Supporting Information

for

Spectral Tuning of Ultraviolet Cone Pigments: An Interhelical Lock Mechanism

Sivakumar Sekharan[†*], Victoria L. Mooney[†], Ivan Rivalta[†], Manija A. Kazmi[‡], Maureen Neitz[§], Jay Neitz[§], Thomas P. Sakmar[‡], Elsa C. Y. Yan[†*], Victor S. Batista[†,*]

[†]Department of Chemistry, Yale University, P.O. Box 208107, New Haven, CT 06520-8107
[†]Laboratory of Chemical Biology and Signal Transduction, Rockefeller University, New York, NY 10065
[§]School of Medicine, Department of Ophthalmology, University of Washington, Seattle, WA 98195

Total Number of Pages: S11

Table of Contents

I. Experimental Section

Nucleic acid extraction, polymerase chain reaction, and DNA sequencing: Nucleic acids were extracted by mincing tissue from *Phodopus sungorus* liver or retina and placing it in 200 µl of 10 mM Tris-HCl, pH 8.0, 0.1 mM EDTA, pH 8.0, 1% (w/v) sodium dodecyl sulfate, and 4 mg/ml proteinase K. The mixture was incubated for 8-12 hours at 55℃. Samples were then extracted with phenol and chloroform, followed by nucleic acid precipitation from ethanol.[1] The *P. sungorus* retinal RNA was reverse transcribed (ABI), resulting in two separate, overlapping segments of the UV pigment cDNA. Both segments were amplified and sequenced. *P. Sungorus* liver nucleic acid was used directly in the PCR reaction to amplify a segment extending from exon 4 into exon 5, which included intron 4. A Kozak sequence and KpnI site were inserted at the 5' end of the UV pigment gene (GenBank ID: JQ036217).[2,3] To the 3' end of the SHUV gene, a 1D4 tag (the last 9 a.a. of the bovine rhodopsin C-terminus, TETSQVAPA) was added and a NotI restriction site was inserted immediately following the stop codon. The UV pigment gene was then cloned into the tetracycline inducible pACMV-TetO vector.[4,5] DNA sequencing was used to confirm the final sequence of the pACMV-TetO-UV plasmid. The UV gene contained the silent mutations Q233Q (CAA → CAG) and A310A (GCC → GCA). F86Y, S90A, and S90C SHUV mutants were prepared via point mutations.

Protein expression & purification: *GnTI*- HEK293S cells were transfected with the pACMV-TetO-UV plasmid using lipofectamine (Invitrogen). The UV stable cell line was made as previously reported.[6] UV pigment expression was induced upon addition of 5 mM sodium butyrate and 2 µg/mL tetracycline; cells were harvested 48 hr after induction.[4] The protein was prepared under dim red light, using procedures reported previously.[2,7] The cells were washed with Buffer A (50 mM HEPES, pH 6.6; 140 mM NaCl, 3 mM MgCl2) before being regenerated overnight in Buffer B (5 µg/mL each of aprotinin and leupeptin and 7.15 µM 11-*cis* retinal in Buffer A) at 4°C. After centrifugation, the cell pellet was resuspended in solubilization buffer (1% w/v n-dodecyl-β-D-maltoside (DM), 20% w/v glycerol, and 0.1 mM phenylmethanesulfonyl fluoride in 75% v/v buffer A) for ~5 hr. Immunoaffinity chromatography was used to purify the SHUV pigment[2,8,9], wherein the 1D4 epitope cloned into the C-terminus of the UV pigment was recognized by the 1D4 monoclonal antibody bound to agarose resin. The resin was washed thrice

with Buffer C (47 mM HEPES, pH 6.6; 132 mM NaCl, 2.8 mM $MgCl_2$, 0.02% DM, and 7.5% w/v glycerol). The UV pigment was eluted into Buffer C upon addition of 0.726 mg/mL 1D5.

Pigment characterization: After purification of the SHUV pigment and its mutants, UV-visible spectra were collected to ensure that their dark states were functional. To produce the photoproducts, the UV pigment was illuminated for 60 s with 365-nm light. This resulted in a shift in the pigments' λ_{max}s to ~379 nm, consistent with the photoconversion of retinal from an 11-*cis* to all-*trans* isomer and the formation of an active Meta II-like product. The extinction coefficient for dark-state, wild-type SHUV was determined using an acid denaturation assay.[10] Addition of hydrochloric acid to the dark-state pigment caused a red shift from 359 nm to 440 nm, consistent with the formation of a protonated Schiff base linkage between the denatured opsin and the retinylidene chromophore (Fig. S1). From three separate trials, a value of 43,300 ± 700 M^{-1} cm^{-1} was calculated for the 359 nm extinction coefficient of the dark-state UV pigment using the known ε_{440} of 30,800 $M^{-1}cm^{-1}$ for a protonated Schiff base in solution. A spectral ratio (A_{280}/A_{359}) of 2.1 was calculated for the dark-state SHUV. This acid denaturation technique was also used for the F86Y mutant to determine whether that 376 nm UV peak present in the pigment's dark-state is due to free retinal. After pigment denaturation, the 376 nm and 432 nm peaks both shifted to ~440 nm, indicating the presence of an intact Schiff base for both species (Fig. S2). All UV-visible spectroscopy was carried out in an UV-2450 Shimadzu spectrophotometer.

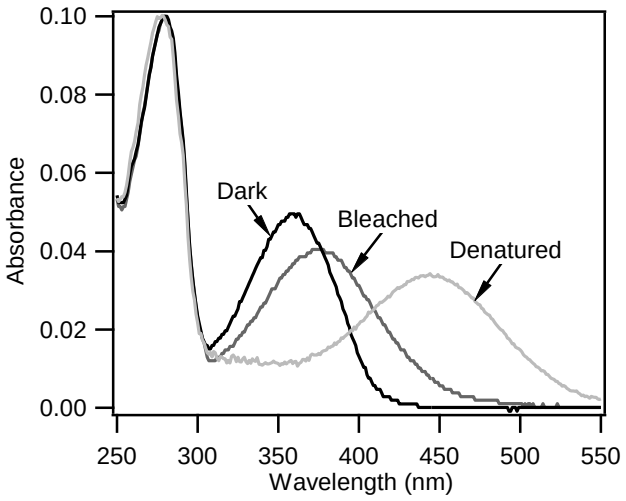

Figure S1. *UV-Visible spectra of the wild-type Siberian hamster UV pigment before (Dark, 359 nm) and after (Bleached, 379 nm) being bleached for 1 min with 365 nm light. The dark-state*

SHUV was acid denatured (440 nm) in order to determine the extinction coefficient of the dark-state peak at 359 nm.

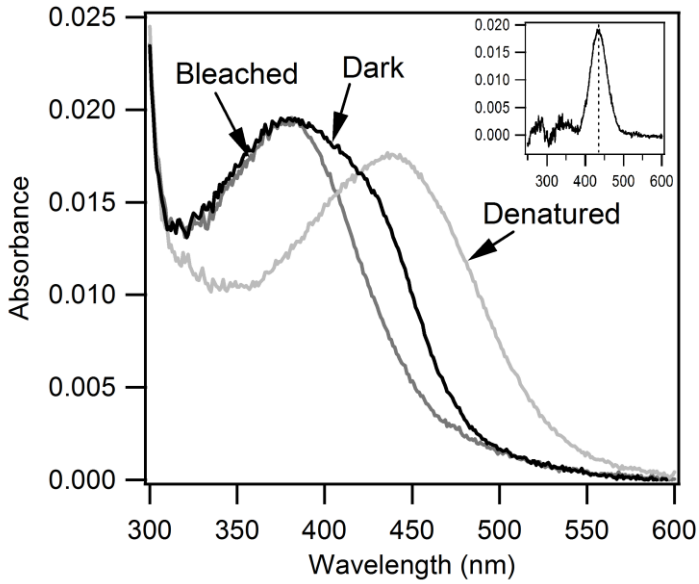

Figure S2. *UV-Visible spectra of the Siberian hamster UV pigment F86Y mutant. In all three spectra, the absorbance was normalized to 0.1 at 280 nm. In the dark-state, there are two species present, one absorbing in the UV and the other in the visible region (Dark). Upon photoactivation with white light for 10 min, the* λ_{max} *shifts to 380 nm (Bleached). The dark-state pigment was also acid denatured to ensure that the UV peak in the dark-state was not due to retinal that had hydrolyzed from the pigment. The inset shows the dark minus bleached difference spectrum with a vertical line is centered at 435 nm.*

II. Computational Details

Primary sequence and secondary structure of the SHUV pigment: We developed a molecular model of the dark-state SHUV pigment to gain a better understanding of the role of counterion (E113), water molecules and H-bonding networks in the active site. Alignment of the primary sequences of SHUV pigment and bovine rhodopsin using the clustalW[11] revealed that there is 43% identity between the two sequences and 193 non-conserved amino acid residues. We used the crystal structure of rhodopsin (PDB code: 1U19 at 2.2 Å resolution)[12] as template to create the homology model of the SHUV pigment (Fig. S3A, S3B). Nonconserved residues have been mutated using the standard mutation procedure implemented in the Schrödinger's Maestro suite[13] (the standard mutate.py script uses the standard Maestro rotamer libraries) and after each mutation the overall structure was relaxed by using the impref Maestro utility. Protonation states of all titratable residues are assigned based on PROPKA pK_a calculations[14] and visual inspection.

Geometry minimizations were based on the impact molecular mechanics engine[15] using the OPLS_2005 force field[16,17] (Fig. S3C) and setting the maximum RMSD to 0.30 Å. No simulated annealing/molecular dynamics refinement was carried out because the proposed homology model was simply used as an initial structure for the QM/MM refinement. The RMSD cutoff is to ensure a minimal movement of the system with respect to the initial Rhodopsin template during the preliminary MM refinement (i.e. the impref optimization), which is aimed at removing close contacts at the mutation sites. The disulfide bond between Cys110 in transmembrane helix 3 and Cys187 in extracellular loop 2 (EL2) was maintained in the homology model.

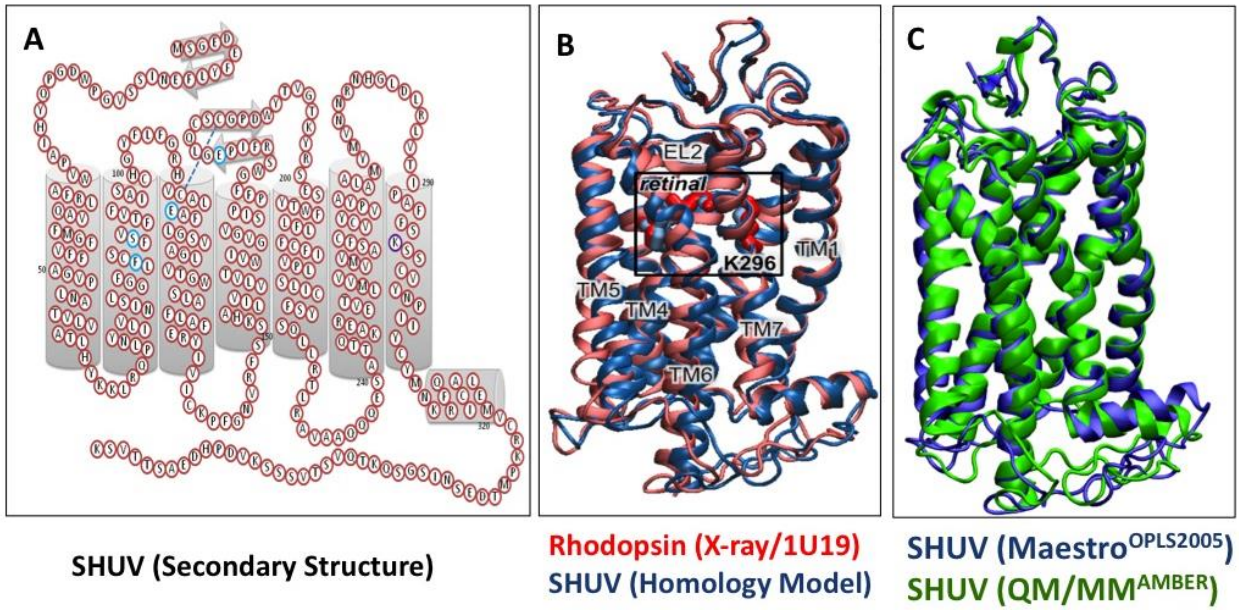

Figure S3. *(A) The numbering system used in the secondary structure of the Siberian hamster UV (SHUV) pigment correspond to that of rhodopsin; there is a 5-residue shift in the N-terminus such that from H1-H8, rhodopsin residue numbers are 5 greater than those of UV residues. The dashed line connecting C110 to C187 represents a disulfide bond. The purple circle surrounds K296, the lysine to which 11-cis retinal binds. Residues that are in blue circles are known to interact with the Schiff base (E113 and E181) or are important for spectral tuning (F86 and S90) in vertebrate visual pigments. (B) Superposition of the X-ray structure of bovine rhodopsin (red) with Maestro-MM optimized starting structure of the SHUV homology model (blue). The pigments are orientated such that the extracellular side is above and the cytoplasmic side is below the transmembrane region. The secondary structure elements (cartoon), including the transmembrane helices (TM) and extra-cellular loop 2 (EL2), as well as the K269 and 11-cis-retinyl chromophore (sticks) are shown. (C) Superposition of the Maestro MM (OPLS2005) optimized starting structure (blue) with the ONIOM QM/MM (Amber) optimized fully relaxed homology model of the SHUV pigment. The positions of the seven TM α-helices are shown to be*

very well conserved during the geometry optimizations and the overall RMSD of the fully optimized QM/MM structure with respect to Maestro optimized starting structure is 1.46 Å.

Chromophore Binding Pocket: Within 3 Å of the chromophore in dark-state SHUV there are 10 non-conserved residues present, including F86(M) and S90(G), where rhodopsin residues are indicated in parentheses (Fig. S4). There are also 8 key active site residues that are conserved, including E113, E181, and Y268. Amino acid residues Y192 and T94 have been proposed to be also part of the extended hydrogen-bonding network in rhodopsin. While amino acid residue Y192 is conserved in the SHUV pigment, the hydrophobic residue V94 substitutes the polar residue T94, disrupting one of the possible hydrogen bonds for the E113 counterion. On the other hand, the rhodopsin non-polar residue G90 is substituted in the SHUV pigment by the polar residue S90, which can establish a hydrogen-bond interaction with the close E113 side chain. Therefore, the SHUV pigment and rhodopsin hydrogen-bonding networks near E113 are expected to be significantly different.

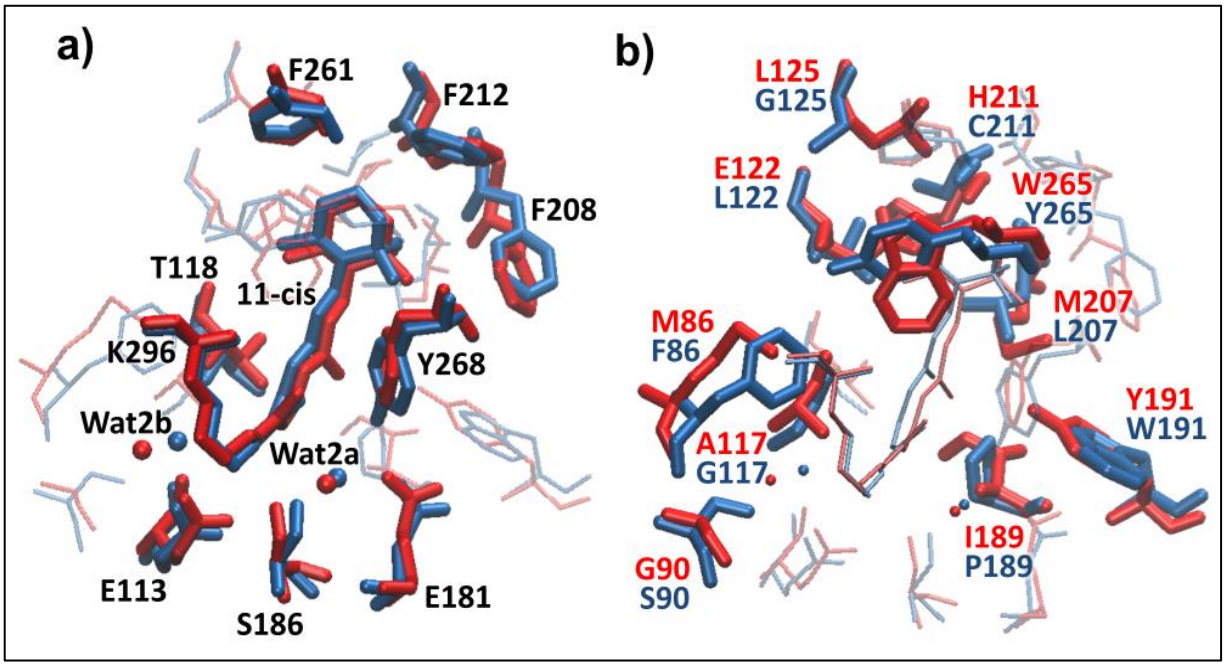

Figure S4. *Superposition of the retinal binding pockets of bovine rhodopsin (red) and the SHUV pigment homology model (blue) after alignment of α-carbon atoms of amino acid residues within 3 Å of the bound retinal. Oxygen atoms belonging to water molecules (w2a and w2b) in the binding pockets are shown as spheres. a) Identity elements conserved among the two pigment types (including the water molecules) in the retinal binding pockets are highlighted, as well as the 11-cis retinal/K296 Schiff bases. b) Non-conserved amino acid residues in the binding pockets are highlighted.*

MM and Hybrid QM/MM Calculations: The pre-refined homology model obtained using the Maestro suite was used as a starting structure to perform a series of initial pure AMBER geometry optimizations to remove close contacts and find more optimum positions for flexible

groups in the entire protein. All MM calculations in the ensuing the QM/MM calculations were performed with AMBER96 force field and TIP3P water model. A series of initial pure AMBER geometry optimizations was performed to remove close contacts and find more optimum positions for flexible groups in the protein: (i) First, only the long flexible chains at the C and N terminals of the protein were optimized. (ii) Second, the flexible chains and water molecules in the protein were optimized. (iii) Finally, the full protein was optimized without any constraints. The resulting coordinates were fully optimized with a hybrid QM/MM (QM=B3LYP/6-31G*; MM=AMBER) method in ONIOM (Our own *N*-layer Integrated molecular Orbital + molecular Mechanics)[18] using electronic embedding scheme[19], which incorporates the partial charges of the MM region into the quantum mechanical Hamiltonian. This technique provides a better description of the electrostatic interaction between the QM and MM regions (as it is treated at the QM level) and allows the QM wavefunction to be polarized as described in our previous works on rhodopsin.[20-24] The positions of the seven α–helices were very well conserved during the geometry optimizations. The QM model part includes full unprotonated or protonated Schiff base of 11-*cis*-retinyl chromophore (USB11 for WT, F86Y, S90A, S90C models and PSB11 for F86Y) with side-chain N–H moiety of Lys296 along with a hydrogen link atom[25] at the QM/MM boundary plus full E113 counterion residue in the protonated (when USB11 is present) and in unprotonated forms (when PSB11 is present). All single-point *ab initio* QM/MM calculations on the resulting coordinates were carried out with the ORCA 2.6.19 program package.[26] We applied three-root spectroscopy oriented configuration interaction (SORCI) method with Davidson correction (+Q) for excitations higher than doubles on complete active space self-consistent field (CASSCF) wave functions to calculate absorption and circular dichroism (CD) spectra. All multireference (MR) *ab initio* calculations were carried out using the 6-31G* basis set[27-30] plus an auxiliary basis set (SV/C)[31-33] to speed-up the calculations. The active space encompasses 6 electrons in 6 orbitals. Computational efficiencies of SORCI+Q calculations were enhanced by setting T_{pre}, T_{nat} and T_{sel} thresholds to 10^{-4}, 10^{-6} and 10^{-6} *E*h, respectively.[34-38] The core orbitals with energies of less than -4 E_h were frozen and a level shift of 0.4 E_h was applied in all perturbative treatments.

III. Energetics

Table S1: The energy values presented in Table S1 correspond to the calculated ONIOM QM/MM extrapolated energies and it shows that the USB form is more stable in the WT, S90A and S90C models, whereas the PSB form is more stable in the F86Y model. The findings are attributed to the absence or presence of an Inter-Helical Lock (IHL) involving TM6—TM2—TM3 helices in the active site.

SHUV Models	***ONIOM QM/MM Extrapolated Energy (kcal.mol^{-1})***		***Inter-Helical Lock between TM6—TM2—TM3***
	USB	***PSB***	
WT	*0.000*	*+6.780*	*Absent*
S90A	*0.000*	*+4.687*	*Absent*
S90C	*0.000*	*+5.838*	*Absent*
F86Y	*+10.392*	*0.000*	*Present*

IV. Comparative analysis of w2c position in the active site.

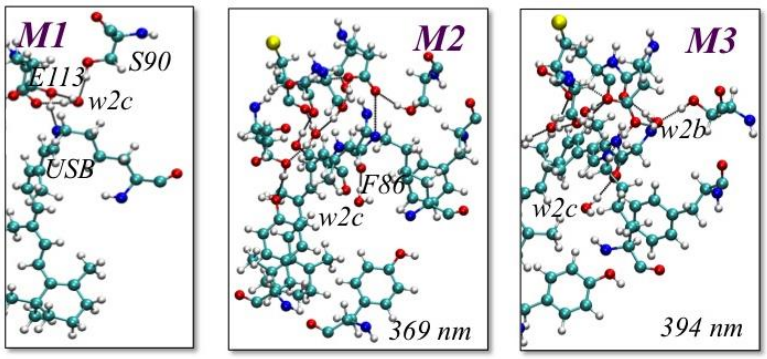

Figure S5. *QM/MM analysis of three different SHUV models (M1, M2, M3) that differs in the position of w2c in the active site. In the case of M1, w2c is present between S90 and E113 and occupies the position corresponding to w2b in rhodopsin. In the case of M2, w2c is present near F86 and as a result, there is a direct hydrogen bond between S90 and E113. Comparison of the*

ONIOM extrapolated energies reveal that, M1 is energetically unstable by more than 15 kcal/mol compared to M2. In the case of M3, the active site involves both w2b and w2c water molecules (w2b between S90 and E113 and w2c near F86) and the calculated λ_{max} is red shifted by almost 40 nm compared to the experimental value (359 nm). Therefore, we rule out M1 and M3 models as possible wildtype structural models for the SHUV pigment.

V. Correlation between the intermolecular hydrogen bonding networks and frontier molecular orbital analysis of the electronic transition responsible for the spectral shifts in the SHUV mutant pigment models.

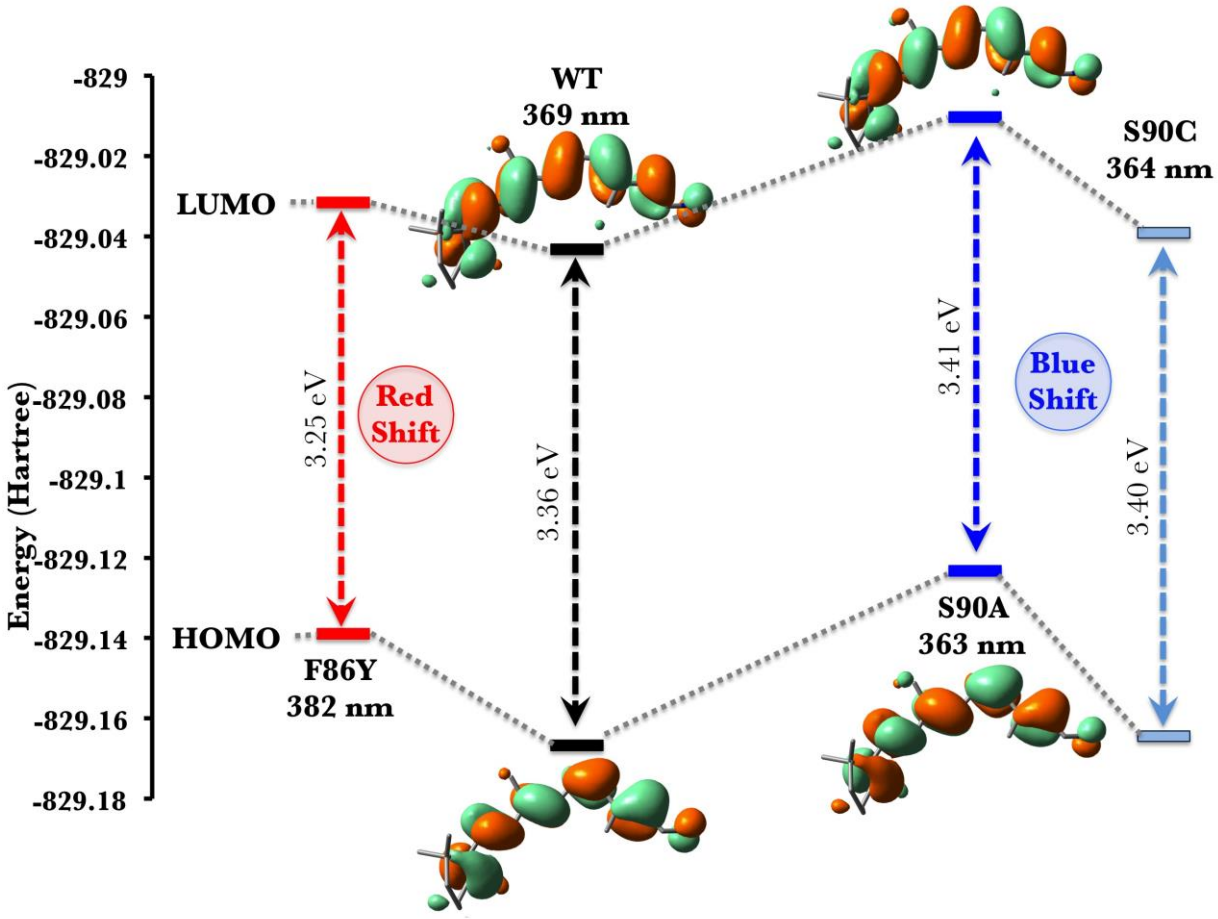

***Figure S6**. The analysis shows that breaking or weakening of the hydrogen bond (between S90 and E113 residues) due to S90A and S90C mutations increases the energy gap and induces a spectral blue shift of ~6 nm, whereas formation of a hydrogen bond involving F86Y mutant decreases the energy gap and induces a spectral red shift of 13 nm relative to the WT (Ref-39,40).*

VI. References

1. Cobb, J. K.; Bialozynski, C.; Neitz, J.; Jacobs, G. H.; Neitz, M. *Invest. Ophthalmol. Vis. Sci.* **1999**, *40*, S353
2. Oprian, D. D.; Molday, R. S.; Kaufman, R. J.; Khorana, H. G. *Proc. Acad. Natl. Sci. U.S.A.* **1987**, *84*, 8874.
3. Kozak, M. *Nucleic Acids Res.* **1984**, *12*, 857.
4. Reeves, P. J.; Kim, J. M.; Khorana, H. G. *Proc. Acad. Natl. Sci. U.S.A.* **2002**, *99*, 13413.
5. Chang, B. S. W.; Kazmi, M. A.; Sakmar, T. P. *Methods Enzymol*. **2002**; *343*, 274.
6. Yan, E. C. Y.; Epps, J.; Lewis, J. W.; Szundi, I.; Bhagat, A.; Sakmar, T. P.; Kliger, D. S. *J. Phys. Chem. C* **2007**, *111*, 8843.
7. Babu, K. R.; Dukkipati, A.; Birge, R. R.; Knox, B. E. *Biochemistry* **2001**, *40*, 13760.
8. Franke, R. R.; Sakmar, T. P.; Graham, R. M.; Khorana, H. G. *J. Biol. Chem.* **1992**, *267*, 14767.
9. Chan, T.; Lee, M.; Sakmar, T. P. *J. Biol. Chem.* **1992**, *267*, 9478.
10. Kito, Y.; Suzuki, T.; Azuma, M.; Sekoguti, Y. *Nature* **1968**, *218*, 955.
11. Larkin, M. A.; *Blackshields, G.; Brown N. P.; Chenna, R.; McGettigan, P. A.; McWilliam, H.; Valentin, F.; Wallace, I. M.; Wilm, A.;* Lopez, R.; Thompson, J. D.; Gibson, T. J.; Higgins, D. G. Clustal W and Clustal X version 2.0. *Bioinformatics* **2007**, *23*, 2947.
12. Okada, T.; Sugihara, M.; Bondar, A.N.; Elstner, M.; Entel, P. Buss, V. *J. Mol. Biol.* **2004**, *342*, 571.
13. Maestro, version 9.0, Schrödinger, LLC, New York, NY, **2009**.
14. Li, H.; Robertson, A. D.; Jensen, J. H. *Proteins: Struct. Funct. Bioinf.* ***2005***, *61*, 704.
15. Banks, J.; Beard, H. S.; Cao, Y. X.; Cho, A. E.; Damm, W.; Farid, R.; Felts, A. K.; Halgren, T. A.; Mainz, D. T.; Maple, J. R.; Murphy, R.; Philipp, D. M.; Repasky, M. P.; Zhang, L. Y.; Berne, B. J.; Friesner, R. A.; Gallicchio, E.; Levy, R. M. *J. Comp. Chem.* **2005**, 1752.
16. Jorgensen, W. L.; Maxwell, D. S.; TiradoRives, J. *J. Am. Chem. Soc.* **1996**, *118*, 11225.
17. Kaminski, G. A.; Friesner, R. A.; Tirado-Rives J, Jorgensen, W. L. *J. Phys. Chem.* **2001**, *105*, 6474.
18. Vreven, T.; Morokuma, K. *Ann. Rep. Comp. Chem.* **2006**, *2*, 35.
19. Antes, I.; Thiel, W. Hybrid Quantum Mechanical and Molecular Mechanical Methods; Gao, J., Ed.; ACS Symposium Series 712, American Chemical Society; Washington DC, **1998**; pp 50-65.
20. Gascón, J. A.; Batista, V.S. *Biophys. J.* **2004**, *87*, 2931.
21. Gascón, J. A.; Sproviero, E.M.; Batista, V. S. *J. Chem. Theory Comput.* **2005**, *1*, 674.
22. Gascón, J. A.; Sproviero, E.M.; Batista, V. S. *Acc. Chem. Res.* **2006**, *39*, 184.
23. Sekharan, S.; Wei, J. N; Batista, V. S. *J. Am. Chem. Soc.* **2012**, *134*, 19536.
24. Pal, R.; Sekharan, S.; Batista, V. S. *J. Am. Chem. Soc.* **2013**, *135*, 9624.
25. Bakowies, D.; Thiel, W. *J. Phys. Chem.* **1996**, *100*, 10580.
26. Neese F, (2007) *ORCA-an ab initio, DFT and semiempirical electronic structure package*, version 2.6, revision 19; Institut für Physikalische und Theoretische Chemie, Universita¨t Bonn, Germany, 2007. Latest version available from http://www.thch.uni-bonn.de/tc/orca
27. Ditchfield, R.; Hehre, W. J.; Pople, J. A. *J. Chem. Phys.* **1971**, *54*, 724–728.

28. Hehre, W. J.; Ditchfield, R.; Pople, J. A. *J. Chem. Phys.* **1972**, *56*, 2257–2261.
29. Hariharan, P. C.; Pople, J. A. *Theor. Chim. Acta* **1973**, *28*, 213–222.
30. Clark, T.; Chandrasekhar, J.; Spitznagel, G. W.; Schleyer, P. v. R. J. Comput. Chem. **1983**, *4*, 294–301.
31. The SV/C auxiliary basis set was obtained from the TURBOMOLE basis set library under ftp.chemie.uni-karlsruhe.de/pub/cbasen.
32. Eichkorn, K.; Weigend, F.; Treutler, O.; Ahlrichs, R. *Theor. Chem. Acc.* **1997**, *97*, 119–124.
33. Weigend, F.; Häser, M. *Theor. Chem. Acc.* **1997**, *97*, 331–340.
34. Neese, F. *J. Chem. Phys.* **2003**, *119*, 9428–9443.
35. Wanko, M.; Hoffmann, M.; Strodel, P.; Koslowski, A.; Thiel, W.; Neese, F.; Frauenheim, T.; Elstner, M. *J. Phys. Chem. B* **2005**, *109*, 3606–3615.
36. Hoffmann, M.; Wanko, M.; Strodel, P.; König, P. H.; Frauenheim, T.; Schulten, K.; Thiel, W.; Tajkhorshid, E.; Elstner, M. *J. Am. Chem. Soc.* **2006**, *128*, 10808–10818.
37. Altun, A.; Kumar, D.; Neese, F.; Thiel, W. *J. Phys. Chem. A* **2008**, *112*, 12904–12910.
38. Altun, A.; Yokoyama, S.; Morokuma, K. *J. Phys. Chem. B* **2008**, *112*, 16883–16890.
39. Zhao, G. J.; Liu, J. Y.; Zhou, L. C.; Han, K. L. *J. Phys. Chem. B.* **2007**, *111*, 8940.
40. Zhao, G. J.; Han, K. L. *Biophys, J.* **2008**, *94*, 38.

VII. Complete References from the Manuscript

R18. 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, Jr., J. A.; 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. Gaussian, Inc., Wallingford CT, 2009.

R32. Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman. P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179.

R35. Birge, R. R.; Murray, L. P.; Pierce, B. M.; Akita, H.; Balogh-Nair, V.; Findsen, L. A.; Nakanishi, K. *Proc. Natl. Acad. Sci. U.S.A* **1985**, *82*, 4117.