

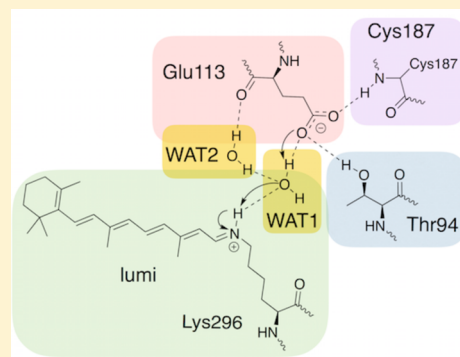
How Rhodopsin Tunes the Equilibrium between Protonated and Deprotonated Forms of the Retinal Chromophore

Siri C. van Keulen,¹ Alicia Solano, and Ursula Rothlisberger*

Institut des Sciences et Ingénierie Chimiques, École Polytechnique Fédérale de Lausanne (EPFL), CH-1015 Lausanne, Switzerland

S Supporting Information

ABSTRACT: Rhodopsin is a photoactive G-protein-coupled receptor (GPCR) that converts dim light into a signal for the brain, leading to eyesight. Full activation of this GPCR is achieved after passing through several steps of the protein's photoactivation pathway. Key events of rhodopsin activation are the initial *cis*–*trans* photoisomerization of the covalently bound retinal moiety followed by conformational rearrangements and deprotonation of the chromophore's protonated Schiff base (PSB), which ultimately lead to full activation in the meta II state. PSB deprotonation is crucial for achieving full activation of rhodopsin; however, the specific structural rearrangements that have to take place to induce this pK_a shift are not well understood. Classical molecular dynamics (MD) simulations were employed to identify intermediate states after the *cis*–*trans* isomerization of rhodopsin's retinal moiety. In order to select the intermediate state in which PSB deprotonation is experimentally known to occur, the validity of the intermediate configurations was checked through an evaluation of the optical properties in comparison with experiment. Subsequently, the selected state was used to investigate the molecular factors that enable PSB deprotonation at body temperature to obtain a better understanding of the difference between the protonated and the deprotonated state of the chromophore. To this end, the deprotonation reaction has been investigated by applying QM/MM MD simulations in combination with thermodynamic integration. The study shows that, compared to the inactive 11-*cis*-retinal case, *trans*-retinal rhodopsin is able to undergo PSB deprotonation due to a change in the conformation of the retinal and a consequent alteration in the hydrogen-bond (HB) network in which PSB and the counterion Glu113 are embedded. Besides the retinal moiety and Glu113, also two water molecules as well as Thr94 and Gly90 that are related to congenital night blindness are part of this essential HB network.



1. INTRODUCTION

In vertebrates' eyes, rod cells are activated when only dim light is available. The protein that is able to convert these photons into a signal for the brain is rhodopsin. Rhodopsin is a protein with seven trans-membrane domains (TMs) and is part of class A in the G-protein-coupled receptor (GPCR) family. For several decades rhodopsin has been intensely studied as a prototype system to understand the activation process of GPCRs. Besides being viewed as an example protein for other GPCRs in class A, rhodopsin bears also other properties of interest, such as low basal activity and a high yield of around 65% for the photoconversion of its inactive 11-*cis*-retinal state (Figure 1) to its *trans* configuration.⁴ However, rhodopsin's activation pathway can be perturbed via mutations in the active site, such as Thr94Ile and Gly90Asp. These mutations lead to a constitutively active form of the GPCR, causing congenital stationary night blindness (CSNB).^{5–7}

The photoactivation of rhodopsin has been studied extensively over the years,^{8–10} and several intermediates with distinct optical properties have been identified. In addition, X-ray data collected at low temperatures have provided the 3D structures of some of these intermediate configurations such as the batho¹¹ and lumi¹² forms. However, different experimental

studies have not always provided fully consolidated views about the details of the activation process, and data collected at low temperatures seem to suggest a different activation pathway than room temperature measurements. For example, depending on the temperature of the system, the deprotonation of the protonated Schiff base (PSB) is shifted to an earlier (metarhodopsin I) or later step (metarhodopsin II) in the photoactivation pathway (Figure 1b).^{1,2,13} Besides the change in PSB deprotonation, the blue-shifted intermediate (BSI) is absent at low temperatures, while at room temperature, BSI is in equilibrium with the batho state (Figure 1b).^{3,14} In view of the fact that in nature rhodopsin functions at body temperature and the observed temperature dependence of the activation path, the actual intermediates under physiological conditions might differ from the configurations found via X-ray crystallography on cold samples.

At body temperature as well as at low temperatures, rhodopsin's activation pathway starts with inactive, dark state, rhodopsin in which Glu113 forms a salt bridge with the protonated Schiff base. After photon exposure, relaxation of the

Received: March 4, 2017

Published: July 21, 2017

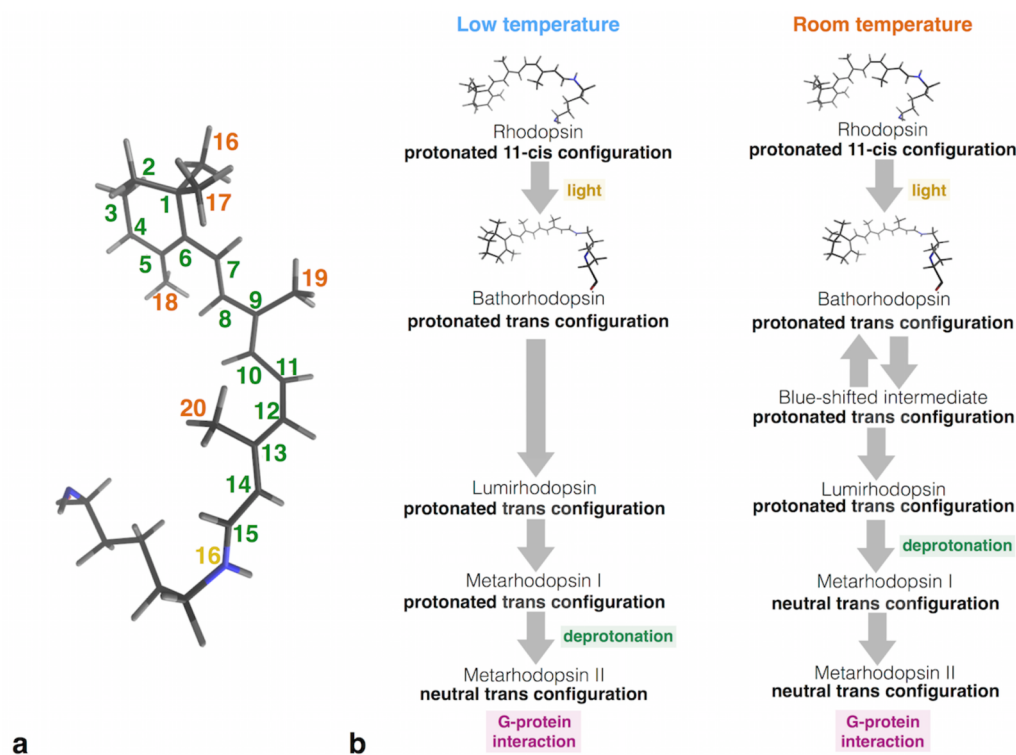


Figure 1. Rhodopsin's photoactivation pathway. (a) Structure and numbering scheme of the retinal moiety in the 11-*cis* configuration covalently linked to the protonated Schiff base (Lys296). The number of the methyl groups is depicted in orange, while the number of the nitrogen is shown in yellow. The numbers of the residual carbons of the retinal moiety are represented in green. (b) Schematic representation of the photoactivation pathway of rhodopsin at low (between 10 K and 273 K) and room temperature.^{1–3}

retinal moiety and deprotonation of the chromophore, the salt bridge is broken and remains broken in the long-lived later intermediates such as meta II (Figure 1b).¹⁵ During the deprotonation step, a proton is transferred from the protonated Schiff base, formed by the retinal and the covalently bound Lys296, to Glu113, the counterion.^{16,17} What causes this pK_a shift is not clear, but it is known that PSB's pK_a , which is estimated to be greater than 15 in dark state bovine rhodopsin,^{18,19} can be decreased to 6 by mutating Glu113 to a glutamine residue,²⁰ which is close to PSB's pK_a of 7.5 ± 0.1 in solution.²¹ The effect of the Glu113Gln mutation can thus trigger deprotonation of the Schiff base and shows that the change from a charged residue to a polar one near the PSB can drastically affect the chromophore's protonation state.^{18,21,22}

As temperature dependence plays a significant role, the available low-temperature X-ray structures of the early intermediates might not be a good starting point for a molecular dynamics study of the deprotonation reaction.^{11,12} A feature of the active site that can be affected by a temperature increase, for example, is the sampling of different type of configurations during the relaxation of the *trans* retinal configuration.^{3,14} The positioning and the number of water molecules is a second important factor that can be altered by the temperature of the system as well as by the crystallization process.

Therefore, in order to investigate PSB deprotonation at body temperature, a dark state X-ray structure of rhodopsin was used to build an initial homodimer structure to minimize temperature and crystallization effects on the activation mechanism of the protein. This is because dark state rhodopsin is not a metastable state like the short-lived intermediates and can thus be equilibrated at body temperature over longer periods of time

than configurations that are sampled just after *cis*–*trans* isomerization. A homodimeric form of rhodopsin was employed in the classical MD simulations in order to mimic the native environment of rhodopsin in the retina.²³ Besides in retina, homodimerization has also been observed for rhodopsin:G-protein complexes for which a stoichiometry of 2:1 could be determined.^{24,25} After classical MD simulations, QM/MM MD simulations were used to follow the relaxation process during the early intermediates (batho, BSI, lumi) as well as to investigate the deprotonation reaction in the lumi state via thermodynamic integration (TI) along a suitably chosen reaction coordinate. The validity of the configurations of the early intermediates was checked through the comparison of their calculated vertical excitation energies with the corresponding experimental absorption maxima. This comparison showed that the found configurations were in good agreement with experiment.

To anticipate our results, the QM/MM MD at body temperature shows that in the early intermediates the relaxation of the retinal conformation mainly takes place close to the PSB where space in the active site is present for methyl relocation. The comparison of PSB deprotonation in the dark and the lumi state suggests that besides the orientation of the retinal moiety also the hydrogen-bond network around the PSB is important for the feasibility of the reaction. At body temperature, water molecules in the lumi state appear to impact this HB network severely by changing the salt bridge between PSB and Glu113 to a bridged water form thereby decreasing Glu113's hydrogen-bond accepting interactions from four to three. After the deprotonation reaction, Glu113's hydrogen bonds are even decreased to two, for one of which Glu113 acts as donor and in the other as acceptor. This investigation suggests that the

retinal orientation within the active-site pocket as well as the rearrangement of the HB network that takes place during the relaxation of the retinal moiety are important to favor PSB deprotonation. The simulation results also show the role Gly90 and Thr94 play in the deprotonation mechanism of the chromophore, which could provide an explanation for their involvement in CSNB after mutation.

2. METHODS

2.1. Classical Molecular Dynamics. The periodically repeated simulation box has a size of approximately $120 \times 115 \times 126 \text{ \AA}^3$, which includes around 40 000 water molecules, 400 lipids, and four Na^+ ions to neutralize the rhodopsin dimer system (Figure S1). The lipids are 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphoethanolamines (POPEs) for which the force field parameters of Jämbek et al. were used.²⁶ The rhodopsin homodimer system was simulated at 310 K and at a pressure of 1 bar using a Nosé–Hoover thermostat and a semiisotropic Parrinello–Rahman barostat, respectively. The initial structure of the rhodopsin monomer was taken from the protein data bank (PDB),²⁷ PDB code 1U19,⁸ while the missing residues 214–246 were taken from PDB code 1GZM.⁹ The intradimeric TM4/TM5 interface was used for the homodimer, extracted from PDB code 1N3M.²³ Default protonation states were used for the residues in the monomers, except for Asp83 and Glu122 that were protonated.²⁸ A disulfide bond was formed between Cys110 and Cys187. Two palmitate groups, for which force field parameters were taken from the literature,^{29,30} are bound to each monomer via Cys322 and Cys323. The force field parameters for the retinal and Lys296 used in this study were obtained by Saam et al.³¹

The force fields used for the rest of the protein dimer and the water molecules are AMBER99SB³² and TIP3P,³³ respectively. All MD runs were performed using GROMACS 4.5.5.³⁴ Electrostatic interactions were calculated with the Ewald's particle mesh method with a real space cutoff of 12 Å. Bonds involving hydrogen atoms were constrained using the LINCS algorithm. The integration time step was set to 2 fs.

The production run of the rhodopsin dimer system in the dark state was performed for 700 ns after 10 ns of equilibration. Subsequently, the classical MD simulation of the dark state dimer was used to initiate *cis*–*trans* isomerization of a single retinal. The use of classical MD simulations to perform the restrained conversion from *cis* to *trans* was previously validated by Röhrig et al.^{35,36} and shows that excited state calculations lead to a very similar configuration of the isomerized retinal as restrained classical MD simulations. The *cis*–*trans* isomerization was performed in one monomer by putting restraints on the C10–C11–C12–C13 and C11–C12–C13–C14 dihedrals to secure the desired *trans* configuration (Figure 1a). After the isomerization has taken place in around 200 fs, the restraints were released, and snapshots were taken that were used as initial structures for the QM/MM simulations.

2.2. Quantum Mechanics/Molecular Mechanics Molecular Dynamics. The QM part of the QM/MM MD simulations for the spectra calculations included the retinal moiety, Lys296, Glu113, and two nearby water molecules (Figure S2). The CPMD 3.15 package³⁷ was employed to perform the QM/MM MD simulations with Car–Parrinello molecular dynamics (CPMD) using the BLYP functional and a plane wave energy cutoff for the wave function of 70 Ry. Martins–Troullier pseudopotentials were used for the QM atoms. In order to cut the QM residues from the rest of the

protein, monovalent pseudopotentials³⁸ were employed. QM/MM simulations were performed at an average temperature of 310 K using a Nosé–Hoover thermostat. A time step of 0.12 fs and a fictitious electronic mass of the CP Lagrangian of 700 au were used. The AMBER force field was used to account for interactions among the atoms of the MM region. In case of the simulations performed for thermodynamic integration, the QM region was extended, compared to the QM/MM MD for the spectra calculations, by the incorporation of the backbone of Glu113, Thr94, and the backbone of Cys187 (Figure S2). These residues were added to the QM region as Thr94 and Cys187 are hydrogen bonded to Glu113. The backbone of Glu113 was included because it interacts with the water molecules in the QM region (Figure 4c).

In the TI simulations, the constrained distance in the dark state was defined as the distance between the proton of the protonated Schiff base and the oxygen of Glu113 that forms a salt bridge with the proton (Figure 4a). In the lumi state the constrained distance was described by the distance between a hydrogen of a water molecule that bridges the PSB and Glu113 and the oxygen of Glu113 that forms a hydrogen bond with the bridging water molecule (Figure 4c). The interval in the constraint distance between each window was 0.1 Å for the dark and lumi state except for three windows, 1, 1.05, and 1.15 Å, in which an interval of 0.05 Å was applied. Each constraint window for both dark and lumi states was sampled for a minimum length of 2.4 ps, unless stated otherwise in the Supporting Information.

2.3. pK_a Calculations. pK_a values for Glu113, Asp83, and Glu122 were calculated via the H++ Web server.^{39–41} The input file for the pK_a calculations contained a dark, lumi, or premeta-I configuration of the rhodopsin monomer that was isomerized and included water molecules within a 7 Å radius around the retinal moiety of the rhodopsin monomer. All aspartates and glutamates were not protonated. The retinal moiety was not included in the pK_a calculations as it could not be taken into account by the H++ server. In the performed calculations, the salinity was kept at 0.15 M, and the external dielectric was set to 80. Values for the interior dielectric were varied from 4, 6, 8, 10, to 20.

2.4. Spectra Calculations. The ZINDO/S method implemented in Gaussian 09⁴² was used to calculate the vertical excitation energies for the dark state rhodopsin, bathorhodopsin, BSI, lumirhodopsin, and premetarhodopsin I configurations. One hundred configurations were used per intermediate with a time interval of 12 fs, except for BSI which included 44 configurations.

Previously obtained vertical excitation energies via ZINDO/S of retinal and related polyenals have been shown to be in line with experiments.⁴³ The four lowest roots were taken into account together with their corresponding oscillator strengths in order to obtain the absorption spectra. The QM cluster model that was used to calculate the vertical excitation energies included the retinal moiety and residues within a distance of 7 Å from the retinal group. The sufficiency of a 5 Å radius has been previously determined by Campomanes et al.⁴⁴ In addition, the backbone of each residue was removed in the model, and the position of the $C\alpha$ carbons was substituted for a hydrogen atom to saturate the system.

2.5. Correlation Feature Selection and Causality Inference Analysis. Similar to our previous study,⁴⁴ a machine learning analysis was performed to identify the relevant features influencing shifts in the optical properties of

the early intermediates. In particular, a two-step protocol was employed to identify the set of electronic and geometrical features responsible for the shifts in the batho, lumi, and premeta-I absorption spectra with respect to the dark conformation. To perform an unbiased analysis, the initial set of features was composed of the bond length alternation (BLA)⁴⁵ as well as all intramolecular and intermolecular distances, angles, and dihedrals in and around the retinal moiety with a distance cutoff of 7 Å. The BLA is defined as the difference between the sum of all the single-bond lengths minus the sum of all the double-bond lengths along the chromophore's conjugated polyene chain, from C5 to the N16 atom of the protonated Schiff base (Figure 1a). Within the selected region all atoms were taken into account except for the hydrogen atoms and the protein backbone. For each photo-intermediate, the 100 snapshots used in the ZINDO/S calculations were included and resulted in an initial set of over 48×10^6 features. To identify which ones of these millions of features had significant impact on the optical properties, the dimensionality of the initial set of descriptors has to be reduced to a smaller subset of features, which is still indispensable and sufficient to describe the target parameter, the spectral shift. The spectral shift is defined as the maximum of the calculated vertical absorption spectra of the excitation energy of the intermediate in each snapshot minus the one of the dark state (ZINDO value 2.39 eV) (Table 1). Once the dimensionality problem is sufficiently reduced, causality between the features in the reduced set and the spectral shift can be inferred.

Table 1. Absorption Spectra Obtained via ZINDO/S for Dark State Rhodopsin, Bathorhodopsin, BSI, Lumirhodopsin, and Premetarhodopsin I

intermediates	absorption maxima (eV)		
	ZINDO/S ^a	exp ^b	difference
dark state rhodopsin	2.39 ± 0.15	2.48	0.09
bathorhodopsin	2.17 ± 0.16	2.34	0.17
blue-shifted intermediate	2.46 ± 0.12	2.60	0.14
lumirhodopsin	2.34 ± 0.15	2.54	0.20
premetarhodopsin I	2.94 ± 0.19	n/a	n/a
metarhodopsin I	n/a	3.29	n/a

^aOne hundred configurations were used with a time interval of 12 fs to calculate the absorption spectra for dark, batho, lumi, and premetarhodopsin I. Forty-four frames with a time interval of 12 fs were isolated for BSI. In the vertical-excitation energy calculations residues and water molecules closer than 7 Å to the retinal moiety were taken into account. Premetarhodopsin I represents the protein structure after deprotonation of the lumi state via QM/MM MD with TI. ^bThe experimental absorption spectra were obtained at room temperature by Lewis et al.²

Hence, the first step toward performing the correlation and causality analysis includes reducing the number of features by removing features that are irrelevant, redundant, and noisy with respect to the target property, the spectral shift. Two feature selection algorithms implemented in the WEKA 3.8 machine learning package⁴⁶ were employed to perform this task. Previously, WEKA has already been used successfully for an analysis of the spectral features of rhodopsin⁴⁴ to detect common factors that influence the spectral shifts of all intermediates up to the lumi state. Here a similar analysis has been applied to the individual shifts for each intermediate including also the premeta-I configuration. The Regression

ReliefF (RReliefF) attribute evaluator algorithm⁴⁷ together with a ranker search method was used first to decrease the dimensionality of the initial set of features to a reduced set. Subsequently, the Correlation Based Feature Selection (CBFS) algorithm,⁴⁸ using the BestFirst search method, was employed to perform a correlation analysis on the reduced feature set to select features that are highly correlated with the spectral shift. CBFS is a filter algorithm that evaluates the heuristic merit of a subset of features considering the correlation to the target and the correlation within features. Therefore, the selected subset, showing the highest merit, contains highly correlated features with the spectral shift, which are minimally correlated with each other. The heuristic merit is defined as

$$M_s = \frac{k \langle r_{ft} \rangle}{\sqrt{k + k(k-1) \langle r_{ff} \rangle}} \quad (1)$$

where M_s is the heuristic merit of a subset s , containing k features, $\langle r_{ft} \rangle$ is the mean feature-target correlation, and $\langle r_{ff} \rangle$ is the mean feature-feature intercorrelation. Hence, the merit numerator provides a measure of the predictive ability of the subset of features to the target, whereas the denominator gives an indication of the redundancy among the features.

After applying the CBFS algorithm on the reduced set of features, the PC-LiNGAM algorithm,⁴⁹ implemented in the TETRAD V causal inference package,⁵⁰ was used to deduce a causal relationship between the spectral shift and the features included in the set provided by the results of the CBFS algorithm. By combining two different approaches, the PC (Peter-Clark) algorithm⁵¹ and the LiNGAM (Linear Non-Gaussian Acyclic Model) algorithm,⁵² the PC-LiNGAM method accepts an arbitrary distribution, meaning that Gaussian and non-Gaussian data sets can be used to obtain the structural equation model (SEM). The SEM represents the causal structure of the model in the form of linear equations and is graphically modeled by a direct acyclic graph (DAG).⁵³ The inference of this approach is based on conditional independences between the variables. Conditional independences between variables were estimated via the Fisher's z -transform, using a significance level, α , of 0.001. Moreover, the Markov blanket of the spectral shift was built, and the values for the edge coefficients in the SEM were estimated using a regression optimizer after standardization (zero mean and unit variance) of the data set variables.

2.6. Structure Superpositions and Images. Multiprot⁵⁴ and VMD⁵⁵ were used to align protein structures. Images were prepared with VMD.⁵⁵

3. RESULTS AND DISCUSSION

3.1. Determination of Intermediates. In order to obtain a configuration in which PSB deprotonation can take place, a dark state rhodopsin dimer system was first simulated for 700 ns at body temperature to generate an initial rhodopsin conformation in which *cis-trans* isomerization can be performed. Subsequently, isomerization was induced in a single monomer's retinal group, followed by unrestrained relaxation of the system for 1 μ s. Snapshots were taken from the unrestrained classical MD simulations before and after *cis-trans* isomerization based on the orientation of C20 in the active site with respect to the dark state conformation (Figure 1a). Next, these structures were used in unconstrained QM/MM MD simulations to identify the dark state as well as the short-lived intermediates batho and BSI via their structural

properties as well as the calculation of absorption spectra (Table 1). The BSI configuration was more difficult to extract from the QM/MM MD simulations compared to the batho and the dark state as the configuration is in equilibrium with bathorhodopsin (Figure S5).^{3,14} The spectral and structural analyses also lead to the identification of lumirhodopsin, the intermediate state that occurs at room temperature before PSB deprotonation in rhodopsin's activation pathway (Figure 1).¹ The calculated excitation energies of dark, batho, BSI, and lumirhodopsin closely reproduce the experimental data, which suggests that the obtained structures are good representatives of the experimentally observed intermediates at body temperature (Table 1). Deprotonation of the PSB in the identified lumirhodopsin configuration was obtained after applying TI to the protein system (see Section 3.2), which resulted in the formation of premetarhodopsin I, a state in which a neutral but still strained *trans*-retinal moiety is present in rhodopsin's active site, while the majority of the protein has not yet adapted its conformation to the new charge distribution in the active site pocket.

Besides the calculated excitation energies, the QM/MM structures of the simulated batho and lumi intermediates are also in good agreement with the (low temperature) X-ray structures of these states (Figures S3, S4). The obtained bathorhodopsin conformation is in good agreement with previous computational studies as well.^{44,56} Remarkable about the alignment of the lumi state (Figure S4) is that, while the overlap between the lumi structures is not as good as for the batho conformation (Figure S3), space for a bridging water molecule, WAT1, is present in the X-ray structure, although the water molecule itself is absent in the experimental data in contrast to the QM/MM results. The absence of the water molecule could be due to the crystallization techniques that were used to obtain the lumi form.

The QM/MM results of dark, batho, BSI, and lumi also show an interesting trend in the relaxation of retinal's conformation. After *cis*–*trans* isomerization, the methyl group on C13 (C20) reorients from a conformation that is more or less parallel to the membrane in the dark state to one that points toward the extracellular region in the lumi state (Figure 2). The relaxation of the *trans* configuration via C20 is not surprising as the volume for the retinal moiety in the mainly hydrophobic/nonpolar active site is tightly fitted around the β -ionone ring

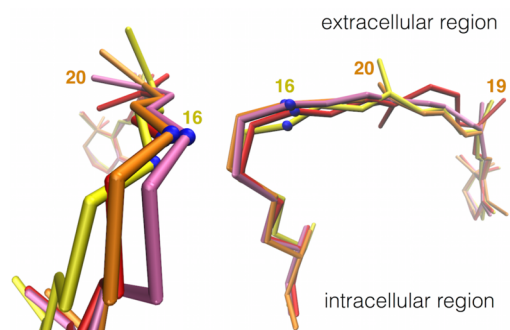


Figure 2. Overlay of rhodopsin structures in the dark (red) and three short-lived intermediate states: bathorhodopsin (pink), BSI (orange), lumirhodopsin (yellow). These structures were obtained via QM/MM MD and aligned on the retinal moiety and Lys296. The carbons of the methyl groups connected to the conjugated chain are depicted by numbers 19 and 20. PSB's nitrogen is labeled as well as highlighted through a blue ball representation.

and the methyl group on C9 (C19) (Figure 3). The only region in the active site that shows a relatively looser fit is the region around PSB, close to C20 (Figure 3).

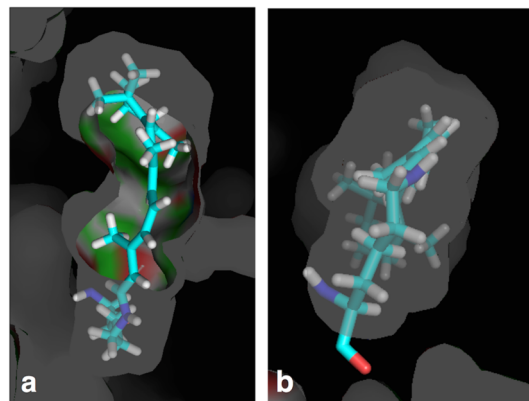


Figure 3. Volume of rhodopsin's active site showing the 11-*cis* configuration of the retinal moiety. (a) View of the retinal moiety from the extracellular region. (b) View of the retinal moiety, turned 90° compared to image (a), with the Lys296 part located in front.

Together with the change in C20's location, the protonated nitrogen of Lys296 also alters its orientation in the active site due to retinal's change in conformation (Figure 2). From the dark to the lumi state, the nitrogen moves from a position in which the atom points in a direction parallel to the membrane to a position in which it points more to the intracellular region. In addition to alterations in PSB's orientation, a shift in the hydrogen-bond network around PSB occurs, which takes place between the BSI and the lumi state. Consequently, batho and BSI contain a similar HB arrangement as the dark state (Figure 4). The change in HB network in the lumi state implies the breakage of the direct salt bridge between PSB and Glu113 and the formation of an indirect hydrogen bond between Glu113 and PSB which is bridged by a water molecule (WAT1) (Figure 4). The alteration in the hydrogen-bond network also contains a relocation of WAT2 and a decrease of hydrogen bonds from three to two for Glu113's O1 (Figure 4). The changes in the hydrogen-bond network of premeta-I in comparison to lumirhodopsin are shown in Figure 7 and are addressed in detail in Section 3.2.

Although some structural rearrangements have already taken place in the premeta-I conformation and the optical properties have undergone a significant shift toward the meta-I form, the chromophore and the protein surrounding are not yet fully relaxed (Table 1). While retinal's carbon chain has straightened compared to the lumi structure on the picosecond time scale, the β -ionone ring is still in the same location in the deprotonated *trans* conformation as prior to deprotonation (Figure 5). This location of the β -ionone ring is not in line with experimental studies that show that the ring alters its location slightly in the meta-I state due to a fully straightened conjugated chain compared to retinal's conformation in the early intermediates.⁵⁷

The conformational changes of the active site across the different intermediates are clearly evident (Figure 4). However, the interplay between the structural changes and the accompanying changes in the optical properties is not obvious. To identify the factors that are responsible for the spectral shifts of batho, lumi, and premeta-I, a correlation and causality

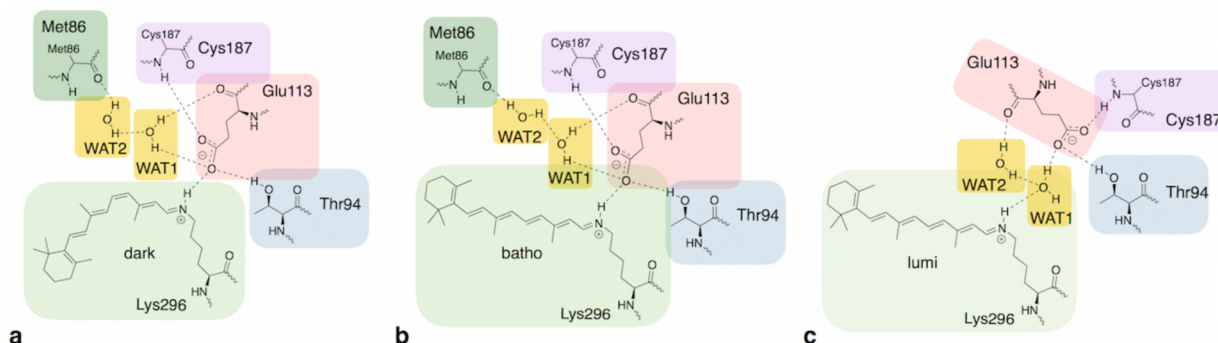


Figure 4. Difference in the hydrogen-bond network around the protonated Schiff's base between the dark (a), batho (b), and lumi state (c). The hydrogen-bond network in the BSI state is the same as in bathorhodopsin.

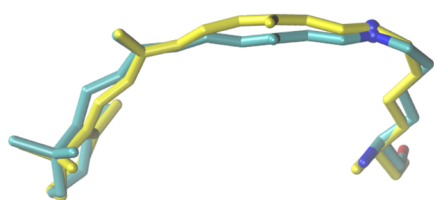


Figure 5. Overlay of the chromophore structures in the lumi state (yellow) and premeta-I configuration (cyan) viewed from the extracellular region. These snapshots were extracted from unconstrained runs obtained before TI and as a result of TI.

inference analysis has been performed, similar to the study of Campomanes et al.⁴⁴ This former study investigated the common factors that determine all shifts for the early

intermediates up to lumirhodopsin by performing a single causality inference analysis over all snapshots. Here, factors that influence individual shifts of the intermediates before deprotonation are studied as well as the premeta-I form. The BSI state was not taken into account because this state is in equilibrium with the batho state (Figure S5), and therefore only relatively few configurations can be unambiguously assigned to this intermediate.

The input for the correlation analysis of batho, lumi, and premeta-I included 48×10^6 geometrical features per intermediate, which were composed of distances, angles, dihedrals, and the BLA. The correlation profiles show that the 48×10^6 features can be decreased to roughly 20 or 30 features having the highest heuristic merit by using Correlation-Based Feature Selection⁴⁸ (Figure S6). Remarkable is the fact

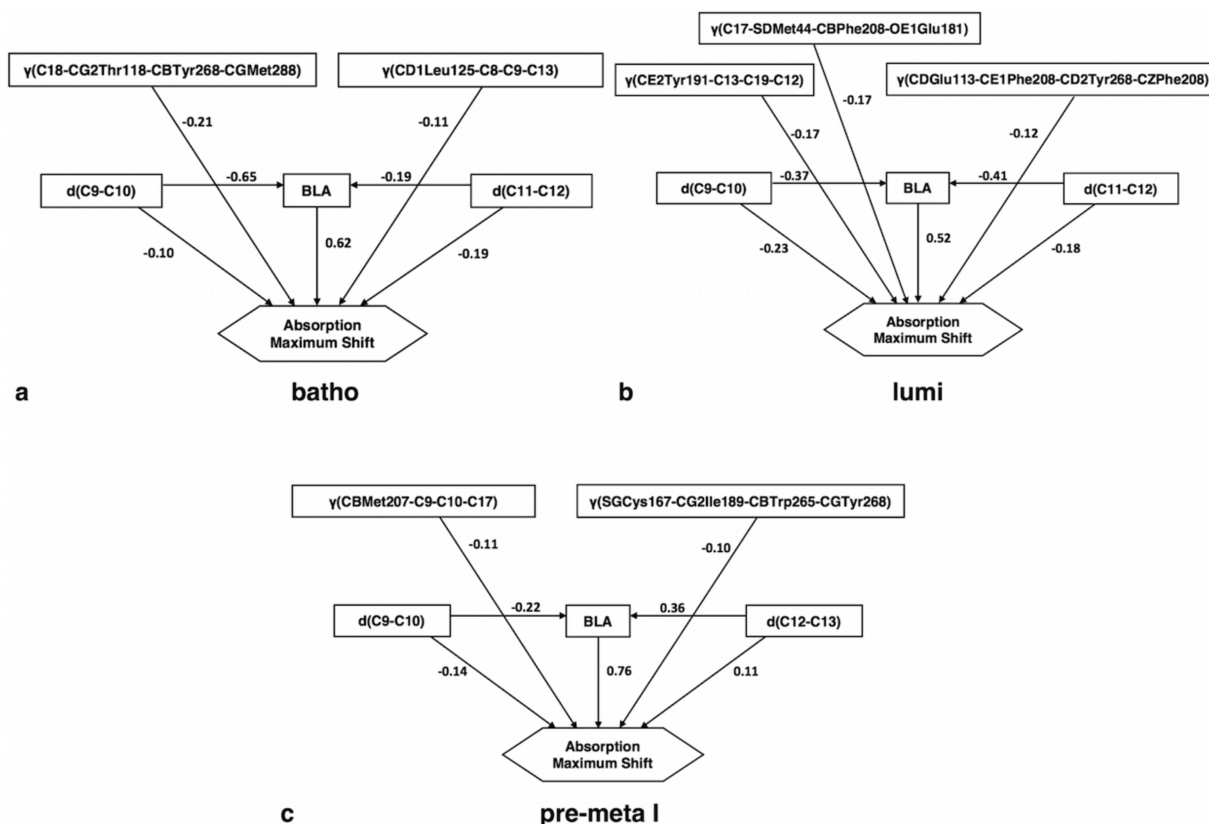


Figure 6. Markov blankets of the spectral shifts for the batho (a), lumi (b), and premeta-I (c) configurations. The edge strength indicates the relative magnitude of the dependency between two variables, given the other inter-relationships.

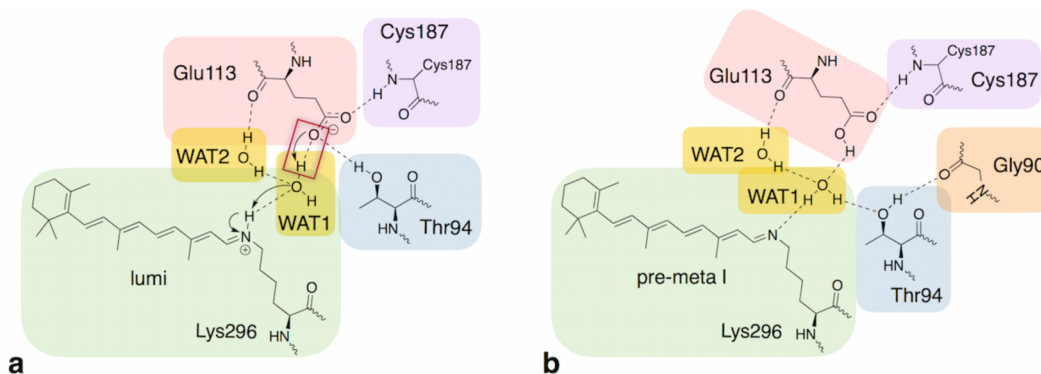


Figure 7. Conformational and configurational changes in the lumi state during and after deprotonation of retinal's positively charged nitrogen. (a) Deprotonation mechanism that takes place during the TI simulations (Figure 8). The red rectangle depicts the distance constraint between H(WAT1) and O1(Glu113) that was used during the TI simulations of the lumi state. (b) Change in the hydrogen-bond network after deprotonation of the protonated Schiff base in the lumi state in the last window constraint distance (1.00 Å) of the TI QM/MM MD simulations and the unconstrained QM/MM MD simulation.

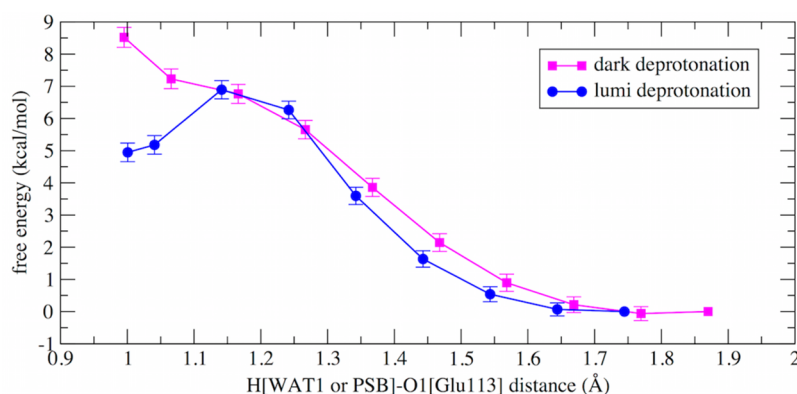


Figure 8. Free energy profile of the deprotonation of the chromophore in rhodopsin in the dark and lumi state. For the thermodynamic integration of the lumi state, the distance between one hydrogen of WAT1 and an oxygen of Glu113, O1, was used to perform the deprotonation reaction (see the red rectangle in image Figure 7a). In the dark state, the distance between the hydrogen on the nitrogen of the PSB, H(PSB), and Glu113's O1 oxygen was used as the reaction coordinate.

that the shifts for batho and lumi are defined by around 30 features, while the premeta-I conformation (Figure 7) only includes 19. The BLA is the number one correlated feature in all configurations (Figure S6). Distance $d(\text{C9}-\text{C10})$ is also present in all top ten features of each intermediate but can vary its position. The correlation of these two descriptors with the spectral shift has also been found in our previous spectral analysis performed over the conjoint data set of intermediates up to the lumi configuration.⁴⁴ Distance $d(\text{C11}-\text{C12})$, the bond around which the *cis-trans* isomerization takes place, appears to be correlated with the spectral shift in batho and lumi, but is replaced in the premeta-I state by $d(\text{C12}-\text{C13})$ (Figure S6). Dihedrals are the most represented in the correlation profiles of the three intermediates, but they are not the same for different intermediates. Residues included in the dihedrals that describe the spectral shifts of batho, lumi, and premeta-I are, for example, Trp265 and Tyr268, which are both located in close proximity to $d(\text{C11}-\text{C12})$. Hence, Tyr268 and Trp265, which have been reported to play an important role in rhodopsin's activation mechanism^{58,59} are sensitive to the location of the retinal in the active site (Figure S6).

Investigating the inference of the causal relationship between features and the spectral shift provides a way to understand which features directly affect the spectral shift and which features only influence other features with a direct impact. The

results for the three intermediates show that all intermediates are highly influenced by the BLA, premeta-I > batho > lumi, as well as $d(\text{C9}-\text{C10})$, lumi > premeta-I > batho (Figure 6 and Figures S6, S7). Similar to the correlation profiles, $d(\text{C11}-\text{C12})$ is substituted in the causality profile by $d(\text{C12}-\text{C13})$ when moving from lumi to premeta-I. The causality between the spectral shift and the dihedrals is remarkable as well because, although specific dihedrals have a similar impact on the spectral shifts for each intermediate as the distances, the dihedral selection per intermediate is unique (Figure 6). This could be due to the fact that the dihedrals are more sensitive to the orientation of the retinal and the conformation of the active site than the BLA and the distances. Hence, according to the Markov blankets of the spectral shifts, all shifts can be described with a handful of parameters, including the BLA, two C-C bonds, and 2–3 dihedral features. However, significant environmental alterations that take place from dark to premetarhodopsin I, such as the change in counterion distance or the rearrangement of water molecules in the active site, do not appear in the causality analysis or the correlation profiles (Figure 6 and Figure S6).

Overall, the causality and correlation results appear to suggest that the descriptors with the most impact on the spectral shifts are intramolecular parameters such as the BLA and distances in the conjugated retinal chain. In addition,

Table 2. pK_a Calculations of Glu113 in Different Rhodopsin Intermediates^a

model	dark			lumi			premeta-I		
	E113	D83	E122	E113	D83	E122	E113	D83	E122
all	4.7	10.9	8.9	4.8	10.7	8.7	8.5	11.2	7.8
no G90	4.5	11.0	8.9	4.8	10.8	8.7	8.1	11.3	8.1
no T94	5.7	10.9	8.9	5.8	10.7	8.7	8.4	11.2	8.0
no WAT1	5.5	10.8	8.9	5.9	10.6	8.7	7.5	11.2	8.2
no WAT1, WAT2	5.6	10.8	8.8	6.4	10.6	8.7	8.0	11.1	8.1
no T94, waters	6.5	10.8	8.9	7.4	10.6	8.6	7.7	11.1	8.2
no G90, T94, waters	6.3	11.0	8.9	7.3	10.8	8.7	7.3	11.2	8.3
shift	1.6	0.1	0.0	2.5	0.1	0.0	-1.2	0.0	0.5

^aThe pK_a values for Asp83 and Glu122 are shown as reference values since these residues are believed to be protonated all throughout rhodopsin's activation.²⁸ pK_a calculations were performed with internal dielectrics of 8, an external dielectric of 80, and a salinity of 0.15 M. The calculated shift is the total difference between the full (all) and the minimal (no Gly90, Thr94, WAT1, WAT2) model system, called *no G90, T94, waters*. Each time a molecule or residue is removed, the charges of the group are set to zero, but the radii are maintained. The full model includes rhodopsin without the chromophore and all water molecules present in the active site.

unique sets of dihedrals that characterize the chromophore's orientation in the active site pocket influence the optical properties of each intermediate.

3.2. Deprotonation Pathway of the Chromophore's Protonated Schiff Base. During rhodopsin's activation pathway, chromophore deprotonation is crucial for conformational changes in the protein that take place before reaching the metarhodopsin II state, a conformation of rhodopsin that is able to activate G proteins on the intracellular side through protein–protein interaction. However, it remains unclear which changes in the active site pocket induce PSB deprotonation after *cis*–*trans* isomerization has occurred.

The equilibrated QM/MM lumi configuration described in Section 3.1 (Figure 4) was used as the starting structure to investigate PSB deprotonation via thermodynamic integration. For the sake of comparison, analogous calculations were performed for the dark state. In both configurations, the distance between Glu113's O1, which is part of the salt bridge with PSB in the dark state (Figure 4), and the hydrogen to which it is hydrogen bonded was used as a reaction coordinate. Hence, in the lumi state the distance between H(WAT1) and O1(Glu113) was used as a constraint, while for the dark state O1(Glu113) and the hydrogen on the nitrogen of the PSB, H(PSB), were used (Figure 4 and Figure 7a). During the thermodynamic integration in the lumi state, the protonation of Glu113 via the proton of WAT1 results in the spontaneous deprotonation of retinal's PSB and reprotonation of WAT1 (Figure 7 and Figures S7, S10, S11), which suggests that the used reaction coordinate is a valid one.

Figure 8 shows that a significant difference can be found in the free energy barrier of the two states (Figure S9). While the lumi state reaches a plateau after deprotonation and its conformation is stable in unconstrained MD simulations following TI (Figure S7), the dark state does not reach a minimum after deprotonation, and the Schiff base is spontaneously reprotonated after the constraint is released (Figure 8 and Figures S8, S12, S13). When the conformational rearrangements upon deprotonation are compared between the dark and lumi state, it becomes apparent that in the dark state, the three hydrogen bonds, present between O1(Glu113) and the active-site environment (Figure 4a), do not change during the TI simulations or in the subsequent unconstrained simulation (Figures S8, S12, S13). In the lumi state, however, the initial number of HBs of Glu113's O1 with the protein environment is two (Figure 4), due to the presence of a

bridging water (WAT1), which is decreased by one in the deprotonated lumi state or premeta-I configuration (Figure 4 and Figure 7). The premeta-I state is a configuration obtained after 6 ps of unconstrained QM/MM MD after deprotonation of the chromophore. The decrease in the number of HBs between the lumi and the premeta-I state is due to the reorientation of Thr94, which is initially interacting with Glu113, but forms a HB with the backbone oxygen of Gly90, Obb(Gly90), in the deprotonated form (Figure 7, Figures S7, S10, S11).

Although the deprotonated retinal is stable in the lumi state (Figure S7), the protonated Schiff base still has a lower free energy compared to the deprotonated form (Figure 8). Therefore, thermodynamically, the deprotonated structure would still appear to be much less populated than the protonated chromophore. However, long-time scale rearrangements that occur upon proton transfer, which are not captured in the picosecond time scale QM/MM MD simulations, could further stabilize the deprotonated form and even make the deprotonation irreversible by increasing the activation barrier for back reaction. Notwithstanding the fact that later rearrangements can lead to additional stabilization of the deprotonated form, the simulations already indicate that the pK_a of the dark state chromophore has already significantly changed in the lumi configurations. In fact, the calculated change in free energy of ca. 4 kcal/mol (Figure 8) corresponds to a lowering of the pK_a by ca. 3 units at body temperature.

To elucidate which molecular factors are responsible for this shift, we performed a semiquantitative analysis of the pK_a shift using pK_a estimates calculated with the H++ server. The effect of the change in electrostatics of the active site upon deprotonation of the chromophore was studied by using the pK_a value of the counterion Glu113. This choice was motivated by the fact that the charge state of Glu113 can highly affect the pK_a of the nearby retinal moiety²⁰ and constitutes therefore a sensitive probe for the changes in the electrostatic environment that favor/accompany the deprotonation reaction.

The pK_a value of Glu113 has been calculated with several internal dielectric constants (Table 2 and Table S2) for several protein conformations and distributions of charge in the active site to provide a qualitative evaluation of the changes in the electrostatic environment when changing from the dark to the premeta-I conformation. In order to investigate the direct effect of the environment on Glu113, the presence/absence of the charge of all moieties that are hydrogen bonded to Glu113's O1

– the oxygen that is protonated during the chromophore's deprotonation – are taken into account (Figure 4 and Figure 7). Also the effect of Gly90 was studied because the backbone carbonyl group of this residue substitutes Thr94's hydrogen-bond donor Glu113 and stabilizes Thr94's change in orientation after deprotonation of the chromophore (Figure 7).

The results in Table 2 suggest that although the pK_a of Glu113 does not change significantly when moving from the dark state to the lumi conformation, the environmental dependence increases. In the dark state, Glu113's pK_a value is lowered by two units by the immediate HB environment (mainly due to Thr94 and WAT1), while lumirhodopsin is lowered by three units (ca. one unit for Thr94, WAT1, and WAT2). Gly90 is not directly hydrogen bonded to Glu113 or a group in Glu113's HB network in the dark and lumi conformation and does not seem to have a significant effect on Glu113's pK_a value. However, in premeta-I, Gly90 appears to have an impact on the pK_a value of Glu113, due to the fact that upon deprotonation of the nitrogen of the PSB the carbonyl oxygen of Gly90 is able to stabilize Thr94's change in orientation via forming a hydrogen bond with the hydroxyl moiety of Thr94, leading to the removal of a hydrogen-bond donor from Glu113's HB network (Figure 7).

While the change in the conformation of the active site from dark to lumirhodopsin is not reflected in the pK_a value of Glu113 in the largest (all) model (Table 2), the nearby environment appears to have a strong effect on tuning the Glu113's pK_a value when moving from the dark to the premeta-I state as the pK_a value is increased by four units in premeta-I compared to the dark form. When removing charge from nearby moieties of Glu113, premeta-I appears to be less affected than the dark state because premeta-I is only increased by ca. one unit, mainly due to WAT1 and Gly90, slightly compensated by WAT2.

Hence, these pK_a calculations appear to suggest that the protein environment has a significant effect on Glu113's protonation state. Moreover, after deprotonation, the HB network around Glu113 seems to cause Glu113's pK_a to increase drastically, leading to values that favor a protonated form of Glu113 (Table 2, Table S2).

4. CONCLUSIONS

The protonation/deprotonation equilibrium of rhodopsin's protonated Schiff base at body temperature is influenced by the conformational change of retinal after *cis-trans* isomerization as well as the induced active site alterations, in particular the presence/absence of water molecules and accompanying rearrangements of the HB network. The first step in the relaxation pathway of the *trans* retinal group is governed by the anchoring of the β -ionone ring in the hydrophobic/nonpolar pocket, the only available additional volume around C20 and the salt bridge between Glu113 and the chromophore. This step includes a clockwise rotation of the C13–C14 bond, which induces the formation of the batho and BSI state, resulting in a C20 oriented toward the extracellular region together with a hydrogen-bond rearrangement near the PSB, leading to the lumi state. This HB rearrangement includes a water molecule that starts disrupting the salt bridge between Glu113 and the PSB in the lumi state as well as a decrease in hydrogen-bond donors for Glu113. Besides the structural changes that are observed upon relaxation of the chromophore after *cis-trans* isomerization, the spectral shifts of batho, BSI, lumi, and premetarhodopsin I are mainly influenced by 5–6 factors: the

BLA, two C–C bonds of retinal's conjugated chain as well as 2–3 specific dihedrals that characterize the orientation of the chromophore with respect to the active site pocket.

Deprotonation in the lumi state results in a stable deprotonated form of the Schiff base, which neutralizes the active site. As an effect of the proton transfer, further changes in the hydrogen-bond network take place that lead to a state in which the proton is stabilized on Glu113. Groups that appear to play a particularly important role in the deprotonation of PSB are two water molecules, WAT1 and WAT2, Gly90 and Thr94. The influence of the conformational rearrangements in the active site that occur during the transition from the dark to the premeta-I state is reflected by pK_a estimates of Glu113.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jctc.7b00229.

Figures and tables (PDF)

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: ursula.roethlisberger@epfl.ch.

ORCID

Siri C. van Keulen: 0000-0001-6995-8389

Author Contributions

S.C.v.K. ran the simulations, prepared the setup for the simulations and alignments, calculated the pK_a values, performed the analysis of the TI trajectories, and wrote the text. A.S. performed the correlation and causality analysis and wrote the methods section for these calculations. U.R. assisted in analysis of the data, supervised the project, and also wrote the text.

Funding

This work was supported by the Swiss National Science Foundation Grant No. 200020_165863, computer time on the DENE SCITAS cluster, and computer resources from CSCS.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors thank Dr. Pablo Campomanes for fruitful discussions.

■ REFERENCES

- (1) Sandberg, M. N.; Greco, J. A.; Wagner, N. L.; Amora, T. L.; Ramos, L. A.; Chen, M.-H.; Knox, B. E.; Birge, R. R. Low-Temperature Trapping of Photointermediates of the Rhodopsin E181Q Mutant. *SOJ. Biochem.* **2014**, *1*, 12.
- (2) Lewis, J. W.; Szundi, I.; Kazmi, M. A.; Sakmar, T. P.; Kligler, D. S. Time-resolved photointermediate changes in rhodopsin glutamic acid 181 mutants. *Biochemistry* **2004**, *43*, 12614–12621.
- (3) Hug, S. J.; Lewis, J. W.; Einterz, C. M.; Thorgeirsson, T. E.; Kligler, D. S. Nanosecond photolysis of rhodopsin: Evidence for a new blue-shifted intermediate. *Biochemistry* **1990**, *29*, 1475–1485.
- (4) Kim, J. E.; Tauber, M. J.; Mathies, R. A. Wavelength dependent *cis-trans* isomerization in vision. *Biochemistry* **2001**, *40*, 13774–13778.
- (5) Singhal, A.; Guo, Y.; Matkovic, M.; Schertler, G.; Deupi, X.; Yan, E. C.; Standfuss, J. Structural role of the T94I rhodopsin mutation in congenital stationary night blindness. *EMBO Rep.* **2016**, *17*, 1431.
- (6) Singhal, A.; Ostermaier, M. K.; Vishnivetskii, S. A.; Panneels, V.; Homan, K. T.; Tesmer, J. J. G.; Vepintsev, D.; Deupi, X.; Gurevich, V. V.; Schertler, G. F. X.; Standfuss, J. Insights into congenital stationary

night blindness based on the structure of G90D rhodopsin. *EMBO Rep.* **2013**, *14*, 520–526.

(7) Zeitz, C.; Gross, A. K.; Leifert, D.; Kloeckener-Gruissem, B.; McAlear, S. D.; Lemke, J.; Neidhardt, J.; Berger, W. Identification and functional characterization of a novel rhodopsin mutation associated with autosomal dominant CSNB. *Invest. Ophthalmol. Visual Sci.* **2008**, *49*, 4105–4114.

(8) Okada, T.; Sugihara, M.; Bondar, A.-N.; Elstner, M.; Entel, P.; Buss, V. The retinal conformation and its environment in rhodopsin in light of a new 2.2 Å crystal structure. *J. Mol. Biol.* **2004**, *342*, 571–583.

(9) Li, J.; Edwards, P. C.; Burghammer, M.; Villa, C.; Schertler, G. F. Structure of bovine rhodopsin in a trigonal crystal form. *J. Mol. Biol.* **2004**, *343*, 1409–1438.

(10) Deupi, X.; Edwards, P.; Singhal, A.; Nickle, B.; Oprian, D.; Schertler, G.; Standfuss, J. Stabilized G protein binding site in the structure of constitutively active metarhodopsin-II. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 119–124.

(11) Nakamichi, H.; Okada, T. Crystallographic analysis of primary visual photochemistry. *Angew. Chem., Int. Ed.* **2006**, *45*, 4270–4273.

(12) Nakamichi, H.; Okada, T. Local peptide movement in the photoreaction intermediate of rhodopsin. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103*, 12729–12734.

(13) Thorgeirsson, T. E.; Lewis, J. W.; Wallace-Williams, S. E.; Kliger, D. S. Effects of temperature on rhodopsin photointermediates from lumirhodopsin to metarhodopsin II. *Biochemistry* **1993**, *32*, 13861–13872.

(14) Lewis, J. W.; Kliger, D. S. Photointermediates of visual pigments. *J. Bioenerg. Biomembr.* **1992**, *24*, 201–210.

(15) Cohen, G. B.; Oprian, D. D.; Robinson, P. R. Mechanism of activation and inactivation of opsin: role of Glu113 and Lys296. *Biochemistry* **1992**, *31*, 12592–12601.

(16) Jaeger, F.; Fahmy, K.; Sakmar, T. P.; Siebert, F. Identification of glutamic acid 113 as the Schiff base proton acceptor in the metarhodopsin II photointermediate of rhodopsin. *Biochemistry* **1994**, *33*, 10878–10882.

(17) Sakmar, T. P.; Franke, R. R.; Khorana, H. G. Glutamic acid-113 serves as the retinylidene Schiff base counterion in bovine rhodopsin. *Proc. Natl. Acad. Sci. U. S. A.* **1989**, *86*, 8309–8313.

(18) Ebrey, T. G. [14] pKa of the protonated schiff base of visual pigments. *Methods Enzymol.* **2000**, *315*, 196–207.

(19) Steinberg, G.; Ottolenghi, M.; Sheves, M. pKa of the protonated Schiff base of bovine rhodopsin. A study with artificial pigments. *Biophys. J.* **1993**, *64*, 1499.

(20) Sakmar, T. P.; Franke, R. R.; Khorana, H. G. The role of the retinylidene Schiff base counterion in rhodopsin in determining wavelength absorbance and Schiff base pKa. *Proc. Natl. Acad. Sci. U. S. A.* **1991**, *88*, 3079–3083.

(21) Gat, Y.; Sheves, M. A mechanism for controlling the pKa of the retinal protonated Schiff base in retinal proteins. A study with model compounds. *J. Am. Chem. Soc.* **1993**, *115*, 3772–3773.

(22) Zhukovsky, E. A.; Robinson, P. R.; Oprian, D. D. Changing the location of the Schiff base counterion in rhodopsin. *Biochemistry* **1992**, *31*, 10400–10405.

(23) Fotiadis, D.; Jastrzebska, B.; Philippsen, A.; Müller, D. J.; Palczewski, K.; Engel, A. Structure of the rhodopsin dimer: a working model for G-protein-coupled receptors. *Curr. Opin. Struct. Biol.* **2006**, *16*, 252–259.

(24) Jastrzebska, B.; Ringler, P.; Lodowski, D. T.; Moiseenkova-Bell, V.; Golczak, M.; Müller, S. A.; Palczewski, K.; Engel, A. Rhodopsin–transducin heteropentamer: Three-dimensional structure and biochemical characterization. *J. Struct. Biol.* **2011**, *176*, 387–394.

(25) Jastrzebska, B.; Orban, T.; Golczak, M.; Engel, A.; Palczewski, K. Asymmetry of the rhodopsin dimer in complex with transducin. *FASEB J.* **2013**, *27*, 1572–1584.

(26) Jämbeck, J. P.; Lyubartsev, A. P. An extension and further validation of an all-atomistic force field for biological membranes. *J. Chem. Theory Comput.* **2012**, *8*, 2938–2948.

(27) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. The protein data bank. *Nucleic Acids Res.* **2000**, *28*, 235–242.

(28) Fahmy, K.; Jäger, F.; Beck, M.; Zvyaga, T. A.; Sakmar, T. P.; Siebert, F. Protonation states of membrane-embedded carboxylic acid groups in rhodopsin and metarhodopsin II: a Fourier-transform infrared spectroscopy study of site-directed mutants. *Proc. Natl. Acad. Sci. U. S. A.* **1993**, *90*, 10206–10210.

(29) Röhrig, U. F.; Guidoni, L.; Rothlisberger, U. Early steps of the intramolecular signal transduction in rhodopsin explored by molecular dynamics simulations. *Biochemistry* **2002**, *41*, 10799–10809.

(30) Dickson, C. J.; Rosso, L.; Betz, R. M.; Walker, R. C.; Gould, I. R. GAFFlipid: a General Amber Force Field for the accurate molecular dynamics simulation of phospholipid. *Soft Matter* **2012**, *8*, 9617–9627.

(31) Saam, J.; Tajkhorshid, E.; Hayashi, S.; Schulten, K. Molecular dynamics investigation of primary photoinduced events in the activation of rhodopsin. *Biophys. J.* **2002**, *83*, 3097–3112.

(32) Case, D. A.; Berryman, J. T.; Betz, R. M.; Cerutti, D. S.; Cheatham, T. E., III; Darden, T. A.; Duke, R. E.; Giese, T. J.; Gohlke, H.; Goetz, A. W. *AMBER 2015*; University of California: San Francisco, CA, 2015.

(33) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, *79*, 926–935.

(34) Hess, B.; Kutzner, C.; Van Der Spoel, D.; Lindahl, E. GROMACS 4: algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J. Chem. Theory Comput.* **2008**, *4*, 435–447.

(35) Röhrig, U. F.; Guidoni, L.; Laio, A.; Frank, I.; Rothlisberger, U. A molecular spring for vision. *J. Am. Chem. Soc.* **2004**, *126*, 15328–15329.

(36) Röhrig, U. F.; Guidoni, L.; Rothlisberger, U. Solvent and protein effects on the structure and dynamics of the rhodopsin chromophore. *ChemPhysChem* **2005**, *6*, 1836–1847.

(37) CPMD, cpmd package. Copyright IBM Corp 1990–2015, Copyright MPI für Festkörperforschung Stuttgart 1997–2001. <http://www.cpmd.org/> (accessed July 24, 2017).

(38) von Lilienfeld, O. A.; Tavernelli, I.; Rothlisberger, U.; Sebastiani, D. Variational optimization of effective atom centered potentials for molecular properties. *J. Chem. Phys.* **2005**, *122*, 014113.

(39) Gordon, J. C.; Myers, J. B.; Folta, T.; Shoja, V.; Heath, L. S.; Onufriev, A. H++: a server for estimating pKas and adding missing hydrogens to macromolecules. *Nucleic Acids Res.* **2005**, *33*, W368–W371.

(40) Myers, J.; Grothaus, G.; Narayanan, S.; Onufriev, A. A simple clustering algorithm can be accurate enough for use in calculations of pKs in macromolecules. *Proteins: Struct., Funct., Genet.* **2006**, *63*, 928–938.

(41) Anandakrishnan, R.; Aguilar, B.; Onufriev, A. V. H++ 3.0: automating pK prediction and the preparation of biomolecular structures for atomistic molecular modeling and simulations. *Nucleic Acids Res.* **2012**, *40*, W537–W541.

(42) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian09*, Revision D.01; Gaussian Inc.: Wallingford, CT, 2009.

- (43) López, C. S.; Faza, O. N.; Estévez, S. L.; de Lera, A. R. Computation of vertical excitation energies of retinal and analogs: scope and limitations. *J. Comput. Chem.* **2006**, *27*, 116–123.
- (44) Campomanes, P.; Neri, M.; Horta, B. A. C.; Röhrig, U. F.; Vanni, S.; Tavernelli, L.; Rothlisberger, U. Origin of the spectral shifts among the early intermediates of the rhodopsin photocycle. *J. Am. Chem. Soc.* **2014**, *136*, 3842–3851.
- (45) Giesecking, R. L.; Risko, C.; Brédas, J.-L. Distinguishing the Effects of Bond-Length Alternation versus Bond-Order Alternation on the Nonlinear Optical Properties of π -Conjugated Chromophores. *J. Phys. Chem. Lett.* **2015**, *6*, 2158–2162.
- (46) Witten, I. H.; Frank, E.; Hall, M. A.; Pal, C. J. *The WEKA Workbench. Online Appendix for "Data Mining: Practical Machine Learning Tools and Techniques"*, 4th ed.; Morgan Kaufmann: 2016.
- (47) Robnik-Šikonja, M.; Kononenko, I. An adaptation of Relief for attribute estimation in regression. *Machine Learning: Proceedings of the Fourteenth International Conference (ICML97)*. 1997; pp 296–304.
- (48) Hall, M. A. Correlation-based feature selection for machine learning. Ph.D. Thesis, The University of Waikato, 1999.
- (49) Hoyer, P. O.; Hyvarinen, A.; Scheines, R.; Spirtes, P. L.; Ramsey, J.; Lacerda, G.; Shimizu, S. Causal discovery of linear acyclic models with arbitrary distributions. *arXiv preprint arXiv:1206.3260* **2012**.
- (50) Glymour, C.; Scheines, R. Causal modeling with the TETRAD program. *Synthese* **1986**, *68*, 37–63.
- (51) Spirtes, P.; Glymour, C. N.; Scheines, R. *Causation, prediction, and search*; MIT Press: 2000.
- (52) Shimizu, S.; Hoyer, P. O.; Hyvärinen, A.; Kerminen, A. A linear non-Gaussian acyclic model for causal discovery. *J. Mach. Learn. Res.* **2006**, *7*, 2003–2030.
- (53) Lacerda, G.; Spirtes, P. L.; Ramsey, J.; Hoyer, P. O. Discovering cyclic causal models by independent components analysis. *arXiv preprint arXiv:1206.3273* **2012**.
- (54) Shatsky, M.; Nussinov, R.; Wolfson, H. J. A method for simultaneous alignment of multiple protein structures. *Proteins: Struct., Funct., Genet.* **2004**, *56*, 143–156.
- (55) Humphrey, W.; Dalke, A.; Schulten, K. VMD: visual molecular dynamics. *J. Mol. Graphics* **1996**, *14*, 33–38.
- (56) Lemaitre, V.; Yeagle, P.; Watts, A. Molecular dynamics simulations of retinal in rhodopsin: from the dark-adapted state towards lumirhodopsin. *Biochemistry* **2005**, *44*, 12667–12680.
- (57) Struts, A. V.; Salgado, G. F. J.; Martínez-Mayorga, K.; Brown, M. F. Retinal dynamics underlie its switch from inverse agonist to agonist during rhodopsin activation. *Nat. Struct. Mol. Biol.* **2011**, *18*, 392–394.
- (58) Crocker, E.; Eilers, M.; Ahuja, S.; Hornak, V.; Hirshfeld, A.; Sheves, M.; Smith, S. O. Location of Trp265 in metarhodopsin II: implications for the activation mechanism of the visual receptor rhodopsin. *J. Mol. Biol.* **2006**, *357*, 163–172.
- (59) Standfuss, J.; Edwards, P. C.; D'Antona, A.; Fransen, M.; Xie, G.; Oprian, D. D.; Schertler, G. F. X. The structural basis of agonist-induced activation in constitutively active rhodopsin. *Nature* **2011**, *471*, 656–660.