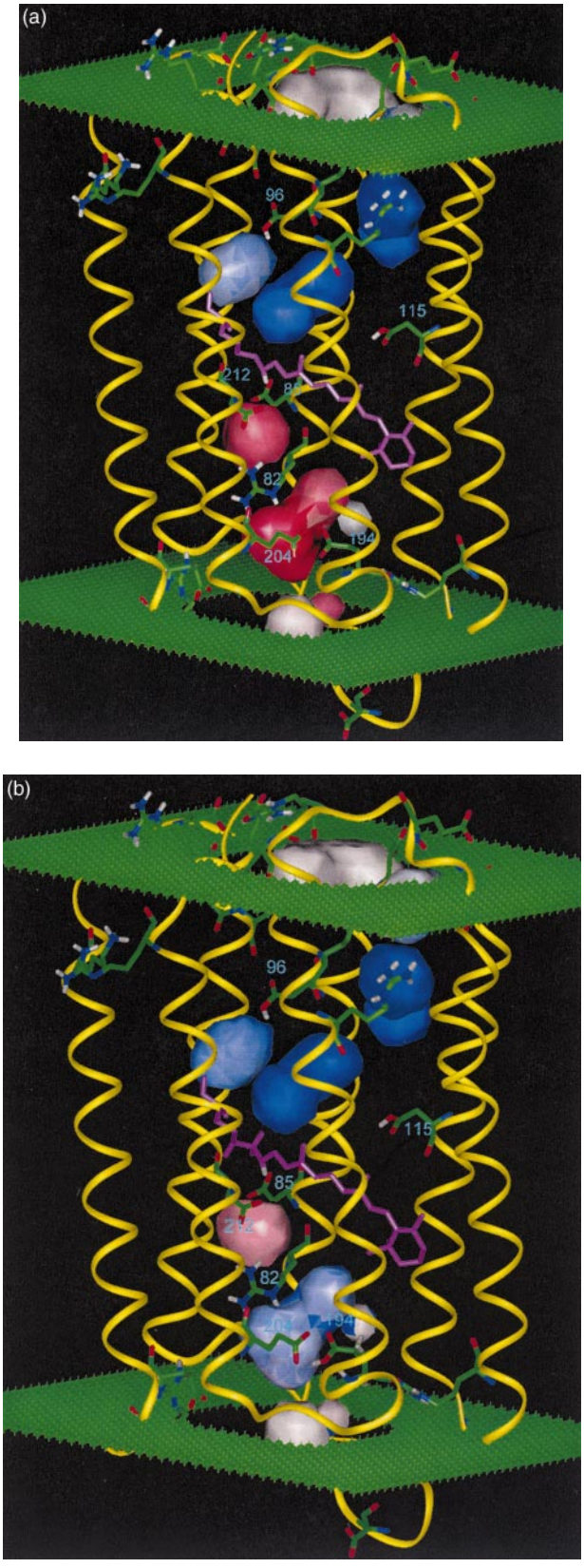


Figure 1. Internal cavities in bacteriorhodopsin and their electrostatic potential. The protein backbone is shown in yellow and the membrane surfaces as green meshes. The extracellular side is at the bottom. Selected side-chains are also shown. The irregular shapes are the surfaces of internal cavities found by the methods described in the text. These surfaces are colored from red to blue for more negative to more positive electrostatic potential, respectively. The potential is calculated by the MEAD potential program, according to the calculated state of protonation at neutral pH (see Table 1), except as noted for the sub-figures. The EC_d cavity is just below Arg82 and the EC_u cavity is just above it. (a) Both Glu194 and Glu204 are assumed to be deprotonated. (b) Like (a) except Glu194 is protonated.

forms of ionizable groups in the low-dielectric protein interior, and tend to be larger for more localized charges. Thus, these terms are consistently large and positive for buried carboxylic acid groups, while for the buried arginine residues and the Schiff base they are negative and of lesser magnitude. The ΔpK_{back} terms reflect electrostatic interactions with non-titrating polar groups, such as the backbone amides, and often compensate for the Born terms of buried groups, particularly in more hydrophilic regions of the protein. Finally the interactions between ionizable groups produce further effects, which may be quite strong in regions where several such groups are clustered together, to produce the individual site titration curves from which the pK_{half} values are extracted (see Theory, Methods and Structural Models).

The resulting pK_{half} values are in good general agreement with the available experimental measurements. The Schiff base and Asp96 have very high calculated pK_{half} values, which are consistent with the finding that in the ground state these groups remain protonated until the alkaline denaturation point. The high pK_{half} value for Asp96, which must be in the protonated state in order to act as a proton donor to the Schiff base in the M-to-N transition, is mainly the result of the large, positive ΔpK_{Born} term which is not significantly compensated by the background term in the relatively hydrophobic region where Asp96 is buried. In contrast, the Schiff base has a relatively small negative ΔpK_{Born} term, because the positive charge of the protonated state is delocalized over the retinal, which tends to reduce the penalty of burial. Its high pK_{half} is due to favorable interactions between its positive charge and nearby negative charges, primarily Asp85 and Asp212. Asp85, the recipient of a proton in the L-to-M transition, has a calculated pK_{half} of 1.7, which is somewhat lower than the experimentally measured pK_a of the acid-induced purple to blue transition (2.6). Asp212 and Arg82, which are part of the counterion complex just below (on the extracellular side of) the Schiff base, are predicted to be in their charged forms throughout the pH range in which the protein is stable, in agreement with experiment. The charged forms of the Schiff base and its counterion complex (glutamate residues 85 and 212, and Arg82) are stabilized by the strong site-site interactions between these groups, which is more than sufficient to offset the Born terms. A large upward pK_a shift is predicted for Asp115, in agree-

Table 1. pK values calculated for single conformation without explicit water (model S)

Site[a]	pK_{mod}	ΔpK_{Born}[b]	ΔpK_{back}[b]	pK_{intr}	pK_{half}	Experiment[55–58]
Arg7	12.0	−3.0	2.1	11.1	14.6	
Glu9	4.4	5.2	−5.6	4.0	0.2	
Lys30	10.4	−10.3	11.7	11.8	14.1	
Asp36	4.0	3.6	−5.4	2.2	<0.0	
Asp38	4.0	1.3	0.2	5.5	3.5	
Lys40	10.4	−4.0	1.0	7.4	7.7	
Lys41	10.4	−2.7	0.8	8.5	12.9	
Glu74	4.4	0.2	0.4	5.0	4.6	
Arg82	12.0	−6.0	−4.1	1.9	>15.0	>13.8
Asp85	4.0	10.3	−6.5	7.8	1.7	2.6
Asp96	4.0	10.1	−1.7	12.4	>15.0	>12.0
Asp102	4.0	0.4	0.4	4.8	3.2	
Asp104	4.0	1.8	−1.2	4.6	< 0.0	
Asp115	4.0	11.6	−5.1	10.5	8.4	> 9.5
Lys129	10.4	−0.6	0.1	9.9	10.0	
Arg134	12.0	−6.4	11.3	16.9	>15.0	
Lys159	10.4	−3.7	−0.6	6.1	10.5	
Glu161	4.4	0.4	0.2	5.0	3.7	
Arg164	12.0	−1.7	4.2	14.5	>15.0	
Glu166	4.4	2.7	−0.1	7.0	< 0.0	
Lys172	10.4	−2.3	0.2	8.3	7.1	
Arg175	12.0	−6.3	4.5	10.2	12.8	
Glu194	4.4	9.6	−2.5	11.5	>15.0	[c]
Glu204	4.4	9.6	−6.0	8.0	< 0.0	[c]
Asp212	4.0	10.5	−8.5	6.0	< 0.0	< 2.5
Arg225	12.0	−6.5	3.5	9.0	0.5	
Arg227	12.0	−0.2	0.9	12.7	14.0	
SB (216)	7.0	−3.3	−1.7	2.0	>15.0	>12.0

[a] Tyrosine residues were also included, but all were found to have pK_{half} values well above 15.
[b] The ΔpK are the pK equivalents of $\Delta\Delta G_{Born}$ and $\Delta\Delta G_{back}$ of equation (2). pK_{int} = pK_{mod} + ΔpK_{Born} + ΔpK_{back}.
[c] The ''release group'' is believed to involve Glu194 or Glu204 or a network of water molecules near these residues. It is certainly protonated at neutral pH. Experimental work indicates a release-group pK_a value of 9.5,[32] or 9.0 (R. Rammelsberg & K.G., unpublished results).

ment with experiment, but the magnitude of the shift is somewhat under-predicted. The calculations predict very high pK_{half} values for all tyrosine residues, which is consistent with the finding that no tyrosinate is seen in the ground state at pH values below 12.[25]

As for the glutamic acid residues 194 and 204, which have been suggested as candidates for the proton release group, 194 is predicted to be protonated throughout the pH range in which the protein is stable, while 204 is predicted to be deprotonated. Thus, the S model is consistent with 194 acting as the release group, but it should be pointed out that an excess proton in the internal water is not included as a possibility in this model. FTIR experiments exclude a protonation change of either 194 or 204 during the photocycle of the wild-type.[13] Furthermore, if either of these groups is protonated in the BR state at neutral pH, it is unlikely that the H-bonding is not changed during the photocycle. Such H-bonding changes would have been observed by FTIR, as in the case of D115, which is seen undergoing changes in the H-bonding environment in the K, L, M and N intermediates. It therefore seems very unlikely that either Glu204 or Glu194 are protonated at neutral pH.[13] In summary, if the possibility of $H_5O_2^+$ as a proton storage site is excluded, Glu194 is predicted to be the proton-release group, but this contradicts the experimental results.

To explore whether flexibility with respect to hydrogen-bonding arrangements could influence the results, calculations were performed that allowed for variable hydrogen atom positions in three distinct regions of the protein: Asp115/Asp96; Asp85/Asp212; and Glu194/Glu204. In each case, four possible hydrogen positions were included for each carboxylic acid: protonation on either of the two oxygen atoms, and for each case, the *syn* or *anti* orientation for the proton. For the states in which both members of a close pair were protonated, all 16 possible arrangements were included, so that correlations between proton positions within close pairs was fully accounted for. The multi-conformational formulation described in Theory, Methods and Structural Models was used. We denote this set of calculations as the MC model.

The pK_{half} values calculated using the MC model (Table 2) are very similar to those from the single-conformer calculation for most residues, the exception being Asp115, whose pK_{half} value is now in significantly better agreement with experiment. Examination of the calculated populations of the various conformers shows that in most cases, one conformation is strongly preferred over others, and that the preferred conformer is the same one found

Table 2. pK values for multi-conformer calculation without explicit water molecules (model MC)

Site	pK_{half}	Experiment
Asp96	>15.0	>12.0
Asp115	9.7	>9.5
Asp85	1.8	2.6
Asp212	<0.0	<2.5
Glu194	>15.0	[a]
Glu204	<0.0	[a]

The calculation includes all the same residues shown in Table 1, but here only sites treated by the multi-conformational formalism are shown. Results for the other sites are essentially identical to Table 1.
[a] See footnote [c] of Table 1.

in the H-placement scheme used to prepare the single-conformer calculation. In effect, the multi-conformer calculation has provided a test of the initial proton-placement scheme, and the results indicate that the initial placements were correct in the sense of obtaining optimal electrostatic interactions.

Models with crystallographic water molecules

Calculations were also carried out in which all 23 crystallographic water oxygen positions were retained. These include seven water molecules in internal cavities between the Schiff base and the extracellular membrane surface, and three in cavities between the Schiff base and the cytoplasmic surface.[17] Hydrogen atom positions were built onto the crystallographic oxygen atom positions for all of these water molecules, by the procedure described in Theory, Methods and Structural Models. Regions occupied by these water molecules are treated as ''protein'' as far as the dielectric model is concerned, which means that the interior region now is nearly all assigned a dielectric constant of 4, with only a few small interior pockets of high-dielectric continuum. Therefore, the partial stabilization that was previously provided to buried charges by the high-dielectric pockets in the above calculations (generally reducing the magnitude of the ΔpK_{Born} term), is now removed, but interactions of the charged form with the explicit water dipoles (now entering in the ΔpK_{back} term) may provide compensating stabilization.

The results of a single-conformer calculation with explicit water, which we denote as model WS (Table 3) are largely similar to the results without explicit water, but there is a striking reversal of the roles of glutamic acid residues 194 and 204. Glu204 now remains protonated throughout the pH range in which the protein is stable, while Glu194 remains deprotonated. Thus, the Glu194/Glu204 pair still stores one proton, but the proton now resides on Glu204 rather than Glu194. As before, the possibility of the release proton as an excess proton in the internal water network is not included in this model.

The pK_{half} value of Asp115 calculated using model WS is significantly lower than that calculated using the corresponding model with no explicit water (model S) and thus farther from the experimental result. The MC results, discussed above, suggest that this may be related to the choice of the Asp115 proton position. The WS model correctly predicts the Asp85 pK_a value to be shifted downward, but overestimates the magnitude of the shift. This is in contrast to the S model, which gives better agreement with experiment. On the whole, the single-conformer model with explicit water predicts the same protonation states as the model without explicit water, apart from the 194/204 reversal and the borderline protonation state of Asp115; but the quantitative agreement with available experimental pK_a data is not as good. This may be because the WS model does not allow the explicit water molecules to re-orient in response to protonation changes, while the high-dielectric interior cavities of the S model are able to respond by polarization changes.

The next model, WMC, allows for flexibility with respect to hydrogen-bonding arrangements, and limited re-orientation of the explicit water molecules. It was designed to focus on the influence of water molecules in the EC_d pocket on the position-

Table 3. Single-conformer calculation with explicit water (model WS)

Site	ΔpK_{Born}	ΔpK_{back}	pK_{intr}	pK_{half}	Experiment
Lys30	−10.3	11.3	11.4	13.2	
Arg82	−7.4	5.8	10.4	>15.0	>13.8
Asp85	11.5	−11.8	3.8	<0.0	2.6
Asp96	10.5	−1.6	12.9	>15.0	>12.0
Asp115	11.7	−6.3	9.4	7.0	>9.5
Lys159	−3.8	6.2	12.8	>15.0	
Arg175	−6.3	4.5	10.2	11.8	
Glu194	10.3	−8.5	6.2	<0.0	[a]
Glu204	10.6	−7.6	7.4	>15.0	[a]
Asp212	11.5	−11.5	4.0	<0.0	< 2.5
SB (216)	−3.4	0.4	4.0	>15.0	>12.0

The calculation includes all the same residues shown in Table 1, but here only sites which are either known to be involved in proton transport, or whose calculated pK_{half} value differs significantly from that calculated in the S model are shown.
[a] See footnote [a] of Table 1.

ing of the putative release proton on either Glu194 or Glu204, or the possibility of the two residues sharing the proton. As in the MC model, all 16 possible ways of positioning two protons on the side-chains of residues 194 and 204, and all eight ways of positioning one proton on the pair are included in a multi-conformer calculation. For each of these proton configurations, as well as for the doubly deprotonated state, water hydrogen positions were re-built using the HBUILD and minimization procedure described in Theory, Methods and Structural Models. This allows the water molecules to re-orient in response to changes in the protonation state. In order to avoid biasing the calculations in favor of more protonated states, an equal number of overall conformers (16) was included for the doubly protonated state, the state with Glu194 alone protonated, the state with Glu204 alone protonated and the doubly deprotonated state. In the doubly protonated case, there were already 16 conformers, but in the other cases, additional conformers were created by combining glutamic acid proton configurations for that particular state with water configurations that had been generated for other protonation states. As a result, a total of 64 conformations are included. The multi-conformer formalism is then applied. The energetic cost of this water re-orientation is accounted for by the reference-state energy differences calculated by the CHARMM22 energy function and MEAD-based solvation effects (see Theory, Methods and Structural Models). Again, the water molecules were treated only as H_2O, and not as possible sites of protonation.

The calculations using this model predict a pH-dependent sharing of a proton between glutamic acid residues 194 and 204 (Figure 2). The fraction of the proton on 204 is higher, but this result is sensitive to the details of the water positions and orientations, and thus to the fine details of the crystallographic structure determination and the hydrogen-building procedures. Essen *et al.*[26] have proposed that both 194 and 204 could serve jointly as a release group, with nearby water molecules exerting a significant influence.

In both this calculation and the WS model calculation, the inclusion of explicit water molecules rather than a pocket of high-dielectric continuum tends to shift the equilibrium towards the protonation of Glu204 rather than Glu194. Examination of the preferred water orientations shows a semi-linear chain of hydrogen bonds (Arg82-W403-W404-Glu194) connecting the negatively charged Glu194 with the positively charged side-chain of Arg82, leading to enhanced stabilization of the state in which Glu204, rather than 194, holds a proton. This effect occurs in both the WS and WMC models but its energetic consequences in the latter are not so strong as to cause complete occupancy of the state with Glu204 protonated.

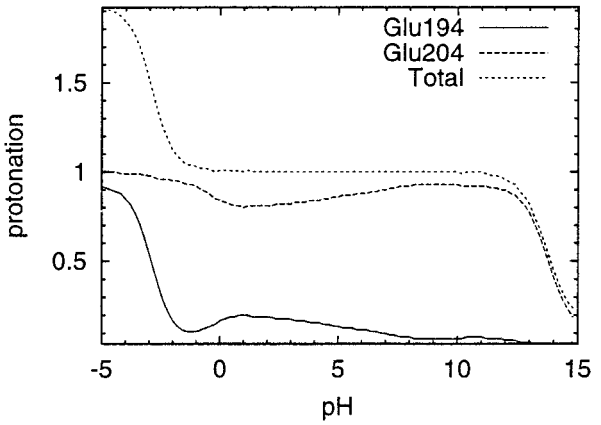


Figure 2. pH titration of Glu194 and Glu204 in model WMC (explicit water, multiple conformations). The curve for the sum of the protonation of Glu194 and Glu204 shows that throughout the normal pH range, one proton is present on the pair. This proton is shared between Glu194 and Glu204. This is in contrast to preceding calculations where either 194 or 204 holds the proton exclusively (Tables 1, 2 and 3).

Models with $H_5O_2^+$ in release channel

A model that allows for the possibility that the release proton is stored in a hydrogen-bonded water network has been prepared by building an $H_5O_2^+$ molecule near the Glu194/Glu204 pair. $H_5O_2^+$ is the minimal model for the storage of a proton in a low-barrier, water-water hydrogen bond of the type that has been proposed to be responsible for the observed continuum band in the FTIR difference spectra.[13] We note that some previous computational studies[22,27] included H_3O^+ as an ionizable site within the Schiff base-counterion complex, but this does not pertain to the question of the release proton or the continuum band.

The placement of the $H_5O_2^+$ is the main difficulty in constructing such a model. The EC_d pocket shown in Figure 1(a) was chosen as a likely location since it is near residues 194 and 204 on the extracellular side, and mutations of 204 and 194 have significant influence on the continuum band (R. Rammelsberg *et al.*, unpublished results). This pocket has a strong negative potential when residues 194 and 204 are deprotonated. The EC_u pocket can be excluded because the mutation, Asp212Asn, has no influence on the continuum band.[13] The EC_d pocket is bounded by the glutamic acid dyad on its extracellular side, and by Arg82 on the opposite side. In the crystal structure, this pocket contains three water molecules denoted W403, W404 and W405, and it is sufficiently large to accommodate a fourth water molecule as well.

Initially, we explored the possibility that the $H_5O_2^+$ molecule could be formed by using the oxygen positions from W403 and W404. These two water molecules form a hydrogen-bonded bridge

between the side-chains of Glu194 and Arg82. Preliminary protonation-state calculations showed that this $H_5O_2^+$ readily converts to $2H_2O$ throughout the pH range, because of its interaction with the positive charge of Arg82. However, a CHARMM energy minimization of this initial $H_5O_2^+$ position with the water oxygen atom near Glu194 (the W404 oxygen) restrained, caused the other oxygen atom (from W403,) which has a high *B*-factor and elongated density in the X-ray crystallographic study, to move away from the arginine and into an unoccupied region of the pocket between the carboxylate groups of glutamate residues 194 and 204. (During this minimization the position of other water and protein atoms was restrained, and residues 194 and 204 were deprotonated.) It was then possible, without steric clashes, to re-fill the vacated W403 position with a water molecule having the original W403 oxygen coordinates. In effect, the empty fourth water position noticed in this pocket in the crystal structure, a position very near the dyad, became occupied by the extra oxygen atom needed to change W404 from H_2O to $H_5O_2^+$. The resulting model is shown in Figure 3. It should be emphasized that as far as heavy-atom positions are concerned, none of the crystallographically determined coordinates have been altered; the model only introduces one new oxygen position.

The methods described in Theory, Methods and Structural Models for treating $H_5O_2^+$ as a titrating ''site'' in a protein were then applied to this structural model. We refer to this as the PW (protonated water) model. The resulting titration curves for the $H_5O_2^+$ and the two glutamic acid residues are shown in Figure 4. Below pH 0, the two glutamic acid residues are protonated but $H_5O_2^+$ is in the $2H_2O$ form. As the pH increases from 0 to 3, this two-proton state of the complex changes to a one-proton state, and the remaining proton is shared between $H_5O_2^+$ and Glu204, with the $H_5O_2^+$ receiving between 75 and 85% of the proton. This situation persists throughout the range between pH 4 and 15. To explore the robustness of these results with respect to the choice of the pK_{mod} value of $H_5O_2^+$ (see Theory, Methods and Structural Models), a calculation was done at pH 7.0 with the pK_{mod} value lowered by 1. The resulting fractional protonations of $H_5O_2^+$ and Glu204 were 0.35 and 0.65, respectively. Thus, in a calculation that allows for the possibility of a proton in an internal hydrogen-bonded water network in the proton release region, such a proton does indeed occur preferentially to a proton on the glutamic acid dyad, although uncertainty as to pK_{mod} values can reduce, but not eliminate the fraction residing on $H_5O_2^+$. The pK_{half} values of other groups are similar to those of the S model which is the non-$H_5O_2^+$ model most similar to this one (Table 4).

The region of Glu204/294 is somewhat disordered in the crystal structure.[17] The side-chains of Glu194, Glu204, Ser193, and water molecule 403 have high *B*-factors. Furthermore, the $H_5O_2^+$ position modeled by the above procedure produced a number of unusually short oxygen-oxygen distances between the extra oxygen atom and heavy atoms from the crystallographic structure, although strong electrostatic interactions more than offset unfavorable van der Waals interactions. This motivated additional calculations to explore the sensitivity of the results to structural variations.

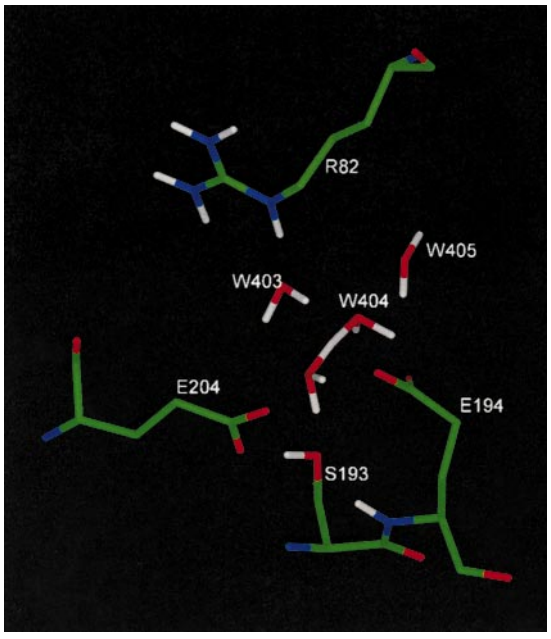


Figure 3. The $H_5O_2^+$ model and its environment in the lower extracellular side pocket, EC_d.

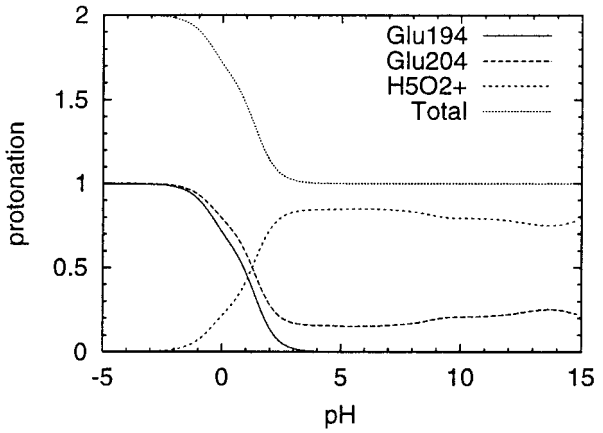


Figure 4. Individual site titration curves from model PW, which includes the possibility of $H_5O_2^+$ as the proton release group. The three lower curves pertain to the individual sites listed in the legend. The upper curve is the sum of the three lower curves. Above pH 4, one proton is present and resides mostly in the $H_5O_2^+$ form, though it is partly shared with Glu204.

Table 4. Models allowing $H_5O_2^+$

Site	PW model		PWR (relaxed) model	
	pK_{int}	pK_{half}	pK_{intr}	pK_{half}
Glu9	4.5	0.2	4.6	1.0
Lys30	11.0	12.3	11.2	12.7
Asp85	3.9	<0	3.9	<0
Asp115	10.7	8.8	10.6	8.8
Glu194	11.6	0.9	10.9	<0
Glu204	12.3	1.3	10.4	<0
Arg225	6.9	< 0	6.9	<0
Arg227	12.7	14.3	12.7	14.3
$H_5O_2^+$	−8.8	[a]	−7.9	[a]

The sites shown in Table 1 were included in the calculations, but only results differing by >0.2 pK unit from Table 1 (model S) are shown.

[a] Result is a complex titration curve (see Figures 4 and 5).

A structural model was therefore prepared by an energy minimization that allowed apparently disordered (high *B*-factor) heavy atoms to relax. Starting with the above PW model, energy minimizations were performed in which the heavy atoms of the side-chains of Ser193, Glu194 and Glu204 and the oxygen atoms of water 403 and the $H_5O_2^+$ were allowed to move with only weak harmonic constraint. As before, all hydrogen atoms were allowed to move without constraint. The protonation states were the same ones used for the PW model preparation. The largest heavy-atom movements were seen for the O^{ε} atoms of Glu194 (0.46 Å r.m.s.d.). The O^{ε} atoms of Glu204 moved by 0.28 Å; the O^{γ} atom of Ser193 moved by 0.17 Å; one of the oxygen atoms of $H_5O_2^+$ moved by 0.32 Å; and water molecule W403 moved by 0.15 Å. We designate this relaxed model as PWR. The calculations for the ionization states of this model (Table 4, relaxed-model column, and Figure 5) give the result that the proton remains on the $H_5O_2^+$ throughout the pH range, while effects on other groups are relatively small. The tendency of the release proton to prefer the $H_5O_2^+$ form rather than a glutamic acid is even stronger in this calculation than in the PW model. There is no discernable sharing of the proton in Figure 5. Sharing of the proton with Glu204 occurs only if the pK_{mod} for $H_5O_2^+$ is artificially lowered to the unrealistic value of −4.7, as compared to the pK_{mod} value of −1.7 adopted here (see Theory, Methods and Structural Models).

Discussion

In general, good agreement was obtained between the calculations and previously published experimental results for the protonation states and, where available, the pK_a values of ionizable residues on, or closely involved in the ion transport pathway. Inclusion of conformational variability with respect to proton placement improved the accuracy of the calculated pK_a values in a few cases, and in other cases served as a check on the initial proton placements while leaving the calculated pK_a values unchanged. It might be thought that inclusion of explicit water molecules in internal cavities should be preferred over the modeling of water-filled cavities as pockets of high dielectric constant, particularly when water positions are available from X-ray crystallographic studies. However, in the usual formulation of electrostatic continuum models for pK_a calculations, this amounts to treating the water molecules as dipoles that do not re-orient even when adjacent side-chains change their ionization state. On the other hand, a high dielectric cavity is able to polarize in response to such events, although this can only be a crude model of the actual response of the cavity water. In the present work, calculations of both kinds were tried, including calculations allowing for limited explicit water reorientation through the minimization protocol. The results were generally similar, but inclusion of explicit water molecules in a single-conformer calculation led to somewhat worse agreement with experimental pK_a values than the treatment of internal water as high-dielectric cavities.

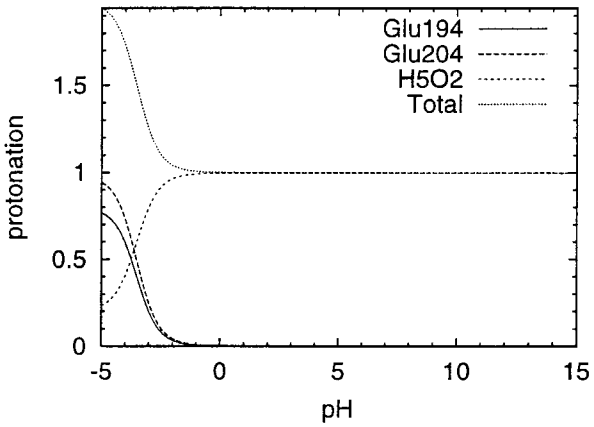


Figure 5. Individual site titration curves from model PWR, a relaxed variant of model PW (see Figure 4). In this case, the proton resides entirely on $H_5O_2^+$ throughout the normally accessible pH range.

In early computational studies of the protonation states of the ground state of bacteriorhodopsin[18,21] based on a more approximate structural model derived from cryo-electron microscopy data,[28] satisfactory agreement with experiment could only be obtained by manipulation of the Arg82 side-chain position. Since the data did not provide density corresponding to this side-chain, it was initially modeled in a ''down'' orientation away from the Schiff base. In order to provide enough stabilization of aspartate negative charges to predict both Asp85 and Asp212 to be deprotonated, Arg82 was moved into an equally plausible ''up'' orientation near these aspartate side-chains.[18,21] In some cases it was also necessary to add an *ad hoc* positive term to the intrinsic p*K* of the Schiff base to force it to remain protonated.[18] From this work it was obvious that the p*K* value of Asp85 depends on the position of Arg82, and that movement of Arg82 can influence the p*K* of Asp85 and could be part of the proton release trigger. Such movement is actually found now in the M state.[29,30] It is gratifying that in the present work, using a high resolution crystallographic structure, and a Schiff-base charge model based on high quality density-functional calculations, no manipulations of heavy-atom coordinates or pK_{intr} values were required to obtain good agreement with experiment.

The actual position of the guanidinium group, as revealed by the new crystal structure, is midway between the ''down'' and ''up'' positions. The EC_u cavity (Figure 1) corresponds approximately to the old ''up'' position, and the EC_d cavity corresponds to the old ''down'' position. In the present calculations, the stability of the release proton in the hydrogen-bonded network involves a balance between the stabilizing effect of negative charge from Glu194 and Glu204, and the destabilizing effect of the positive charge of Arg82. Indeed, Arg82 is thought to serve as a crucial link in a chain of conformational changes that trigger proton release.[3,22,31]

As for the groups which have previously been proposed as the group X, which releases a proton to the extracellular side during the L-to-M transition, the computational models that do not allow for the proton to be stored in water predict that either Glu204 and Glu194 holds the proton, or that the proton is shared between them, depending on the model details. A previous MEAD-type study[23] of various side-chains as possible proton release groups found that Glu204, but not Glu194, could act as the release group; but this study was based on older structural models in which Glu194 was positioned differently, and not in such close association with Glu204. In the development of a minimal model for a proton stored in a hydrogen-bonded water network, we found that the most plausible location was the lower part of the EC_d cavity shown in Figure 1. A position for an $H_5O_2^+$ molecule was found that allowed for favorable interactions with Glu204 and Glu194 in their carboxylate forms, while avoiding direct contact with the positively charged side-chain of Arg82 (Figure 3). When this $H_5O_2^+$ site was included in the protein titration model and allowed to compete for the release proton with the two glutamic acid residues, it was found that the proton did indeed reside preferentially on the $H_5O_2^+$. The calculations therefore indicate that a proton in a hydrogen-bonded network of internal water molecules is an energetically plausible candidate for the storage of the release proton as proposed on the basis of observation of a broad continuum band in the FTIR difference spectra.[13]

In all of the computational models, the proton release group, be it Glu194, Glu204 or $H_5O_2^+$, continues to hold the proton up to a pH value of 15 or more. This is in contradiction to experimental findings that the release group has a pK_a of 9-9.5[32] (R. Rammelsberg *et al.*, unpublished results). In the calculations, the fully deprotonated state is highly unfavorable because the two negatively charged glutamate residues are very close together, whereas the protonation of one of the glutamate residues, or of a nearby water dimer either removes this repulsive interaction or provides a compensating attractive interaction, respectively. It seems likely that in the real protein, deprotonation of the release group under alkaline conditions would cause Glu194 and Glu204 to move apart to lower the energy of the fully deprotonated state. Since none of the model calculations allow for pH-induced changes of heavy-atom positions, this avenue is closed, and the calculations over-predict the pK_a. The crystal structure of the ground state suggests a fairly high degree of conformational mobility in this region,[17] and in the recent crystallographic structure of the M state the region undergoes significant conformational change relative to the ground state, including a movement of the Arg82 guanidinium group towards the extracellular side.[29,30] It therefore seems plausible that the deprotonation of the release proton at around pH 9 in the ground state is accompanied by a shift to a more locally M-like structure in the region of the EC_d pocket.

The EC_d pocket in which the $H_5O_2^+$ is located is large enough to accommodate four water molecules, and three are seen in the crystallographic structure. Our $H_5O_2^+$ positioning is consistent with one of these three observed crystallographic water oxygen positions, combined with a fourth oxygen position in a vacant region of the pocket. In protein X-ray crystallography, typically only well-ordered water molecules show up in the density. Remaining cavities may or may not be filled with statically or dynamically disordered water molecules. (An example of a borderline case is water 403 in the EC_d pocket, which, as evidenced by its *B* factor and elongated density, is partially disordered.[17]) Thus, while there is no positive crystallographic evidence for the added oxygen atom position that we have modeled, it is not inconsistent with the crystallographic data.

Examination of the electrostatic potentials in the pocket, and preliminary titration calculations indicated that both of the oxygen atoms of $H_5O_2^+$ must be within hydrogen bonding distance of the carbonyl oxygen atoms of the glutamate side-chains 194 or 204, in order to be stable as $H_5O_2^+$ (rather than as $2H_2O$). A hydrogen bond with the Ser193 side-chain also provides crucial stabilization. Very recently, mutation of Ser193 to Cys has been found to abolish the continuum band observed in the wild-type, and to delay proton release (R. Rammelsberg *et al.*, unpublished results). $H_5O_2^+$ models with oxygen atom positions ''higher'' in the pocket are unsustainable because of unfavorable interaction with the positively charged Arg82 at the top of the pocket. We also made preliminary calculations based on a model in which the release proton was stored as H_3O^+; but in titration calculations this model collapsed to H_2O and a proton on either Glu194 or Glu204 (results not shown). In any case, a release proton stored as H_3O^+ would not account for the continuum band, while $H_5O_2^+$ is the minimal model that does account for it. A more extended ionized water system, such as $H_7O_3^+$, could also account for a continuum band, but it appears from the crystal structure that this could only occur by incorporating a third oxygen position that would be farther from the glutamate residues and closer to guanidinium group of Arg82. This would seem to be less energetically favorable than the $H_5O_2^+$ case, so no attempts were made to model ionized water complexes larger than $H_5O_2^+$. The electrostatic and steric barrier imposed by the Arg82 guanidinium group appears to provide a ''ceiling'' on any ionized water network storing the release proton, while the negative charges of the glutamate 194 and 204 carboxylate groups provide a ''floor.'' In view of these structural constraints, and the lack of any other pockets with strong negative potential in the putative proton-releasing region of the protein, we believe that if the release proton is indeed stored in an internal water network, as suggested by FTIR spectroscopy[13] and the present results, it must be very close to the $H_5O_2^+$ position that we have modeled. Finally, in the recently obtained M-state structures,[29,30] EC_d no longer exists as a closed pocket, but is opened to the extracellular side, and the arrangement of the polar and charged groups lining it is substantially changed. The identification of this pocket as the proton storage site is consistent with its disruption in the state from which the proton has presumably been released.

As mentioned in the Introduction, it would at first seem implausible that $H_5O_2^+$ would be preferred over a glutamic acid as the location of a proton, because the typical pK_a value of a glutamic acid, 4.4, is higher than that of $H_5O_2^+$ (approximately −1.7). Burial of these groups would make the pK_a difference even greater, by penalizing the formation of both glutamate and $H_5O_2^+$ relative to the neutral forms. Two effects compensate for these tendencies in the calculations. First, because the charge of $H_5O_2^+$ is delocalized across two water molecules, it does not suffer as large a Born desolvation penalty as a more compact species, such as H_3O^+ would do. Second, the $H_5O_2^+$ has very strong favorable interactions with the negative charges of the two glutamate residues: its interaction energies with Glu194 and Glu204 are equivalent to 17.4 and 14.4 pK units, respectively. If the proton were to instead reside on either Glu194 or Glu204, these favorable interactions would be lost, although the unfavorable interaction equivalent to 8.4 pK units between the two glutamate residues would also be lost. The net interaction energy terms among these three groups (Glu194,204 and $H_5O_2^+$) thus favors protonation of the $H_5O_2^+$ by 23.4 pK units, which is enough to compensate for the pK_{intr} difference of 21.1 between Glu204 and $H_5O_2^+$ (see Table 4).

Theory, Methods and Structural Models

The calculations are based on the idea that the difference between the titration behavior of ionizable groups in a protein, and a group of the same kind in a model compound is caused mainly by differences in the electrostatic effects in the protein *versus* the model compound. It is further supposed that these electrostatic effects are adequately described by the following semi-macroscopic model: the solvent region is represented by a continuum having the bulk dielectric constant of water ε_w; the region inside the protein, membrane or model compound has a dielectric constant ε_{in}, which is a parameter of the model; the charge distributions of the protein or model compound, including both forms of the titratable groups, are modeled by atomic partial charges. The boundary between the interior and exterior of the molecule is the Connolly molecular surface[33] which is defined by the atomic radii and coordinates through a hypothetical process of rolling a solvent-sized probe sphere over the atomic spheres. The electrostatic potential ϕ is then determined by the Poisson equation:

$$\nabla\varepsilon(r)\nabla\phi(r) = -4\pi\rho(r) \qquad (1)$$

where the charge distribution ρ is defined by atomic partial charges, and the dielectric constant $\varepsilon(r)$ is ε_{in} or ε_w according to the region in which r falls. This equation can be solved by the finite-difference method.[34] We refer to this electrostatic model as MEAD (macroscopic electrostatics with atomic detail). The use of the MEAD model for calculations of ionization states in proteins has been described;[18,24] but in the present work some variations from the usual methods, particularly the inclusion of $H_5O_2^+$ as a titratable site, allowance for variable hydrogen-atom positions, and the presence of a membrane, require a fairly detailed description of the theoretical methods.

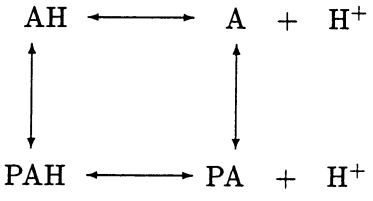
A side-chain as a titratable site in protein

Because the present study includes consideration of $H_5O_2^+$ as a titrating site, it is necessary to describe the ideas and assumptions underlying the standard treatment of ionizable side-chains to justify the extension of these methods to an $H_5O_2^+$ molecule in a protein cavity. The standard scheme starts with Tanford & Kirkwood's[35] observation that if one has a matrix of site-site interactions W_{ij}, and a set of "intrinsic pK_a" values, where the pK_{intr} of a site is defined as the pK_a that it would have if all other sites were neutralized, it is then possible in principle to calculate the full ionization behavior of the protein as a function of pH. The calculations of the pK_{intr} values are based on the assumption that any difference between the pK_{intr} and the pK_a of a corresponding model compound pK_{mod} is due purely to electrostatic effects.

The usual formulation, which is incorporated into several computer programs for protein electrostatics, including the MEAD suite used here[18,36] is:

$$pK_{intr} = pK_{mod} - \frac{\Delta\Delta G_{Born} + \Delta\Delta G_{back}}{2.303RT} \qquad (2)$$

where $\Delta\Delta G_{Born}$ is the difference between the Born-like solvation energy of the partial charges of the titrating group in its protonated *versus* deprotonated form in the protein *versus* the model compound, and $\Delta\Delta G_{back}$ is the corresponding difference of differences for the interaction of the titrating site with non-titrating charges in the protein or model compound (e.g. the peptide backbone). To put the assumptions involved in arriving at equation (2) more formally, let PAH and PA denote the protein with the site under consideration protonated or deprotonated, respectively, and all other sites neutralized. Let AH and A denote the protonated and deprotonated forms of the corresponding model compound for which pK_{mod} is presumed to be known. Consider the thermodynamic cycle:



Scheme 1.

where the vertical lines represent "alchemical" processes corresponding to "cutting out" the model compound from the protein, as suggested in Figure 1 of Bashford & Karplus.[24] Because 2.303 RT, $pK_{intr} = \mu^\circ_{PAH} - \mu^\circ_{PA} - \mu^\circ_{H+}$, for the protein and a similar expression applies to the model compound, we can write:

$$pK_{intr} = pK_{mod} - \frac{\Delta\mu^\circ_{prot} - \Delta\mu^\circ_{deprot}}{2.303RT} \qquad (3)$$

where we have introduced $\Delta\mu^\circ_{prot} = \mu^\circ_{PAH} - \mu^\circ_{AH}$ and $\Delta\mu^\circ_{deprot} = \mu^\circ_{PA} - \mu^\circ_{A}$, the differences in chemical potential associated with the vertical lines of the thermodynamic cycle. These differences can be expressed as:

$$\Delta\mu^\circ = \Delta G_{elec} + \Delta G_{ne} + \Delta G_{chem} \qquad (4)$$

where the elec component is the change in electrostatic interactions with the surroundings in the protein *versus* the model compound, the ne component is the change in non-electrostatic nonbonded interactions, and the chem component is for any changes not accounted for by the first two, such as the alteration of chemical bonds implied by the process of "cutting out" the model compound from the protein. Clearly, ΔG_{elec} will be quite different for the protonated and deprotonated forms. But ΔG_{ne} should be nearly the same, since the protonated and deprotonated forms are sterically very similar; and ΔG_{chem} should be nearly the same provided the cut is made far enough from the titrating functionality that there is no through-bond influence on its pK_a value. (In this work, for ionizable side-chains of amino acid residues, we follow the common practice of taking the model compound to be the residue and its two flanking backbone peptide groups.[24] In Asp, for example, the cuts are then four covalent bonds away from the carboxylic acid group.) The essential approximation of the standard method for calculating pK_{intr} is that the ΔG_{ne} and ΔG_{elec} terms will be the same for μ°_{prot} and μ°_{deprot}, so that equation (3) becomes:

$$pK_{intr} = pK_{mod} - \frac{\Delta G_{elec,prot} - \Delta\ G_{elec,deprot}}{2.303RT} \qquad (5)$$

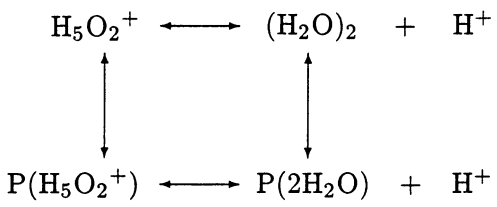
The usual expression for pK_{intr} (equation (2)) is then essentially equation (5), re-written in terms of Born and background electrostatic contributions (see Bashford & Gerwert[18] for the specific details of these terms).

The same solutions of equation (1) from which the $\Delta\Delta G_{Born}$ and $\Delta\Delta G_{back}$ terms are obtained can also be used to obtain the matrix of site-site interactions W_{ij}. The fractional protonation of any particular site in a protein with N titratable sites can then be calculated by a Boltzmann-weighted average over the 2^N possible protonation states, or by a suitable approximation of such an average.[37] Having determined the protonation as a function of pH for each site, the pK_{half} of a site is defined as the pH at which that site is half protonated. The pK_{half} value is thus roughly analogous to the pK_a value of a molecule with a single ionizable group.

$H_5O_2^+$ as titratable site within protein

Consider an $H_5O_2^+$ molecule confined within a protein cavity as a titrating site which, upon depro-

tonation, becomes a pair of water molecules within the protein cavity. We wish to treat this as a ''site'' within the framework of the usual methods of calculating ionization states of proteins. In this case, the thermodynamic cycle of Scheme (1) becomes:



Scheme 2

.

where P(. . .) denotes a species confined in the protein cavity, and in the upper reaction, the species are in bulk water. Here, the two vertical arrows represent not the alchemical ''cutting out'' process of Scheme (1), but rather the transfer of an $H_5O_2^+$ molecule, or a dimer of water molecules, respectively, between the protein cavity and the bulk. The chemical potential differences between the species in protein and in solution can again be divided into electrostatic and other components, as in equation (4), but in the present case there is no covalent bond breaking as previously represented by ΔG_{chem}. If we once again assume that the nonelectrostatic components of $\Delta\mu^\circ$ do not change between the protonated and deprotonated forms (i.e. the difference between the interaction of a water dimer and $H_5O_2^+$ with the protein cavity is purely electrostatic) then we again obtain equation (5), which is equivalent to the standard formula, equation (2).

It remains to choose an appropriate pK_{mod} value corresponding to the top equilibrium in Scheme (2). The difficulty is that in the deprotonated form, our ''model compound'' becomes a water dimer in bulk water, that is, K_{mod} is the dissociation constant for the reaction $H_5O_2^+ \leftrightarrow (H_2O)_2 + H^+$:

$$K_{mod} = \frac{[(H_2O)_2][H^+]}{[H_5O_2^+]} \tag{6}$$

It would not be correct to use $H_5O_2^+ \leftrightarrow 2H_2O + H^+$, as the top reaction, because in the protein reaction (bottom row of Scheme (2)) the proton is accepted by a *pre-formed* water dimer within the cavity. If we suppose that an excess proton in bulk water exists partly in the form H_3O^+ or oxonium (probably coordinated by other water molecules to make larger complexes, such as $H_9O_4^+$, as proposed by Eigen[38]), and partly in the form $H_5O_2^+$, as proposed by Zundel,[14] we can express the overall proton activity as $[H^+] = [H_3O^+] + [H_5O_2^+]$. Then equation (6) becomes:

$$\begin{aligned} K_{mod} &= \frac{[(H_2O)_2]([H_5O_2^+] + [H_3O^+])}{[H_5O_2^+]} \\ &= [(H_2O)_2](1 + 1/\alpha) \end{aligned} \tag{7}$$

where $\alpha = [H_5O_2^+]/[H_3O^+]$. In these expressions $[(H_2O)_2]$ refers to the concentration of water molecule pairs that have a suitable hydrogen-bonded geometry to be converted to $H_5O_2^+$. For a given microscopic state of a sample of N water molecules, one can, in principle, describe them as a collection of $N/2$ water pairs, and then enquire as to what fraction of them, f, meet the criteria of being readily convertible to $H_5O_2^+$. Using a total water concentration of 55 mole/l, and taking the negative logarithm of equation (6):

$$\begin{aligned} pK_{mod} &= -\log f(55/2)(1 + 1/\alpha) \\ &= -1.7 - \log f(1 + 1/\alpha)/2 \end{aligned} \tag{8}$$

Given the strongly H-bonded structure of water, it is likely that a high fraction of the water molecules are participating in dimers that have a suitable geometry for protonation to $H_5O_2^+$, therefore f should be close to 1. If one supposes that the proton in water occurs as a roughly equal mixture of $H_5O_2^+$ and oxonium,[39] then $\alpha \approx 1$. Combining these suppositions about f and α, one finds $\log f(1 + 1/\alpha)/2 \approx 0$, and $pK_{mod} \approx -1.7$. Admittedly, the above derivation of the pK_{mod} value is approximate, but uncertainties corresponding to even an order of magnitude in f or α correspond to uncertainties in pK_{mod} of only approximately 1. Therefore, in using this value one should check whether the final results are sensitive to pK_{mod} variations of order 1.

Membrane, dielectric boundary and charges

In this work, we model the membrane simply as an infinite dielectric slab having the same dielectric constant ε_{in} as the protein interior. In using the finite-difference method to solve equation (1), it is necessary to specify ϕ on the outer boundary of the finite-difference lattice (i.e. Dirichlet boundary conditions[40]). In many applications to biomolecules it is common to set this potential to zero, or to use a simple analytical approximation, such as Coulomb's law. However, these approximations may lead to significant error in the case of an infinite dielectric slab. We therefore use the method of images[40] to find the lattice boundary potential of a set of atomic partial charges in or near an infinite dielectric slab. Since the membrane in the present case is transverse to the coordinate z-axis, the obvious definition for $\varepsilon(r)$ for a protein-membrane system is that it is ε_{in} if r is either in the protein or if the z coordinate is between the two values that define the two surfaces of the membrane. However, bacteriorhodopsin, like many membrane spanning proteins, is known to have internal water-filled pockets or channels that lie between the two membrane surfaces, and in some cases, we

wish to be able to model these pockets as having the bulk water dielectric constant ε_w. Therefore, we define the ε_{in} region as follows: all points which lie within the protein interior, and all points which are between the top and bottom surfaces of the membrane and outside of a 10 Å cylinder whose axis is transverse to the membrane and passes through the center of the protein. The protein boundary used in this definition is the molecular surface,[33] including the surface of interior pockets. In other words, a point is outside the protein if it is possible to find a position for a solvent-sized (1.4 Å) spherical probe such that the point falls inside the probe, but the probe does not overlap any atoms of the molecule; otherwise the point is inside the protein.

The use of the method of images described above, which is incorporated into the software used here, is valid only for the Poisson equation, and not for the Poisson-Boltzmann equation, which would be needed in order to include salt effects. Therefore, we had to choose between including salt effects or including the membrane. Preliminary calculations including salt, but no membrane, or a membrane but no salt, showed that the membrane effects on calculated pK_{half} values were much larger, so we chose to include the membrane effects. The effects of the membrane's lipid head-group charges are also not included. It is expected that these would be largely cancelled by salt effects as an electrical double-layer is formed, and that calculations with head groups, but no salt, would be worse than calculations without head groups at all.

In early applications of the MEAD method to protein pK_a calculations the charge distributions used included the full set of atomic partial charges for the non-titrating charges, but titrating groups were modeled as having their formal charge concentrated on a single point.[24,41] Later, this single-site-charge restriction was lifted[18] so that the titrating site could also be described by atomic partial charges, but uncertainty as to proton placement was reflected by ''smearing'' the hydrogen-atom charges between heavy atoms. For example, the protonated form of glutamic or aspartic acid was modeled by distributing the hydrogen partial charge between the two oxygen atoms of the carboxylic acid group. In the present work we make an explicit placement of all hydrogen atoms and use atomic partial charges without smearing.

Multi-conformer case

In several cases (see Results) ensembles of structural models were generated which differed from one another as to the positions of the labile hydrogen atoms in the ionizable groups. Some cases also included variation of water hydrogen-atom positions as well. Calculations were then done in a way that allowed the system to select which members of this ensemble were most populated at various pH values. This was done by calculating the protonation of each site as a function of pH for each member of the ensemble, using the single-conformer methods described above, and then combining the results using integrals over proton binding isotherms, which allow the pH-dependence of the free energy of a conformational change between two members, A and B, of the conformational ensemble to be expressed as:

$$\Delta G_{AB}(\mathrm{pH}) - \Delta G_{AB}(\mathrm{pH_o}) = 2.302RT \int_{\mathrm{pH_o}}^{\mathrm{pH}} [Q_B(\mathrm{pH}') - Q_A(\mathrm{pH}')]d\mathrm{pH}' \quad (9)$$

where the Q are either the total charge or total number of protons bound in state A or B.[42–44] Using this formula, and the values of Q provided by the calculations for the individual members of the ensemble, the relative population of all members of the ensemble can be found at any pH, provided that the ratios are known at some reference pH, $\mathrm{pH_o}$. The average ionization of individual sites as a function of pH can then be obtained by weighting the results of the single-conformer calculations according to these ratios.

This leaves the problem of selecting a reference pH, and finding the relative populations of conformers at that pH. In calculations including only the conformational variability of the labile protons, the solution is simple. The reference pH is selected sufficiently high that all of the labile protons (typically carboxylic acid protons) are dissociated. Then all members of the ensemble become identical and have identical free energies and populations ($\Delta G_{AB}(\mathrm{pH_o}) = 0$). If the hydrogen atom positions of water molecules are also considered as flexible, it is again convenient to choose high pH so that all relevant ionizable groups are deprotonated, but in this case, variation as to water hydrogen atoms remains. Here, we calculate these energy differences using the CHARMM22 force field, with the dielectric constant set to 4.0 to maintain consistency with the MEAD calculations, and solvent effects are treated by the MEAD model.

Structural models, parameters and programs

All calculations are based on the 1.55 Å X-ray crystallographic structure of Luecke *et al.*[17] (PDB[45,46] entry 1C3W). With a very small number of exceptions, noted in Results, all heavy-atom positions from this structure were left unchanged. The missing loop, Thr157-Glu161, was constructed using the loop-building facility of the SYBYL molecular modeling system† Hydrogen-atom positions were initially generated using the HBUILD algorithm[47] as implemented in the CHARMM computer program,[48] and then energy minimized using the CHARMM22 force field[49] and the ABNR module of CHARMM with all heavy atoms fixed.

† Tripos Inc., SYBYL, St. Louis, Missouri, USA.

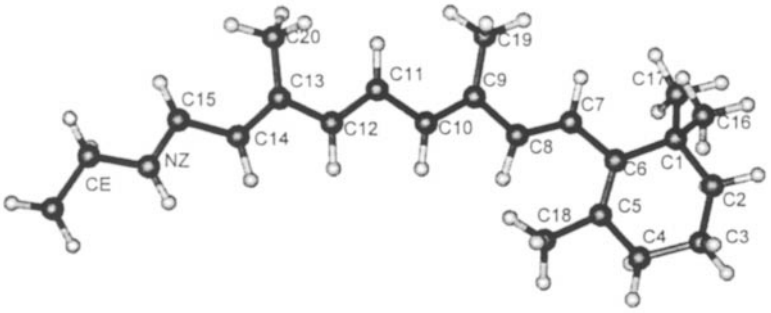


Figure 6. Structure of lysine-retinal with atom names used in Table 5.

In addition to atomic coordinates, MEAD calculations require atomic charges and radii. (The radii are used to define the dielectric boundaries.) The radii used are taken from Bondi,[50] and where available, the atomic partial charges are taken from the CHARMM22 force field. However, this force field does not have partial charges for either the protonated or deprotonated form of the retinal Schiff base, for H_3O^+, or for $H_5O_2^+$, so charges were determined by density functional calculations for these species. These calculations were carried out using the Amsterdam Density Functional program (ADF).[51] A triple-zeta basis set size and frozen core orbitals are used for all atoms. (In ADF this is ''basis set IV.'') For the exchange and correlation part of the density functional, the local density approximation of Vosko *et al*.[52] (VWN) with the the Becke gradient correction[53] and Perdew correlation term[54] were used. The atomic partial charges are then obtained using electrostatic potential fitting. The resulting charges for the retinal Schiff base are shown in Table 5 and Figure 6. The charges for $H_5O_2^+$ are shown in Table 6.

The electrostatic calculations are carried out using the MEAD computer program suite[18,36] which uses the finite-difference method was used to solve the Poisson equation. The finite-difference lattices were arranged according to a focusing scheme: a 65^3 grid with 1.5 Å spacing centered on the molecular center was followed by a 65^3 grid with a 0.25 Å spacing centered on each particular titrating site. The interior regions were assigned a dielectric constant of 4.0. The model-compound pK$_a$ values used were as follows: retinal Schiff base, 7; Asp, 4.0; Glu, 4.4; Arg, 12.0; Lys, 10.4. All residues of the above types were included as ionizable groups in the calculations. Tyrosine residues were considered as ionizable with a model compound pK_a of 9.6, in the first set of calculations, but since it was found here and in previous work[18]

Table 5. Lysine-retinal charges calculated from DFT and used in calculations

Atom[a]	Charge	
	Protonated	Deprotonated
CE	−0.102	−0.020
HE1(2)	0.146	0.089
NZ	−0.260	−0.474
HZ	0.346	0.000
C15	0.054	0.368
H15	0.205	0.057
C14	−0.487	−0.507
H14	0.213	0.193
C13	0.305	0.244
C20	−0.358	−0.321
H20(3)	0.136	0.108
C12	−0.413	−0.373
H12	0.227	0.224
C11	0.032	−0.053
H11	0.157	0.170
C10	−0.345	−0.349
H10	0.204	0.188
C9	0.173	0.138
C19	−0.263	−0.219
H19(3)	0.116	0.087
C8	−0.300	−0.330
H8	0.205	0.208
C7	−0.076	−0.129
H7	0.172	0.189
C6	−0.243	−0.301
C5	0.081	0.024
C18	−0.329	−0.269
H18(3)	0.128	0.095
C4	−0.108	−0.085
H4(2)	0.092	0.064
C3	−0.089	−0.081
H3(2)	0.066	0.047
C2	−0.192	−0.188
H2(2)	0.057	0.040
C1	0.604	0.690
C16	−0.454	−0.436
H16(3)	0.101	0.088
C17	−0.427	−0.436
H17(3)	0.101	0.088

[a] Notations like H19(3) indicate that there are three hydrogen atoms attached to the heavy atom, C19, and each has the indicated charge.

Table 6. $H_5O_2^+$ charges calculated from DFT and used in calculations

Atom	Charge	
	Protonated	Deprotonated[a]
O(2)	−0.654	−0.834
Bridge H	0.432	0.417
Periph. H	0.469(4)	0.417(3)

[a] Charges for the deprotonated form are taken from the TIP3P water model.

that tyrosine side-chains were always predicted to remain protonated, subsequent calculations treated them as fixed in their protonated state.

Acknowledgments

This work was supported by grants from the National Institutes of Health (GM45607 and GM59970) and the Deutsche Forschungsgemeinschaft (SFB-394 and SFB-480). We are grateful to Professor Lou Noodleman for discussion and assistance with DFT calculations.

References

1. Oesterhelt, D. (1998). The structure and mechanism of the family of retinal proteins for halophilic archaea. *Curr. Opin. Struct Biol.* **8**, 489-500.
2. Wikström, M. (1998). Proton translocation by bacteriorhodopsin and heme-copper oxidases. *Curr. Opin. Struct. Biol.* **8**, 480-488.
3. Haupts, U., Tittor, J. & Oesterhelt, D. (1999). Closing in on bacteriorhodopsin: progress in understanding the molecule. *Annu. Rev. Biophys. Biomol. Struct.* **28**, 367-399.
4. Okamura, M. Y., Paddock, M. L., Graige, M. S. & Feher, G. (2000). Proton and electron transfer in bacterial reaction centers. *Biochim. Biophys. Acta,* **1458**, 148-163.
5. Gerwert, K., Hess, B., Soppa, J. & Oesterhelt, D. (1989). Role of Asp96 in proton translocation by bacteriorhodopsin. *Proc. Natl Acad. Sci. USA,* **86**, 4943-4947.
6. Gerwert, K., Souvignier, G. & Hess, B. (1990). Simultaneous monitoring of light-induced changes in protein side-groups' protonation, chromophore isomerization and backbone motion of bacteriorhodopsin by time-resolved FTIR. *Proc. Natl Acad. Sci. USA,* **87**, 9774-9778.
7. Lanyi, J. K. & Varo, G. (1995). The photocycles of bacteriorhodopsin. *Israel J. Chem.* **35**, 365-385.
8. Heßling, B., Souvignier, G. & Gerwert, K. (1993). A model-independent approach to assigning bacteriorhodopsin's intramolecular reactions to photocycle intermediates. *Biophys. J.* **65**, 1929-1945.
9. Balashov, S. P., Imasheva, E. S., Govindjee, R. & Ebrey, T. G. (1996). Titration of aspartate-85 in bacteriorhodopsin: what it says about chromophore isomerization and protein release. *Biophys. J.* **70**, 473-481.
10. Richter, H.-T., Brown, L. S., Needleman, R. & Lanyi, J. K. (1996). A linkage of the pK$_a$s of Asp85 and Glu204 forms part of the reprotonation switch in bacteriorhodopsin. *Biochemistry,* **35**, 4054-4062.
11. Dioumaev, A. K., Richter, H. T., Brown, L. S., Tanio, M., Tuzi, S. & Saitô, H. *et al*. (1998). Existence of a proton transfer chain in bacteriorhodopsin: participation of Glu194 in the release of protons to the extracellular surface. *Biochemistry,* **37**, 2496-2506.
12. Brown, L. S., Sasaki, J., Kandori, H., Maeda, A., Needleman, R. & Lanyi, J. K. (1995). Glutamic acid 204 is the terminal proton release group at the extracellular surface of bacteriorhodopsin. *J. Biol. Chem.* **270**, 27122-27126.
13. Rammelsberg, R., Huhn, G., Lübben, M. & Gerwert, K. (1998). Bacteriorhodopsin's intramolecular proton-release pathway consists of a hydrogen-bonded network. *Biochemistry,* **37**, 5001-5009.
14. Zundel, G. (1999). Hydrogen bonds with large proton polarizability and proton transfer processes in electrochemistry and biology. *Advan. Chem. Phys.* **111**, 1-217.
15. Lobaugh, J. & Voth, G. A. (1996). The quantum dynamics of an excess proton in water. *J. Chem. Phys.* **104**, 2056-2069.
16. Wang, J. & El-Sayed, M. A. (2000). Proton polarizability of hydrogen-bonded network and its role in proton transfer in bacteriorhodopsin. *J. Phys. Chem. ser. A,* **104**, 4333-4337.
17. Luecke, H., Schobert, B., Richter, H.-T., Cartailler, J.-P. & Lanyi, J. K. (1999). Structure of bacteriorhodopsin at 1.55 Å resoultion. *J. Mol. Biol.* **291**, 899-911.
18. Bashford, D. & Gerwert, K. (1992). Electrostatic calculations of the pK$_a$ values of ionizable groups in bacteriorhodopsin. *J Mol. Biol.* **224**, 473-486.
19. Yang, A.-S., Gunner, M. R., Sampogna, R., Sharp, K. & Honig, B. (1993). On the calculations of pK$_a$s in proteins. *Proteins: Struct. Funct. Genet.* **15**, 252-265.
20. Sampogna, R. V. & Honig, B. (1994). Environmental effects on the protonation states of active site residues in bacteriorhodopsin. *Biophys. J.* **66**, 1341-1352.
21. Engels, M., Gerwert, K. & Bashford, D. (1995). Computational studies on bacteriorhodopsin: conformation and proton transfer energetics. *Biophys. Chem.* **56**, 95-104.
22. Scharnagl, C., Hettenkofer, J. & Fischer, S. F. (1995). Electrostatic and conformational effects on the proton translocation steps in bacteriorhodopsin: analysis of multiple M structures. *J. Phys. Chem.* **99**, 7787-7800.
23. Sampogna, R. & Honig, B. (1996). Electrostatic coupling between retinal isomerization and the ionization state of Glu204: a general mechanism for proton release in bacteriorhodopsin. *Biophys. J.* **71**, 1165-1171.
24. Bashford, D. & Karplus, M. (1990). pK$_a$'s of ionizable groups in proteins: atomic detail from a continuum electrostatic model. *Biochemistry,* **29**, 10219-10225.
25. Herzfeld, J., Das Gupta, S. K., Farrar, M. R., Harbison, G. S., McDermott, A. E. & Pelletier, S. L. *et al*. (1990). Solid-state ^{13}C NMR study of tyrosine protonation in dark-adapted bacteriorhodopsin. *Biochemistry,* **29**, 5567-5574.
26. Essen, L.-O., Siegert, R., Lehmann, W. D. & Oesterhelt, D. (1998). Lipid patches in membrane protein oligomers: crystal structure of the bacteriorhodopsin-lipid complex. *Proc. Natl Acad. Sci. USA,* **95**, 11673-11678.
27. Sandberg, L. & Edholm, O. (1997). pK$_a$ calculations along a bacteriorhodopsin molecular dynamics trajectory. *Biophys. Chem.* **65**, 189-204.
28. 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.
29. Luecke, H., Schobert, B., Richter, H.-T., Cartailler, J. P. & Lanyi, J. K. (1999). Structural changes in bacteriorhodopsin during ion transport at 2 Å resolution. *Science,* **286**, 255-261.
30. Luecke, H., Schobert, B., Cartailler, J.-P., Hans-Thomas, R., Rosengarth, A., Needleman, R. & Lanyi, J. K. (2000). Coupling photoisomerization of retinal

to directional transport in bacteriorhodopsin. *J. Mol. Biol.* **300**, 1237-1255.
31. Scharnagl, C. & Fischer, S. F. (1996). Conformational flexibility of arginine-82 as source for the heterogeneous and pH-dependent kinetics of the primary proton transfer step in the bacteriorhodopsin photocycle: an electrostatic model. *Chem. Phys.* **212**, 231-246.
32. Govindjee, R., Misra, S., Balashov, S. P., Ebrey, T. G., Crouch, R. K. & Menick, D. R. (1996). Arginine-82 regulates the pK_a of the group responsible for the light-driven proton release in bacteriorhodopsin. *Biophys. J.* **71**, 1011-1023.
33. Connolly, M. L. (1983). Solvent-accessible surfaces of proteins and nucleic acids. *Science,* **221**, 709-713.
34. Press, W. H., Flannery, B. P., Teukolsky, S. A. & Vetterling, W. T. (1986). *Numerical Recipes. The Art of Scientific Computing*, Cambridge University Press, Cambridge, UK.
35. Tanford, C. & Kirkwood, J. G. (1957). Theory of protein titration curves. I. General equations for impenetrable spheres. *J. Am. Chem. Soc.* **79**, 5333-5339.
36. Bashford, D. (1997). An object-oriented programming suite for electrostatic effects in biological molecules. In *Scientific Computing in Object-Oriented Parallel Environments* (Ishikawa, Y., Oldehoeft, R. R., Reynders, J. V. W. & Tholburn, M., eds), vol. 1343 of *Lecture Notes in Computer Science*, pp. 233-240, ISCOPE97, Springer, Berlin.
37. 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.
38. Eigen, M. & De Maeyer, L. (1958). Self-dissociation and protonic charge transport in water and ice. *Proc. Roy. Soc. ser. A,* **247**, 505-533.
39. Marx, D., Tuckerman, M., Hutter, J. & Parrinello, M. (1999). The nature of the hydrated excess proteon in water. *Nature,* **397**, 601-604.
40. Jackson, J. D. (1975). *Classical Electrodynamics*, Wiley and Sons, New York.
41. Antosiewicz, J., McCammon, J. A. & Gilson, M. K. (1994). Prediction of pH-dependent properties of proteins. *J. Mol. Biol.* **238**, 415-436.
42. Tanford, C. (1970). Protein denaturation. C. theoretical models for the mechanism of denaturation. *Advan. Protein Chem.* **24**, 1-95.
43. Schellman, J. A. (1975). Macromolecular binding. *Biopolymers,* **14**, 999-1018.
44. Yang, A.-S. & Honig, B. (1994). Structural origins of pH and ionic strength effects on protein stability. Acid denaturation of sperm whale apomyoglobin. *J. Mol. Biol.* **237**, 602-614.
45. Abola, E. E., Sussman, J. L., Prilusky, J. & Manning, N. O. (1997). Protein Data Bank archives of three-dimensional macromolecular structures. In *Methods in Enzymology* (Carter, C. W., Jr & Sweet, R. M., eds), vol. 277, pp. 556-571, Academic Press, San Diego.
46. Abola, E. E., Bernstein, F. C., Bryant, S. H., Koetzle, T. F. & Weng, J. (1987). Protein Data Bank. In *Crystallographic Databases-Informaton Content, Software Systems, Scientific Applications* (Allan, F. H., Bergerhoff, G. & Sievers, R., eds), pp. 107-132, Data Commission of the International Union of Crystallography, Bonn. Cambridge, Chester.
47. Brünger, A. T. & Karplus, M. (1988). Polar hydrogen positions in proteins: empirical energy placement and neutron diffraction comparison. *Proteins: Struct. Funct. Genet.* **4**, 148-156.
48. 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.
49. MacKerell, A. D., Jr, Bashford, D., Bellott, M., Dunbrack, R. L., Jr., Evanseck, J. D. & Field, M. J. *et al*. (1998). All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. ser. B,* **102**, 3586-3616.
50. Bondi, A. (1964). Van der Waals volumes and radii. *J. Chem. Phys.* **68**, 441-451.
51. te Velde, G., Bickelhaupt, F. M., Baerends, E. J., Fonesca Guerra, C., van Gisbergen, S. J. A., Snijders, J. G. & Ziegler, T. (2001). Chemistry with ADF. *J. Comput. Chem.* **22**, 931-967.
52. Vosko, S. H., Wilk, L. & Nusair, M. (1980). Accurate spin-dependent electron liquid correlation energies for local spen density calculations: a critical analysis. *Canad. J. Phys.* **58**, 1200-1211.
53. Becke, A. D. (1986). Density functional calculations of molecular bonding energies. *J. Chem. Phys.* **84**, 4524-4529.
54. Perdew, J. P. (1986). Density-functional approximation of the correlation energy of the inhomogeneous electron gas. *Phys. Rev. ser. B,* **33**, 8822-8824.
55. Balashov, S. P., Govindjee, R., Imasheva, E. S., Misra, S., Ebrey, T. G. & Feng, Y. *et al*. (1995). The two pK_a's of aspartate-85 and control of thermal isomerization and proton release in the arginine-82 to lysine mutant of bacteriorhodopsin. *Biochemistry,* **35**, 8820-8835.
56. Zscherp, C., Schlesinger, R., Tittor, J., Oesterhelt, D. & Heberle, J. (1999). *In situ* determination of transient pK_a changes of internal amino acids of bacteriorhodopsin by using time-resolved attenuated total reflection Fourier-transform infrared spectroscopy. *Proc. Natl Acad. Sci. USA,* **96**, 5498-5503.
57. Metz, G., Siebert, F. & Engelhard, M. (1992). Asp^{85} is the only internal aspartic acid that gets protonated in the M intermediate and the purple-to-blue transition of bacteriorhodopsin. *FEBS Letters,* **303**, 237-241.
58. Brown, L. S., Bonet, L., Needleman, R. & Lanyi, J. K. (1993). Estimated acid dissociation constants of the Schiff base, Asp85 and Arg82 during the bacteriorhodopsin photocycle. *Biophys. J.* **65**, 124-130.

Edited by G. von Heijne

(Received 1 March 2001; received in revised form 1 July 2001; accepted 1 July 2001)