

# Mechanism of Spectral Tuning in Green-Absorbing Proteorhodopsin<sup>†</sup>

Rekha Rangarajan,<sup>‡</sup> Jhenny F. Galan,<sup>‡</sup> Gregg Whited,<sup>§</sup> and Robert R. Birge<sup>\*,‡</sup>

Departments of Chemistry and of Molecular and Cell Biology, University of Connecticut, 55 North Eagleville Road, Storrs, Connecticut 06269-3060, and Genencor International, Inc., 925 Page Mill Road, Palo Alto, California 94304-1013

Received May 18, 2007; Revised Manuscript Received July 26, 2007

**ABSTRACT:** The absorption spectrum of green proteorhodopsin (GPR) is pH-dependent, exhibiting either red-shifted (low pH) or blue-shifted (high pH) absorption maxima. We examine the molecular basis of the pH-dependent spectral properties of green proteorhodopsin by using homology modeling and molecular orbital theory. Bacteriorhodopsin (BR) and sensory rhodopsin II (SRII) are compared as homology templates. The model of GPR generated by using BR as the homology parent is better than that generated by using SRII on the basis of the potential energy, relative stability to dynamics, and ability to rationalize pH effects. MNDO-PSDCI (molecular neglect of differential overlap with partial single- and double-configuration interaction) calculations provide insight into the spectroscopic properties of GPR and help rule out the viability of the SRII-based model. The proximity of His 75 to the quadrupole residues (LYR, D97, D227, and R94) in the BR-based model provides a good model for both the low- and high-pH spectral states of GPR. The observation that BR is a better structural model for GPR than SRII is in contrast to our previous study of BPR, which observed that SRII was the better homology parent [Hillebrecht, J. R. (2006) *Biochemistry* 45, 1579–1590]. The implications of this observation are discussed.

Proteorhodopsin (PR)<sup>1</sup> is a light-driven proton pump found in the cell membrane of an uncultivated marine bacterium of the SAR86 group of  $\gamma$ -proteobacteria. Following its initial discovery off Monterey Bay (California) (1), several variants of proteorhodopsin have been identified throughout the world (2, 3). The amino acid sequence of PR shows a high degree of sequence similarity to the archaeal rhodopsins (1, 3). Proteorhodopsin-expressing variants can be classified into two major groups depending on their absorption maxima as the green-absorbing (520 nm) and blue-absorbing (490 nm) forms. The green-absorbing variants (GPR) are surface dwellers in comparison to the blue forms (BPR), which are bottom dwellers (1, 4, 5). Light absorption by proteorhodopsin triggers a series of conformational changes associated with the chromophore and the protein environment and results in proton translocation from the cytoplasmic side to the extracellular surface (1, 3, 6). The electrochemical gradient generated by this process is used to drive ATP synthesis for fueling metabolic processes within the cell (7).

The photocycle of GPR is similar to that of BR (8–10), and the photocycle kinetics of GPR and BR are both fast (10–15 ms), a property characteristic of transport proteins (3, 10).

Previous homology models of GPR were generated on the basis of the BR crystal structure (5). The present study examines the relative viability of the GPR models based on BR and SRII as structural templates. We find that BR-based models show lower energies and enhanced stability to long-term molecular dynamics (MD) simulations relative to SRII-based models. The BR-based structure provides a better model of the pH dependence on the absorption spectrum. We conclude that the BR-based model is more realistic in elucidating the spectral tuning and pH dependence properties of GPR (Figure 1).

## MATERIALS AND METHODS

**Expression, Isolation, and Purification of GPR.** The GPR gene (Bac31A08) was PCR amplified and cloned into a C-terminal His-tag-carrying plasmid pTrcHis2 TOPO TA expression vector (Invitrogen, Carlsbad, CA). The GPR-carrying expression vector was transformed into BL21-Codonplus RIL *Escherichia coli* cells (Stratagene, La Jolla, CA) using standard protocols. *E. coli* transformants were grown in LB medium containing 0.5% glucose, 100  $\mu$ g/mL carbenicillin, and 10  $\mu$ M *all-trans*-retinal at 37 °C. Proteorhodopsin expression was induced at 37 °C for an additional 4 h with the addition of 0.5 mM isopropyl  $\beta$ -D-thiogalactopyranoside and 10  $\mu$ M *all-trans*-retinal. The cells were harvested via centrifugation at 3400g for 10 min and stored at –80 °C. The proteorhodopsin-containing membrane fraction was isolated by resuspending the cell pellets in lysis buffer containing 50 mM Tris–HCl (pH 7.7), 0.5 M ethylenediaminetetraacetic acid (EDTA) (pH 8.0), and 0.2 mg/mL DNase. The resulting cell lysates were further

<sup>†</sup> This work was supported in part by grants from the National Institutes of Health (GM34548), the National Science Foundation (BES-0412387, CCF-0432151), and DARPA (HR0011-05-1-0027). Computational facilities were provided by the Harold S. Schwenk Sr. Distinguished Chair in Chemistry.

\* To whom correspondence should be addressed. Phone: (860) 486-6720. Fax: (860) 486-2981. E-mail: rbirge@uconn.edu.

<sup>‡</sup> University of Connecticut.

<sup>§</sup> Genencor International, Inc.

<sup>1</sup> Abbreviations: PR, proteorhodopsin; GPR, green proteorhodopsin; BPR, blue proteorhodopsin; BR, bacteriorhodopsin; SRII, sensory rhodopsin II; BLOSUM40, block substitution matrix 40; MM2, molecular mechanics 2; MM3, molecular mechanics 3; SD, steepest descent; CG, conjugate gradient; ABNR, adopted basis Newton–Raphson; MD, molecular dynamics; CHARMm, chemistry at Harvard molecular mechanics; TM, transmembrane; MNDO-PSDCI, molecular neglect of differential overlap with partial single- and double-configuration interaction; PSB, protonated Schiff base.

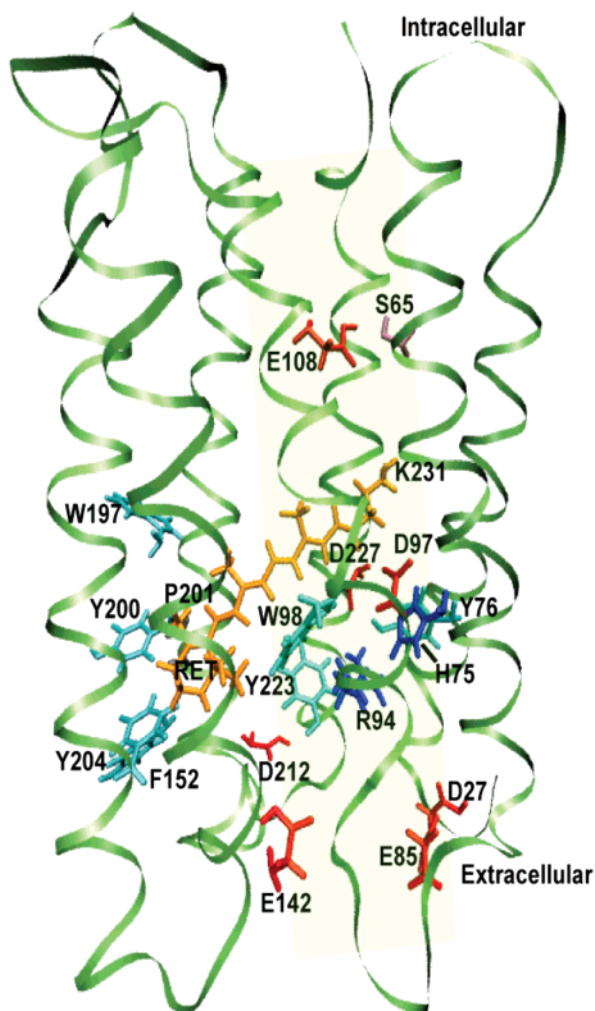


FIGURE 1: Three-dimensional model of GPR based on the 1C3W (PDB ID) crystal structure (1.55 Å) generated using homology modeling and MD simulations (CHARMM). The protein is characterized by the presence of seven transmembrane  $\alpha$ -helical segments with the prosthetic group, an *all-trans*-retinal bound via K231 (shown in orange). Highlighted are several key residues in the protein (D97, D227, R94) which are conserved among most retinylidene proteins.

homogenized using a French press. The proteorhodopsin-containing membrane fraction was purified via a series of low-speed (10000g) and high-speed (200000g) centrifugation spins for 30 min at 4 °C. The pellet was resuspended in a solubilization buffer containing 50 mM Tris (pH 7.7) and 2% dodecyl  $\beta$ -D-maltoside (D $\beta$ M) prior to a second round of homogenization using the Dounce tissue homogenizer. The protein isolated from the membrane was purified via fractionation using a nickel–nitrilotriacetic acid (Ni–NTA) resin column (Amersham, Piscataway, NJ). The protein was applied to the resin, which was equilibrated in buffer containing 50 mM Tris–HCl (pH 7.7) and 0.05% D $\beta$ M in a Buchner funnel. The resin was washed extensively to remove unbound and weakly bound proteins. The proteorhodopsin was eluted with the same buffer containing 250 mM EDTA. The proteorhodopsin-containing fractions were pooled and concentrated using an Amicon stirred cell concentrator with a PM-30 membrane. The concentrated protein was dialyzed using a 12000–14000 MW cutoff membrane against 3 100-fold volumes of 10 mM Tris–HCl (pH 7.7) buffer to remove salts and eluted nickel from the

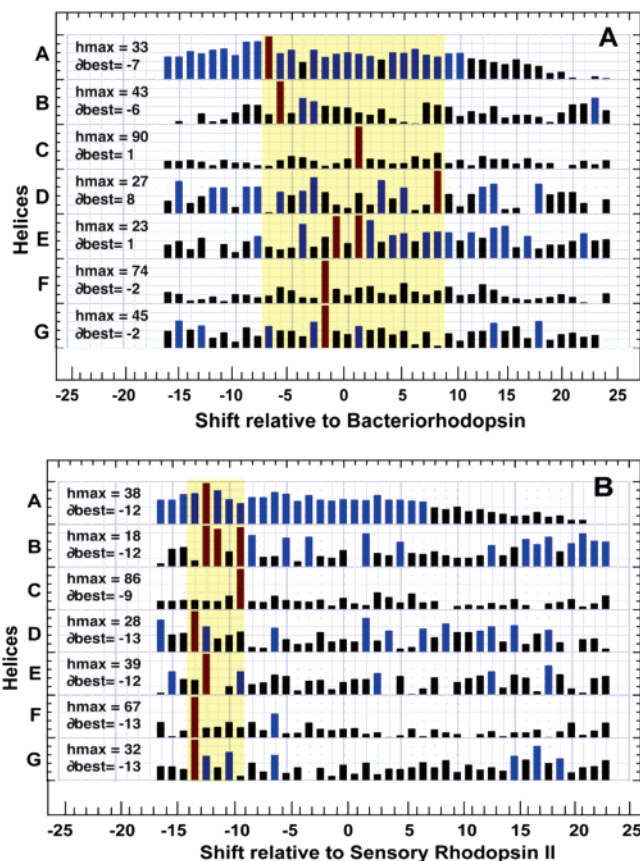


FIGURE 2: Homology analysis of the primary protein sequence of GPR (Swiss-Prot accession number Q9F7P4) with BR (PDB ID 1C3W) in panel A and SRII (PDB ID 1H68) in panel B using the BLOSUM40 scoring matrix. The different regions of the protein are shown along the y axis. To achieve maximum homology, the residues are shifted by a certain number of residues ( $\delta_{\text{best}}$ ) relative to the parent sequence represented along the x axis. Red bars indicate maximum homology, blue bars indicate good homology, and black bars indicate bad homology. The yellow box is representative of the number of shifts needed to achieve maximum homology. The net shift ( $\Sigma h_{\text{max}}$ ) was calculated for the BR- and SRII-based models.

resin. The proteorhodopsin concentration was determined from the absorption spectra assuming the values for  $\epsilon_{280}$  and  $\epsilon_{551}$  were the same as for the bR solubilized in a DMPC/cholate/SDS mixture at pH 8.0 (11, 12).

**Absorption Spectroscopy.** Absorption spectra of GPR were collected at pH values of 6 and 10 by resuspending the protein (200  $\mu$ L of 4.5 mg/mL Bac31A08) in 400  $\mu$ L of 0.1 M phosphate buffer (pH 6.0) and 0.1 M CAPS buffer (pH 10.0). Ultraviolet–visible spectra were collected at ambient temperature from 250 to 750 nm using a Cary 50 spectrophotometer (Varian, Inc., Palo Alto, CA).

**Homology Modeling.** The primary peptide sequence of GPR (Swiss-Prot accession number Q9F7P4) was aligned against the peptide sequences of the BR (PDB ID 1C3W) and SRII (PDB IDs 1H68 and 1JGJ) crystal structures using the BLOSUM40 scoring matrix (13–18). On the basis of the BLOSUM40 alignment, homology scores were assigned to each individual transmembrane region as shown in Figure 2. To achieve maximum homology, the initial 25 residues (MKLLLLILGSLVIALPTFAAGGGDLDA) or 26 residues (MKLLLLILGSLVIALPTFAAGGGDLDA) of the GPR sequence were deleted to align the sequence against BR and SRII. The total homology score was calculated for all the

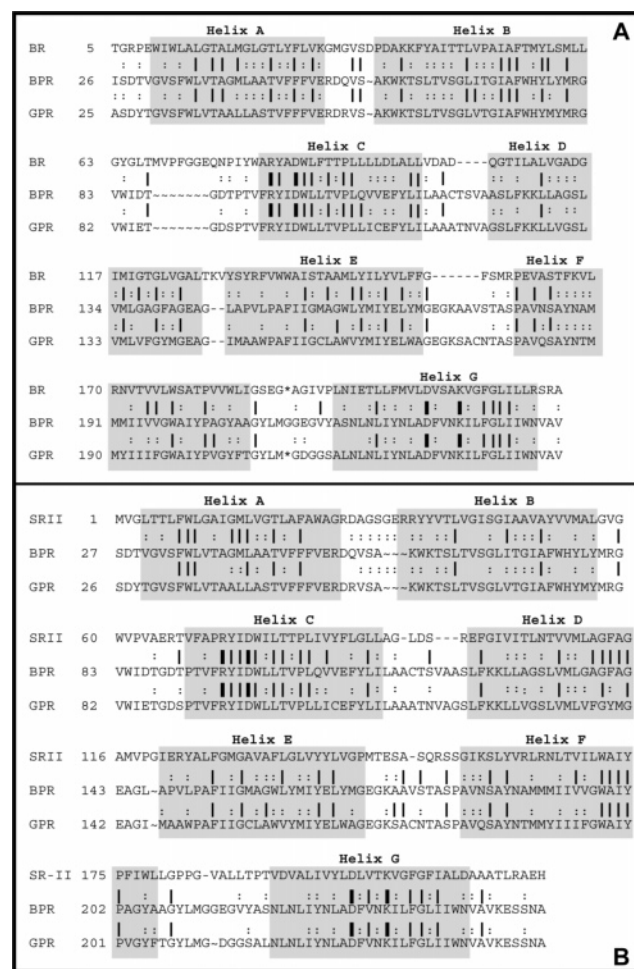


FIGURE 3: Linear sequence alignment of GPR and BPR with BR (A) and SR11 (B). The gray blocks represent the transmembrane regions of the protein based on the BR (PDB ID 1C3W) and SR11 (PDB ID 1H68) crystal structures. Linear alignment file between the template and the PR sequences: Highly conserved residues in retinylidene proteins are denoted with a bold vertical dash, conserved residues are represented with a vertical dash, and similar residues are represented with a colon. Residue deletions (~) and insertions (-) are introduced to maximize the homology score for the transmembrane regions.

transmembrane regions of the protein ( $\Sigma A-H$ ). Residue insertions into and deletions from the GPR sequence were consistent with the linear sequence alignment file. Bold vertical lines in the alignment file are indicative of highly conserved amino acid residues among all retinal binding proteins as shown in Figure 3. The Kyte–Doolittle plots generated for GPR relative to the template sequences revealed the presence of seven hydrophobic transmembrane regions (Figure 4) connected by hydrophilic loops. The GPR models (BR- and SR11-based) were generated following an initial sequence alignment in Cache 6.1.1 (Fujitsu, Inc., Beaverton, OR). Protonation states of key residues in GPR were maintained consistent with those observed in the BR resting state: Arg 94 (+), Asp 97 (-), Asp 227 (-), and Lys 231 (+). All ionic amino acid residues in GPR lacking accurate  $pK_a$  values were neutralized to avoid imprecise charge assignment (20). Energy minimizations (MM3 force field) employing steepest descent (SD) and conjugate gradient (CG) routines were carried out to a convergence criterion of 0.001 kcal/(mol Å) as described by Allinger (19, 20).

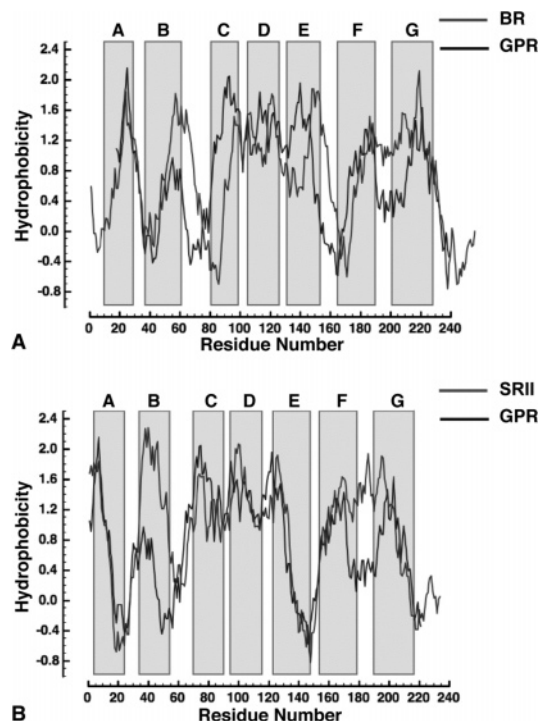


FIGURE 4: Kyte–Doolittle hydrophobicity plots generated for GPR relative to the BR (A) and SR11 (B) coding sequences. Gray blocks represent the predicted transmembrane regions for the GPR sequence, which correlates with that of the BR (PDB ID 1C3W) and SR11 (PDB ID 1H68) crystal structures. The initial 25 residues for the GPR sequence were omitted to maximize the homology score in alignment with the template sequence.

MM3 calculations were carried out on structures with a fixed peptide backbone and a cutoff distance of 9 Å.

Following an initial MM3 minimization, the loop regions of the GPR models were restored to the native GPR peptide sequence. Molecular mechanics based on the MM2 force field were carried out on the loop regions with the rest of the protein locked in a fixed state to ensure that the loops correspond to a stable conformation (19, 21, 22). GPR models were further optimized using ABNR (adopted basis Newton–Raphson) methods with harmonic constraints on the peptide backbone. The system was heated to 300 K for 20 ps followed by equilibration for 100 ps at 300 K. The system was then subjected to an NVT molecular dynamics (MD) simulation under vacuum for a total of 2 ns with a time step of 0.001 ps. Harmonic constraints were applied to the peptide backbone during the entire MD simulation. A nonbonded cutoff of 15 Å with a switch nonbonding smoothing function between 11 and 14 Å and a dielectric constant of 1.0 was used for all calculations. All molecular dynamics simulations were carried out using the CHARMm (c29b2) CHARMm22 parameter set for proteins and TIP3P for water molecules. Chromophore parameters were adapted from the optimized *all-trans*-retinyl-protonated Schiff base (PSB) parameters as published by Hermone and Kuczera (23).

The relative energies of the TM regions for the BR- and SR11-based minimized GPR models were calculated. A parameter structure file containing the required information for energy evaluation was created separately for the TM regions of each model using the same coordinates as for the 2 ns structure. The potential energies of the TM regions of



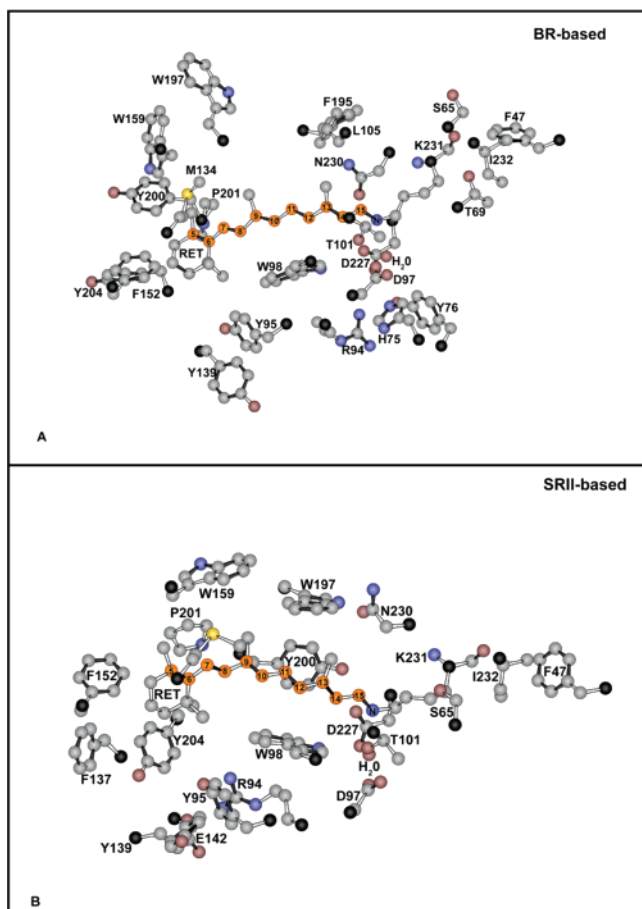


FIGURE 5: Retinal binding pocket of GPR based on the BR (A) and SR II (B) homology models. Shown are several aromatic residues lining the binding pocket extracted 5 Å from the chromophore retinal with several key residues, K231, D227, D97, and R94.

the different systems were calculated using the CHARMM energy function (24). H75K variant models of GPR were generated using InsightII (Accelrys, Inc., San Diego, CA). The models were optimized using the ABNR method with harmonic constraints on the peptide backbone. MD simulations were carried out on the variant models using the same procedure as the wild-type models.

**MNDO-PSDCI MO Theory.** The spectroscopic properties for the BR- and SR II-based retinal binding sites of GPR were calculated using the modified neglect of differential overlap with partial single- and double-configuration interaction (MNDO-PSDCI) molecular orbital theory (25, 26). PM3 parametrization with full single- and double-configuration interaction including the retinal  $\pi$  system, water molecules, and key residues (H75, R94, D97, D227, E142) 5 Å from the chromophore were included in the binding site simulation as shown in Figure 5 (26, 27).

## RESULTS

**Homology Modeling.** Aligning the primary sequence of GPR relative to BR and SR II using an in-house sequence alignment program, Protein 3.6 (available upon request from the corresponding author), yields a homology score of 0.27 for the transmembrane regions (BLOSUM40 scoring matrix). Due to the variable number of loop residues in the template sequence, the transmembrane regions of the protein are

Table 1: Relative Energies of GPR Calculated for BR-Based and SR II-Based Homology Models<sup>a</sup>

model	total energy	internal energy	van der Waals	electrostatic
GPR_BR (TM)	−13264 <sup>b</sup>	−13379 <sup>c</sup>	−187	−604
GPR_SR II (TM)	0 <sup>b</sup>	0 <sup>c</sup>	−264	−643
GPR_BR (min, TM)	−16645 <sup>b</sup>	−16764 <sup>c</sup>	−157	−630
GPR_SR II (min, TM)	−274 <sup>b</sup>	−442 <sup>c</sup>	−152	−586
GPR_BR (min, all)	−2042	2716	−396	−5081
GPR_SR II (min, all)	−1828	2706	−374	−4906

<sup>a</sup> All energies are in kilocalories per mole and were calculated after a 2 ns molecular dynamics simulation using CHARMM. When “TM” is indicated, the energies are calculated only for those residues in the transmembrane segments, whereas “all” indicates the entire protein. When “min” is indicated, the energies are for a 2 ns molecular dynamics simulation followed by a minimization. Models of GPR based on BR are listed as GPR\_BR, and those based on SR II are listed as GPR\_SR II. <sup>b</sup> Total energies relative to that of the GPR\_SR II (TM) model, which had the largest total energy of the TM structures. <sup>c</sup> Internal energies calculated relative to that of the GPR\_SR II (TM) model, which had the largest internal energy of the TM structures.

aligned individually. Residues in the GPR sequence are shifted ( $\delta_{\text{best}}$ ) relative to the template sequence to maximize the homology score ( $h_{\text{max}}$ ) as shown in Figures 2 and 3. The net shift for the BR- and SR II-based models shows that the BR-based model requires lesser shifts than the SR II-based model although the overall bandwidth is broader in the BR-based model as shown in Figure 2. Kyte–Doolittle plots generated for GPR (Figure 4) reveal the presence of seven hydrophobic transmembrane regions consistent with other retinylidene proteins.

Following an initial sequence alignment in Cache 6.1.1, the models are subjected to a series of energy minimization and molecular dynamics simulations. The relative potential energies for the TM regions of the 2 ns dynamics and 2 ns minimized structures are calculated for the BR- and SR II-based models. The energies calculated for the BR-based structure are significantly lower in comparison to those calculated for the SR II-based structure (Table 1). Careful examination of the BR- and SR II-based GPR binding pocket further demonstrates that the SR II-based model collapses following a 2 ns dynamics simulation. The collapse of the SR II-based model is attributed to the formation of a strong salt bridge between the chromophore and Asp 227. Distances measured between the Schiff base nitrogen (NZ) of the chromophore and the carboxyl oxygen atoms on D97 and D227 (OD1 and OD2) show that Asp 227 (OD1) is separated from the PSB linkage by 2.95 Å in contrast to the same atom in the BR-based model (4.13 Å). Molecular dynamics simulations on the 1H68 crystal structure of SR II produces a more accurate estimation of the binding pocket with Asp 201 (OD1) separated from NZ by a distance of 4.56 Å (Table 2). The calculated absorption spectra of GPR models based on BR and SR II support the fact that Arg 94 in GPR is in an upward orientation. Another major observation in the BR-based model is the proximity of His 75 (~5 Å) to the quadrupole residues contributed by PSB, Asp 97, Asp 227, and Arg 94 relative to that of the SR II-based model (10 Å) as shown in Figure 5 and Table 2. Previous mutagenesis studies on BAC31A8 have identified that replacement of His 75 with lysine eliminated the pH dependence property of the protein on the absorption maxima (pH range studied 4.9–9.1). H75K-GPR models generated in InsightII further

Table 2: Distances Measured between NZ and the Carboxyl Oxygens OD1 and OD2 of the Counterion Complex<sup>a</sup>

protein/atoms	distance (Å)		stability
	crystal structure	after a 2 ns MD simulation	
BR (1C3W)			+
NZ (Ret)—OD1 D85	3.79	3.44	
NZ (Ret)—OD2 D85	3.38	4.55	
NZ (Ret)—OD1 D227	3.74	4.61	
NZ (Ret)—OD2 D227	4.87	5.97	
SRII (1H68)			+
NZ (Ret)—OD1 D75	4.07	5.03	
NZ (Ret)—OD2 D75	4.28	3.78	
NZ (Ret)—OD1 D201	4.05	4.56	
NZ (Ret)—OD2 D201	4.63	5.28	
BR-GPR-E142 (+)			+
NZ (Ret)—OD1 D97		4.34	
NZ (Ret)—OD2 D97		3.78	
NZ (Ret)—OD1 D227		4.13	
NZ (Ret)—OD2 D227		6.08	
SRII-GPR-E142 (—)			—
NZ (Ret)—OD1 D97		5.49	
NZ (Ret)—OD2 D97		4.81	
NZ (Ret)—OD1 D227		3.02	
NZ (Ret)—OD2 D227		4.23	
SRII-GPR-E142 (+)			—
NZ (Ret)—OD1 D97		5.43	
NZ (Ret)—OD2 D97		4.97	
NZ (Ret)—OD1 D227		2.95	
NZ (Ret)—OD2 D227		4.36	
1JGJSRII-GPR-E142 (—) <sup>b</sup>			—
NZ (Ret)—OD1 D97		5.39	
NZ (Ret)—OD2 D97		7.43	
NZ (Ret)—OD1 D227		2.66	
NZ (Ret)—OD2 D227		4.87	

<sup>a</sup> Interatomic residue distances were calculated for the crystal structures prior to and after the 2 ns molecular dynamics simulation. The model stability is represented with a “+” and instability due to the collapse of the binding site as “—”. <sup>b</sup> Interatomic residue distances for the 1JGJ SRII-based GPR model were calculated after a 600 ps molecular dynamics simulation.

corroborate this observation (details below). From our theoretical models we conclude that the SRII-based model is neither adequate in terms of stability nor rational in explaining the spectral tuning and pH dependence of the protein.

**Molecular Orbital Calculations.** The electronic transitions observed for the BR- and SRII-based wild-type GPR and the H75K-GPR models are superimposed on the absorption spectra of GPR at low (6) pH (Figure 7A,B) and high (10) pH (Figure 7C). The calculated absorption spectra predict a strongly allowed state that is blue-shifted relative to the experimental  $\lambda_{\text{max}}$ . This observation is ascribed to the inadequacy of the semiempirical methods in the treatment of water molecules (position and orientation) in the binding pocket. MNDO-PSDCI calculations on the BR-based model provide a level ordering consistent with most retinal polyenes (26, 28). However, the calculated energy of the strongly allowed  $^1\text{B}_u^{*+}$ -like state and the level ordering in the SRII-based GPR model are unrealistic (Figure 7A).

## DISCUSSION

**BR- and SRII-Based Homology Models.** Relative stabilities of the BR- and SRII-based GPR models are analyzed by comparing the potential energies calculated for the 2 ns structures and 2 ns minimized structures. As shown in Table

1, the minimized transmembrane region of the BR-based model possesses a lower potential energy in contrast to that of the SRII-based model. The transmembrane regions of the BR- and SRII-based structures were utilized in the energy calculations taking into account the variable number of loop residues in the BR (1C3W) and SRII (1H68) crystal structures. We conclude that BR provides a more realistic three-dimensional homology parent for GPR than that provided by SRII. The 1H68 structure of sensory rhodopsin II was used as a template in the homology modeling study because of the high resolution of the crystal structure (2.1 Å) relative to that of the 1JGJ structure of SRII (2.4 Å) (29). However, we also note that there are issues with the 1H68 structure. The hypothetical chloride ion binding site in the 1H68 crystal structure is absent from the 1JGJ structure of SRII. This difference is attributed to the protonation of Asp 193 in the 1H68 crystal structure with altered hydrogen-bonding interaction with the carbonyl oxygen on Ile 177. Protonated Asp 193 facilitates the binding of a chloride ion to the guanidinium side chain of Arg 72. These structural changes alter the local environment around Arg 72 between the two crystal structures (18, 30). Therefore, we also carried out simulations based on the 1JGJ structure for comparative purposes. As noted below, we found no significant differences in the simulations. This observation can be traced to the fact that the backbones have rms deviations of less than 0.44 Å. In contrast, our previous homology modeling study of BPR based on BR (1C3W) and SRII (1H68 and 1JGJ) yielded collapse when using BR but stability when using either of the two SRII structures (28). Thus, the collapse of both SRII-based models in the present study reflects on their inappropriateness for use as homology models for GPR, not on any inherent problems with the crystal structures. More importantly, there are additional advantages of the BR-based model in terms of explaining the spectroscopic and pH dependence properties of this protein. We explore this issue below.

Analysis of the retinal binding pockets in the two GPR models revealed significant differences regarding the position and orientation of the charged residues Asp 227 and Arg 94 relative to the chromophore. After a 2 ns dynamics simulation on the SRII-based model, Asp 227 and the Schiff base collapsed into each other to form a close-contact salt bridge. The nitrogen atom in the Schiff base is separated from the carboxyl oxygen atom on Asp 227 by a distance of only 2.95 Å. This distance is significantly shorter than those observed between NZ and the carboxyl oxygen atoms on the counterion residues in the 1C3W and 1H68 crystal structures of BR and SRII, respectively (data shown in Table 2). The GPR model generated on the basis of the 1JGJ structure behaved similarly to that of the 1H68-based structure and was unstable to extended MD simulation. The binding site of the 1JGJ-based GPR model collapsed with the carboxyl oxygen on Asp 227 moving closer (2.66 Å) to the protonated Schiff base linkage (NZ) relative to the same atom in the BR-based GPR model. This collapse occurred despite the fact that water was included in the binding site. MNDO-PSDCI calculations on the BR- and SRII-based GPR models provide additional perspective. The calculated transition energy for the  $^1\text{B}_u^{*+}$ -like state for the SRII-based GPR model is unrealistic (Figure 7) and can be traced to the collapse of the binding pocket in the SRII-based structure.

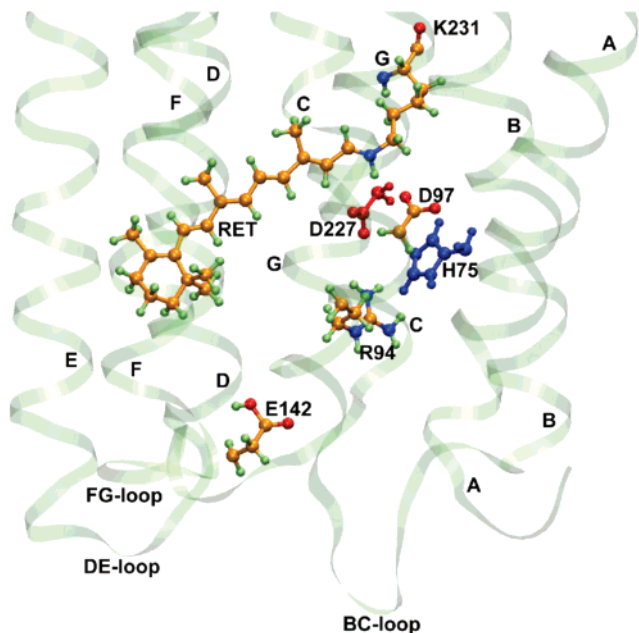


FIGURE 6: Ribbon representation of the BR-based GPR model. Highlighted are several key residues. Note that H75 is located in close proximity to the quadrupole residue complex (Schiff base, D97, D227, and R94).

In contrast, the BR-based homology model provides a more realistic binding pocket with Asp 227 separated from the Schiff base nitrogen by a distance of 4.13 Å. The calculated absorption maximum of the BR-based model ( $\lambda_{\text{max}} = 482$  nm) is in reasonable agreement with the experimental data ( $\lambda_{\text{max}} = 520$  nm). The calculated  $\lambda_{\text{max}}$  for the SR11-based structure is considerably blue-shifted (363 nm) relative to that of the BR-based model (482 nm), suggesting the collapse of the binding pocket, and is similar to the absorption properties of an unprotonated Schiff base (31). The SR11-based models generated by using both 1H68 and 1JGJ were qualitatively identical and unrealistic.

**Orientation of the Guanidinium Side Chain of Arg 94 in GPR.** The two GPR models also differed regarding the orientation of Arg 94. Previous studies have identified that the position of the positively charged amino acid residue in the quadrupole complex of BR (Arg 82) and SR11 (Arg 72) and the electrostatic environment influenced by its interaction with the charged counterion complex are responsible for the wavelength regulation of the protein (26, 32). Following a 2 ns molecular dynamics simulation, the guanidinium side chain of Arg 94 stabilizes with an orientation toward the Schiff base linkage of the chromophore in the BR-based model. In contrast, the SR11-based GPR model yields a guanidinium side chain pointed away from the Schiff base. This orientation is comparable to that found in the BPR model proposed by Hillebrecht et al. (28) where the side chain of the arginine residue under alkaline conditions is oriented away from the Schiff base. In particular, Arg 95 in BPR (Arg 94 in GPR) under alkaline conditions interacts with Glu 143 (Glu 142 in GPR), forming a salt bridge, a situation that closely resembles that of the binding pocket environment in NpSR11 (Figures 5A and 6). However, our study on GPR indicates that a comparable arrangement in GPR leads to instability and collapse of the binding site. This finding further supports our conclusion that Arg 94 in GPR is pointing toward the PSB.

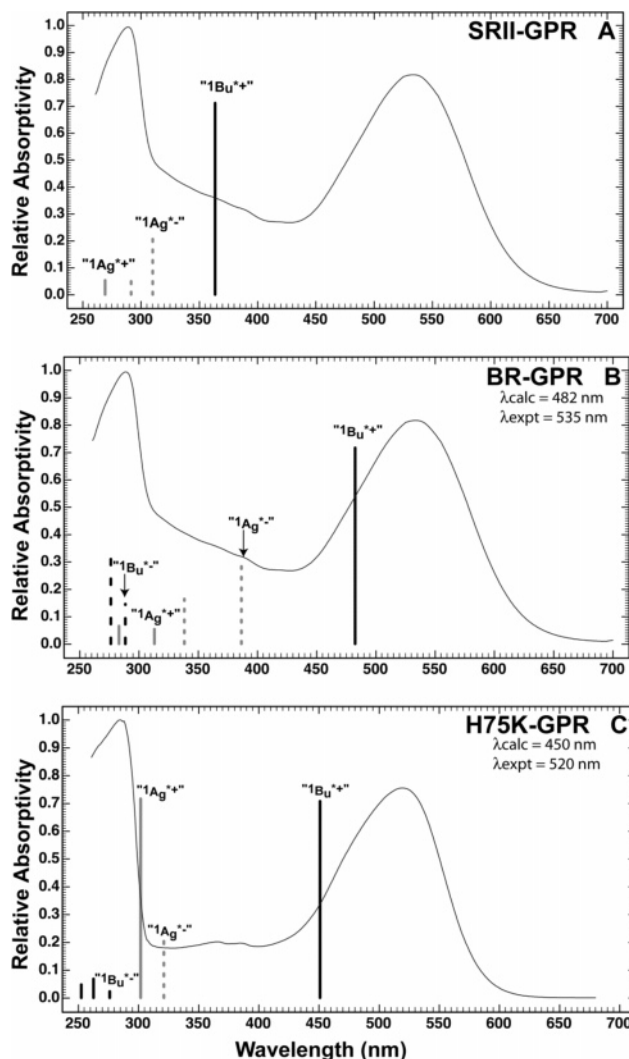


FIGURE 7: Calculated energies, relative oscillator strengths, and approximate symmetries of the low-lying  $\pi\pi^*$  states of the chromophore have been superimposed on the absorption spectra at low pH (B) and high pH (C). The height of the bars is representative of the relative oscillator strength. The  $1B_u^{*+}$ -like state for the SR11-based model (A) does not accurately predict the absorbance maximum of the protein in comparison to that of the BR-based model (B). The  $1B_u^{*+}$ -like state for the H75K variant is considerably blue-shifted relative to that of the WT GPR (C).

**Role of His 75 in Spectral Tuning in GPR.** The absorption spectra of GPR are pH-dependent with a red-shifted spectral species absorbing at  $\sim 540$  nm at low pH (acidic) and a blue-shifted species absorbing at  $\sim 520$  nm in a high-pH (basic) solution. The histidine residue at position 75 is near the quadrupole residue complex in the BR-based GPR model as shown in Figure 6 and Table 3. On the basis of our study, we predict that histidine 75 in GPR is primarily unprotonated under physiological pH conditions ( $pK_{a,2} = 6.0$ ). Titrating GPR produces a spectral shift of  $\sim 18$  nm consistent with the protonation and reprotonation of H75 in the pH range studied (pH 4.9–9.1). Previous experimental data suggested that replacement of His 75 with lysine in GPR completely abolished the pH dependency of the protein on its absorption maximum (33). The absorption maximum of the H75K-GPR variant at  $\sim 522$  nm is pH-independent (pH 4.5–9.2). Our calculations on the H75K-GPR model predict that the butyl ammonium side chain of lysine is predominantly protonated ( $pK_{a,2} = 9.0$ ). This finding supports the pH independence



Table 3: Calculated Distances between Residue 75 and the Counterion Complex (D97, D227) in the BR- and SRII-Based GPR Models

protein/atoms	distance (Å)	
	WT GPR	H75K-GPR
BR-GPR <sup>a</sup>		
75-OD1 D97	4.96	4.97
75-OD2 D97	3.82	4.51
75-OD1 D227	4.71	3.20
75-OD2 D227	5.16	3.58
SRII-GPR <sup>b</sup>		
75-OD1 D97	10.08	10.76
75-OD2 D97	9.66	9.98
75-OD1 D227	14.82	14.45
75-OD2 D227	13.08	12.40

<sup>a</sup> GPR generated on the basis of BR as the homology parent and after a 2 ns molecular dynamics simulation. <sup>b</sup> GPR generated on the basis of SRII as the homology parent and after a 2 ns molecular dynamics simulation.

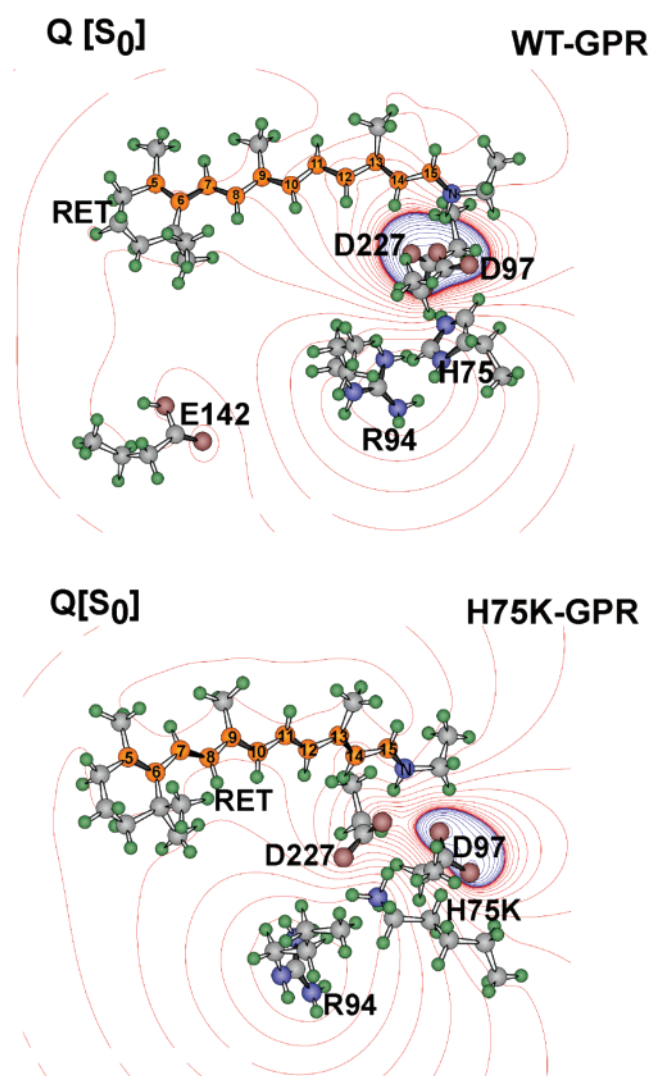


FIGURE 8: Potential model for the change in absorption as a function of pH. In wild-type GPR, His 75 is predominantly neutral, and hence, WT GPR exists in two spectral forms, a red-shifted acidic form and a blue-shifted basic form. Replacement of H75 with lysine abolishes the pH dependence, causing the protein to exist exclusively in the basic form.

exhibited by this variant form of GPR in the pH range studied. MNDO-PSDCI calculations of the GPR and H75K-

GPR binding sites reveal that the absorption maximum in the H75K variant is blue-shifted (450 nm) relative to that of wild-type GPR (480 nm) and that the H75K variant predominantly exists as a single spectral form (Figure 7). This finding is consistent with experimental studies indicating that the wild-type form of GPR absorbs maximally at 540 nm under acidic conditions and at 522 nm under basic conditions.

The electrostatic environment of the retinal binding pocket is significantly altered in the H75K variant in comparison to wild-type GPR (Figure 8). Distances measured between His 75 and Arg 94 and the counterion residues Asp 97 and Asp 227 predict that, in the H75K variant, Asp 97 moves closer to the basic amino acid residues (His 75 and Arg 94) while Asp 227 moves closer to the protonated Schiff base, altering the electrostatic environment in the binding site (data not shown). In the SRII-based H75K variant, His 75 is separated from the counterion complex by 10.00 Å (Figure 5B and Table 3). Because we know from the site-directed studies that His 75 is a critical residue in facilitation of a pH dependence of the absorption maximum of GPR, we conclude that only the BR-based model of GPR is capable of providing insight into the structure–function relationship responsible for the pH-dependent spectral properties.

## COMMENTS AND CONCLUSIONS

From this study, we conclude that the BR-based model of GPR is significantly more realistic than the SRII-based model in terms of structural stability, spectroscopic properties, and the pH dependence of the spectroscopic properties. The position of His 75 is predicted to be in close proximity to the binding site in the BR-based model, and this model is thus capable of rationalizing the molecular origins of the pH dependence and the reason H75K has pH-independent spectral properties. Another key difference between the BR-based and SRII-based models involves the orientation of Arg 94, and the molecular dynamics simulations conclude that this residue is pointing up into the binding site to yield a counterion quadrupole environment similar to that observed in BR. Our present study on GPR is in contrast to our previous conclusions regarding BPR, which was modeled much more effectively by using SRII as the homology parent. Thus, GPR and BPR must be viewed as distinctly different proteins despite significant homology. This observation reopens the question of whether both are proton pumps or whether BPR may be a photosensory pigment in vivo. Alternatively, the mechanism of spectral tuning may be the defining variable, and the fact that BR is a better model for GPR and SRII a better model for BPR may reflect a common mechanism for spectral tuning in both bacterial and archaeal proton pumps. These interesting and important issues remain open to interpretation and should stimulate further study.

## REFERENCES

1. Béjà, O., Aravind, L., Koonin, E. V., Suzuki, M. T., Hadd, A., Nguyen, L. P., Jovanovich, S. B., Gates, C. M., Feldman, R. A., Spudich, J. L., Spudich, E. N., and DeLong, E. F. (2000) Bacterial rhodopsin: Evidence for a new type of phototrophy in the sea, *Science (Washington, D.C.)* 289, 1902–1906.
2. Sharma, A. K., Spudich, J. L., and Doolittle, W. F. (2006) Microbial rhodopsins: functional versatility and genetic mobility, *Trends Microbiol.* 14, 463–469.

3. B  j  , O., Spudich, E., Spudich, J., Leclerc, M., and DeLong, E. (2001) Proteorhodopsin phototrophy in the ocean, *Nature* **411**, 786–789.
4. Man, D., Wang, W., Sabehi, G., Aravind, L., Post, A. F., Massana, R., Spudich, E. N., Spudich, J. L., and Beja, O. (2003) Diversification and spectral tuning in marine proteorhodopsins, *EMBO J.* **22**, 1725–1731.
5. Sabehi, G., Beja, O., Suzuki, M. T., Preston, C. M., and DeLong, E. F. (2004) Different SAR86 subgroups harbour divergent proteorhodopsins, *Environ. Microbiol.* **6**, 903–910.
6. Friedrich, T., Geibel, S., Kalmbach, R., Chizhov, I., Ataka, K., Heberle, J., Englehard, M., and Bamberg, E. (2002) Proteorhodopsin is a light-driven proton pump with variable vectoriality, *J. Mol. Biol.* **321**, 821–838.
7. Gomez-Consarnau, L., Gonz  lez, J. M., Coll-Llado, M., Gourdon, P., Pascher, T., Neutze, R., Pedros-Alio, C., and Pinhassi, J. (2007) Light stimulates growth of proteorhodopsin-containing marine flavobacteria, *Nature* **445**, 210–213.
8. Bergo, V., Amsden, J. J., Spudich, E. N., Spudich, J. L., and Rothschild, K. J. (2004) Structural changes in the photoactive site of proteorhodopsin during the primary photoreaction, *Biochemistry* **43**, 9075–9083.
9. Dioumaev, A., Brown, L., Shih, J., Spudich, E., Spudich, J., and Lanyi, J. (2002) Proton transfers in the photochemical reaction cycle of proteorhodopsin, *Biochemistry* **41**, 5348–5358.
10. Wang, O. A. S., Wei-Wu, Spudich, Elena N., and Spudich, John L. (2003) Spectroscopic and photochemical characterization of a deep ocean proteorhodopsin, *J. Biol. Chem.* **278**, 33985–33991.
11. Krebs, R. A., Alexiev, U., Partha, R., Devita, A. M., and Braiman, M. S. (2002) Detection of fast light-activated H<sup>+</sup> release and M intermediate formation from proteorhodopsin, *BMC Physiol.* **2**, No. 5.
12. London, E., and Khorana, H. G. (1982) Denaturation and renaturation of bacteriorhodopsin in detergents and lipid-detergent mixtures, *J. Biol. Chem.* **257**, 7003–7011.
13. Henikoff, S., and Henikoff, J. G. (1992) Amino acid substitution matrices from protein blocks, *Proc. Natl. Acad. Sci. U.S.A.* **89**, 10915–10919.
14. Henikoff, S., Henikoff, J. G., and Pietrokovski, S. (1999) Blocks+: a non-redundant database of protein alignment blocks derived from multiple compilations, *Bioinformatics* **15**, 471–479.
15. Teodorescu, O., Galor, T., Pillardy, J., and Elber, R. (2004) Enriching the sequence substitution matrix by structural information, *Proteins: Struct., Funct., Bioinf.* **54**, 41–48.
16. Luecke, H., Schobert, B., Richter, H. T., Cartailler, J. P., and Lanyi, J. K. (1999) Structure of bacteriorhodopsin at 1.55   resolution, *J. Mol. Biol.* **291**, 899–911.
17. Royant, A., Nollert, P., Edman, K., Neutze, R., Landau, E. M., Pebay-Peyroula, E., and Navarro, J. (2001) X-ray structure of sensory rhodopsin II at 2.1-  resolution, *Proc. Natl. Acad. Sci. U.S.A.* **98**, 10131–10136.
18. Spudich, J., and Luecke, H. (2002) Sensory rhodopsin II: functional insights from structure, *Curr. Opin. Struct. Biol.* **12**, 540–546.
19. Allinger, N. L., and Burkert, U. (1982) *Molecular Mechanics*, American Chemical Society, Washington, DC.
20. Allinger, N. L., Yuh, Y. H., and Lii, J. H. (1989) Molecular mechanics. The MM3 force field for hydrocarbons, *J. Am. Chem. Soc.* **111**, 8551–8566.
21. Allinger, N. L. (1977) Conformational analysis. 130. MM2. A hydrocarbon force field utilizing V1 and V2 torsional terms, *J. Am. Chem. Soc.* **99**, 8127–8134.
22. Allinger, N. L., Kuang, J., and Thomas, H. D. (1990) Molecular mechanics (MM2 and MM3) calculations on aliphatic and aromatic nitro compounds, *THEOCHEM* **68**, 125–148.
23. Hermone, A., and Kuczera, K. (1998) Free-energy simulations of the retinal cis-trans isomerization in bacteriorhodopsin, *Biochemistry* **37**, 2843–2853.
24. Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S., and Karplus, M. (1983) CHARMM: a program for macromolecular energy, minimization, and dynamics calculations, *J. Comput. Chem.* **4**, 187–217.
25. Martin, C. H., and Birge, R. R. (1998) Reparameterizing MNDO for excited state calculations using ab initio effective Hamiltonian theory: Application to the 2,4-pentadien-1-iminium cation, *J. Phys. Chem. A* **102**, 852–860.
26. Ren, L., Martin, C. H., Wise, K. J., Gillespie, N. B., Luecke, H., Lanyi, J. K., Spudich, J. L., and Birge, R. R. (2001) Molecular mechanism of spectral tuning in sensory rhodopsin II, *Biochemistry* **40**, 13906–13914.
27. Shima, S., Ilagan, R. P., Gillespie, N., Sommer, B. J., Hiller, R. G., Sharples, F. P., Frank, H. A., and Birge, R. R. (2003) Two-photon and fluorescence spectroscopy and the effect of environment on the photochemical properties of peridinin in solution and in the peridinin-chlorophyll-protein from *Amphidinium carterae*, *J. Phys. Chem. A* **107**, 8052–8066.
28. Hillebrecht, J. R., Galan, J., Rangarajan, R., Ramos, L., McCleary, K., Ward, D. E., Stuart, J. A., and Birge, R. R. (2006) Structure, function and wavelength selection in blue-absorbing proteorhodopsin, *Biochemistry* **45**, 1579–1590.
29. Luecke, H., Schobert, B., Lanyi, J. K., Spudich, E. N., and Spudich, J. L. (2001) Crystal structure of sensory rhodopsin II at 2.4   resolution: insights into color tuning and transducer interaction, *Science (Washington, D.C.)* **293**, 1499–1503.
30. Gordeliy, V. I., Labahn, J., Moukhametzianov, R., Efremov, R., Granzin, J., Schlesinger, R., Buldt, G., Savopol, T., Scheidig, A. J., Klare, J. P., and Englehard, M. (2002) Molecular basis of transmembrane signalling by sensory rhodopsin II-transducer complex, *Lett. Nat.* **419**, 484–487.
31. Kusnetzow, A., Dukkupati, A., Babu, K. R., Ramos, L., Knox, B. E., and Birge, R. R. (2004) Vertebrate ultraviolet visual pigments: protonation of the retinylidene Schiff base and a counterion switch during photoactivation, *Proc. Natl. Acad. Sci. U.S.A.* **101**, 941–946.
32. D  r, A., Sz  raz, S., T  th-Bocon  di, R., Tokaji, Z., Keszthelyi, L., and Stoeckenius, W. (1991) Alternative translocation of protons and halide ions by bacteriorhodopsin, *Proc. Natl. Acad. Sci. U.S.A.* **88**, 4751–4755.
33. Jensen, R. B., and Kelemen, B. (2005) Freepatentsonline pp 1–241, Ellicott City, MD.

BI700955X