Deconstructing activation events in rhodopsin

Elena N. Laricheva,[†] Karunesh Arora,[†] Jennifer L. Knight,[†] Charles L. Brooks III[*,†,‡]

[†]Department of Chemistry and [‡]Biophysics Program, University of Michigan, Ann Arbor, Michigan, 48109, United States

SUPPORTING INFORMATION

S1. Structure preparation: Rhodopsin was modeled in its putative semi-active (Meta II_a) and fully active (Meta II) forms using the 4.15Å and 3.0Å X-ray structures with PDB IDs 2I37[1] and 3PXO,[2] respectively. In each case, chain A was selected and the protein was prepared using the CMAP-corrected[3] CHARMM22 force field for proteins[4] in a manner similar to that reported by Provasi and Filizola in their recent work.[5] Specifically, loop residues 230–238 and 311–313 missing from the 2I37 crystal structure were generated using MODELLER[6] as implemented in the loopModel.pl script of the MMTSB toolset.[7] Retinal was built into the 2I37 structure by mutating K296 into unprotonated all-*trans* retinal Schiff base using the parameters from Saam and co-workers[8] modified to account for the loss of the proton from the Schiff base. Missing hydrogen atoms were added to the retinal-bound protein using the HBUILD facility in the CHARMM simulation package.[9] Protonation states of all titratable residues, except for those that are involved in the two protonation-dependent switches[10] critical for rhodopsin activation, were assigned according to the computational predictions reported in ref.[11] In particular, histidines H65, H152, H195, and H278 were protonated at the ε-carbon, whereas H100 and H211 were protonated at the δ-carbon. Residues E122, and D83 were protonated, while the protonation states of the functionally important E113, E181, and E134 were determined using constant pH molecular dynamics (CPHMD)[12,13] with pH-based replica exchange (pH-REX)[14] assisted by the parallel distributed replica code (REPDSTR)[15] in CHARMM. Based on the information obtained during the pH-REX simulations, E113 and E181 were subsequently protonated in both Meta II_a and Meta II states, while E134 was kept charged in order to generate two equilibrium configurations between which a pathway was further constructed using the Harmonic Fourier Beads (HFB)[16] method following the procedure outlined in S5. The change of the E134 protonation state at each point of the generated pathway was then assessed using pH-REX, as described in S4.

S2. Choice of the putative Meta II_a structure: Recently solved crystallographic structures of the active state of rhodopsin, including that of the constitutively active E113Q mutant,[17] all-*trans* retinal soaked opsin,[2] and ligand-free opsin at low pH,[18] all display a pronounced rigid body movement of the cytoplasmic end of transmembrane helix H6, tilted by *ca.* 6Å outwards with respect to its position in the dark state. In contrast, crystal structures of the early rhodopsin intermediates bathorhodopsin (Batho)[19] and lumirhodopsin (Lumi),[20] as well as the electron microscopy map of the inactive Meta I state,[21] show only small rearrangement of the α-helical bundle. In this respect, a structure of the photoactivated deprotonated intermediate of rhodopsin (PDI),[1] containing an unprotonated all-*trans* retinal Schiff base (as

experimentally confirmed by the characteristic change of absorption maximum from 480 to 380 nm), but lacking any significant changes in the H6 position and featuring only small alterations in the intracellular loops, can tentatively be assigned to the semi-active Meta II_a state – a conformational substate of the sequential reaction scheme introduced by Hoffman and Hubbell and colleagues[22–24] and later extended by Vogel and co-workers.[25–27] Assuming that both 2I37 and 3PXO crystal structures are constituent parts of the rhodopsin activation pathway, we can generate a conformational transition pathway between them and consider it a reasonable representation of the Meta II_a to Meta II activation event in rhodopsin.

S3. Identification of the optimal position in the membrane: In all simulations, the Generalized Born with a simple Switching (GBSW) implicit membrane model[28] was used within the all-atom representation of rhodopsin. Initial optimal positioning of both Meta II_a and Meta II states inside the membrane was identified by systematically translating each constructed structure along the membrane normal (Z-axis). Default GBSW parameters with 50 angular integration points were used. The membrane thickness was 30Å with membrane switching length set to 5.0 Å. The optimized GBSW intrinsic radii[29,30] and a nonpolar surface tension coefficient γ of 0.005 kcal/mol/Å^2 were used. The center of mass of each protein was positioned at the origin of the Cartesian coordinate system and its principal axis was oriented perpendicular to the plane of the membrane. Subsequently, the structures were briefly minimized with backbone restraints using a soft force constant of 20 kcal/mol/Å^2 and then translocated up and down along Z-axis in 1Å increments. The lowest energy configurations obtained in this way were chosen as the starting points for the pH-REX simulations and generation of the HFB pathway (Figure S1). Each structure was then minimized with both protonated and unprotonated E134 for 10,000 steps of Adopted Basis Newton-Raphson (ABNR) further subjected to a 5 ns molecular dynamics run, as described in S4.

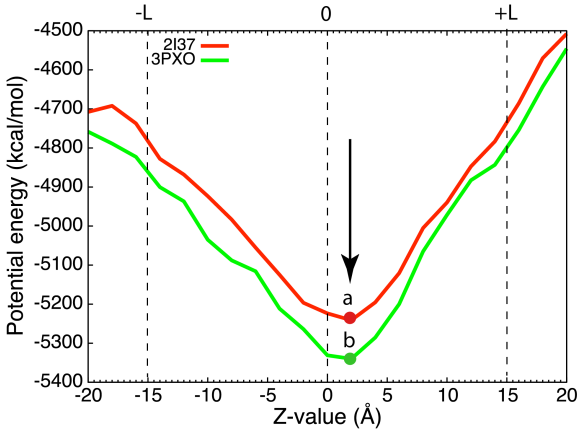

Figure S1. Identification of the optimal positions of the Meta II_a (2i37) and Meta II (3PXO) conformations in the 30Å-thick low-dielectric membrane slab (positions *a* and *b* corresponding to Z-values = 2.0 in each case, respectively). +L and –L indicate extracellular and cytoplasmic membrane planes, respectively.

S4. Molecular dynamics simulations of the end states: To test the stability of the constructed Meta II_a and Meta II structures, 5 ns long MD trajectories were generated using the GBSW implicit membrane representation using CHARMM, version c38a1.[9,31] The leapfrog Verlet algorithm was used to integrate the equations of motion and to propagate atomic coordinates and velocities. A non-bonded cutoff of 20 Å was used and van der Waals and electrostatic switching functions were implemented between 16 Å and 20 Å unless otherwise specified. Hydrogen bonds were restrained using the SHAKE algorithm[32] and the time step was 2.0 fs. During 100 ps heating phase all backbone atoms were harmonically restrained with a mass-weighted force constant of 50 kcal/(mol·Å^2). Four equilibration phases of 50 ps each were performed with force constraints of 50, 25, 10 and 0 kcal/(mol·Å^2). The final production run was 5.0 ns. All simulations were performed on dual 2.66 GHz Intel Quad Core Xeon CPUs. Stability of the 2I37 and 3PXO structures during simulation runs was monitored by the backbone root-mean-square deviation (RMSD) of each protein with respect to the corresponding reference state (Figure S2).

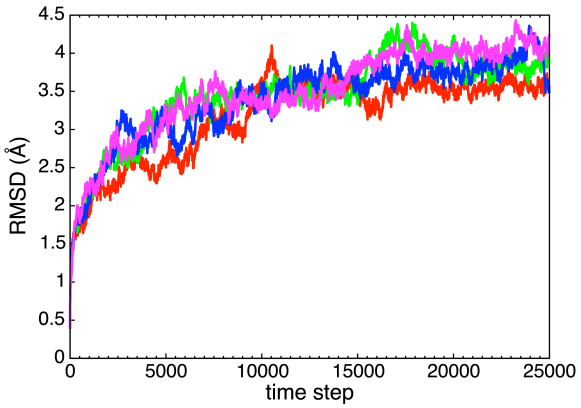

Figure S2. Backbone RMSD of the end states during 5 ns long MD simulations: Meta IIa (2I37) with unprotonated (red) and protonated (green) E134; Meta II (3PXO) with unprotonated (blue) and protonated (magenta) E134, respectively. Each time step is equal 2 fs.

S5. Constant pH molecular dynamics simulations with pH-based replica exchange: The pK_a values of residues E113, E181, and E134, involved in the two protonation dependent switches controlling rhodopsin activation, were determined using constant pH molecular dynamics (CPHMD)[12,13] with pH-based replica exchange (pH-REX)[14] through the parallel distributed replica code (REPDSTR).[15] The simulations were first performed at the 2I37 and 3PXO end states for all three residues subject to titration, and then, taking into account the determined protonation states of E113 and E181, for residue E134 only at every point of the transition pathway constructed between the two states using the HFB method, details of which are described in S6. A particular CPHMD variant used here is based on a continuous λ-dynamics approach[33,34] to free energy calculations, as originally implemented in CHARMM by Lee et al.[13] and later extended by Khandogin and Brooks[12] to account for the effect of proton tautomerism. This method allows propagating both spatial and titration coordinates simultaneously, with the latter de-

fined as a dynamic λ-variable with fictitious mass that continuously changes within the [0,1] interval between the protonated (λ=0) and unprotonated (λ=1) states at a specified external pH. It is worth noting that the residual charge of the system, arising during the change in protonation state, does not cause any artifacts in the simulations due to the instantaneous dielectric response of the solvent defined by the underlying GBSW implicit solvent model. In such simulations, acid-based and proton tautomeric equilibria of His, Glu and Asp residues are usually accounted for by an additional parameter *x*. In particular, two λ–driven processes are initiated at one end point (λ=0 or λ=1) and further coupled to one another via a parameter *x* at the other end point (λ=1 or λ=0). In our case, to improve a sampling in λ-space and achieve better convergence of the protonation states, pH-REX was introduced. Unlike parallel tempering, in pH-REX the simulations of different system replicas are run at the same temperature but different pH conditions with an exchange of those between a pair of adjacent replicas according to the Metropolis criterion. We exchanged adjacent replicas every 2 ps, with the acceptance ratio, monitored by the number of accepted versus total Hamiltonian exchanges, in the 40-60% range. The fractional population of the unprotonated states (S^U) was obtained for 16 pH windows (from 0 to 15) during 1.5ns long simulations and is defined as follows:

$$S^U \approx \frac{N^U}{N^U + N^P}$$

where N^U and N^P are populations of the unprotonated and protonated states, respectively, and are defined in our analysis using the following λ and *x* cutoffs:

$$N^U = N(\lambda > 0.9; x < 0.1 \; or \; x > 0.9)$$

$$N^P = N(\lambda < 0.1; x < 0.1 \; or \; x > 0.9)$$

The obtained S^U values at different external pH windows were further fitted into the Henderson-Hasselbach equation to obtain pK_a values (*n* is the Hill coefficient):

$$S^U = \frac{1}{1 + 10^{n(pKa - pH)}}$$

Convergence of the pH-REX simulations was monitored by following the dynamics of λ values for each titratable residue (E113, E181, and E134) at the pH value nearest to the corresponding pK_a calculated using a procedure outlined above, at the same time assessing the stability of the cumulative S^U values (Figure S3).

Our results indicate that E113, involved in the first of the two protonation-dependent switches during rhodopsin activation, is protonated in both the Meta II_a and Meta II states, while E181 is charged. At the same time, the protonation state of E134, participating in the second protonation-dependent switch, changes going from one state to the other. Notably, protonation states of E113 and E181 have long been a subject of great debate.[35–38] According to one model, known as the counterion switch,[35] E181 is protonated in the dark state of rhodopsin where E113 serves as a primary counterion stabilizing positive charge on the nitrogen of the 11-cis retinal protonated Schiff base (11-PSB) formed by the chromophore retinal covalently bound to K296. Upon ultrafast light-induced *cis–trans* isomerization of the latter and transition to the Meta I state though a series of inactive intermediates (bathorhodopsin, blue-shifted intermediate, and lumirhodopsin), the proton is transferred from one residue to the other resulting in the breakage of the K296-E113 interactions and formation of the new K296-E181 salt bridge. In contrast, the complex counterion model, introduced by Lüdeke et al.,[36] argues that both residues are initially unprotonated equally stabilizing the PSB in the dark and changing their relative positions and functional roles upon attainment of the Meta I state, with E113 becoming a primary counterion located closer to the retinal. While the counterion switch model is supported by the resonance Raman, FTIR, and UV-visible spectroscopy of the site-directed rhodopsin mutants,[35] a number of computational studies and most recent ^{2}H NMR data suggest the complex counterion model as most plausible.[36,37] Our pH-REX simulations offer and additional support to the latter model, as observed protonated E113 and unprotonated E181 upon attainment of the Meta II_a state are likely to originate from both of the residues being charged in the dark. Accordingly, E113 and E181 protonation states were further fixed as neutral for the construction of the HFB pathway, while E134 was left charged and its protonation state was probed along the computed path using pH-REX protocol as described above.

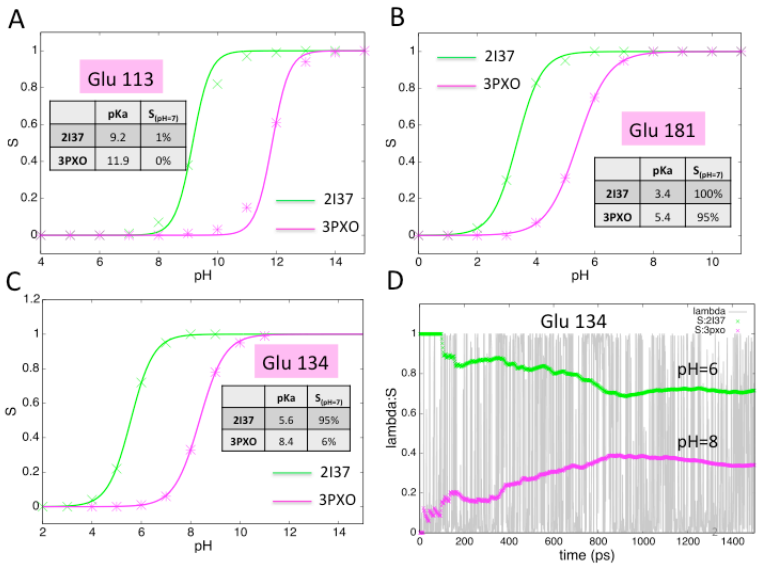

Figure S3. Protonation states of the E113 (A), E181 (B), and E134 (C) residues at the Meta II_a (2I37) and Meta II (3PXO) states, initially constructed, minimized, and briefly equilibrated prior to generation of the HFB transition pathway. Both pK_a values and unprotonated fraction S are reported at pH =7 window. Convergence of λ-values and cumulative S^U fraction at pH closest to the determined pK_a values are shown for E134 at both Meta II_a and Meta II forms (D).

S6. Generation of transition pathway using the Harmonic Fourier Beads method: The Harmonic Fourier bead (HFB) method,[16] implemented in CHARMM, was used to evolve an initial naive pathway between Meta II_a and Meta II states to the lowest energy one. The models of Meta II_a and Meta II were prepared as described in S1. Initial pathways were generated by interpolation between two structures in the reactive coordinate space (RCS). The RCS is the set of atoms that characterize the transition of interest. In our case, the RCS was chosen as the backbone C_α, C, N and O atoms of the entire protein. The initial pathway consisted of 64 structures (beads) including the endpoint Meta II_a and Meta II structures. The initial string was then evolved for 82 cycles with the RCS atoms harmonically restrained using 20 kcal/mol/Å^2 force constant to their current position and then minimized for 1000 steepest descent (SD) steps, followed by 2000 Adopted Basis Newton-Raphson (ABNR) steps. An additional 68 HFB cycles were carried out with reduced (10 kcal/mol/Å^2) harmonic restraints on the backbone atoms. Minimization was performed using the GBSW implicit membrane model and the non-bonded interactions were switched off between 12 and 14 Å. Following the minimization procedure, the string was re-parameterized using 48 Fourier basis functions, in order to keep the beads equally spaced in the RCS. The cycles of minimization and re-parameterization were repeated until there was no significant change in the RMSD of structures along the pathway with respect to their configurations in the initial string (Figure S4).

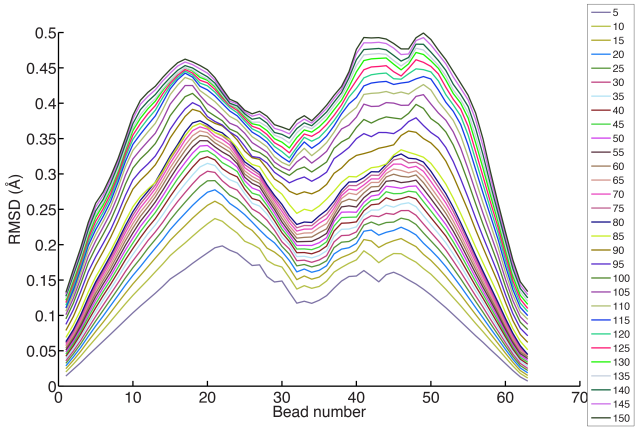

Figure S4. RMSD of each bead with respect to corresponding bead in the initial pathway. Results are shown for 150 HFB iteration cycles with an interval of 5 cycles.

As shown in Figure S4, the RMSD of each bead with respect to corresponding bead in the initial pathway increases rapidly till cycle 40 and then levels off. Then again after cycle 82, where the harmonic restraining force on RCS atoms is halved, there is an increase in the RMSD for few more cycles. There is no significant increase in RMSD beyond cycle 130 of HFB iterations suggesting that the path has converged. Notably, the two peaks in Figure S4 with the biggest RMSD change may correspond to metastable intermediate states undergoing substantial conformational relaxation in search for a local minimum. Additionally, as Figure S5 and S6 illustrate, for each (i+1) cycle of the HFB the RMSD was computed between the minimized configuration and the corresponding configuration in the i^{th} cycle, and then summed over all beads in the string. The alignment and the RMSD calculations were performed over the RCS atoms. Figure S5 shows the RMSD metric for the first 82 cycles of minimization with a 20 kcal/mol/Å^2 force constant applied to harmonically restrain the RCS atoms, while Figure S6 illustrates the other 68 cycles after the force constant was halved. As shown, based on the RMSD metric, the system eventually displays convergence, despite of the somewhat increased RMSD of the 83rd cycle with respect to cycle 82 associated with the decrease of the force constant in the simulations. The lowest energy pathway after 150 iterations of HFB was subsequently used in the pH-REX simulations.

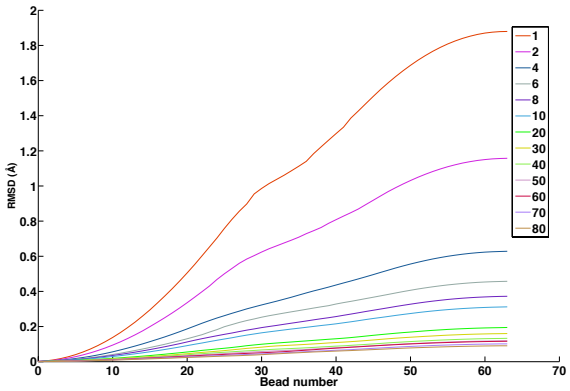

Figure S5. RMSD–based convergence of the first 82 cycles of the HFB calculations. For each iteration of the HFB, all beads in the string were aligned to its corresponding bead in the initial pathway and the RMSD was summed over all beads. The RMSD based on the RCS is shown in blue**.**

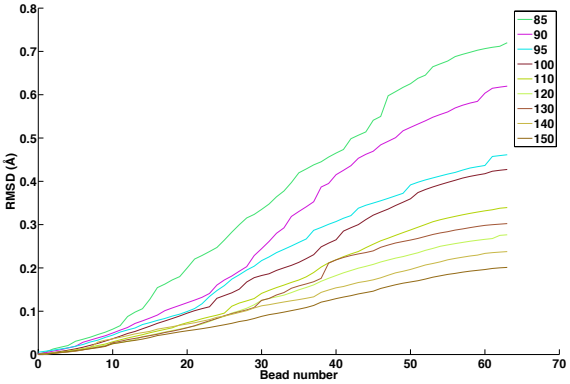

Figure S6. RMSD–based convergence of the 68 cycles of the HFB calculations performed with reduced 10 kcal/mol/Å^2 force constant.

S7. Computing the contacts between V227 and H6 along the HFB pathway: The number of contacts between the side chain of residue V227 located on the cytoplasmic side of H5 and residues of H6 was computed along the Meta II$_a$ to Meta II transition pathway using the contacts.pl script of the MMTSB toolset[7] with an arbitrary distance cutoff of 8 Å (Figure S7). In this Figure, H6 moving outwards along the pathway is marked in red while the position of V277 on a neighboring helix H5 is marked in cyan. Residues of H6 coming into contact with V227 within the 8 Å distance are depicted in red as well. As shown, the number of contacts is small along the initial portion of the pathway corresponding to segment *a-b'* of Figure 2B and associated with a large shift of the E134 pK$_a$ and presumably significant pH dependence of the H6 tilt. However, along the second portion of the path, associated with a small E134 pK$_a$ shift and, thus, not significant pH dependence of the H6 tilt, this number grows from 2 to 9. This supports our conclusions that along the first segment a motion of the H6 is not enough to saturate the EPR signal arising from the interaction between spin-labeled V227 and H6, while along the second segment of the path it is capable of doing so.

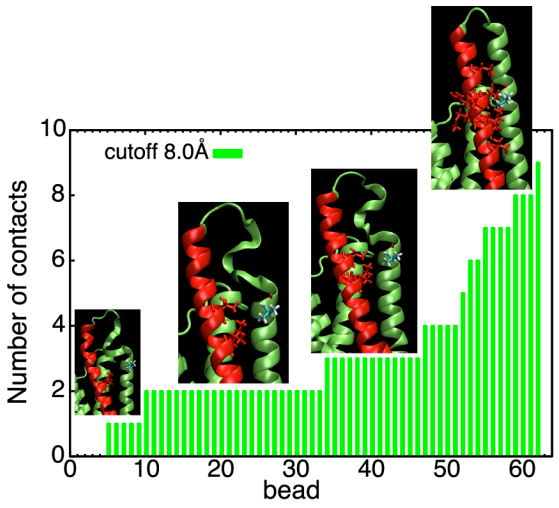

Figure S7. Evolution of the number of contacts between V227 belonging to H5 and residues of H6 along the computed Meta II_a to Meta II transition pathway

REFERENCES

(1) Salom, D.; Lodowski, D. T.; Stenkamp, R. E.; Trong, I. Le; Golczak, M.; Jastrzebska, B.; Harris, T.; Ballesteros, J. A.; Palczewski, K. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 16123–16128.

(2) Choe, H.; Kim, Y.; Park, J.; Morizumi, T. *Nature* **2011**, *471*, 651–655.

(3) Mackerell, A. D.; Feig, M.; Brooks, C. L., III *J. Comput. Chem.* **2004**, *25*, 1400–1415.

(4) MacKerell A. D.; Bashford, D.; Bellott; Dunbrack R. L.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E.; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiórkiewicz-Kuczera, J.; Yin, D.; Karplus, M. *J. Phys. Chem. B* **1998**, *102*, 3586–3616.

(5) Provasi, D.; Filizola, M. *Biophys. J.* **2010**, *98*, 2347–2355.

(6) Fiser, A.; Do, R. K. G.; Šali, A. *Protein Sci.* **2000**, *9*, 1753–1773.

(7) Feig, M.; Karanicolas, J.; Brooks III, C. L. *J. Mol. Graphics Modell.* **2004**, *22*, 377–395.

(8) Saam, J.; Tajkhorshid, E.; Hayashi, S.; Schulten, K. *Biophys. J.* **2002**, *83*, 3097–3112.

(9) Brooks, B. R.; Brooks, C. L., III; Mackerell, A. D.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; Caflisch, A.; Caves, L.; Cui, Q.; Dinner, A. R.; Feig, M.; Fischer, S.; Gao, J.; Hodoscek, M.; Im, W.; Kuczera, K.; Lazaridis, T.; Ma, J.; Ovchinnikov, V.; Paci, E.; Pastor, R. W.; Post, C. B.; Pu, J. Z.; Schaefer, M.; Tidor, B.; Venable, R. M.; Woodcock, H. L.; Wu, X.; Yang, W.; York, D. M.; Karplus, M. *J. Comput. Chem.* **2009**, *30*, 1545–1614.

(10) Mahalingam, M.; Martinez-Mayorga, K.; Brown, M. F.; Vogel, R. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 17795–17800.

(11) Periole, X.; Ceruso, M.; Mehler, E. *Biochemistry* **2004**, *43*, 6858–6864.

(12) Khandogin, J.; Brooks, C. L., III Constant pH Molecular Dynamics with Proton Tautomerism. *Biophysical journal* **2005**, *89*, 141–157.

(13) Lee, M. S.; Salsbury, F. R.; Brooks, C. L., III *Proteins: Struct., Funct., Bioinf.* **2004**, *56*, 738–752.

(14) Sugita, Y.; Okamoto, Y. *Chem. Phys. Lett.* **1999**, *314*, 141–151.

(15) Jiang, W.; Hodoscek, M.; Roux, B. *J. Chem. Theory Comput.* **2009**, *5*, 2583–2588.

(16) Khavrutskii, I. V; Arora, K.; Brooks, C. L., III; III *J. Chem. Phys.* **2006**, *125*, 174107–174108.

(17) Standfuss, J.; Edwards, P. C.; D'Antona, A.; Fransen, M.; Xie, G.; Oprian, D. D.; Schertler, G. F. X. *Nature* **2011**, *471*, 656–660.

(18) Park, J. H.; Scheerer, P.; Hofmann, K. P.; Choe, H.-W.; Ernst, O. P. *Nature* **2008**, *454*, 183–187.

(19) Nakamichi, H.; Okada, T. *Angew. Chem., Int. Ed.* **2006**, *45*, 4270–4273.

(20) Nakamichi, H.; Okada, T. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 12729–12734.

(21) Ruprecht, J.; Mielke, T.; Vogel, R. *The EMBO journal* **2004**, *23*, 3609–3620.

(22) Arnis, S.; Hofmann, K. P. *Proc. Natl. Acad. Sci. USA* **1993**, *90*, 7849–7853.

(23) Knierim, B.; Hofmann, K. P.; Ernst, O. P.; Hubbell, W. L. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 20290–20295.

(24) Farrens, D. L.; Altenbach, C.; Yang, K.; Hubbell, W. L.; Khorana, H. G. *Science* **1996**, *274* , 768–770.

(25) Vogel, R.; Sakmar, T. *Photochem. Photobiol.* **2007**, *83*, 286–292.

(26) Vogel, R.; Mahalingam, M.; Lüdeke, S.; Huber, T.; Siebert, F.; Sakmar, T. P. *J. Mol. Biol.* **2008**, *380*, 648–655.

(27) Zaitseva, E.; Brown, M. F.; Vogel, R. *J. Am. Chem. Soc.* **2010**, *132*, 4815–4821.

(28) Im, W.; Lee, M. S.; Brooks, C. L., III *J. Comput. Chem.* **2003**, *24*, 1691–1702.

(29) Nina, M.; Beglov, D.; Roux, B. *J. Phys. Chem. B* **1997**, *101*, 5239–5248.

(30) Chen, J.; Im, W.; Brooks, C. L., III *J. Am. Chem. Soc.* **2006**, *128*, 3728–3736.

(31) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187–217.

(32) Van Gunsteren, W. F.; Berendsen, H. J. C. *Mol. Phys.* **1977**, *34*, 1311–1327.

(33) Kong, X.; III, C. L. B. *J. Chem. Phys.* **1996**, *105*, 2414–2423.

(34) Knight, J. L.; Brooks, C. L., III *J. Comput. Chem.* **2009**, *30*, 1692–1700.

(35) Yan, E. C. Y.; Kazmi, M. A.; Ganim, Z.; Hou, J.-M.; Pan, D.; Chang, B. S. W.; Sakmar, T. P.; Mathies, R. A. *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 9262–9267.

(36) Lüdeke, S.; Beck, M.; Yan, E. C. Y.; Sakmar, T. P.; Siebert, F.; Vogel, R. *J. Mol. Biol.* **2005**, *353*, 345–356.

(37) Martínez-Mayorga, K.; Pitman, M. C.; Grossfield, A.; Feller, S. E.; Brown, M. F. *J. Am. Chem. Soc.* **2006**, *128*, 16502–16503.

(38) Mertz, B.; Struts, A.; Feller, S.; Brown, M. *Biochim. Biophys. Acta: Biomembr.* **2012**, *1818*, 241–51.