



Published in final edited form as:

*J Am Chem Soc.* 2009 May 27; 131(20): 7064–7078. doi:10.1021/ja809767v.

## Long-distance proton transfer with a break in the bacteriorhodopsin active site

Prasad Phatak<sup>1</sup>, Jan S. Frähmcke<sup>1</sup>, Marius Wanko<sup>2</sup>, Michael Hoffmann<sup>2</sup>, Paul Strodel<sup>3</sup>, Jeremy Smith<sup>5,6,7</sup>, Sándor Suhai<sup>4</sup>, Ana-Nicoleta Bondar<sup>4,5,8,#,\*</sup>, and Marcus Elstner<sup>1,4,†</sup>

<sup>1</sup>Institute of Physical and Theoretical Chemistry, TU Braunschweig, D-38106 Braunschweig, Germany

<sup>2</sup>BCCMS, Universität Bremen, D-28334 Bremen, Germany

<sup>3</sup>Accelrys Ltd., Cambridge CB4 0WN, United Kingdom

<sup>4</sup>Molecular Biophysics Department, German Cancer Research Institute, Im Neuheimer Feld 280, D-69120, Heidelberg, Germany

<sup>5</sup>Computational Molecular Biophysics, IWR, University of Heidelberg, Im Neuenheimer Feld 368, D-69120, Heidelberg, Germany

<sup>6</sup>Center for Molecular Biophysics, Oak Ridge National Laboratory, PO BOX 2008 MS6164, Oak Ridge, Tennessee 37831, USA

<sup>7</sup>Department of Biochemistry and Molecular Biology, University of Tennessee, M407 Walters Life Sciences, 1414 Cumberland Ave, Knoxville Tennessee 37996, USA

<sup>8</sup>Department of Physiology and Biophysics and the Center for Biomembrane Systems, University of California at Irvine, Med. Sci. I, D-347, Irvine, CA 92697, USA

### Abstract

Bacteriorhodopsin is a proton-pumping membrane protein found in the plasma membrane of the archaeon *Halobacterium salinarium*. Light-induced isomerization of the retinal chromophore from all-*trans* to 13-*cis* leads to a sequence of five conformation-coupled proton transfer steps and the net transport of one proton from the cytoplasmic to the extracellular side of the membrane. The mechanism of the long-distance proton transfer from the primary acceptor Asp85 to the extracellular proton release group during the O→bR is poorly understood. Experiments suggest that this long-distance transfer could involve a transient state [O] in which the proton resides on the intermediate carrier Asp212.<sup>1</sup> To assess whether the transient protonation of Asp212 participates in the deprotonation of Asp85, we performed hybrid Quantum Mechanics/Molecular Mechanics proton transfer calculations using different protein structures, and with different retinal geometries and active site water molecules. The structural models were assessed by computing UV-Vis excitation energies and C=O vibrational frequencies. The results indicate that a transient [O] conformer with protonated Asp212 could indeed be sampled during the long-distance proton transfer to the proton release group. Our calculations suggest that in the starting proton transfer state O, the retinal is strongly twisted, and at least three water molecules are present in the active site.

\*Corresponding author. nicoleta.bondar@uci.edu. †Corresponding author. m.elstner@tu-bs.de.

#Present address: Department of Physiology and Biophysics, University of California at Irvine, Med. Sci. I, D-347, Irvine, CA 92697, USA

**A. Supplementary Information Available:** The Supplementary Information provides data of all PT pathways. This material is available free of charge via the Internet at <http://pubs.acs.org>.

## I. INTRODUCTION

The light-induced proton-pumping cycle of bacteriorhodopsin consists of five sequential proton transfer (PT) steps (Figure 1). A proton is first transferred from the chromophore's Schiff base to the nearby Asp85. The Asp85 site is electrostatically coupled to an extracellular group (denoted as the proton release group (PRG)),<sup>2,3</sup> which releases a proton to the bulk once Asp85 is protonated. The Schiff base is then reprotonated from the cytoplasmic group Asp96, which will in turn receive a proton from the cytoplasmic bulk (Figure 1). The last PT step (the O→bR transition in Figure 1a), and the focus of the present work, is from the primary acceptor Asp85 to the extracellular PRG over a distance of ~12 Å (Figure 1b). The mechanism of the Asp85 deprotonation and the structural details of the conformer ready for the last PT step are poorly understood. Key open questions are: (i) does the PT occur in a single step, from Asp85 to the PRG, or it involves intermediate(s), and (ii) what are the structural elements critical for PT from Asp85 to the PRG. To address these questions, we performed an extensive set of computations of the PT paths, excitation energies, and vibrational spectra.

A recent investigation in which vibrational spectra calculated with various structural models indicated that the PRG likely consists of Glu194 and Glu204 sharing a proton.<sup>4</sup> The combined quantum mechanics/molecular mechanics (QM/MM) simulations<sup>4</sup> demonstrated that the excess proton binds to Glu194/Glu204 instead of the water cluster. The structure with the excess proton shared by the negatively charged Glu194 and Glu204 is stable on the nanosecond timescale of the QM/MM simulations, and gives rise to a continuum band similar to that observed in previous studies.<sup>5</sup> This role of Glu194/Glu204 as the PRG is consistent with previous interpretations,<sup>6,7</sup> but contradicts a more recent proposal that the excess proton is stored on a water cluster.<sup>8,9</sup>

The path from Asp85 to Glu194/Glu204 is curved, and is interrupted by the side-chain of Arg82 (Figure 1b) in the light-adapted bacteriorhodopsin resting state (bR). The coupling between the deprotonation of Asp85 and protein conformational changes is unclear. It has been suggested that the intermediate states of the second half of the photocycle (i.e. late-M, N and O; Figure 1a) can be approximated by a single protein conformation.<sup>10</sup> Although several crystal structures exist for M, the assignment of these structures to a specific M substate is difficult. Moreover, it has been questioned whether the crystal environment allows the full extent of protein conformational changes associated with late-M *in vivo*.<sup>11,12</sup> Deprotonation of Asp85 is presumably triggered by the thermal isomerization of the retinal from 13-*cis* back to all-*trans* during the transition between the N and O intermediate states (Figure 1a). Experiments have indicated that the retinal back isomerization is coupled to protein structural rearrangements: mutations of amino-acid residues close to the retinal such as Leu93 (Figure 1b), dramatically delays retinal isomerization by 250-fold.<sup>13</sup>

Rearrangement of water molecules is important for the coupling between Asp85 and the PRG;<sup>14,15</sup> but the details of how the long-distance PT from Asp85 to the PRG is coupled to local structural rearrangements of the protein and water molecules are not clear.

Intermediate conformers distinguished by the geometry of the retinal chain and by the protonation states of key protein groups may accumulate during the O→bR step. The rate of the deprotonation of Asp85 is influenced by the local environment around Glu204.<sup>16</sup> The sensitivity of the O→bR transition to temperature and pH<sup>17,18</sup> indicates that this step of the photocycle is rather complex. It has been suggested<sup>16</sup> that the thermal back isomerization of the retinal triggers the formation of a transient O intermediate in which retinal is twisted and Asp85 is deprotonated. The transient O intermediate then converts into the bR resting state after relaxation of the retinal twist. Experiments on the Glu194Gln<sup>1,19</sup> and Glu204Gln<sup>1</sup> mutant phenotypes provided further evidence for the presence of a transient O state (denoted here as

[O]) between the O and the bR states. According to ref<sup>1</sup>, the transfer of the proton from Asp85 to the PRG may involve an intermediate step [O] in which Asp85 proton had been transferred to Asp212.<sup>1</sup> These experiments on bacteriorhodopsin mutants<sup>1,19</sup> raise the intriguing question of whether the transient [O] intermediate could also occur in the wild-type photocycle. However, the site-directed mutagenesis data must be interpreted with caution because the coupling between the retinal binding site and the PRG demonstrated by experiments<sup>20</sup> and theory<sup>3</sup> could lead to structural changes of the active site in the mutants, and a PT mechanism different from that in the wild-type.

Because it is difficult to separate O from the previous N intermediate, and the O state does not accumulate in significant amounts in the wild-type photocycle,<sup>21</sup> a crystal structure of the wild-type O state has not been solved. Two putative O-like structures have been proposed: for the chloride pumping Asp85Ser mutant (pdb code:1JV7),<sup>22</sup> and for the acid-blue bacteriorhodopsin (pdb code:1XOI).<sup>23</sup> The assignment of these structural models to an O-like state was based on the main characteristics of the O state - neutral Asp85, optical absorption maximum red-shifted relative to the wild-type bR resting state, and all-*trans* retinal. Compared to the bR resting state,<sup>24</sup> both O-like structures<sup>22,23</sup> exhibit significant conformational differences at the extracellular side of the protein (Figure 2a). However, it is unclear to what extent differences in the crystal packing (the space-group is C222<sub>1</sub> for the Asp85Ser mutant, and P622 for acid-blue bacteriorhodopsin) contribute to the structural differences among the two models.

A recent EPR study<sup>25</sup> using site-directed spin labels illustrates the specific structural changes (relative to the bR resting state) at the extracellular side of the O state of the Glu204Gln mutant - which is thought to be almost identical to the O state of the wild-type bacteriorhodopsin.<sup>25</sup> The changes include an inward movement of helix D towards the proton channel, and an increase of the distance between helices C and F (Figure 2a). The structural characteristics of the O state protein in the purple membrane environment are better reproduced<sup>25</sup> by the acid-blue bacteriorhodopsin<sup>23</sup> than by Asp85Ser<sup>22</sup>. Although the EPR study is a valuable test of the protein conformation in O,<sup>25</sup> it does not provide direct information about essential details of structural elements important for PT.

Structural details of the protein groups close to Asp85 can have a significant effect on the energetics of the Asp85 deprotonation. Previous QM/MM computations demonstrated that the proton affinity of Asp85 is greatly influenced by hydrogen bonding water molecules from the active site.<sup>26</sup> Water molecules hydrogen bonding to Asp85 and Asp212 could thus be an important determinant of the energy difference between the O and the putative [O] intermediates.

Water molecules close to the retinal Schiff base may also be important for the structure and spectroscopic properties of the active site. Of the three active site water molecules observed in the bR resting state,<sup>24,27-29</sup> only one is present in the O-like structural models of Asp85Ser mutant or ref<sup>22</sup> and acid-blue bacteriorhodopsin of ref<sup>23</sup>. It is unclear whether indeed there is only one water molecule in the O state active site or, the absence of additional water molecules is due to water molecules being disordered.

Due to the proximity between Asp85/Asp212 and the retinal Schiff base (Figure 2c), the geometry of the retinal could also influence PT. Resonance Raman studies have suggested that the O state retinal chromophore is twisted all-*trans*, with a magnitude of twisting similar to that of the K state structure.<sup>30</sup> The retinal twist relaxes as the photocycle completes with the O→bR transition. Both O-like structures<sup>22,23</sup> have all-*trans* retinal, possibly mixed with structurally similar 13-*cis*, 15-*syn* retinal.<sup>22,23</sup>

Given the electrostatic coupling between the active site and the PRG,<sup>3,20</sup> the geometry of the positively charged Arg82 side-chain may be another important determinant of the energetics of PT. In the bR resting state, the guanidinium group of Arg82 is oriented towards the retinal Schiff base region. In contrast, O-like structural models<sup>22,23</sup> indicate an orientation of the Arg82 side-chain towards the extracellular Glu194/Glu204 (Figure 2c). If the guanidinium group of Arg82 indeed adopts an extracellular orientation in O,<sup>22,23</sup> it must then return to its original cytoplasmic orientation towards the Schiff base during the O→bR transition. The orientation of Arg82 could also influence the dynamics of the water molecules visiting transiently the protein interior, and whether or not they can reach the active site. It is not known at what stage during the O→bR transition the side-chain of Arg82 changes its orientation, and whether or not the movement of Arg82 is part of the PT reaction.

In the absence of accurate structural information on the wild-type O intermediate, it is necessary to understand how key structural features affect the deprotonation of Asp85, and to use a comparison between theory and experiments to validate (or invalidate) structural characteristics of the O and the putative [O] intermediates. The structural features assessed here are: (i) the role of the retinal geometry; (ii) the number of active site water molecules, and (iii) the geometry of Arg82 side-chain. We computed reaction paths using seven different protein models. To validate specific structural characteristics of the models, we computed vertical excitation energies and vibrational spectra. The calculations indicate that an [O] intermediate with deprotonated Asp85/protonated Asp212 could indeed be sampled along the O→bR path. In O state retinal is twisted, and three water molecules are present in the Schiff base region.

## II. COMPUTATIONAL DETAILS

### A. QM/MM setup

The protein is divided in a quantum mechanics (QM) region that contains the groups involved in the chemical reaction and the remaining protein is treated with classical molecular mechanics (MM). The total energy is given by,

$$E = E_{\text{QM}} + E_{\text{MM}} + E_{\text{QM:MM}} \quad (1)$$

where  $E_{\text{QM}}$  is the interaction energy between the electrons and the nuclei of the QM region,  $E_{\text{MM}}$  describes the interactions between the MM atoms, and  $E_{\text{QM:MM}}$  represents the interaction energy between the quantum and classical subsystems. The  $E_{\text{QM:MM}}$  term can be further decomposed in,

$$E_{\text{QM:MM}} = E_{\text{QM:MM}}^{\text{b}} + E_{\text{QM:MM}}^{\text{Coul}} + E_{\text{QM:MM}}^{\text{vdw}} \quad (2)$$

where  $E_{\text{QM:MM}}^{\text{b}}$  gives the bonded interactions at the frontier bonds between the QM and MM atoms.  $E_{\text{QM:MM}}^{\text{Coul}} + E_{\text{QM:MM}}^{\text{vdw}}$  give the Coulombic and Van der Waals interactions respectively between the QM and MM particles.

The interaction energy between the classical particles are calculated using the CHARMM molecular force field.<sup>31</sup> The QM and QM:MM interactions are computed with an approximate DFT method, the self-consistent charge density functional tight binding (SCC-DFTB).<sup>32</sup> The implementation of SCC-DFTB in CHARMM has been described in ref<sup>33</sup>. SCC-DFTB is about

three orders of magnitude faster as compared to full DFT methods.<sup>34</sup> The accuracy of SCC-DFTB QM/MM approach for studies of biological systems has been documented in refs 33-41. As a prerequisite for the application to retinal proteins, we demonstrated that SCC-DFTB is in good agreement with full DFT methods for describing the ground state properties of protonated Schiff base models (torsional barriers, bond length alternation of the polyene chain etc.).<sup>42</sup> In the context of retinal proteins, SCC-DFTB has been applied to study the primary proton transport in bR,<sup>26,43-46</sup> to refine the rhodopsin crystal structure<sup>47</sup> and to generate energy minimized geometries and molecular dynamics (MD) trajectories for excited states calculations.<sup>48-51</sup> A detailed description of a QM/MM setup for bacteriorhodopsin is given in ref<sup>26</sup>.

The combined QM/MM methods<sup>52-55</sup> have greatly extended the usefulness of theoretical approaches for understanding complex biological systems. We can calculate spectroscopic properties and investigate the energetic details of chemical reactions. QM/MM methods had been applied to study the dynamics of the initial photo-isomerization event in bacteriorhodopsin,<sup>56,57</sup> the structure and spectroscopic properties of the active site in the bR resting state,<sup>58</sup> and the early intermediate K structure,<sup>45,59,60</sup> and of protonated water clusters in bacteriorhodopsin.<sup>61,62</sup> The mechanism of the first PT event has been investigated in detail,<sup>26,43-46,63</sup> Excited states properties of bacteriorhodopsin have been studied with QM/MM approaches,<sup>48,49,64-67</sup> several of them including polarization effects from the protein environment.<sup>50,51,56,68</sup>

The QM/MM studies above and of systems such as carbonic anhydrase II<sup>69</sup> or Staphylococcal nuclease<sup>70</sup> have illustrated the strengths and the limitations of the current QM/MM methods.<sup>55,71-73</sup> The QM/MM computations provided an understanding at an unprecedented level of detail of the possible scenarios for the first PT step and of the energetics associated with the PT paths in bacteriorhodopsin.<sup>26,43-45,63</sup> However, modern computational models cannot yet reproduce experimental results within the experimental error - e.g., reaction energies within 1 kcal/mol, vibrational frequencies within a few  $\text{cm}^{-1}$ , and excitation energies within few nanometers. In spite of the limitations regarding the computation of absolute values, QM/MM computations with a proper treatment of the QM region are reliable for the investigation of relative values.

## B. Proton transfer pathways

To investigate the minimum-energy pathways, we used the Conjugate Peak Refinement (CPR)<sup>74</sup> method as implemented in the TRAVEL module of CHARMM. The CPR method does not require an explicit definition of the reaction coordinate, but only the energy-minimized reactant and product structures. We investigated the PT scenario in all the structural models described in section II E below. The intermediate path points between the saddle points of the converged CPR pathway were further refined using the Synchronous Chain Minimization (SCM) algorithm<sup>75</sup> implemented in the TREK module of CHARMM.

## C. Absorption Shifts

We calculated the excitation energies using a high-level *ab initio* method, Spectroscopy Oriented Configuration Interaction (SORCI)<sup>76</sup> as implemented in the ORCA quantum chemical package.<sup>77</sup> SORCI utilizes the division of the first-order interaction space into weakly and strongly perturbing configurations, thereby combining the concepts of the classical multireference configuration interaction and the multireference perturbation theory. While the strongly perturbing configurations are treated variationally, the weakly perturbing configurations are treated with second-order Møller-Plesset perturbation theory. The use of approximate natural orbitals eliminates the problem of choosing a suitable single-particle basis whose quality would affect the final CI result. SORCI gains computational efficiency by use

of several thresholds which have been carefully adjusted for the system under study. The thresholds used are:  $T_{\text{Pre}} = 10^{-3}$ ,  $T_{\text{Nat}} = 10^{-6}$  and  $T_{\text{Sel}} = 10^{-6}E_h$  (see supporting information of ref<sup>48</sup>). Only core orbitals are frozen. We used Ahlrich's SV(P) basis set<sup>78</sup> which is appropriate for the calculations of the complete chromophore. SORCI has been successfully used to calculate vertical excitation energies for retinal proteins previously, based on optimized geometries calculated with SCC-DFTB/CHARMM.<sup>48</sup> The spectral shifts between proteins are described very well e.g. the shift between bR and SRII could be explained on the basis of structural determinants,<sup>49</sup> but the absolute excitation energies are overestimated by approximately 0.15 eV.<sup>48</sup> The overestimation of the excitation energies is partly due to the effects of polarization and dispersion,<sup>50,51</sup> which are not covered in the widely applied standard QM/MM methods. Since calculations that account for these effects are rather costly, we perform standard QM/MM excited states calculations for all the models and recalculate the excitation energies with polarization effects for several selected models. The effect of protein polarization is described using an interactive polarizable model in which we assign atomic polarizabilities to MM atoms of the protein in addition to the fixed atomic charges.<sup>50,51</sup>

#### D. Vibrational Frequencies

The standard normal mode analysis requires diagonalization of the calculated mass-weighted second derivative matrix (Hessian). Since for a system size of several thousand atoms like bacteriorhodopsin, diagonalization of the full Hessian matrix is computationally demanding, we perform the diagonalization in a reduced basis. To calculate the vibrational frequencies we used the VI-BRAN module<sup>31</sup> of the CHARMM in the QM/MM setup.<sup>79</sup> The advantage of such an approach consists in accurate description of the polarization effect by the surrounding MM environment on the QM region.<sup>79</sup> We computed the vibrational frequencies using SCC-DFTB for the QM region. Recent studies have shown that SCC-DFTB is accurate in reproducing DFT results for vibrational frequencies of various systems.<sup>80,81</sup> We carried out additional benchmark calculations for the  $\nu_{\text{COOH}}$  mode of isolated acetic acid molecule, and for strongly hydrogen bonded cyclic acetic acid dimer (see Supplementary Information). The calculations demonstrate that SCC-DFTB can accurately reproduce the  $\nu_{\text{COOH}}$  band for the isolated acetic acid molecule ( $1778 \text{ cm}^{-1}$ ) compared with the experimental value of  $1780 \text{ cm}^{-1}$ . For the strongly hydrogen bonded cyclic dimer, the  $\nu_{\text{COOH}}$  band computed with SCC-DFTB is red-shifted by approximately  $17 \text{ cm}^{-1}$  relative to the experimental value of  $1725 \text{ cm}^{-1}$  (see Table S-1 of Supplementary Information).

#### E. Structural models

The results reported here are based on three structural models (model-A, model-B and model-C; see below) distinguished by the protein structure, the geometry of retinal and the Arg82 side-chain, and by the number and location of water molecules close to Asp85/Asp212 (Table 1). To investigate in more detail the role of water molecules, we performed four additional sets of computations in which one or two water molecules not present in the crystal structure<sup>22,23</sup> were added in the model-A and model-B structures (Table 1). The computations using model-A and model-B with additional water molecules were pursuant to our preliminary observation that the presence of three water molecules in the active site of model-A and model-B leads to an active site geometry similar to that of the bR resting state. For all models used, we denote an O state as neutral Asp85 and negatively charged Asp212 and a [O] state as negatively charged Asp85 and neutral Asp212.

**Models with Acid-blue form of bR (model-A)**—To prepare the model-A we used the crystal structure of acid-blue bacteriorhodopsin.<sup>23</sup> This structure indicates coordinates for one active site water molecule (w603) that bridges Asp85 and Asp212 (Figure 2c). We prepared the following models: (i) model-A<sup>1w</sup> contains the single active site water molecule w603 from the crystal structure (Figure 3a-b); (ii) model-A<sup>2w</sup> contains w603 from the crystal structure and

w401\* (Figure 4a-b); (iii) model-A<sup>3w</sup> contains water molecules w603, w401\* and w406\* (Figure 7a-b).

**Models with Asp85Ser mutant (model-B)**—The model-B structural conformers were prepared starting from the Asp85Ser mutant crystal structure.<sup>22</sup> The crystal structure has a single water molecule w402 in the active site, and w402 is hydrogen bonded to the Schiff base (N-O distance: 2.7 Å). We mutated Ser85 back into Asp85 and prepared end-conformers for: model-B<sup>1w</sup> (one active site water molecule w402; Figure 3c-d), model-B<sup>2w</sup> (two active site water molecules w402 and w401\*; Figure 4c-d), model-B<sup>3w</sup> (three active site waters molecules w402, w401\* and w406\*; Figure 7c-d).

**Model with resting state of bR (model-C<sup>3w</sup>)**—To assess how the PT energetics is influenced by the protein structure, the orientation of the Arg82 side-chain and by the active site water molecules, we performed an additional set of computations starting from the coordinates of the bR resting state (pdb code: 1C3W).<sup>24</sup> The active site of model-C<sup>3w</sup> contains three water molecules: w401, w402 and w406 (Figure 8a-b).

## F. System setup

Asp96 and Asp115 are modeled neutral as indicated by experiments.<sup>1,82</sup> Standard protonation states are used for all remaining titratable protein amino-acids. All buried water molecules present in the crystal structures are included in the calculations. The hydrogen atoms are built using the HBUILD utility of CHARMM package.<sup>31</sup>

The QM region consisted of 78 atoms (for model-A<sup>1w</sup> and model-B<sup>1w</sup> with one active site water molecule), 81 atoms (for the model-A<sup>2w</sup> and model-B<sup>2w</sup> models with 2 active site water molecules), or 84 atoms (for the models with 3 active site water molecules). The QM region comprised Lys216, the complete retinal molecule, the side-chains of Asp85, Asp212, and the active site water molecule(s). The boundary between the quantum and classical subsystems was chosen by adding the link atom at the C<sub>β</sub>-C<sub>γ</sub> covalent bond for Lys216, and at the C<sub>α</sub>-C<sub>β</sub> covalent bond of Asp85 and Asp212. The QM/MM frontier was treated using DIV linking scheme.<sup>83</sup> In the DIV scheme, the partial charge of the MM host atom is removed from the host and distributed evenly over the remaining MM host atoms.<sup>83</sup>

To maintain the shape of the protein close to the crystal structure we fixed a part of the protein backbone to the crystal structure coordinates. The remaining part was mobile and consisted of 1578 atoms including the quantum mechanical region and a layer of protein groups and surrounding water molecules that can be related to the PT path. The SCC-DFTB/CHARMM energy minimizations were performed to a RMS energy gradient of 10<sup>-3</sup> kcal/mol.Å.

## III. RESULTS

In the following, we discuss QM/MM-optimized models and spectral fingerprint calculations of the O (protonated Asp85 and anionic Asp212) and [O] states (anionic Asp85 and protonated Asp212) and reaction pathways that connect the O and [O] structural models. For each structural model described in Table 1, we computed the vertical excitation energies of O and [O] (Tables 3 and 4) and the vibrational frequencies of the  $\nu_{COOH}$  band (Table 5).

### A. Structural characteristics of end-states

**Models with one active site water molecule**—In the crystal structure of the acid-blue bacteriorhodopsin,<sup>23</sup> the distance between Asp85-O<sup>δ1</sup> and Asp212-O<sup>δ2</sup> is shorter (4.2 Å) than in the bR resting state (5.1 Å) of ref<sup>24</sup> The shorter distance between Asp85 and Asp212 in the O state was explained by either Asp85 or Asp212 being protonated. In the QM/MM-optimized

end-conformers of model-A<sup>1w</sup>, retinal is twisted. The Schiff base salt bridges to Asp212 in O (model-A<sup>1w</sup>-O: Figure 3a), and to Asp85 in [O] (model-A<sup>1w</sup>-[O]: Figure 3b).

In the crystal structure of the Asp85Ser mutant<sup>22</sup> water molecule w402 is hydrogen bonded to the Schiff base (O-N distance: 2.7 Å). In the QM/MM-optimized O conformer of model-B<sup>1w</sup> the Schiff base:w402 hydrogen bond is broken and w402 is displaced towards a more extracellular location, where it forms hydrogen bonds with Asp85 and Asp212 (Figure 3c). Retinal is twisted and the Schiff base salt bridges to Asp212 in O (model-B<sup>1w</sup>-O: Figure 3c), and to Asp85 in [O] (model-B<sup>1w</sup>-[O]: Figure 3d). That is, the structures of model-A<sup>1w</sup> and model-B<sup>1w</sup> are similar.

The salt bridge between the retinal Schiff base and Asp212 (in O conformer) and between the retinal Schiff base and Asp85 (in [O] conformer) leads to significant blue-shift of the [O]-O excitation energies relative to experimental [O]-O shift 0.04 eV.<sup>1</sup> The absolute excitation energies in model-A<sup>1w</sup> are 2.09 eV for O, and 2.16 eV for [O]. In model-B<sup>1w</sup>, the excitation energies are 2.20 eV for O, and 2.44 eV for [O] (Table 3). The differences between the excitation energies of model-A<sup>1w</sup> and model-B<sup>1w</sup> models can be explained by differences in the hydrogen bonding pattern of the active site. In model-A<sup>1w</sup> conformers there is a hydrogen bond between Thr89:Asp85, whereas the Tyr57:Asp212 hydrogen bond is absent (Figure 3a-b). In contrast, in model-B<sup>1w</sup> conformers, the Thr89:Asp85 hydrogen bond is not present, but the Tyr57:Asp212 hydrogen bond is present (Figure 3c-d). There are also differences in the retinal geometry (see below).

The pattern of hydrogen bonding in the active site may be particularly important for the vibrational frequencies of the  $\nu_{COOH}$  band. The  $\nu_{COOH}$  vibrational frequencies of the both model-A<sup>1w</sup>-O (1690 cm<sup>-1</sup> and model-A<sup>1w</sup>-[O] (1652 cm<sup>-1</sup>) conformers are largely red-shifted by ~60 cm<sup>-1</sup> relative to the experimental values (Table 5), indicating a more distorted hydrogen bonded network than expected from experiments. The red-shift in the vibrational frequencies computed with model-A<sup>1w</sup> can be attributed (at least in part) to the presence of Thr89:Asp85 hydrogen bond. The  $\nu_{COOH}$  vibrational frequencies computed with model-B<sup>1w</sup>-O and model-B<sup>1w</sup>-[O] conformers in which the Thr89:Asp85 hydrogen bond is absent, are within ~10 cm<sup>-1</sup> from the experimental values (Table 5). Moreover, the O-[O] shifts computed with model-A<sup>1w</sup> (38 cm<sup>-1</sup>) and model-B<sup>1w</sup> (39 cm<sup>-1</sup>) are in good agreement with the experimental value of 42 cm<sup>-1</sup>.

**Models with two active site water molecules**—In model-A<sup>2w</sup> prepared from the acid-blue bacteriorhodopsin structure,<sup>23</sup> we inserted w401\* within hydrogen bonding distance from the retinal Schiff base (Figure 4a-b). The distance between Asp85-O<sup>δ1</sup> and Asp212-O<sup>δ2</sup> is shorter (4.0 Å) than in the bR resting state (5.1 Å of ref<sup>24</sup>). In O conformer of model-A<sup>2w</sup> (model-A<sup>2w</sup>-O), w401\* hydrogen bonds with the Schiff base (2.7 Å), with the negatively charged Asp212 (2.8 Å), and with Asp85 (2.7 Å: Figure 4a). As a result, the presence of the w401\* and w402 in the active site reduces the retinal twist relative to models (model-A<sup>1w</sup> and model-B<sup>1w</sup>) in which only one water molecule was present (Table 6). In [O] conformer of model-A<sup>2w</sup> (model-A<sup>2w</sup>-[O]), the hydrogen bond between the Schiff base and w401\* is broken, the retinal chain is twisted, and the Schiff base salt bridges to Asp85 (2.8 Å: Figure 4b).

The excitation energy of model-A<sup>2w</sup>-O (1.92 eV) is close to the experimental value of 1.94 eV, indicating that binding of the retinal Schiff base to water is an important structural determinant of the electronic structure of the chromophore. Due to the salt bridge between the retinal Schiff base and Asp85, the absolute excitation energy of model-A<sup>2w</sup>-[O] state (2.06 eV) is somewhat blue-shifted relative to the experimental value of 1.98 eV. The O-[O] shift (0.14 eV) of model-A<sup>2w</sup> is much larger than the experimental value of 0.04 eV. Therefore, accurate models of the



O and [O] states may be characterized by the retinal Schiff base being hydrogen bonded to water.

Model-B<sup>2w</sup> was prepared from the Asp85Ser mutant<sup>22</sup> by adding water molecule w401\* (Figure 4c-d). The O state of model-B<sup>2w</sup> (model-B<sup>2w</sup>-O: Figure 4c) has similar structural features of the active site as in model-A<sup>2w</sup>-O: there is a hydrogen bond between the Schiff base and a water molecule (w402 in model-B<sup>2w</sup>-O and w401\* in model-A<sup>2w</sup>-O) and, a hydrogen bonded chain between Asp85 and Asp212 with the participation of the second active site water molecule (w603 in model-A<sup>2w</sup>-O and w401\* in model-B<sup>2w</sup>-O).

The higher excitation energy of model-B<sup>2w</sup>-O (2.17 eV) compared to the excitation energy of model-A<sup>2w</sup>-O (1.92 eV: Table 3) may result from the shorter distance between the Schiff base and the counterion Asp212 (3.4 Å in model-B<sup>2w</sup>-O, compared to 4.1 Å in model-A<sup>2w</sup>-O). The shorter distance in model-B<sup>2w</sup>-O leads to an increased stabilization of the positive charge at the Schiff base in the ground state and consequently to a larger excitation energy.

In contrast to the models containing a single water molecule in the active site, the presence of two water molecules in model-A<sup>2w</sup> and model-B<sup>2w</sup> is associated with the presence of the Tyr57:Asp212 hydrogen bond. The Thr89:Asp85 hydrogen bond is, however absent in the model-B<sup>2w</sup> conformers (Figure 4c-d). The absence of the Thr89:Asp85 hydrogen bond in model-B<sup>2w</sup>-O leads to a higher  $\nu_{COOH}$  stretching frequency (1758 cm<sup>-1</sup>) compared to that computed in model-A<sup>2w</sup>-O (1741 cm<sup>-1</sup>), which contains the Thr89:Asp85 hydrogen bond (Figure 4a). These observations support the conclusion from the calculations based on models with a single active site water molecule that the Thr89:Asp85 hydrogen bond plays an important role in tuning of the C=O stretch frequency.

**Models with three active site water molecules**—The discussion above indicates that the structural models with one or two water molecules in the active site are not consistent with the experimental spectral fingerprints. The formation of a salt bridge between the retinal Schiff base and either Asp85 or Asp212 leads to a very distorted retinal geometry, and to optical excitation energies that are blue-shifted relative to experimental excitation energies of O and [O] states (Table 3).

To further investigate the configuration of the water molecules in the active site, and the relationship between the location of water molecules and the optical excitation spectra, we performed the following two sets of MD simulations (for details of the simulation setup see Supplementary Information). To investigate the dynamics of the active site water molecules on a timescale longer than accessible with QM/MM MD simulations, in the first set of simulations we performed three MM MD simulations (5ns each) for model-A<sup>1w</sup>-O. During these MD simulations w603 moved away from Asp85/Asp212, and the Asp85-O<sup>δ1</sup>:Asp212-O<sup>δ2</sup> distance decreased to ~2.7 Å (Figure 5c), i.e., much shorter than the 4.2 Å distance indicated by the crystal structure.<sup>23</sup> We also observed that water molecules moved from the PRG region closer to the active site (Figure 5b). Therefore one cannot exclude that in O there may be more water molecules in the active site than indicated by the crystal structure.

In the second set of simulations, we performed QM/MM MD for model-A<sup>3w</sup>-O (Figure 6). We observed that the water molecule w401\* that connects the Schiff base to Asp85/Asp212 remained in the active site during the simulations (Figure 6a-b). The Asp85-O<sup>δ1</sup>:Asp212-O<sup>δ2</sup> distance (~4.2 Å; Figure 6c) agrees well with the crystal structure<sup>23</sup> value of 4.2 Å. The good agreement between the location of water molecule w603 in the crystal structure<sup>23</sup> (Figure 6a) and in the QM/MM MD the simulations of model-A<sup>3w</sup>-O suggests that three water molecules could be accommodated in the crystal structure. Pursuant to these considerations, we performed computations on models with three water molecules as described below.

Model-A<sup>3w</sup>-O (Figure 7a) has a hydrogen bonded network in which Asp85 and Asp212 are hydrogen bonded via w603 and w406\*; and w401\* hydrogen bonds with the Schiff base, Asp85, and Asp212. A similar hydrogen bonded network is observed in model-A<sup>3w</sup>-[O] (Figure 7b). The hydrogen bonding pattern of the active site of model-A<sup>3w</sup> probably explains the absolute excitation energies of 1.96 eV and 2.03 eV for O and [O], respectively. The calculated excitation energies are close to the experimental values (1.94 for O and 1.98 for [O]). The calculated shift between the [O] and O models (0.07 eV) is in excellent agreement with the experimental [O]-O shift (Table 3).

The hydrogen bonds between Thr89 and Asp85, and between Asp85 and water molecules (Figure 7a-b), downshift the  $\nu_{COOH}$  band in model-A<sup>3w</sup>-O to 1737 cm<sup>-1</sup>. In model-A<sup>3w</sup>-[O] in which Asp212 shares hydrogen bonds with Trp86, Tyr57 and Tyr185, the  $\nu_{COOH}$  band is at 1700 cm<sup>-1</sup>. Since SCC-DFTB slightly underestimates these carboxyl stretching frequencies when hydrogen bonds are present (see Table S-1 of Supplementary Information), the frequencies computed with model-A<sup>3w</sup> are in excellent agreement with the experimental values. The O-[O] shift of the  $\nu_{COOH}$  band (37 cm<sup>-1</sup>) is also in very good agreement with the experimental O-[O] shift of 42 cm<sup>-1</sup> (see Table 5).

To estimate the influence of the Thr89:Asp85 hydrogen bond, we constructed a model where this hydrogen bond is absent, and Thr89 hydrogen bonds to the backbone carbonyl oxygen of Asp85. The structural model is 1.1 kcal/mol higher in energy than the model in which the Thr89:Asp85 hydrogen bond is present, and the  $\nu_{COOH}$  frequency (1758 cm<sup>-1</sup>) is higher than the experimental value of O state (1752 cm<sup>-1</sup>).

The active site geometry of the mutant O structures (O and [O] conformers of model-B<sup>3w</sup>; Figure 7c-d) is similar to that of the acid-blue model (Figure 7a-b). The main difference between model-A<sup>3w</sup>-O and model-B<sup>3w</sup>-O is in the absence of the Thr89:Asp85 hydrogen bond in model-B<sup>3w</sup>-O (Figure 7c). As discussed above for model-A<sup>3w</sup> conformers without a Thr89:Asp85 hydrogen bond, the absence of this hydrogen bond leads to a blue shifted  $\nu_{COOH}$  frequency in model-B<sup>3w</sup>-O. The  $\nu_{COOH}$  frequency is at 1756 cm<sup>-1</sup> in model-B<sup>3w</sup>-O, and at 1719 cm<sup>-1</sup> in model-B<sup>3w</sup>-[O]. Given that SCC-DFTB slightly underestimates the C=O frequency in strongly hydrogen bonded structures, the  $\nu_{COOH}$  frequencies computed for model-B<sup>3w</sup> conformers are too high relative to the experimental values. Moreover, the excitation energies of 2.15 eV for model-B<sup>3w</sup>-O and 2.29 for model-B<sup>3w</sup>-[O] (Table 3) are somewhat blue-shifted relative to the experimental values. The blue-shift may be due to the missing Thr89:Asp85 hydrogen bond, and to the short distance between the retinal C<sub>e</sub> atom and Asp85-O<sup>δ2</sup> (3.3 Å in model-B<sup>3w</sup>-O in contrast to 3.8 Å in model-A<sup>3w</sup>-O). This supports the proposal that the acid-blue model resembles the conformation of the O state in the purple membrane,<sup>25</sup> and demonstrates that the small spectral shifts calculated between O and [O] for model-A<sup>3w</sup> are reproducible.

The bR resting state<sup>24</sup> has a strong hydrogen bonded network in the active site comprising the Schiff base, the three water molecules w401, w402 and w406, and the two anionic Asp85 and Asp212 groups. The crystal structure of the bR resting state indicates a cytoplasmic orientation of the Arg82 side-chain<sup>24</sup>, which is preserved in the end states of model-C<sup>3w</sup> prepared from the bR resting state<sup>24</sup>. Both model-C<sup>3w</sup>-O (Figure 7a) and model-C<sup>3w</sup>-[O] (Figure 7b) have a twisted all-*trans* retinal. The Thr89:Asp85 hydrogen bond present in the bR resting state crystal structure (2.7 Å) is present in both model-C<sup>3w</sup>-O and model-C<sup>3w</sup>-[O] conformers. In model-C<sup>3w</sup>-O, all water molecules from the active site are hydrogen bonded to Asp212; while in model-C<sup>3w</sup>-[O], the water molecules are oriented towards the donor Asp85, and not hydrogen bonded to Asp212.

In contrast to model-A<sup>3w</sup> and model-B<sup>3w</sup>, the absolute excitation energies of model-C<sup>3w</sup>-O (1.85 eV) and model-C<sup>3w</sup>-[O] (1.91 eV) are red-shifted compared to the experimental values: 1.94 eV for O (by 0.09 eV) and 1.98 eV for [O] (by 0.07 eV) respectively. Qualitatively, this may be expected due to the cytoplasmic orientation of the guanidinium group of Arg82 in model-C<sup>3w</sup>-O. In model-C<sup>3w</sup>-O the positive charge of Arg82 is close to the Schiff base, and thus destabilizes the ground state and decreases the excitation energy relative to ground state. Although the absolute energies of O and [O] states are red-shifted relative to the experimental values (see Table 3), the [O]-O shift (0.06 eV) is very close to the experimental value of 0.04 eV.

## B. Proton Transfer pathways

**PT in models with one active site water molecule**—In model-A<sup>1w</sup>, PT begins with the reorientation of the Schiff base away from Asp212. As a result, the salt bridge between the Schiff base and Asp212 is weakened. Simultaneously, Asp85 moves closer to w603, such that the Asp85:w603 hydrogen bond is strengthened. At the transition state ( $\lambda=0.44$  in Figure 9a), the hydrogen bonding distances between Asp85:w603 and between w603:Asp212 are equal to 2.4 Å. After the transition state, the retinal chain becomes more planar, and the reaction continues with the transfer of the proton from Asp85 to w603, and from w603 to Asp212, followed by the formation of the salt bridge between the Schiff base and Asp85.

The rate-limiting barrier of 5.3 kcal/mol ( $\lambda=0.44$ ; Figure 9a) is dominated by  $E_{QM}$  (4.6 kcal/mol), and corresponds to the twisted retinal geometry (Figure 9a). The favorable contribution of  $E_{QM:MM}^{nb}$  (-1.79 kcal/mol) to the rate-limiting barrier corresponds to w603 being strongly hydrogen bonded to Asp85 and Asp212. Following the transition state, decrease in the retinal twist is associated with a decrease  $E_{QM}$  energy contribution.  $E_{QM:MM}^{nb}$  contribution starts rising from -1.79 kcal/mol at the transition state to 5.3 kcal/mol at  $\lambda = 1$ , and partially compensates for the negative  $E_{QM}$  contribution and leads to an overall exothermic energy profile.

We observed a similar PT mechanism in model-B<sup>1w</sup> (see Supplementary Information). Similar to the results for model-A<sup>1w</sup>, in model-B<sup>1w</sup> the PT energy barrier is small (~5 kcal/mol; Table 2). Within the accuracy of the methods used, the reaction energy is, however, different in model-A<sup>1w</sup> and model-B<sup>1w</sup>, being negative for model-A<sup>1w</sup>, and positive for model-B<sup>1w</sup>. That is, model-A<sup>1w</sup> favors the proton on Asp212, whereas model-B<sup>1w</sup> favors the proton on Asp85. The different reaction energies computed for model-A<sup>1w</sup> and model-B<sup>1w</sup> may be related to the Tyr57:Asp212 hydrogen bond, which is present in model-B<sup>1w</sup>, but absent in model-A<sup>1w</sup>. Test computations with a model-A<sup>1w</sup> conformer in which the Tyr57:Asp212 hydrogen bond is present indicate a reaction energy of 4.5 kcal/mol (Supplementary Information) similar to the value of 3.1 kcal/mol computed for model-B<sup>1w</sup>. It is, however, unlikely that either of the model-A<sup>1w</sup> and model-B<sup>1w</sup> geometries are a good representation of the active site geometry in O.

**PT in models with two active site water molecules**—The mechanism of PT in models with two active site water molecules is similar to that discussed above for models with one active site water molecule (Supplementary Information). In the case where two active site water molecules are present, the proton can be transferred either via the water molecule hydrogen bonded to the retinal Schiff base, or via the water molecule with the more extracellular location in the active site. When only one water molecule is present in the active site, it acts as an intermediate carrier for the proton.

**PT in models with three active site water molecules**—The PT pathway from Asp85 to Asp212 via water molecules w603 and w406\* occurs via a concerted mechanism in model-A<sup>3w</sup> (Figure 9b). The pathway starts with the decrease of the distance between Asp212 and w406\* from 2.7 Å to 2.6 Å, such that the distances between Asp85 and w603, between w603

and w406\*, and between w406\* and Asp212 become equal to 2.6 Å. At the transition state ( $\lambda=0.47$  in Figure 9b), protons are shared between Asp85-w603 (2.5 Å), between w603-w406\* (2.5 Å), and between w406\*-Asp212 (2.4 Å). This gives the rate-limiting energy barrier of 10.9 kcal/mol. The reaction then proceeds with PT from Asp85 to w603, from w603 to w406\*, and from w406\* to Asp212. As a consequence of the transfer, the hydrogen bond between w401\* and Asp212 is broken, and a new hydrogen bond is formed between the w401\* and w406\*. The  $E_{QM}$  and the  $E_{QM:MM}^{nb}$  terms contribute 4.9 kcal/mol and, 5.0 kcal/mol respectively to the rate-limiting barrier (Figure 9b). The increase in  $E_{QM:MM}^{nb}$  from 5.0 kcal/mol at the transition state to 13.5 kcal/mol at the end of PT (Figure 9b) is due to the increase of the Tyr185:Asp212 and Tyr57:Asp212 hydrogen bonding distances from 2.6 Å in the O, to 2.7 Å in [O] state.

Due to the close structural similarity between the active sites in model-B<sup>3w</sup> (Figure 7c) and model-A<sup>3w</sup> (Figure 7a), the pathway and the energetics for PT in model-B<sup>3w</sup> are similar to those obtained for model-A<sup>3w</sup> (see Supplementary Information).

PT in model-C<sup>3w</sup> occurs via water 402 with a large activation barrier of 13.8 kcal/mol and a reaction energy of 3.6 kcal/mol. PT could also occur via water molecules 401 and 406, with a reaction mechanism and associated energetics that are similar to those obtained for model-A<sup>3w</sup>. To further investigate the influence of the Arg82 orientation on the PT energetics and spectral fingerprints, we used model-C<sup>3w</sup> structure to construct a model in which the Arg82 side-chain is oriented towards the extracellular side. This conformer is denoted here as model-D<sup>3w</sup>. The extracellular-oriented Arg82 side-chain was taken from the late M-state structure (pdb-code: 1CWQ)<sup>84</sup>, superimposed onto model-C<sup>3w</sup>. The conformer with an extracellular orientation of the Arg82 side-chain is ~4 kcal/mol lower in energy than model-C<sup>3w</sup>. Due to the larger distance between the Schiff base and the positive charge of Arg82 in model-D<sup>3w</sup>, the absolute excitation energies of O (2.04 eV) and [O] (2.00 eV) states of model-D<sup>3w</sup> model are blue-shifted relative to model-C<sup>3w</sup> (Table 3). The rate-limiting PT barrier computed with model-D<sup>3w</sup> is 8.0 kcal/mol (see Supplementary Information).

### C. Excited states calculated by including protein polarization effects

In a recent study on the influence of the polarizability of the protein/solvent/membrane environment on the calculation of vertical excitation energies, a significant bathochromic shift was found when substituting the fixed CHARMM point-charge model by the polarization model polar.h.<sup>50,51</sup> The latter combines a polarization-free point charge model with an interactive induced-atomic dipole model to treat polarization within the protein environment explicitly. Furthermore, the screening of solvent-exposed charged groups by the solvent/membrane environment may affect the excitation energy. The Poisson - Boltzmann based charge-scaling scheme of Dinner *et al.*<sup>85</sup> represents a simple way to include this effect and is appropriate also for the calculation of excitation energies of retinal proteins.<sup>51</sup>

Table 4 shows the effects of the charge scaling and the self-consistent polarization of the chromophore and its environment employing the polar.h model. In the three O-state models, the net effects are both very small, in contrast to the situation in the bR resting state, where the charge scaling and the polar.h protein electrostatics give rise to bathochromic shifts of 0.09 and 0.17 eV, respectively. The latter is primarily caused by the refined, self-consistent charge distribution in the protein, which includes an electrostatic screening of the doubly charged complex counter ion. This screening reduces both the electrostatic interaction with the chromophore and the charge transfer between the protonated Schiff base and the complex counter ion. In the O state, only a single counter ion is present (Asp212), and these effects are less pronounced.

A minor contribution to the larger bathochromic shift in bR originates from the response of the induced dipoles in the protein to the charge transfer associated with the optical excitation of the chromophore: When neglecting this response, i.e., keeping the ground-state adapted atomic dipoles fixed, the resulting vertical excitation energies are increased by 0.09 and 0.05 eV for bR resting state and the O state models, respectively. As the polarizability of the protein environment is essentially the same in both cases, the difference correlates with a stronger charge transfer in bR resting state: The difference-dipole moment  $|\Delta_{S_1-S_0}\mu|$  is 10.8 debye in bR resting state and 8.4 debye in model-A<sup>3w</sup>-O conformer.

#### D. Retinal chromophore configuration

The crystallographic data has shown that the retinal adopts an almost planar geometry in its isolated form;<sup>86-88</sup> the protein environment induces non-planarity due to steric and electrostatic interactions.<sup>24,89-93</sup> solid-state NMR measurements of the H-C<sub>14</sub>-C<sub>15</sub>-H bond along the photocycle indicated that the retinal twist increases by  $\sim 17^\circ$  from the bR resting state to early-M.<sup>93</sup> FTIR studies on various mutant bR phenotypes suggested that in the O intermediate the retinal is more twisted than in bR, and that the twist decreases during the O $\rightarrow$ bR transition.<sup>16</sup>

Resonance Raman studies indicated that in O, the all-*trans* retinal chromophore is twisted;<sup>30</sup> the twist is delocalized around the C<sub>7</sub>=C<sub>8</sub>, C<sub>11</sub>=C<sub>12</sub>, and C<sub>15</sub>=N<sub>16</sub> bonds.<sup>30,92</sup> The dihedral angles of these retinal double bonds for all O-like models investigated here are given in Table 6. The retinal twist is stronger in O than in bR. For example, the dihedral angle C<sub>12</sub>-C<sub>13</sub>=C<sub>14</sub>-C<sub>15</sub> is  $-151.3^\circ$  in model-A<sup>3w</sup>-O, and  $-156.9^\circ$  in the bR resting state (see Table 6). The stronger twist of the retinal chain in the O models than in the bR resting state is consistent with the suggestions from experiments.<sup>16,30</sup> The maximum twists measured for C<sub>6</sub>-C<sub>7</sub>=C<sub>8</sub>-C<sub>9</sub> and C<sub>10</sub>-C<sub>11</sub>=C<sub>12</sub>-C<sub>13</sub> dihedral angles are  $\pm 10^\circ$ . Larger twists of  $\pm 30^\circ$  are observed for the C<sub>10</sub>-C<sub>11</sub>=C<sub>12</sub>-C<sub>13</sub> and C<sub>14</sub>-C<sub>15</sub>=N<sub>16</sub>-C <sub>$\epsilon$</sub>  dihedral angles.

Model-A<sup>1w</sup>-O has large deviations of up to  $\sim 20^\circ$  of the twists for C<sub>12</sub>-C<sub>13</sub>=C<sub>14</sub>-C<sub>15</sub> and C<sub>14</sub>-C<sub>15</sub>=N<sub>16</sub>-C <sub>$\epsilon$</sub>  dihedral angles relative to the planar all-*trans* value of  $180^\circ$  (see Table 6). Together with the large blue-shift of the excitation energy and the large red-shift of C=O stretch frequency computed with model-A<sup>1w</sup>, this distorted retinal geometry indicates that more water molecules may be present in the active site.

Relative to the bR resting state, large change was observed for C<sub>12</sub>-C<sub>13</sub>=C<sub>14</sub>-C<sub>15</sub> and C<sub>14</sub>-C<sub>15</sub>=N<sub>16</sub>-C <sub>$\epsilon$</sub>  dihedral angles for model-A<sup>3w</sup>-O, model-B<sup>3w</sup>-O, and model-C<sup>3w</sup> (Table 6). The significant deviations from planarity of the retinal chain in these structural models is consistent with the experimental findings.<sup>16,30</sup>

## IV. DISCUSSION

The aim of this work was to investigate structural models of the O state, and the likelihood that an [O] intermediate characterized by deprotonated Asp85 and protonated Asp212 is sampled during the O $\rightarrow$ bR transition. To address these questions we computed PT, excitation energies, and vibrational frequencies using protein models constructed from the crystal structures of acid-blue bacteriorhodopsin,<sup>23</sup> Asp85Ser mutant,<sup>22</sup> and bR resting state<sup>24</sup>. The models used in our computations are distinguished not only by the protein structure, but also by the number of active site water molecules, orientation of the Arg82 side-chain, and by the geometry of the retinal polyene chain.

#### A. Role of active site water molecules

The crystal structures of acid-blue bacteriorhodopsin<sup>23</sup> and Asp85Ser mutant<sup>22</sup> indicate only one water molecule close to the retinal Schiff base. The QM/MM-geometry optimizations

reported here indicate that when only one water molecule is present in the active site, local minima in which the Schiff base salt bridges to either Asp85 or Asp212 are energetically preferred. The strong interaction between the Schiff base and Asp85 (or Asp212) leads to optical excitation energies that are blue-shifted relative to the experimental values. That is, the geometries of the active site obtained in the presence of a single water molecule in the active site give spectral fingerprints inconsistent with experiments. Consequently, we constructed additional structural models with two and three water molecules in the active site. The short Asp85:Asp212 distance of 4.2 Å in the acid-blue bacteriorhodopsin crystal structure<sup>23</sup> relative to 5.1 Å in the bR resting state<sup>24</sup> is compatible with three water molecules in the active site, and does not indicate a direct hydrogen bond between Asp85 and Asp212. The mutual repulsion of the negatively charged Asp85 and Asp212 in the bR resting state is likely reduced in O (where Asp85 is neutral). This lesser repulsion in O allows a decrease of the Asp85:Asp212 distance relative to the bR resting state, as indicated in the acid-blue crystal structure.<sup>23</sup> The QM/MM MD simulations on the acid-blue crystal structure (model-A<sup>3w</sup>-O) indicate that the active site can accommodate three water molecules.

## B. Proton transfer pathways

We calculated PT pathways between the O and [O] states for all investigated models. Since FTIR studies indicate that the protein environment does not change during the O to [O] transition,<sup>19</sup> we calculated minimal energy pathways - i.e., we neglected the entropic contribution to the energy. The entropic contribution is expected to be more important for the PT to the PRG where the proton is transferred over ~11 Å distance. The activation energy of various PT pathways varies between ~4.0 kcal/mol and ~11.0 kcal/mol with the exception of the path in model-C<sup>3w</sup>. The slightly high barrier for model-C<sup>3w</sup> (13.8 kcal/mol, path 3c) corresponds to the PT via w402 which is strongly hydrogen bonded to the Schiff base. Except for the path computed from the acid-blue structure with one water molecule (path 1a), the energy profiles of the paths (path 1b, 2a-b, and 3a-c; Table 2) indicate small endothermic reaction energies of ~4 kcal/mol. The reaction energies and PT barriers obtained for the various PT paths support the idea that a transient [O] state with protonated Asp212 can be sampled during the O→bR transition.

## C. UV-Vis spectral fingerprints

The calculated optical excitation energies allow us to distinguish between the structural models used. Structural models containing either one or two water molecules in the active site also have a salt bridge between the retinal Schiff base and Asp85 or Asp212. Standard QM/MM computations using these models give excitation energies that are strongly blue-shifted relative to the experimental values. However, since the excitation energy of the bR resting state is also highly overestimated (2.50 eV instead of the experimental value of 2.18 eV), the standard QM/MM calculations are not conclusive.<sup>104</sup> Therefore, we estimated the polarization red-shifts by using a polarizable model for the protein that was developed recently.<sup>50,51</sup> Polarization effects largely red-shift the absorption of the bR resting state to 2.24 eV (see Table 4). In contrast, in the case of the O state, the effect of polarization on the excitation energies is very small. From the optical spectra computations we conclude that the structural models with only one or two water molecules in the active site are unlikely, since they would give a strongly blue-shifted absorption spectrum that is incompatible with experiments. The Arg82 side-chain is likely oriented towards the extracellular side, because a cytoplasmic orientation red-shifts the absorption spectrum as in case of model-C<sup>3w</sup>.

The results in Table 4 indicate that a proper description of the protein electrostatics is required for a quantitative description of the excitation energies for bR resting state, and of the shift between the bR resting state and the O state. The theoretical value of 0.54 eV for bR-O shift (model-A<sup>3w</sup>) is still overestimated (compared to the experimental bR-O shift of 0.24 eV).

Although the absolute excitation energies are within 0.06 eV of the experimental absorption maxima. The hypsochromic shift from the O to the transient [O] state is reproduced regardless of the protein electrostatic model employed. The excitation energy for model-C<sup>3w</sup> is underestimated by 0.08 eV, a value that can be attributed to the cytoplasmic orientation of Arg82. With its positive excess charge closer to the Schiff base, and further away from the retinal  $\beta$ -ionone ring. The excitation energy of model-C<sup>3w</sup> is significantly blue-shifted when the Arg82 side-chain extracellularly oriented as in model-A<sup>3w</sup>.

The active site of the structure prepared from the Asp85Ser mutant<sup>22</sup> with three water molecules (model-B<sup>3w</sup>) has a rather unusual interaction between the retinal C<sub>e</sub> atom and Asp85. Due to the interaction between retinal and Asp85, the excitation energy of model-B<sup>3w</sup> is strongly blue-shifted relative to the experimental value. From all structural models investigated here, the experimental excitation energies are best reproduced by model-A<sup>3w</sup>. Model-A<sup>3w</sup> also reproduces the [O]-O shift in the excitation energy. However, the excitation energies are not sensitive to large scale conformational changes that occur in regions remote from the active site.

#### D. IR spectral fingerprints

We used the approximate method SCC-DFTB to compute vibrational frequencies. Using gas-phase models of acetic acid, we benchmarked SCC-DFTB for  $\nu_{COOH}$  for two extremes scenarios, acetic acid in the gas phase model and a strongly hydrogen bonded cyclic dimer (see Supplementary Information). We discuss all vibrational frequencies results with respect to these gas phase reference calculations. SCC-DFTB reproduces the monomer C=O stretch frequencies extremely well (Table S-1 of Supplementary Information), deviating only by 2 cm<sup>-1</sup> relative to the experimental value<sup>94</sup> of 1780 cm<sup>-1</sup>. The agreement between experiments and theory is worse for the cyclic dimer, where we found a deviation of 17 cm<sup>-1</sup>. These two test cases are the two extremes of a free C=O group, and of a strongly hydrogen bonded acetate, that mark all hydrogen bonding possibilities in the protein. We expect SCC-DFTB to produce increasingly red-shifted vibrational frequencies with increasing hydrogen bonding interactions. This allows us to assess the O and [O] models with respect to the C=O stretch frequencies. The C=O stretch frequencies do not change when salt bridge forms between the Schiff base and Asp85 or Asp212, being sensitive only to direct hydrogen bonding interactions of the protonated aspartate. The main structural determinants of the aspartate stretching frequencies are the hydrogen bonds with active site water molecules and to nearby groups such as Thr89 (Asp85) and Tyr57 (Asp212). For most models, the [O] state vibrational frequencies are in good agreement with the experimental values, whereas the O state frequencies vary significantly, mostly due to the Thr89:Asp85 hydrogen bond.

In the models used in our study, we find that Thr89 hydrogen bonds either to the side-chain (in model-A structures derived from acid-blue bacteriorhodopsin<sup>23</sup>) or to the backbone of Asp85 (in model-B structures derived from the Asp85Ser mutant<sup>22</sup>). Since the presence of the Thr89:Asp85 hydrogen bond red-shifts the C=O frequency by approximately 20 cm<sup>-1</sup>, the hydrogen bonding of Thr89 (to the carboxyl or to the carbonyl group of Asp85) is clearly distinguished by the vibrational frequency of Asp85. Based on the critical examination of the SCC-DFTB performance for vibrational frequencies, we expect slightly red-shifted frequencies compared to experiment (1752-1756 cm<sup>-1</sup>). Therefore, computation of vibrational frequencies further supports the model-A<sup>3w</sup> with the Thr89:Asp85 hydrogen bond. The shift in the  $\nu_{COOH}$  band for O and [O] states is best reproduced in models with three active site water molecules.

Resonance Raman studies have indicated that the all-*trans* retinal in the O intermediate has a twisted geometry.<sup>30,95</sup> The O-like models used in our study agree with the experimental findings,<sup>30</sup> indicating a larger twist in the retinal geometry in O as compared to bR. The retinal

twist occurs during the optimization of model-C<sup>3w</sup>. To obtain a twist of the retinal chain it is sufficient to start from the bR resting state and change the protonation states to correspond to the O state.

## V. CONCLUSION

We modeled the active site of the O state structure using QM/MM simulations. Our investigation of the details of the hydrogen bonded network in the active site indicate that water molecules not present in acid-blue bacteriorhodopsin<sup>23</sup>, Asp85Ser mutant<sup>22</sup> crystal structures, have an important structural role in the O intermediate. The geometric and the electronic (absorption energies) features of the retinal chromophore depend on the number of water molecules in the active site. Although the crystal structures of the Asp85Ser mutant<sup>22</sup> and of the acid-blue bacteriorhodopsin<sup>23</sup> indicate only one active site water molecule, the geometry of active site from the crystal structure of acid-blue bacteriorhodopsin<sup>23</sup> is best preserved in computations when three water molecules are present in the active site. Similar to the bR resting state,<sup>24</sup> in O, the Schiff base hydrogen bonds directly to a water molecule. The absence of the water molecule hydrogen bonding to the Schiff base would lead to large blue-shifts in the optical absorption spectra. The Thr89:Asp85 hydrogen bond present in the first half of the photocycle<sup>12,24,26,43,96-98</sup> is absent in some of the O state models investigated here. The agreement between the experimental and the calculated C=O stretch frequency computed with the acid-blue bacteriorhodopsin structure in which the Thr89:Asp85 hydrogen bond is present indicates that the Thr89:Asp85 hydrogen bond is indeed a characteristic of the O state. Consistent with the findings of Resonance Raman experiments, all O state models indicate an increased twist of the retinal chromophore with respect to the bR resting state.

To summarize, we propose that the O state model derived from the acid-blue bacteriorhodopsin structure<sup>23</sup> with three active site water molecules, a hydrogen bond between Thr89 and Asp85, and twisted retinal is a good representation of the wild-type O state.

Standard QM/MM excitation energy calculations seem to be insufficient for describing the change in absorption energies between the bR and O state since protein polarization effects are very different in magnitude in these two states. The dependency of the effect of polarization on the structural models used indicates that the polarization effects must be carefully taken into account in studies of color tuning in rhodopsins.

According to the reaction path calculations, a PT from Asp85 to Asp212 is thermodynamically feasible considering the rate-limiting energy barrier of ~15 kcal/mol for the O→bR reaction.<sup>18,99</sup> The putative [O], as an intermediate between O and bR, shows changes in the spectroscopic properties as reported experimentally.<sup>1,82</sup> The experimental values of the shift in the C=O stretch frequency is ~40 cm<sup>-1</sup> and the shift in the optical excitation energy is ~0.04 eV between O and [O] states. These changes are well reproduced by our structural models of the O and [O] states modeled from the crystal structure of ref<sup>23</sup> (model-A<sup>3w</sup>), thereby confirming the picture of a transient O-intermediate, which is characterized by a protonated Asp212 and deprotonated Asp85. These results indicate that indeed, the spectral shift observed in ref<sup>1</sup> is likely due to Asp212 becoming protonated. The calculations support the proposal of ref<sup>1</sup> that Asp212 is an intermediate carrier of the proton during the O→bR transition.

The mechanism of the complete O→bR will be the focus of future work. Several key issues must be addressed. For example, the present computations indicate that in O, the Arg82 side-chain is oriented towards the extracellular side. For the recovery of the bR resting state, the side-chain must orient back to its cytoplasmic configuration. The rate-limiting energy barrier, the sequence of events associated with the reorientation of the Arg82 side-chain, and whether



or not reorientation of Arg82 is part of the last PT step, are still open questions that will be addressed in the future using the approach presented here.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

This work was supported in part by the Deutsche Forschungsgemeinschaft through Forschergruppe 490 (M.E.). P.P. thanks computational resources from German Cancer Research Center and T. Kubař for stimulating discussions. A.-N.B. is supported by grants GM74637 and GM68002 from the National Institutes of General Medical Sciences. J.C.S. was supported by a United States Department of Energy Laboratory-Directed Research and Development grant.

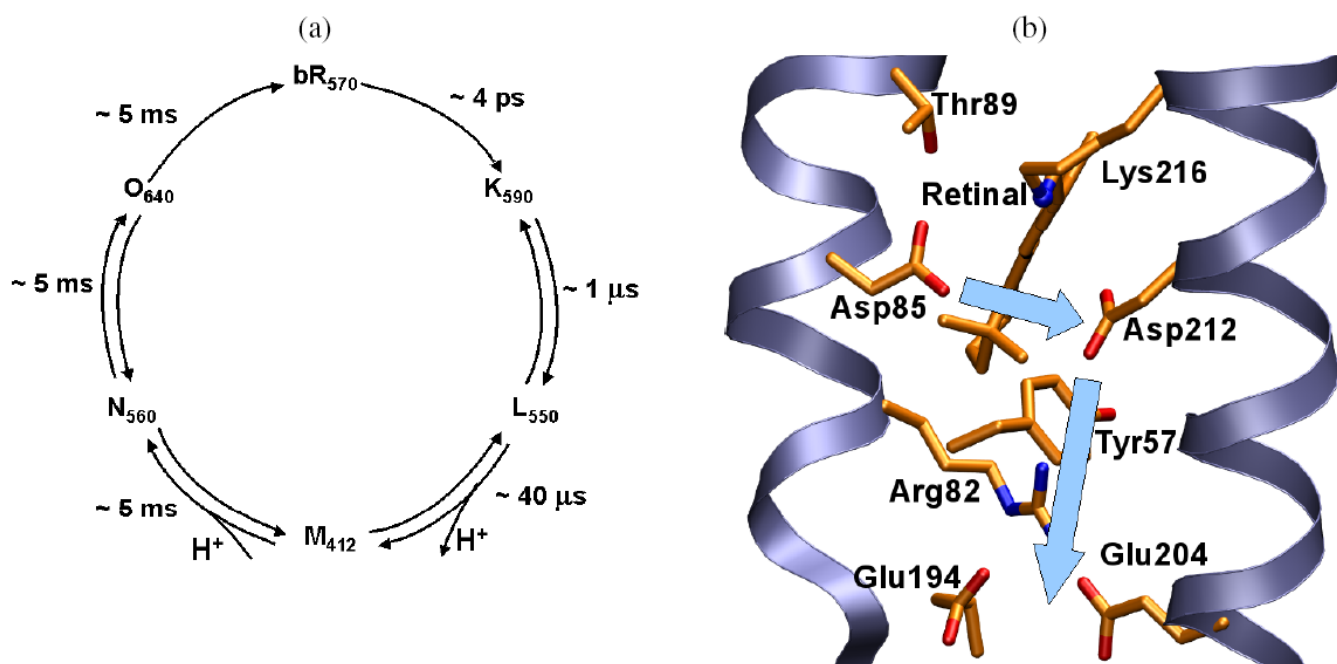
## References

1. Dioumaev AK, Brown LS, Needleman R, Lanyi JK. *Biochemistry* 1999;38:10070–10078. [PubMed: 10433714]
2. Balashov SP, Govindjee R, Imasheva ES, Misra S, Ebrey TG, Feng Y, Crouch RK, Menick DR. *Biochemistry* 1995;34:8820–8834. [PubMed: 7612623]
3. Song Y, Mao J, Gunner MR. *Biochemistry* 2003;42:9875–9888. [PubMed: 12924936]
4. Phatak P, Ghosh N, Yu H, Cui Q, Elstner M. *Proc Natl Acad Sci USA* 2008;105:19672–19677. [PubMed: 19064907]
5. Rammelsberg R, Huhn G, Luebben M, Gerwert K. *Biochemistry* 1998;37:5001–5009. [PubMed: 9538019]
6. Essen LO, Siegert R, Lehmann WD, Oesterhelt D. *Proc Natl Acad Sci USA* 1998;95:11673–11678. [PubMed: 9751724]
7. Zscherp C, Schlesinger R, Tittor J, Oesterhelt D, Heberle J. *Proc Natl Acad Sci USA* 1999;96:5498–5503. [PubMed: 10318912]
8. Garczarek F, Brown LS, Lanyi JK, Gerwert K. *Proc Natl Acad Sci USA* 2005;102:3633–3638. [PubMed: 15738416]
9. Garczarek F, Gerwert K. *Nature* 2006;439:109–112. [PubMed: 16280982]
10. Subramaniam S, Lindahl M, Bullough P, Faruqi AR, Tittor J, Oesterhelt D, Brown L, Lanyi JK, Henderson R. *J Mol Biol* 1999;287:145–161. [PubMed: 10074413]
11. Sass HJ, Schachowa IW, Rapp G, Koch MHJ, Oesterhelt D, Dencher NA, Buldt G. *EMBO J* 1997;16:1484–1491. [PubMed: 9130693]
12. Kouyama T, Nishikawa T, Tokuhisa T, Okumura H. *J Mol Biol* 2004;335:531–546. [PubMed: 14672661]
13. Subramaniam S, Greenhalgh DA, Rath P, Rothschild KJ, Khorana HG. *Proc Natl Acad Sci USA* 1991;88:6873–6877. [PubMed: 1650486]
14. Brown LS, Sasaki J, Kandori H, Maeda A, Needleman R, Lanyi JK. *J Biol Chem* 1995;270:27122–27126. [PubMed: 7592966]
15. Dioumaev AK, Brown LS, Needleman R, Lanyi JK. *Biochemistry* 1998;37:9889–9893. [PubMed: 9665693]
16. Richter HT, Needleman R, Kandori H, Maeda A, Lanyi JK. *Biochemistry* 1996;35:15461–15466. [PubMed: 8952499]
17. Chizhov I, Engelhard M, Chernavskii DS, Zubov B, Hess B. *Biophys J* 1992;61:1001–1006. [PubMed: 19431821]
18. Ludmann K, Gergely C, Váró G. *Biophys J* 1998;75:3110–3119. [PubMed: 9826631]
19. Zscherp C, Schlesinger R, Heberle J. *Biochem Biophys Res Comm* 2001;283:57–63. [PubMed: 11322767]
20. Balashov SP, Imasheva ES, Govindjee R, Ebrey TG. *Biophys J* 1996;70:473–481.
21. Efremov R, Gordeliy VI, Heberle J, Buldt G. *Biophys J* 2006;91:1441–1451.

22. Rouhani S, Cartiailler JP, Facciotti MT, Walian P, Needleman R, Lanyi JK, Glaeser RM, Luecke H. *J Mol Biol* 2001;313:615–628. [PubMed: 11676543]
23. Okumura H, Murakami M, Kouyama T. *J Mol Biol* 2005;351:481–495. [PubMed: 16023672]
24. Luecke H, Schobert B, Richter HT, Cartiailler JP, Lanyi JK. *J Mol Biol* 1999;291:899–911. [PubMed: 10452895]
25. Chen D, Wang JM, Lanyi JK. *J Mol Biol* 2007;366:790–805. [PubMed: 17196982]
26. Bondar AN, Suhai S, Fischer S, Smith JC, Elstner M. *J Struct Biol* 2007;157:454–469. [PubMed: 17189704]
27. Belrhali H, Nollert P, Royant A, Menzel C, Rosenbusch JP, Landau EM, Pebay-Peyroula E. *Structure* 1999;7:909–917. [PubMed: 10467143]
28. Facciotti MT, Rouhani S, Burkard FT, Betancourt FM, Downing KH, Rose RB, McDermott G, Glaeser RM. *Biophys J* 2001;81:3442–3455.
29. Matsui Y, Sakai K, Murakami M, Shiro Y, Adachi S, Okumura H, Kouyama T. *J Mol Biol* 2002;324:469–481. [PubMed: 12445782]
30. Smith SO, Pardo JA, Mulder PPJ, Curry B, Lugtenburg J, Mathies R. *Biochemistry* 1983;22:6141–6148.
31. Brooks BR, Bruccoleri RE, Olafson BD, States DJ, Swaminathan S, Karplus M. *J Comp Chem* 1983;4:187–217.
32. Elstner M, Porezag D, Jungnickel G, Elsner J, Haugk M, Frauenheim Th, Suhái S, Seifert G. *Phys Rev B* 1998;58:7260–7268.
33. Cui Q, Elstner M, Kaxiras E, Frauenheim Th, Karplus M. *J Phys Chem B* 2001;105:569–585.
34. Elstner M. *Theor Chem Acc* 2006;116:316–325.
35. Zhang X, Harrison DVT, Cui Q. *J Am Chem Soc* 2002;124:14871–14878. [PubMed: 12475328]
36. Cui Q, Elstner M, Karplus M. *J Phys Chem B* 2002;106:2721–2740.
37. Li G, Cui Q. *J Am Chem Soc* 2003;125:15028–15038. [PubMed: 14653737]
38. Li G, Zhang X, Cui Q. *J Phys Chem B* 2003;107:8643–8653.
39. Elstner M, Cui Q, Munih P, Kaxiras E, Frauenheim T, Karplus M. *J Comp Chem* 2003;24:565. [PubMed: 12632471]
40. Elstner M, Frauenheim T, Suhái S. *J Mol Struct:THEOCHEM* 2003;632:29–41.
41. Riccardi D, König P, Prat-Resina X, Yu H, Elstner M, Frauenheim T, Cui Q. *J Am Chem Soc* 2006;128:16302–16311. [PubMed: 17165785]
42. Zhou H, Tajkhorshid E, Frauenheim T, Suhái S, Elstner M. *Chem Phys* 2002;277:91–103.
43. Bondar AN, Elstner M, Suhái S, Smith JC, Fischer S. *Structure* 2004;12:1–20. [PubMed: 14725756]
44. Bondar AN, Fischer S, Smith JC, Elstner M, Suhái S. *J Am Chem Soc* 2004;126:14668–14677. [PubMed: 15521787]
45. Bondar AN, Fischer S, Suhai S, Smith JC. *J Phys Chem B* 2005;109:14786–14788. [PubMed: 16852870]
46. Bondar AN, Baudry J, Suhai S, Fischer S, Smith JC. *J Phys Chem B* 2008;112:14729–14741. [PubMed: 18973373]
47. Okada T, Sugihara M, Bondar AN, Elstner M, Entel P, Buss V. *J Mol Biol* 2004;342:571–583. [PubMed: 15327956]
48. Wanko M, Hoffmann M, Strodel P, Koslowski A, Thiel W, Neese F, Frauenheim T, Elstner M. *J Phys Chem B* 2005;109:3606–3615. [PubMed: 16851399]
49. Hoffmann M, Wanko M, Strodel P, Koenig P, Frauenheim T, Schulten K, Thiel W, Tajkhorshid E, Elstner M. *J Am Chem Soc* 2006;128:10808–10818. [PubMed: 16910676]
50. Wanko M, Hoffmann M, Frauenheim T, Elstner M. *J Phys Chem B* 2008;112:11462–11467. [PubMed: 18698712]
51. Wanko M, Hoffmann M, Fraehmcke J, Frauenheim T, Elstner M. *J Phys Chem B* 2008;112:11468–11478. [PubMed: 18729405]
52. Warshel A, Levitt M. *J Mol Biol* 1976;103:227–249. [PubMed: 985660]
53. Field M, Bash P, Karplus M. *J Comp Chem* 1990;6:700–733.
54. Gao J. *Acc Chem Res* 1996;29:298–305.

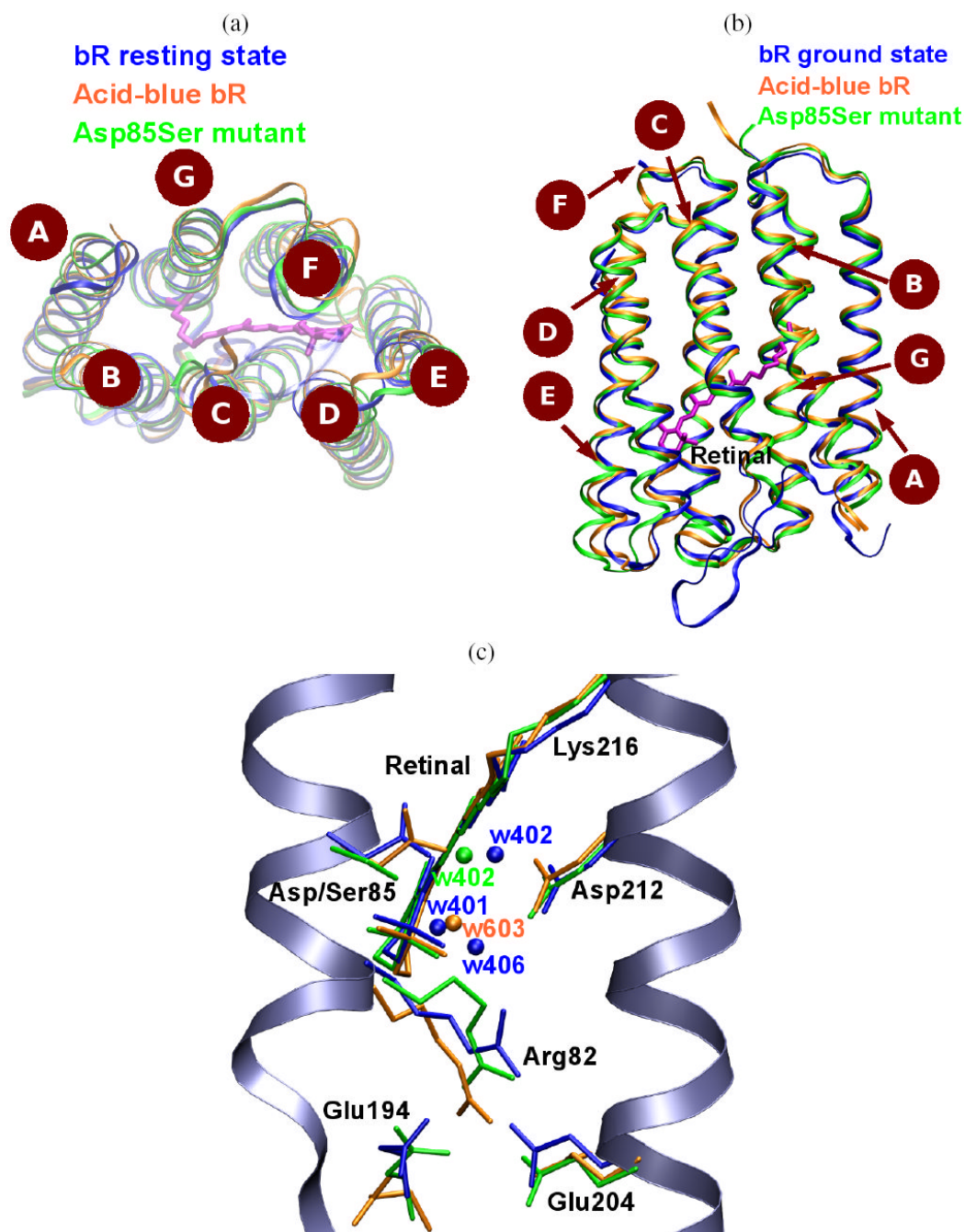
55. Senn HM, Thiel W. *Top Curr Chem* 2007;268:173–290.
56. Warshel A, Chu ZT. *J Phy Chem B* 2001;105:9857–9871.
57. Hayashi S, Tajkhorshid E, Schulten K. *Biophys J* 2003;85:1440–1449. [PubMed: 12944261]
58. Hayashi S, Ohmine I. *J Phys Chem B* 2000;104:10678–10691.
59. Hayashi S, Tajkhorshid E, Schulten K. *Biophys J* 2002;83:1281–1297.
60. Hayashi S, Tajkhorshid E, Schulten K. *J Am Chem Soc* 2004;126:10516–10517. [PubMed: 15327290]
61. Rousseau R, Kleinschmidt V, Schmitt UW, Marx D. *Phys Chem Chem Phys* 2004;6:1848–1859.
62. Mathias G, Marx D. *Proc Natl Acad Sci USA* 2007;104:6980–6985. [PubMed: 17438299]
63. Braun-sand S, Sharma P, Chu Z, Pislakov AV, Warshel A. *Biochim Biophys Acta* 2008;1777:441–452.
64. Ren L, Martin CH, Wise KJ, Gillespie NB, Luecke H, Lanyi JK, Spudich JL, Birge RR. *Biochemistry* 2001;40:13906–13914. [PubMed: 11705380]
65. Hayashi S, Tajkhorshid E, Pebay-Peyroula E, Royant R, Landau E, Navarro J, Schulten K. *J Phys Chem B* 2001;105:10124–10131.
66. Fujimoto K, Hasegawa J, Hayashi S, Kato S, Nakatsuji H. *Chem Phys Lett* 2005;414:239–242.
67. Fujimoto K, Hayashi S, Hasegawa J, Nakatsuji H. *J Chem Th and Comp* 2007;3:605–618.
68. Matsuura A, Sato H, Houjou H, Saito S, Hayashi T, Sakurai M. *J Comp Chem* 2006;27:1623–1630. [PubMed: 16900496]
69. Schaefer P, Riccardi D, Cui Q. *J Chem Phys* 2005;123:014905–014914. [PubMed: 16035867]
70. Ghosh N, Cui Q. *J Phys Chem* 2008;112:8387–9397.
71. Riccardi D, Schaefer P, Yang Y, Yu H, Ghosh N, Prat-Resina X, Koenig P, Li G, Xu D, Guo H, Elstner M, Cui Q. *J Phys Chem B* 2006;110:6458–6469. [PubMed: 16570942]
72. Elstner, M.; Cui, Q. Chapter Combined QM/MM methods for the simulation of condensed phase processes using an approximate DFT approach. In: Canuto, S., editor. *Challenges and Advances in Computational Chemistry and Physics*. Vol. 6. Springer; 2008. p. 381–405.
73. Cui, Q.; Elstner, M. Chapter Multi-scale QM/MM methods with Self-Consistent-Charge Density-Functional-Tight-Binding (SCC-DFTB). In: Lee, T.; York, D., editors. *Multi-scale Quantum Models for Biocatalysis: Modern Techniques and Applications*. Springer; In Press
74. Fischer S, Karplus M. *Chem Phys Lett* 1992;194:252–261.
75. Choi C, Elber R. *J Chem Phys* 1991;94:751–760.
76. Neese F. *J Chem Phys* 2003;119:9428–9443.
77. Neese, F. ORCA - An ab initio, density functional and semiempirical program package; Version 2.3 - Revision 09. Max Planck Institut fuer Strahlenchemie; Muelheim, Germany: 2004.
78. Schäfer A, Horn H, Ahlrichs R. *J Chem Phys* 1992;97:2571–2577.
79. Cui Q, Karplus M. *J Chem Phys* 2000;112:1133–1149.
80. Yang Y, Yu H, York D, Cui Q, Elstner M. *J Phys Chem A* 2007;111:10861–10873. [PubMed: 17914769]
81. Yu H, Cui Q. *J Chem Phys* 2007;127:234504. [PubMed: 18154397]
82. Zscherp C, Heberle J. *J Phys Chem* 1997;101:10542–10547.
83. König PH, Hoffmann M, Frauenheim Th, Cui Q. *J Phys Chem B* 2005;109:9082–9095. [PubMed: 16852081]
84. Sass HJ, Büldt G, Gessenich R, Hehn D, Neff D, Schlesinger R, Berendzen J, Ormos P. *Nature* 2000;406:649–653. [PubMed: 10949308]
85. Dinner AR, Lopez X, Karplus M. *Theor Chem Acc* 2003;109:118.
86. Simmons CJ, Liu RSH, Denny M, Seff K. *Acta Crystallogr B* 1981;37:2197–2205.
87. Hamanaka T, Mitsui T, Ashida T, Kakudo M. *Acta Crystallogr B* 1972;28:214–222.
88. Santarsiero BD, James MNG, Mahendran M, Childs RF. *J Am Chem Soc* 1990;112:9416–9418.
89. Fahmy K, Siebert F, Grossjean MF, Tavan P. *J Mol Struct* 1989;214:257–288.
90. El-Sayed MA, Lin CT, Mason WR. *Proc Natl Acad Sci USA* 1989;86:5376–5379. [PubMed: 16594053]

91. Wu S, El-Sayed MA. *Biophys J* 1991;60:190–197.
92. Tajkhorshid E, Baudry J, Schulten K, Suhai S. *Biophys J* 2000;78:683–693.
93. Lansing JC, Hohwy M, Jaroniec CP, Creemers AFL, Lugtenburg J, Herzfeld J, Griffin RG. *Biochemistry* 2002;41:431–438. [PubMed: 11781081]
94. Dioumaev AK. *Biochemistry (Moscow)* 2001;66:1269–1276. [PubMed: 11743871]
95. Smith SO, Lugtenburg J, Mathies R. *J Memb Biol* 1985;85:95–109.
96. Schober B, Cupp-Vickery J, Hornak V, Smith S, Lanyi JK. *J Mol Biol* 2002;321:715–726. [PubMed: 12206785]
97. Lanyi JK, Schober B. *J Mol Biol* 2002;321:727–737. [PubMed: 12206786]
98. Kandori H, Yamazaki Y, Shichida Y, Raap J, Lugtenburg J, Belenky M, Herzfeld J. *Proc Natl Acad Sci USA* 2001;98:1571–1576. [PubMed: 11171992]
99. Subramaniam S, Faruqi AR, Oesterhelt D, Henderson R. *Proc Natl Acad Sci USA* 1997;94:1767–1772. [PubMed: 9050853]
100. Birge RR, Zhang CF. *J Chem Phys* 1990;92:7178–7195.
101. Váró G, Lanyi JK. *Biochemistry* 1991;30:5008–5015. [PubMed: 1645187]
102. Hessling B, Souvignier G, Gerwert K. *Biophys J* 1993;65:1929–1941.
103. Humphrey W, Dalke A, Schulten K. *J Molec Graphics* 1996;14:33–38.
104. The absolute excitation energy of bR resting state is 0.07 eV higher than that calculated before at SORCI level(2.34 eV). 48 This difference can be attributed to the large QM region used here (retinal, Asp85, Asp212 and three water molecules) whereas in ref 48, only retinal has been treated with QM.

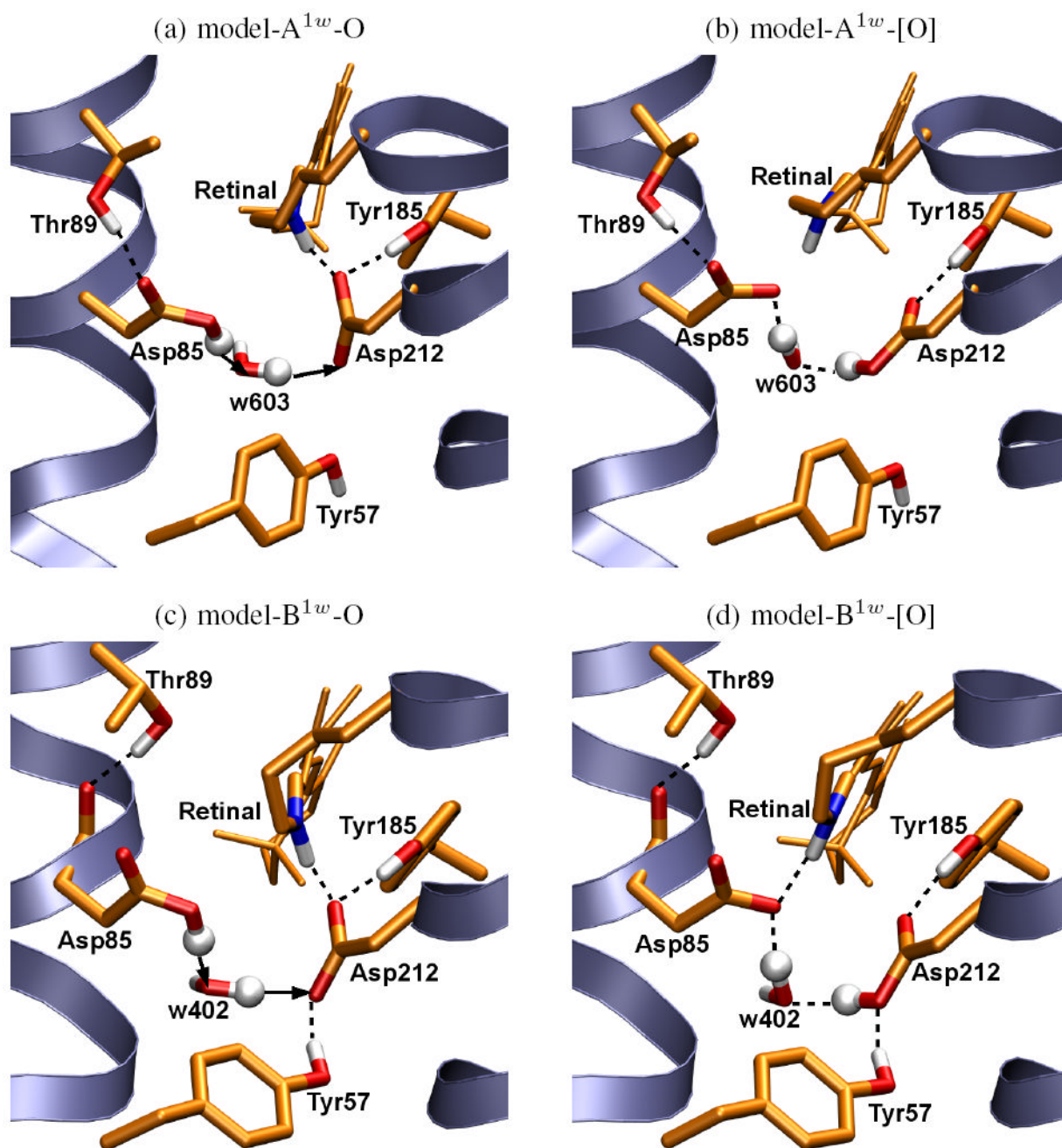


**Figure 1.**

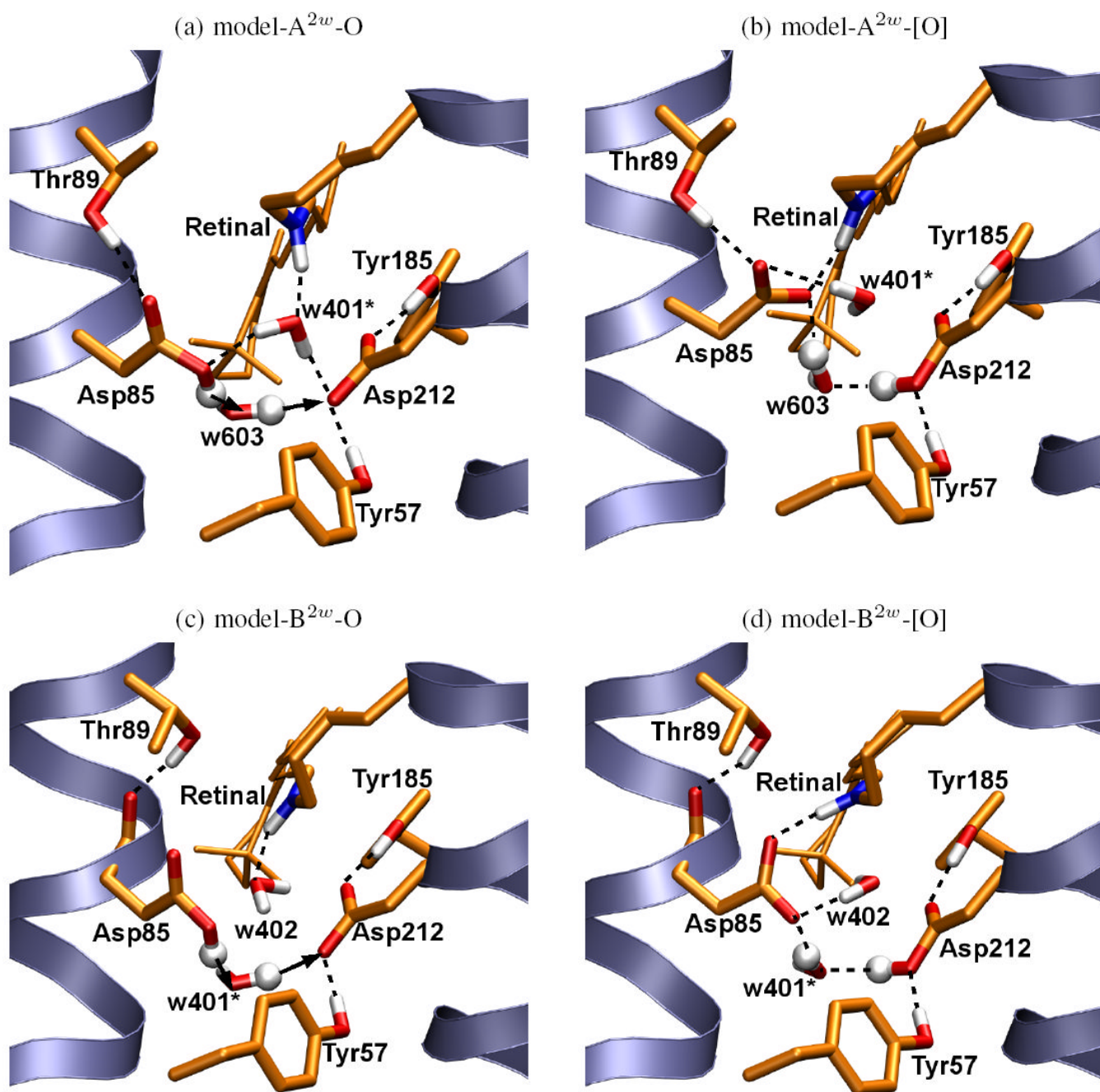
The bacteriorhodopsin proton-pumping cycle. (a) Schematic representation of the bacteriorhodopsin photocycle showing intermediate states characterized by their absorption maxima; (b) Location of protein groups essential for the proton transfer. The arrows indicate the proton transfer (PT) from Asp85 to the PRG during O→bR transition. In the bacteriorhodopsin photocycle, light-activated isomerization of the retinal from all-*trans* to 13-*cis* takes place during bR→K transition, followed by relaxation of the twisted retinal during K→L transition. The first PT takes place from Schiff base to Asp85 along with the release of another proton to the extracellular bulk by PRG during L→M transition. The M→N involves reprotonation of the Schiff base from Asp96. Subsequently Asp96 is reprotonated from the cytoplasmic medium, followed by reisomerization of retinal during N→O transition. The last PT steps takes place during O→bR transition from Asp85 to PRG. The picture is prepared using the crystal structure of bR resting state<sup>24</sup> and using the VMD package.<sup>103</sup>



**Figure 2.** Protein conformational changes in O relative to the bR resting state. Comparison of the crystal structures of the acid-blue bacteriorhodopsin<sup>23</sup>, the Asp85Ser mutant<sup>22</sup>, and the bR resting state<sup>24</sup> of a view from the cytoplasmic side (a), and parallel to the membrane normal (b). (c) Detailed view of the retinal, water molecules and amino-acids important for the last proton-transfer step. The following color codes are used: orange - acid-blue bR; green -Asp85Ser mutant; blue - bR resting state.

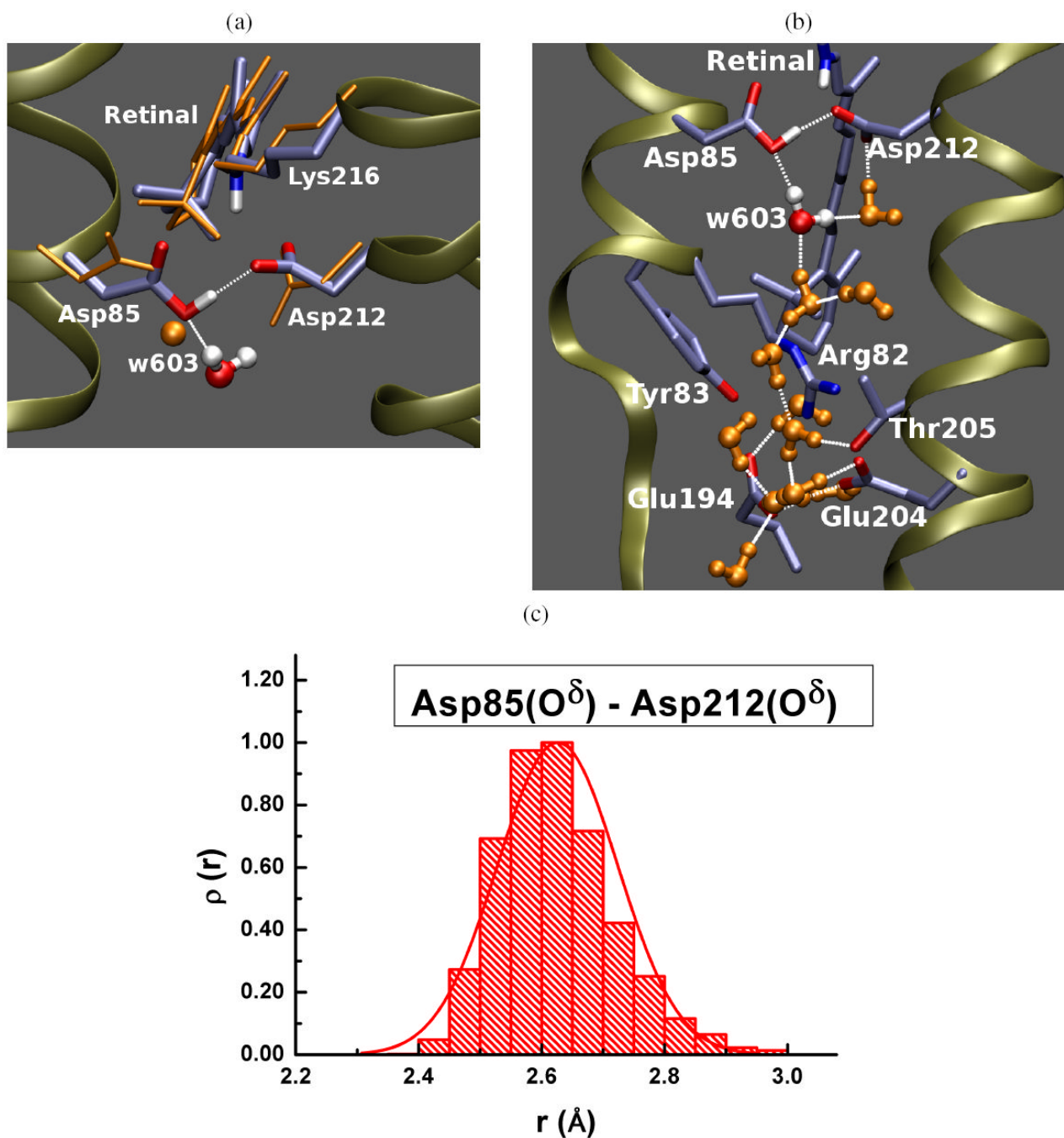


**Figure 3.** Reactant (O-like) and product ([O]-like) states of proton transfer in the presence of a single water molecule in the active site. The Thr89:Asp85 hydrogen bond is present in model-A<sup>1w</sup>-O (a) and model-A<sup>1w</sup>-[O] (b) and absent in model-B<sup>1w</sup>-O (c) and model-B<sup>1w</sup>-[O] (d). The Tyr57:Asp212 hydrogen bond is absent in model-A<sup>1w</sup>-O (a) and model-A<sup>1w</sup>-[O] (b), and is present in model-B<sup>1w</sup>-O (c) and model-B<sup>1w</sup>-[O] (d).

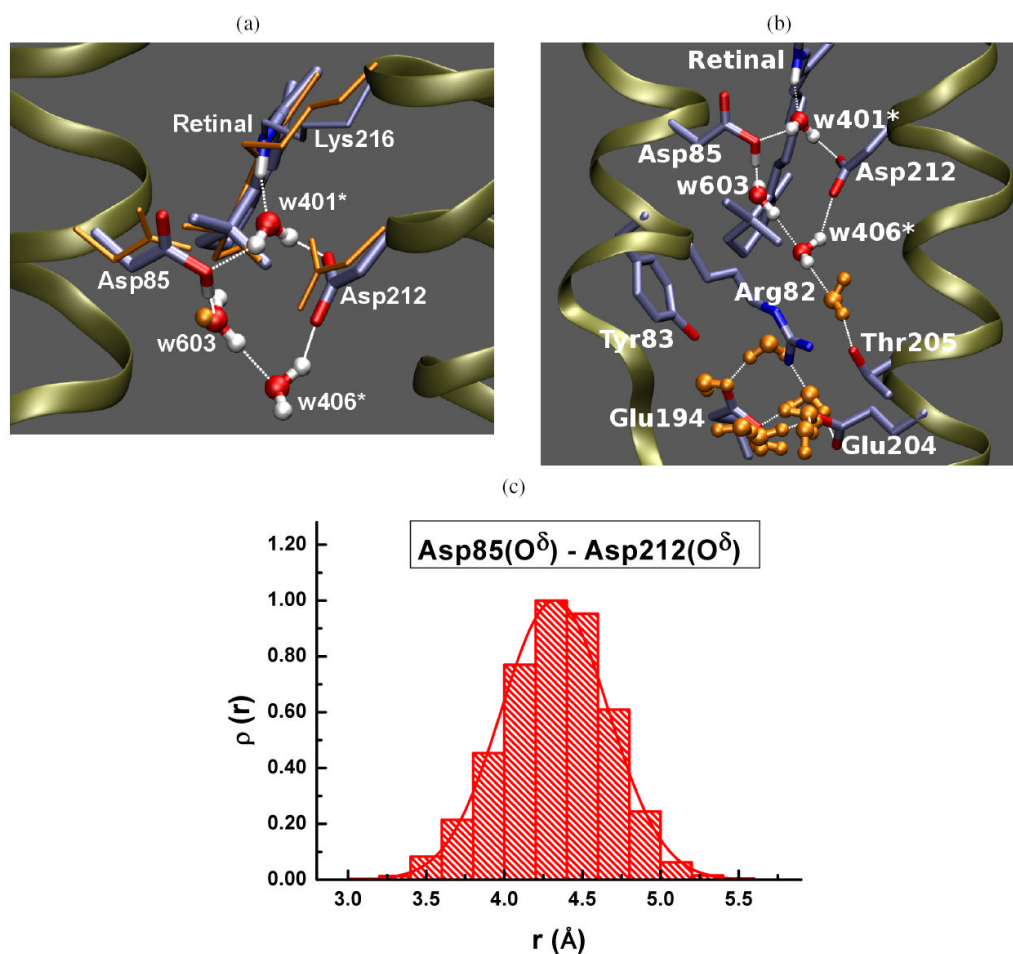


**Figure 4.** Reactant (O-like) and product ([O]-like) states in the presence of two active site water molecules. The retinal Schiff base hydrogen bonds with w401\* in model-A<sup>2w</sup>-O (a), with w402 in model-B<sup>2w</sup>-O (c), and with Asp85 in model-A<sup>2w</sup>-[O] (b) and model-B<sup>2w</sup>-[O] (d).

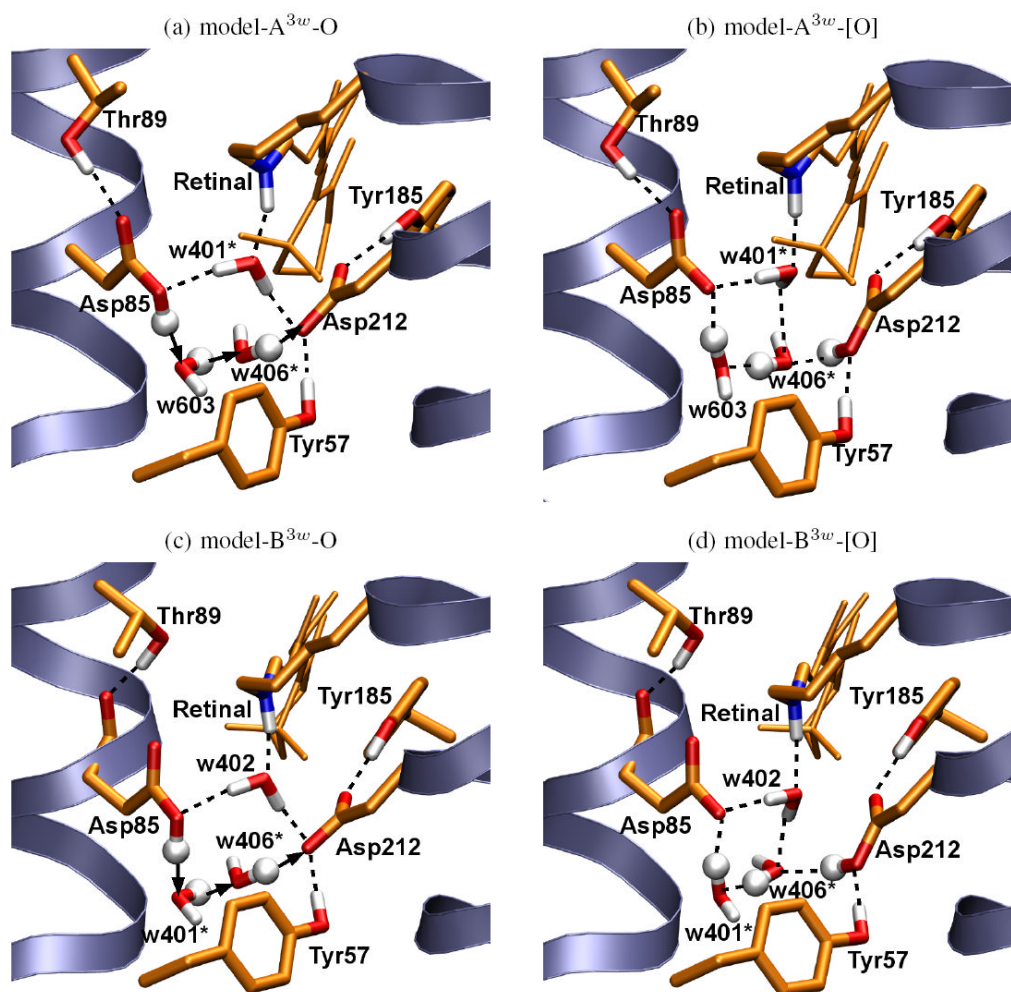




**Figure 5.** Classical molecular dynamics simulations of model-A<sup>1w</sup>-O. (a) Comparison of the active site of the crystal structure<sup>23</sup>(orange) with a snapshot from the simulation shows that the essential structural details of the crystal structure are not preserved during the MD simulations; (b) A hydrogen-bonded network forms that connects the proton donor group Asp85 with the PRG Glu194/Glu204; (c) The distance between Asp85 and Asp212 decreases dramatically from 4.2 Å in the crystal structure<sup>23</sup> to ~2.7 Å in the simulations.

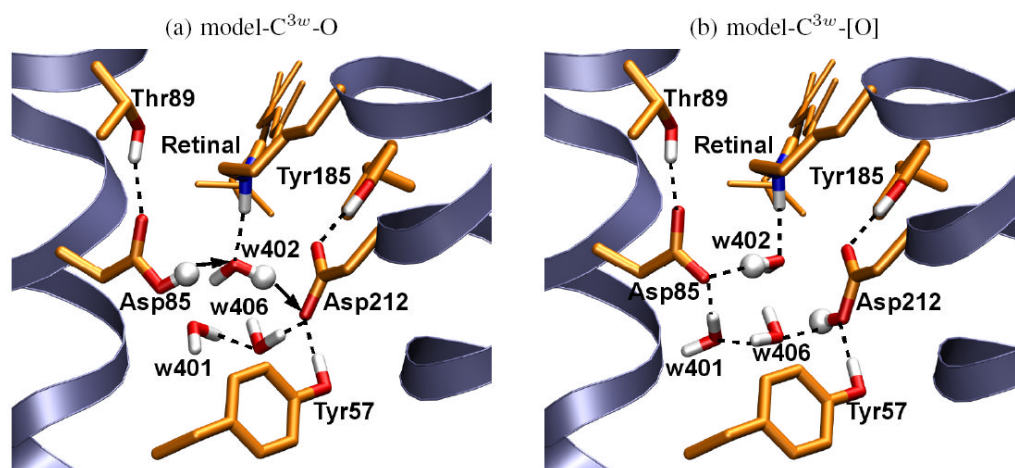


**Figure 6.** The geometry of the active site indicated in ref <sup>23</sup> is well preserved in MD simulation when three water molecules are present in the active site. The good agreement between the crystal structure and the MD simulations with three water molecules (model-A<sup>3w</sup>-O) indicate that three water molecules are likely present in the active site. (a) Comparison of the active site geometry of the crystal structure (orange) with a snapshot from the SCC-DFTB MD simulation; (b) A snapshot from MD simulation starting from model-A<sup>3w</sup>-O; (c) Histogram of the distance between Asp85/Asp212 closest oxygen atoms.

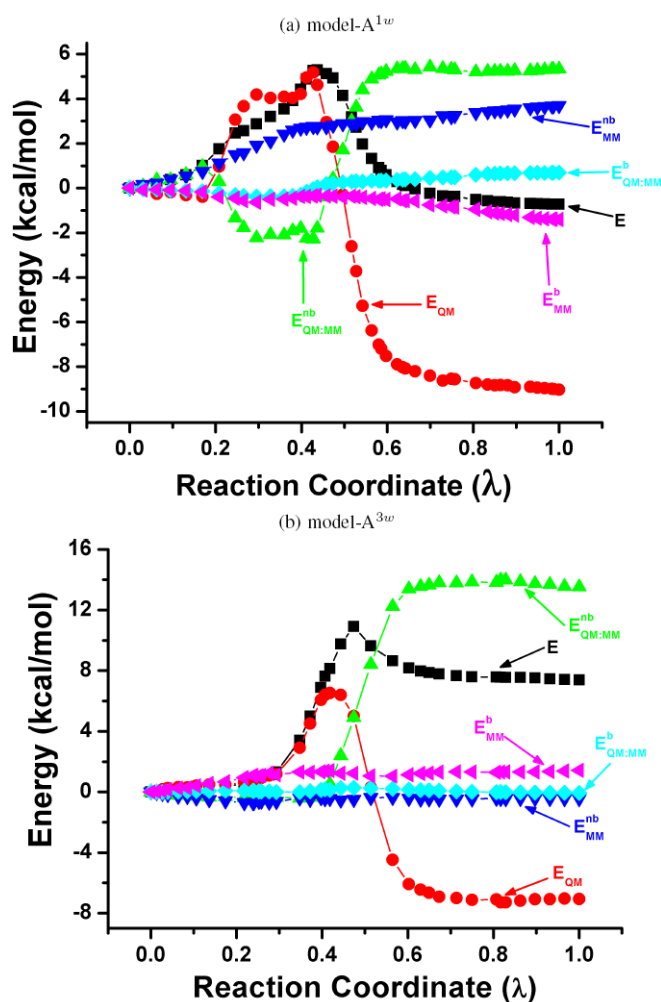


**Figure 7.**

Reactant (O-like) and product ([O]-like) states with three water molecules in the active site. The Schiff base:water hydrogen bond is present regardless of the starting coordinates used for the protein (acid-blue bR<sup>23</sup> in model-A<sup>3w</sup> (panels (a) & (b)), and the Asp85Ser mutant in model-B<sup>3w</sup> (panels (c) & (d)). The Thr89:Asp85 hydrogen bond is present in model-A<sup>3w</sup> structures (panels (a) & (b)), but absent in model-B<sup>3w</sup>. The arrows in panels (a) and (c) indicate the PT path.



**Figure 8.** Reactant (O-like) and product ([O]-like) states derived from the bR resting state. In the both model- $C^{3w}$ -O (panel (a)) and model- $C^{3w}$ -[O] (panel (b)) the Schiff base hydrogen bonds with w402, and Thr89 hydrogen bonds with Asp85. The arrows in panel (a) indicate the PT path.



**Figure 9.** Energy profile of PT from Asp85 to Asp212 calculated with model-A<sup>1w</sup> (a) and model-A<sup>3w</sup> (b). The following color codes are used: E (black); E<sub>QM</sub> (red); E<sub>QM:MM</sub><sup>nb</sup> (green); E<sub>QM:MM</sub><sup>b</sup> (cyan); E<sub>MM</sub><sup>nb</sup> (blue) and E<sub>MM</sub><sup>b</sup> (magenta). At  $\lambda=0$  the proton is on Asp85, and at  $\lambda=1.0$  the proton is on Asp212.

TABLE 1

Details of the end states of the various O models<sup>a</sup>

Model	Starting crystal structure	Protonation state	Active-site water molecules(s)
Model-A <sup>1w</sup> -O	Acid-blue bR <sup>b</sup>	Asp85 <sup>0</sup> /Asp212 <sup>-1</sup> /Glu204 <sup>-1</sup>	w603
Model-A <sup>1w</sup> -[O]	Acid-blue bR	Asp85 <sup>-1</sup> /Asp212 <sup>0</sup> /Glu204 <sup>-1</sup>	w603
Model-A <sup>2w</sup> -O	Acid-blue bR	Asp85 <sup>0</sup> /Asp212 <sup>-1</sup> /Glu204 <sup>-1</sup>	w603, w401*
Model-A <sup>2w</sup> -[O]	Acid-blue bR	Asp85 <sup>-1</sup> /Asp212 <sup>0</sup> /Glu204 <sup>-1</sup>	w603, w401*
Model-A <sup>3w</sup> -O	Acid-blue bR	Asp85 <sup>0</sup> /Asp212 <sup>-1</sup> /Glu204 <sup>-1</sup>	w603, w401*, w406*
Model-A <sup>3w</sup> -[O]	Acid-blue bR	Asp85 <sup>-1</sup> /Asp212 <sup>0</sup> /Glu204 <sup>-1</sup>	w603, w401*, w406*
Model-B <sup>1w</sup> -O	Asp85Ser mutant <sup>c</sup>	Asp85 <sup>0</sup> /Asp212 <sup>-1</sup> /Glu204 <sup>-1</sup>	w402
Model-B <sup>1w</sup> -[O]	Asp85Ser mutant	Asp85 <sup>-1</sup> /Asp212 <sup>0</sup> /Glu204 <sup>-1</sup>	w402
Model-B <sup>2w</sup> -O	Asp85Ser mutant	Asp85 <sup>0</sup> /Asp212 <sup>-1</sup> /Glu204 <sup>-1</sup>	w402, w401*
Model-B <sup>2w</sup> -[O]	Asp85Ser mutant	Asp85 <sup>-1</sup> /Asp212 <sup>0</sup> /Glu204 <sup>-1</sup>	w402, w401*
Model-B <sup>3w</sup> -O	Asp85Ser mutant	Asp85 <sup>0</sup> /Asp212 <sup>-1</sup> /Glu204 <sup>-1</sup>	w402, w401*, w406*
Model-B <sup>3w</sup> -[O]	Asp85Ser mutant	Asp85 <sup>-1</sup> /Asp212 <sup>0</sup> /Glu204 <sup>-1</sup>	w402, w401*, w406*
Model-C <sup>3w</sup> -O	bR resting state <sup>d</sup>	Asp85 <sup>0</sup> /Asp212 <sup>-1</sup> /Glu204 <sup>-1</sup>	w401, w402, w406
Model-C <sup>3w</sup> -[O]	bR resting state	Asp85 <sup>-1</sup> /Asp212 <sup>0</sup> /Glu204 <sup>-1</sup>	w401, w402, w406

<sup>a</sup>Extra added waters are marked with asterisk(\*);<sup>b</sup><sub>ref</sub> 23;<sup>c</sup><sub>ref</sub> 22;<sup>d</sup><sub>ref</sub> 24

**TABLE 2**  
 Activation energies ( $\Delta E_A$ , kcal/mol) and reaction energies ( $\Delta E_R$ , kcal/mol) for proton transfer from Asp85 to Asp212

Pathway	Model	Active-site water molecule(s)	Member of the Proton Wire	$\Delta E_A$ (kcal/mol)	$\Delta E_R$ (kcal/mol)
1a	model-A <sup>1w</sup>	w603	w603	5.3	-0.7
1b	model-B <sup>1w</sup>	w402	w402	4.8	3.1
2a	model-A <sup>2w</sup>	w603, w401*	w603	3.9	0.5
2b	model-B <sup>2w</sup>	w402, w401*	w401*	10.2	5.8
3a	model-A <sup>3w</sup>	w603, w401*, w406*	w603-w406*	10.9	7.4
3b	model-B <sup>2w</sup>	w402, w401*, w406*	w401*-w406*	6.7	3.8
3c	model-C <sup>3w</sup>	w401, w402, w406	w402	13.8	3.6

**TABLE 3**  
 S<sub>1</sub> Excitation energies (in eV) and spectral shifts (in eV) for various models<sup>a</sup>.

Model	State	Excitation Energy	bR-O shift <sup>b</sup>	Experimental bR-O shift	[O]-O shift	Experimental [O]-O shift
model-A <sup>1w</sup>	O	2.09	0.41	0.24		
	[O]	2.16			0.07	0.04
model-B <sup>1w</sup>	O	2.20	0.30	0.24		
	[O]	2.4			0.23	0.04
model-A <sup>2w</sup>	O	1.92	0.58	0.24		
	[O]	2.06			0.14	0.04
model-B <sup>2w</sup>	O	2.17	0.33	0.24		
	[O]	2.32			0.16	0.04
model-A <sup>3w</sup>	O	1.96	0.54	0.24		
	[O]	2.03			0.07	0.04
model-B <sup>3w</sup>	O	2.15	0.35	0.24		
	[O]	2.29			0.14	0.04
model-C <sup>3w</sup>	O	1.85	0.65	0.24		
	[O]	1.91			0.06	0.04

<sup>a</sup> Experimental values: br resting state: 2.18 eV<sup>100</sup>, O state: 1.94 eV<sup>101</sup>, [O] state: 1.98 eV<sup>1</sup>.

<sup>b</sup> Calculated excitation energy of bR: 2.50 eV (Table 4)



**TABLE 4**

Effect of charge scaling (CS) and protein polarization (polar.h) on the excitation energies (eV) of ground-state bR and various O Models.

Model	State	CHARMM		
		no CS	with CS	polar.h with CS
bR	resting state	2.50	2.41	2.24
Model-A <sup>3w</sup>	O	1.96	1.94	1.92
Model-A <sup>3w</sup>	[O]	2.03	1.99	2.00
Model-C <sup>3w</sup>	O	1.85	1.87	1.86

**TABLE 5**C=O stretching mode (in  $\text{cm}^{-1}$ ) in various O-models<sup>a</sup>.

Model	State	IR mode	O-[O] shift <sup>b</sup>
Model-A <sup>1w</sup>	O	1690	38
	[O]	1652	
Model-B <sup>1w</sup>	O	1741	39
	[O]	1702	
Model-A <sup>2w</sup>	O	1741	39
	[O]	1702	
Model-B <sup>2w</sup>	O	1758	54
	[O]	1704	
Model-A <sup>3w</sup>	O	1737	37
	[O]	1700	
Model-B <sup>3w</sup>	O	1756	37
	[O]	1719	
Model-C <sup>3w</sup>	O	1728	24
	[O]	1704	

<sup>a</sup> Experimental values: 1,82,102 Asp85 in O state: 1752-1756  $\text{cm}^{-1}$ , Asp212 in transient O([O]): 1712  $\text{cm}^{-1}$ , 82.

<sup>b</sup> Experimental O-[O] shift  $\approx$  42  $\text{cm}^{-1}$ .

**TABLE 6**

Torsional angles (in degree) around the double bonds of the polyene chain of retinal chromophore in various O models

	C <sub>7</sub> =C <sub>8</sub>	C <sub>11</sub> =C <sub>12</sub>	C <sub>13</sub> =C <sub>14</sub>	C <sub>15</sub> =N <sub>16</sub>
bR resting state <sup>24</sup>	175.1	178.9	-156.9	-162.7
O Model	C <sub>7</sub> =C <sub>8</sub>	C <sub>11</sub> =C <sub>12</sub>	C <sub>13</sub> =C <sub>14</sub>	C <sub>15</sub> =N <sub>16</sub>
Model-A <sup>1w</sup>	176.5	-172.4	-154.6	-155.7
Model-B <sup>1w</sup>	172.9	-174.4	-168.2	-170.8
Model-A <sup>2w</sup>	176.2	-171.2	-150.0	-162.9
Model-B <sup>2w</sup>	173.3	173.8	172.3	178.0
Model-A <sup>3w</sup>	174.8	-170.9	-151.3	-165.6
Model-B <sup>3w</sup>	173.2	-175.4	-158.0	-163.8
Model-C <sup>3w</sup>	170.8	-174.9	-153.4	-162.3