JCTC Journal of Chemical Theory and Computation

Theoretical Studies on the Color-Tuning Mechanism in Retinal Proteins

Kazuhiro Fujimoto,[†] Shigehiko Hayashi,[‡,§] Jun-ya Hasegawa,[†,||] and Hiroshi Nakatsuji*[,†,||]

Department of Synthetic Chemistry and Biological Chemistry, Graduate School of Engineering, Kyoto University, Nishikyou-ku, Kyoto 615-8510, Japan, PRESTO, Japan Science and Technology Agency, Kawaguchi, Saitama, Japan, Department of Chemistry, Graduate School of Science, Kyoto University, Kitashirakawa-Oiwake-cho, Sakyou-ku, Kyoto 606-8520, Japan, and Quantum Chemistry Research Institute (QCRI), 58-8 Mikawa, Momoyama-cho, Fushimi-ku, Kyoto 612-8029, Japan

Received August 16, 2006

Abstract: The excited states of the three retinal proteins, bovine rhodopsin (Rh), bacteriorhodopsin (bR), and sensory rhodopsin II (sRII) were studied using the symmetry-adapted cluster-configuration interaction (SAC-CI) and combined quantum mechanical and molecular mechanical (QM/MM) methods. The computed absorption energies are in good agreement with the experimental ones for all three proteins. The spectral tuning mechanism was analyzed in terms of three contributions: molecular structures of the chromophore in the binding pockets, electrostatic (ES) interaction of the chromophore with the surrounding protein environment, and quantum-mechanical effect between the chromophore and the counterion group. This analysis provided an insight into the mechanism of the large blue-shifts in the absorption peak position of Rh and sRII from that of bR. Protein ES effect is primarily important both in Rh and in sRII, and the structure effect is secondary important in Rh. The quantum-mechanical interaction between the chromophore and the counterion is very important for quantitative reproduction of the excitation energy. These results indicate that the present approach is useful for studying the absorption spectra and the mechanism of the color tuning in the retinal proteins.

1. Introduction

The rhodopsin family of photoreceptors is among the best characterized membrane proteins. These proteins have a seven-transmembrane helical structure and function as photosensing and ion-pumps.[1] In a common visual photoreceptor, rhodopsin (Rh), the retinal chromophore shows the photoisomerization from 11-cis to all-trans forms. This reaction leads the protein to a signaling state, which is amplified biochemically through interaction with the G protein transducin.[2] In bacteriorhodopsin (bR) of *Halobacterium salinarum*, the photoisomerization of retinal from all-trans to 13-cis forms establishes an electrochemical gradient across the membrane and serves as a unidirectional proton transport.

These receptors consist of an apoprotein (opsin) and a retinal chromophore which is covalently bound to the apoprotein via a lysine residue by a protonated Schiff base (PSB) linkage. While the PSB form of retinal absorbs at about 440 nm in organic solvents, its maximal absorption (λ_{max}) drastically changes after binding to the apoprotein (opsin), which is known as "opsin shift".[3] The absorption maxima is regulated by opsin and widely spreads from 360 to 635 nm[4] to furnish the photoreceptors with color sensitivity, whereas the proteins include a common identical chromophore, retinal. In Rh environment, the chromophore has 11-cis form, and it is in an all-trans form in bR and sRII.

* Corresponding author e-mail: hiroshi@sbchem.kyoto-u.ac.jp.
† Department of Synthetic Chemistry and Biological Chemistry, Graduate School of Engineering, Kyoto University.
‡ PRESTO, Japan Science and Technology Agency.
§ Department of Chemistry, Graduate School of Science, Kyoto University.
|| Quantum Chemistry Research Institute (QCRI).

10.1021/ct6002687 CCC: $37.00
Published on Web 01/03/2007

The diversity of absorption maxima of rhodopsin has been investigated extensively,[5] and several explanations have been proposed.[6]

The spectral tuning mechanism can be analyzed in terms of the following three contributions. The first one is the chromophore-structural origin. Retinal would be distorted in the protein in order to accommodate the chromophore itself in the binding pocket. In fact, it has been shown by X-ray crystallographic studies that the polyene chain of the chromophore in Rh is strongly twisted, whereas those in bR and sRII are nearly planer.[7−9] Such torsion of the polyene chain is expected to sensitively alter the absorption energy, since the torsion disrupts conjugation of the π orbitals responsible for the excitation.

Second, ES (electrostatic) interaction between the chromophore and the surrounding protein environment could play a crucial role in the spectral tuning.[10−14] Mutagenesis experiments and theoretical analyses have shown that the absorption energies are strongly affected by the charged residues such as counter negative ion groups in the binding pocket.[15] The ES tuning mechanism exploits a characteristic in the excitation property of the PSB retinal molecule, i.e., significant change in the positive-charge distribution in the PSB along the polyene chain upon the excitation.[16] The positive charge redistribution upon the excitation creates a difference in ES interaction with the surrounding polar groups between the ground and excited states, which gives rise to the spectral shift. Especially, the counter negative ion groups contribute dominantly to the ES energy change upon the excitation so that mutual geometries of the PSB and the counterion groups are suggested to be one of the key factors for determining the absorption maximum.[10,15]

The third contribution includes higher-order interactions with the protein surroundings such as electronic polarization and charge-transfer interactions. The ES interaction described above induces the electronic polarization in both the chromophore and the protein surroundings. The electronic polarization enhances the ES interaction of the chromophore with the surroundings, leading to a shift of the absorption energy.[17−19] In addition to the electronic polarization, there is a strong electronic charge transfer interaction between the PSB and the counterion group,[20] which is suggested to significantly increase the absorption energy.[11,12,20]

In order to identify physical mechanism of the color tuning in the retinal proteins, several computational investigations have been performed by using modern quantum-chemistry methodologies.[10,11,20−23] Recent advances in computational technique have realized to predict the absorption energy of chromophore in protein whose X-ray crystallographic structures[7−9] have been solved. The computational studies mentioned above reported that the absorption maxima are reasonably reproduced with their own approaches.

However, the proposed mechanisms underlying the absorption spectra are still different in those studies, and they have not reached to consensus on the mechanism (see below). This difference arise from the fact that definitive elucidation of the underlying mechanism requires high methodological accuracy for *all* of the contributing factors, i.e., the electronic wave functions of the states involved, the protein structure including the chromophore, and the interaction of the chromophore with the surrounding environment. Errors in any of those contributing factors introduce ambiguity in the definitive determination of the molecular mechanism.

To accurately calculate the electronic energy for the ionic $\pi-\pi^*$ excited state of a polyene-like molecule, dynamic electron-correlation due to strong $\sigma-\pi$ polarization should be included appropriately.[24] In addition, the positive-charge migration on the Schiff base upon the excitation makes the electronic structures more complex. Various calculations have been performed so far by using modern methodologies in quantum chemistry, multireference (MR) perturbation theory (PT),[11,13,20,22,25,26] MR-configuration interaction (CI) method,[12] time-dependent density functional theory (TD-DFT),[23] and symmetry-adapted cluster-CI (SAC-CI) method.[14] Wanko et al. showed that TD-DFT gives qualitatively different excitation energy from the other methods when the C_6-C_7 bond is rotated.[12] This result indicates the rather complex nature of the excited-state wave function, which therefore requires careful and extensive treatment of the electron correlation.

Second, the absorption energy is highly sensitive to the structure of chromophore and protein. Wanko et al.[12] have shown that the absorption energy strongly depends on strength of bond alternation of the polyene chain of the chromophore, in addition to the torsional angle of the polyene chain as mentioned above. For example, the computed absorption energy at the geometry optimized by using the Hartree−Fock (HF) method significantly overestimates the experimental result,[10,20] which is due to too strong bond alternation in the HF geometry.[12] For proper description of the bond alternation of the chromophore, the dynamic electron-correlation should be taken into account. The second-order Moller−Plesset (MP2) perturbation theory or DFT was suggested for the geometry optimization.[12]

Geometry refinement of the protein surroundings is also crucial for the absorption energy calculation. As described above, the interaction of the chromophore with the surrounding protein is essential to the spectral tuning. Especially, it has been suggested that mutual distance between the PSB and its counterion group is one of the main factors for controlling the absorption maximum, and a small difference in the distance less than 1 Å could cause a significant change in the observed spectrum.[10,15]

However, spatial resolutions of X-ray crystallographic structures of the retinal proteins solved so far are more than 1.5 Å and hence are not sufficient to detect decisive structural differences completely. Thus it is necessary to refine the X-ray crystallographic protein structures before the absorption energies are computed. The geometry refinements for the entire protein structures were carried out by means of various computational approaches, such as semiempirical AM1[19,27] and tight binding DFT (DFTB) methods,[12,26,28,29] and hybrid ab initio quantum mechanical/molecular mechanical (QM/MM) methods (see below),[10,14,20] although in some studies the protein surroundings around the reaction center is kept fixed at the original coordinates of X-ray crystallographic models.[13,22]

Finally, the interaction of the chromophore with the surrounding has to be described properly. As described

above, the ES interaction plays a primary role in the color tuning. In order to describe the microscopic ES interaction, the QM/MM method is often employed, in which the ES field of the protein surrounding is represented with the effective point charges of the MM force field. However, the treatment using the point charges lacks the higher-order electronic effects of the protein surroundings such as the electronic polarization and the charge-transfer described above. Especially, the quantum-mechanical interaction of the chromophore with the counterion group contributes considerably to the absorption energies (0.2−0.5 eV)[11,12,20] and therefore has to be taken into account in the quantitative calculation.

In the present study, we have performed ab initio QM/MM and SAC-CI calculations to compute the absorption energies of three retinal proteins, bR, sRII, and Rh. The absorption maxima of sRII (497 nm)[30] and Rh (498 nm)[1] are largely blue-shifted by 70 nm compared with that of bR (568 nm).[31] It is interesting to note that the absorption maxima of sRII and Rh are similar to each other, even though the structures of the chromophores and the protein surroundings are distinctly different.[7,9] There must be different mechanisms for the blue-shift in sRII and Rh. We first carried out QM/MM geometry optimization for the entire protein, bR, sRII, and Rh. The QM/MM method divides the entire system into the QM segment and the MM segment. The QM segment is treated by the quantum-mechanical calculation to describe the electronic ground and excited states involved in the photoabsorption. The MM segment describes the steric and ES effects of the surrounding environment from the rest of the system by means of the molecular mechanics. With the QM/MM refined structures, we next calculated the absorption energies of the QM segment at the SAC-CI level of theory with the MM effective point charges representing the electrostatic field of the surrounding protein.

The SAC[32]/SAC-CI[33] method[34] is an accurate electronic-structure theory for the ground and excited states. This method has been established as a reliable theoretical method for calculating the ground and excited states of atoms and molecules.[34] The SAC method belongs to the cluster expansion method and treats the electron correlation in the ground state. Based on the correlated ground state described by the SAC method, the SAC-CI method was derived as the theory for the excited state.[33] The SAC-CI wave function satisfies the orthogonalities to the ground state SAC wave function. Therefore, the electron correlations in the ground and excited states are treated in a balanced way by the SAC/SAC-CI method. As indicated by the applications to more than 150 systems,[35] the SAC-CI method has been established as a powerful tool for studying the spectroscopy of the atoms and molecules. Owing to the perturbation selection technique of the excitation operators,[36] the computation program of the SAC-CI method is applicable to moderately large molecules. Recently, spectroscopy of biological systems[34] has become our target of applications.

Such advantages in the SAC-CI method realize the quantitative description of the electronic excited states of the retinal proteins. First, it is possible to compute the absorption energies for the whole chromophore molecule without truncating the β-ionone ring and methyl groups. We could further extend the systems to include the counterion and polar groups which have turned out to give important electronic-interaction effects as described above. In previous studies,[11,22,26] the truncated chromophore analogues were often used to reduce the high computational cost of the methods. However, such approximated models lead non-negligible error in the result of the computation.[13]

Moreover, all the valence orbitals were included in the active space of the coupled-cluster expansion in the SAC wave function, and thus the electron correlations among the π and σ orbitals were equivalently treated. This clearly contrasts with the previous calculations using CASSCF (complete active space SCF) or CAS-CI as the reference wave function.[11,13,22,26] Previous MRMP, CASPT2, and spectroscopy oriented CI (SORCI)[12,37] calculations included only π orbitals in the CAS-CI expansion. The multiconfigurational perturbation theories give excellent results, if the reference wave function is enough accurate. However, the first excited-state of retinal PSB involves large $\sigma-\pi$ polarization and large positive-charge migration as described above. In such a situation, it would be desirable to treat the electron correlation of all the valence electrons in equivalent way as in the SAC-CI method.

In our previous study,[10,14,20] the geometry of chromophore was optimized with the HF method, and the QM segment included only the retinal chromophore. In our present study, we improved these inaccurate treatments. Since the HF method emphasizes bond alternation in the retinal chromophore, we used DFT with B3LYP functional for optimizing the structure of retinals. Geometry obtained by DFT is expected to be similar to that obtained by MP2 calculation as shown in a previous report[38] and also in this study. We extended the QM segment to include the counterion residues around the retinal chromophore. The present calculations successfully reproduced the absorption energies of all three proteins, demonstrating accuracy and consistency of the method used. Based on these calculations, we further performed quantitative analysis on the molecular origins of the color tuning among the retinal proteins. The present result is also compared to the previous ones obtained by using the methodologies.[10−14,20,26] We found the source of the discrepancy in the previous studies: treatment of the electron correlation in the excited-state calculation, geometry of retinal and counter residues, and the quantum-mechanical interactions between the Schiff base and the protein residues, especially the counter-residue and the charged residues.

2. Computational Details

Figure 1(a,b) shows computational models of the PSB retinal chromophores in 11-cis and all-trans conformations, respectively. Schiff base (SB) retinal in an all-trans conformation was also examined. We note this structure is the C_6-anti form and is different from the C_6-syn structure used in the gas-phase experiment, the electrostatic ion storage ring in Aarhus (ELISA).[39,40] For the gas-phase calculations, geometry optimizations for the all-trans retinal were performed with the HF, DFT with B3LYP functional (B3LYP), and MP2 methods as well as the semiempirical AM1 method. The

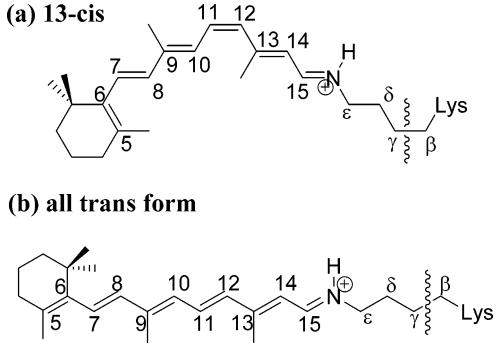

Figure 1. Retinal and all-trans PSB analog structure: (a) 11-cis form and (b) all-trans form.

absorption energies were then calculated at the SAC-CI and TD-DFT with B3LYP functional levels of theory for those optimized structures.

For the protein systems, the entire protein structures including the chromophore molecule were optimized by QM/MM calculation.[10,18,20,41,42] In the QM/MM method, the active site is treated by quantum-mechanical calculation, and the rest of protein is described at the MM level of theory. The QM/MM method used in the present study[20] takes into account ES interaction between the chromophore and the surrounding protein environment through restrained ES potential charge (RESP) operators, permitting us to efficiently determine the optimized structure of the entire protein.

The detail of the QM/MM method was reported elsewhere.[10,20] For visual receptor Rh, we constructed a starting structural model based on an X-ray crystallography structure recently reported by Okada and co-workers (PDB code: 1L9H).[7] For bR and sRII, the initial structures taken from PDB were 1C3W[8] and 1H68,[9] respectively. We carried out a geometry optimization for the whole protein using the QM/MM method.[10,20] The QM segment includes the whole retinal chromophore with the side chains of Lys296 and Glu113 (counterion group) and a proximal water molecule which has hydrogen-bonding to Glu113. The QM segment for bR and sRII was equivalent to that for Rh. The boundaries between the QM and MM segments were at the $C_\beta-C_\gamma$ of Lys296 (Lys216 in bR and Lys205 in sRII) and $C_\alpha-C_\beta$ bonds of Glu113 (Asp85 in bR and Asp75 in sRII). Hydrogen atoms were introduced for the link atom.

For the QM segment, the DFT with B3LYP functional was employed for the structure optimization. The basis functions used were Dunning's double-ζ plus polarization basis sets (D95(d))[43] for C atoms of the retinal π-system, N and H atoms of Schiff base, and O and C atoms of the carboxylate of the counter residue and the water molecule. For the other atoms, Dunning's double-ζ sets (D95)[43] were employed. In addition, single p-type anion functions ($\alpha=0.059$)[43] were augmented on the anionic O atoms of the counterion groups to properly describe the charge transfer interaction between the counterion groups and the chromophore. The coefficients of restraint terms in the RESP method are set to be 10^{-4} au for the atoms near the boundaries and 10^{-5} au for the other atoms.[10,20] The AMBER99 force field[44] was used in the MM calculation to describe the rest of the protein, and TIP3P[45] is used for water molecules. A residue-based 12 Å cutoff was used for van der Waals interactions. The QM/MM code[10,20,42] was incorporated in the QM program package GAMESS.[46]

The rms deviation between the X-ray and QM/MM optimized structures was measured for the residues within 6 Å from the retinal chromophores. The deviation is 0.6−0.9 Å for the three proteins, which is much smaller than the resolution of the X-ray structures. These deviations, however, depend on the origin common to the two structures. For the distance between the N(retinal) and O(glutamate), the deviation is about 0.3 Å.

To evaluate the absorption energies in the opsins, we performed the SAC-CI calculation. CI-singles (CIS) and TD-DFT calculations were also performed for comparison. Figure 2(a) shows the optimized structure of the QM segments: chromophore-Lys296 moiety, the counterion group Glu113, and a proximal molecule, Wat2b,[7] in Rh. In the single point SAC-CI calculations, the absorption energies were computed for the entire PSB retinal chromophore with the side chain of lysine shown in Figure 1(a,b). Furthermore, in order to examine the quantum-mechanical interaction with protein, we also include the counterion group and a proximal water molecule shown in Figure 2 (the retinal PSB active site; Glu113 and Wat2b for Rh,[7] Asp85 and Wat402 for bR,[8] and Asp75 and Wat402 for sRII[9]) which were carried out in the SAC-CI calculation. The ES effect of the other residues was included by the point charges of AMBER99.[44] The QM/MM system termed "AS" includes the chromophore, the counterion group, and the proximal water molecule in the QM segment. For comparison, we also computed the absorption energies for the QM/MM system including only the chromophore in the QM segment ("RET").

The detail of the SAC-CI method can be found elsewhere.[34] For the SAC-CI single-point calculations, we used the same basis functions as those used for the geometry optimizations. The active space of the SAC-CI calculations included all the valence orbitals. The 1s orbitals and the corresponding virtual orbitals were treated as the frozen orbitals. Total 451, 427, and 427 MOs were taken for the active orbitals in Rh, bR, and sRII, respectively. All the single excitation operators and the selected double excitation operators were included in the SAC/SAC-CI wave functions. Perturbation selection technique was used for selecting important double excitation.[36] In the perturbation selection, HF/CIS wave functions were used for the reference states to estimate second-order perturbation energy, and the energy thresholds of 5×10^{-6} and 5×10^{-7} au were used for the ground and excited states, respectively.

For SAC-CI, CIS, and TD-DFT calculations for the excited states and HF, B3LYP, MP2, and AM1 calculations for geometry optimization in the gas phase, we used a development version of Gaussian03 program system.[47] For the SAC-CI calculations, we have improved the computational algorithm for the perturbation selection of the two-electron operators, which realized the large-scale calculations of the PSB retinal active site.

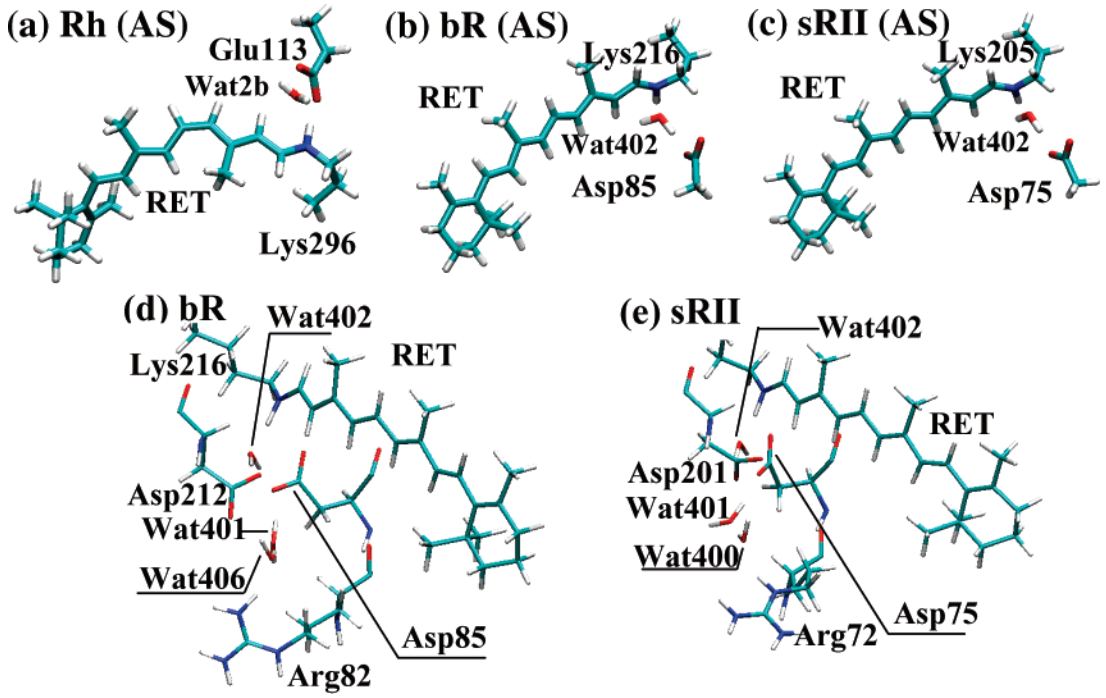

Figure 2. (a−c) QM/MM optimized structure of the active-site (AS) model (the retinal chromophore, counterion group, and a proximal water molecule) for (a) Rh, (b) bR, and (c) sRII. (d,e) The arrangement of the charged residudes close to the chromophore in (d) bR and (e) sRII.

3. Results

The first excited states which exhibit large oscillator strengths are assigned to the first absorption peaks for SAC-CI and CIS. The main character of the first excited states is $\pi-\pi^*$ excitation from the highest occupied molecular orbital (HOMO) to the lowest unoccupied MO (LUMO). In the TD-DFT calculations, it was found that the first excited states of the active site (AS) models do not possess the character of the HOMO−LUMO $\pi-\pi^*$ excitation, and the oscillator strength is rather small. We therefore assign the lowest $\pi-\pi^*$ excited states which exhibit large oscillator strengths to the first absorption peaks.

3.1. Assessment of Methodology for Geometry Optimization. We first carried out several test calculations to evaluate the methodologies for the geometry optimization. As reported by Wanko et al.,[12] the absorption energy is rather sensitive to the computational method used for the geometry optimization. Especially, bond alternation of the polyene chain in the chromophore tends to be overestimated in the HF and CASSCF optimized structure due to the lack of dynamic electron-correlation.[12] This structural deficiency causes overestimation of the excitation energy especially in the HOMO−LUMO transition, because the energy level of HOMO and LUMO are stabilized and destabilized, respectively. As shown in Figure 3 (a), the HOMO has bonding character at C5═C6, C7═C8, and C9═C10 double bonds, while the LUMO is antibonding at C11═C12, C13═C14, and C15═N double bonds.

In order to examine the bond-alternation effect on the SAC-CI absorption energies, we computed the absorption energies for all-trans retinal (Figure 1(b)) using several geometries optimized by HF, B3LYP, MP2, and AM1 methods. As shown in Table 1, HF and AM1 geometries give rather large absorption energies (1.44 and 1.37 eV, respectively) compared to the B3LYP and MP2 ones (1.18 and 1.19 eV, respectively). On the other hand, B3LYP and MP2 give structures very similar to each other.[12] We also confirmed that the SAC-CI absorption energy using the B3LYP geometry is in good agreement with that using an MP2 one, validating the use of B3LYP for the geometry optimization in the present study. These results are very similar to those obtained by Wanko and co-workers[12] and thus confirmed their previous study.

3.2. Absorption Energies of the Retinal Proteins. In Table 2, SAC-CI results are summarized and compared with experimental[1,30,31,48,49] and previous theoretical results.[10,13,20,22,26] The calculated excitation energy using the active site QM model ("AS") "in opsin" environment are 2.45 eV for Rh, 2.23 eV for bR, and 2.53 eV for sRII, which nicely agree with the experimentally observed absorption energies (2.49, 2.18, and 2.49 eV, respectively) for *all* the systems studied. The present results significantly improved the previous theoretical ones.[10,13,14,20,22,26] The root-mean-square error is around 0.04 eV, and the maximum error is 0.05 eV for bR.

It should be noted that the present result has been much improved over our previous one.[14] In the present study, the QM/MM geometry optimizations were carried out at the B3LYP/MM level, whereas the HF/MM method was employed for the optimizations in the previous study.[14] Moreover, in the previous calculations, we included only retinal PSB in the QM segment. The present model includes the counterion group and a water molecule close to the Schiff base in the QM segment in addition to the chromophore. The computed excitation energies are quite sensitive to these two factors. Although our previous results[14] were very close to the present results, this is actually due to a lucky

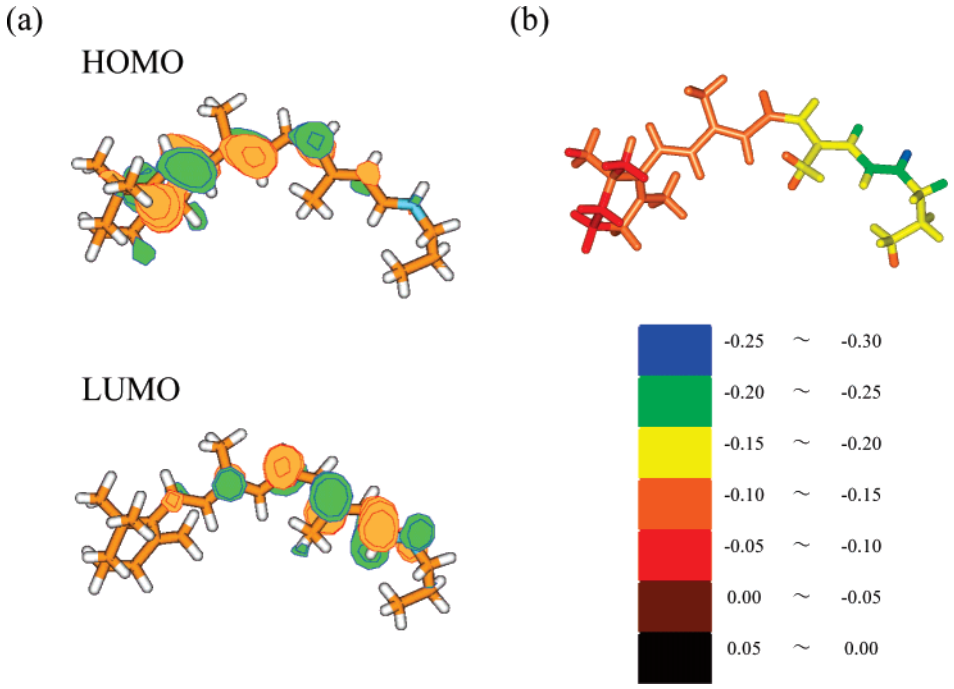

Figure 3. (a) HOMO and LUMO distribution for Rh and (b) ES potential by protein environment (opsin) in atomic unit.

Table 1. SAC-CI Absorption Energies and Oscillator Strength Calculated with the Gas-Phase Optimized Geometries

optimize	main config (|C|>0.3)	E_{ex}(eV)	f(au)
B3LYP	0.92(H → L)	1.18	1.01
MP2	0.92(H → L)	1.19	0.97
HF	0.92(H → L)	1.44	0.91
AM1	0.92(H → L)	1.37	0.81

cancellation between the two factors. The introduced error also affects the accuracy of the analysis on the color tuning mechanism. Detailed analysis is given in a later section.

The absorption energies of the chromophore are strongly affected by the interaction with the surrounding protein environments. In order to extract the effect of the protein environments, we computed the absorption energies of the chromophore in the absence of the protein environment (referred to as "bare" chromophore) and compared with those in the opsins. Note that the structures of the bare chromophore for the calculations are the same as those in the proteins. Thus, the differences in the absorption energies in the opsins from those of the bare chromophore represent the effects of the interaction with the protein environment but do not include contributions of changes in the chromophore structure upon the bindings to the opsins (see below). Table 2 lists the absorption energies of the bare chromophore. The absorption energies of the bare chromophore (1.36 eV for Rh, 1.30 eV for bR, and 1.31 eV for sRII) are considerably lower than those in the opsins by more than 1 eV, indicating that the interactions with the protein surroundings give rise to the large blue-shifts of the absorption energies.

The contributions of the chromophores' structural changes upon the bindings are also significant. We compared the results for "bare" chromophores with an all-trans chromophore (Figure 1(b)) whose structure was optimized in vacuo. The result shows that the absorption energies of the bare chromophore also exhibit blue-shifts by 0.18 eV for Rh, 0.12 eV for bR, and 0.13 eV for sRII from that for all-trans chromophore, 1.18 eV. Therefore, both the structural changes of the chromophore and the interaction with the protein surroundings cause blue-shift in the absorption energies under the opsin environment.

Table 3 lists oscillator strengths, f, calculated by the SAC-CI method for all the proteins studied. It was found that the oscillator strengths in the opsins are also significantly larger than those of the bare chromophore, indicating larger absorbances in the opsins. Note that the oscillator strength is proportional to the product of absorption energy, E_{ex}, and square of transition dipole moment, $|\mu_{eg}|$.[2] We therefore computed the transition dipole moments as well in order to identify the source of the increase of the oscillator strength. Table 3 also lists the computed transition dipole moments. In contrast to the oscillator strengths, the transition dipole moments of the chromophore in the opsins are smaller than those of the bare chromophore. Hence the larger oscillator strengths in the opsins are attributed to the large blue-shifts of the absorption energies in the opsins.

3.3. Spectral Blue-Shift Caused by the Protein Electrostatic Environment. The interaction of the chromophore with the surrounding protein environment contributes to the large blue-shift of the absorption energy as shown in Table 2. This is indicated by the significant decrease in the absorption energy of the bare chromophore. As suggested in the previous studies,[10,14] the main cause of the large blue-shift is difference in the ES interaction energy between the ground and excited states of the PSB retinal. The migration of the positive charge of PSB along the polyene chain upon the electronic excitation largely reduces the interaction with the protein environment, especially with the negatively

Table 2. Comparison of SAC-CI Excitation Energies with Other Results (eV)

				SAC-CI								
protein	model		environment	main config (|C|>0.3)	*f* (au)	$\Delta\mu$ (debye)	E_{ex} (eV)	exptl (eV)	CASPT2 E_{ex}(eV)	MRMP E_{ex}(eV)	SORCI E_{ex}(eV)	TD-B3LYP E_{ex}(eV)
Rh	WT[f]	AS	in opsin	0.94(H → L)	1.03		2.45	2.49[i]	2.86[a]			2.52
		RET		0.93(H → L)	0.88	−13.25	2.06		2.78,[b] 2.59[c]			2.44
		RET	bare	0.91(H → L)	0.63	−12.69	1.36	-	2.72,[b] 2.72[c]			2.53
bR	WT[f]	AS	in opsin	0.94(H → L)	1.29		2.23	2.18[j]				2.57
		RET		0.92(H → L)	1.15	−14.98	1.88			2.75[d]	2.34[e]	2.49
		RET	bare	0.92(H → L)	0.91	−13.29	1.30	-		2.05[d]	1.86[e]	2.31
	R82A[g]	AS	in opsin	0.94(H → L)	1.33		2.34	2.23[k]				
sRII	WT[f]	AS	in opsin	0.94(H → L)	1.42		2.53	2.49[l]				2.68
		RET		0.93(H → L)	1.27	−13.88	2.17					2.58
		RET	bare	0.92(H → L)	0.89	−14.00	1.31	-				2.30
	R72A[h]	AS	in opsin	0.94(H → L)	1.46		2.58	2.48[m]				

[a] CASPT2 result described in ref 26. [b] CASPT2 result described in ref 22. [c] CASPT2 result described in ref 13. [d] MRMP result described in ref 20. [e] SORCI result described in ref 12. [f] Shows "wild type". [g] Shows "R82A" mutant. [h] Shows "R72A" mutant. [i] Reference 1. [j] Reference 31. [k] Reference 48. [l] Reference 30. [m] Reference 49.

Table 3. Decomposition of Oscillator Strength[a]

protein	model	environment	*f* (au)	E_{ex} (eV)	$\|\mu_{eg}\|^2$
Rh	AS	in opsin	1.03 (1.63)	2.45 (1.80)	17.12 (0.92)
	RET		0.88 (1.40)	2.06 (1.51)	17.37 (0.93)
	RET	bare	0.63 (1.00)	1.36 (1.00)	18.65 (1.00)
bR	AS	in opsin	1.29 (1.42)	2.23 (1.72)	23.68 (0.85)
	RET		1.15 (1.26)	1.88 (1.45)	25.08 (0.90)
	RET	bare	0.91 (1.00)	1.30 (1.00)	28.01 (1.00)
sRII	AS	in opsin	1.42 (1.59)	2.53 (1.94)	22.95 (0.82)
	RET		1.27 (1.43)	2.17 (1.66)	23.91 (0.86)
	RET	bare	0.89 (1.00)	1.31 (1.00)	27.91 (1.00)

[a] The values in parentheses stand for the ratio, (in opsin)/(bare).

charged counterion groups in the vicinity of PSB[10] shown in Figure 2.

In order to estimate the contribution of the ES interaction energy, we computed the SAC-CI absorption energy of the QM/MM system where the QM segment includes only the chromophore (referred to as "RET"). The rest of the environmental effect was included as a point-charge model. Thus the interaction between the chromophore and the protein surroundings is only taken into account by the electrostatic potential. Table 2 lists the SAC-CI absorption energies of RET. The absorption energies were computed to be 2.06, 1.88, and 2.17 eV for Rh, bR, and sRII, respectively, and exhibit large blue-shifts from those of the bare chromophore by 0.58−0.86 eV, indicating the large contributions of the ES interaction.

Figure 3(a) depicts distributions of HOMO and LUMO of the chromophore in Rh, which are mainly responsible for the electronic excitation in the first excited state. As clearly seen, the distributions of the π orbitals are localized in the β-ionone ring and PSB halves of the polyene chain in HOMO and LUMO, respectively. Since the main configuration of the excited-state is an excitation from HOMO to LUMO, the excitation has charge-transfer character from the β-ionone ring side to the other side. In other words, the positive charge migration occurs from the PSB side to the other side, as mentioned above. The large changes in dipole moment of the chromophore upon the excitations, $\Delta\mu$ (−12.69 to −14.98 debye), listed in Table 2 clearly indicate the positive charge migrations of PSB. Figure 3(b) shows ES potential by the protein environment of Rh acting on atoms of the chromophore. Strong negative ES potential due to the counterion negative group Glu113 is observed in the PSB region. The ES potential gradually increases toward the β-ionone ring side, creating a gradient of the ES potential along the polyene chain. The large contribution of the ES interaction to the absorption energy is therefore clearly explained by the positive charge migration along the polyene chain against the gradient of the ES potential produced by the protein surroundings, especially the counterion group.

3.4. Impact of the Counterion Group on the Absorption Energy. It should be noted that the interactions of the chromophore with the protein surroundings are not fully recovered by the ES interactions described with the point-charge model. The ES treatment covers only 60−70% of the spectral blue-shift. As seen in Table 2, the absorption energies for the RET systems (1.88−2.17 eV) which only consider the ES interactions approximated by the effective point charges are still considerably lower than those of the experiments (2.18−2.49 eV) and for the AS systems (2.23−2.53 eV) where the quantum-mechanical interactions with the counterion and water molecules in the vicinity of PSB are taken into account by explicitly including them in the SAC-CI absorption energy calculations. The quantum-mechanical and ES interaction therefore provides large contributions to the blue-shifts, as suggested previously.[20] The large quantum-mechanical interaction due to the electronic polarization and the charge transfer in the ground state of bR was identified by Morokuma-Kitaura decomposition analysis in a previous study by Hayashi and Ohmine.[20] Thus the positive charge migration toward the β-ionone ring upon the excitation significantly diminishes the electronic interaction, leading to the blue-shift of the absorption energy.

The contributions of the quantum-mechanical interaction to the absorption energies, ΔE_{ex}^{ele}, are estimated with differences between the absorption energies of RET, E_{ex}(RET), and AS, E_{ex}(AS)

$$\Delta E_{ex}^{ele} = E_{ex}(\mathrm{AS}) - E_{ex}(\mathrm{RET}) \quad (1)$$

Table 4. Electronic Interaction Energies, ΔE_{ex}^{ele} (eV)

protein	ΔE_{ex}^{ele} SAC/SAC-CI	ΔE_{ex}^{ele} HF/CIS	ΔE_{ex}^{ele}(SAC/SAC-CI) − ΔE_{ex}^{ele}(HF/CIS)
Rh	0.386	0.057	0.329
bR	0.353	0.085	0.268
sRII	0.354	0.088	0.266

Table 4 lists the contributions estimated for the SAC/SAC-CI absorption energies. The contributions, 0.39, 0.35, and 0.35 eV for Rh, bR, and sRII, respectively, are considerably large and are more than one-third of the total blue-shifts by the interaction with the protein surroundings. Interestingly, dynamic electron-correlation is crucial for accurately describing the interaction. To see this, we also computed ΔE_{ex}^{ele} for the absorption energies at the HF/CIS level of theory. As seen in Table 4, the HF/CIS contributions of the electronic interaction were estimated to be 0.06−0.09 eV and are rather underestimated by 0.27−0.33 eV compared to the SAC/SAC-CI ones.

In addition, we checked the convergence of absorption energy with respect to the size of the QM segment. Since, SAC-CI calculation could not be performed due to the limitation in the computational resource, an ONIOM-like analysis was performed to examine the effect of the other charged residues close to the retinal chromophore.

$$E_{ex}^{Large,SAC-CI} \cong E_{ex}^{Small,SAC-CI} + (E_{ex}^{Large,CIS} - E_{ex}^{Small,CIS}) \quad (2)$$

$E_{ex}^{X,Y}$denotes the excitation energy for the X (X = Small or Large) QM region using the Y (Y = SAC-CI or CIS) method. The "Small" QM region is identical to the "AS" models as shown in Figure 2. The "Large" QM region additionally includes one arginine, one aspartate, and three water molecules in bR and sRII as shown in Figure 2(d,e). For the bR case, Arg82, Asp212, Wat401, Wat402, and Wat406 were included. In the sRII case, Arg72, Asp201, Wat400, Wat401, and Wat402 were added to the QM segment. With this estimation, we obtained the correction $E_{ex}^{Large,CIS} - E_{ex}^{Small,CIS}$ of 0.00 and 0.01 eV for bR and sRII, respectively. Therefore, the AS systems are suitable choices for the QM segment during the absorption energy calculation. This result indicates the quantum-mechanical effect from these secondary residues is negligible.

3.5. Effect of the Chromophore Structure on the Absorption Energy. As mentioned before, the change in the bond alternation significantly affects the absorption energy of the chromophore. With stronger bond alternation, more spectral blue-shift is expected as shown in Table 1. It is noteworthy that the bond alternation becomes strong in opsin due to the interaction with the protein surroundings.

Figure 4 illustrates the bond-alternation patterns of the all-trans PSB retinal chromophore in vacuo and in the bR opsin. The bond-alternation pattern of SB retinal is also shown in Figure 4 for comparison. It is clearly seen that the bond alternation in the bR opsin is more enhanced than that in vacuo. The result is consistent with a previous study by Sugihara et al.[28] using the DFTB method. The strongest bond alternation was observed in the SB retinal. The stronger bond

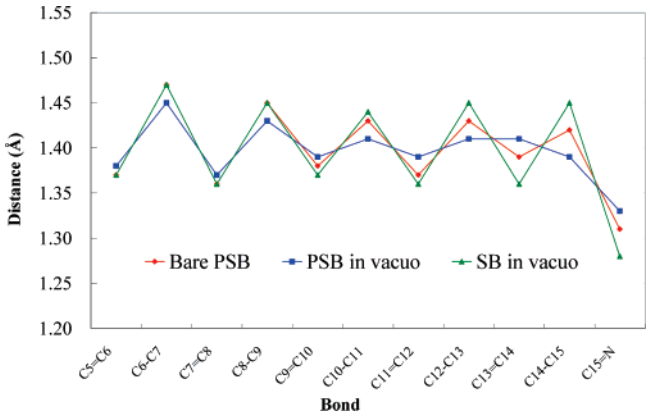

Figure 4. Bond length on the π-chain of PSB with B3LYP geometry (Å).

alternation of the PSB retinal in the opsin originates from the interaction of PSB with the counterion groups that suppresses resonance structures of the PSB retinal involving migrations of the positive charge along the polyene chain.[28] The pronounced bond alternation in the SB retinal is attributed to the absence of the resonance structures.

As shown above, the stronger bond alternation of the PSB retinal leads to the larger absorption energy. The main origin of the larger absorption energies of the bare chromophores than that of the stable chromophore in vacuo are therefore the stronger bond alternations of the bare chromophores induced by the interaction with the counterion groups in the opsins. Since the increases of the absorption energies of three proteins, 0.12−0.18 eV, are larger than the differences of the increases (maximally 0.06 eV), the stronger bond alternation is suggested to be a common mechanism giving rise to the main increases of the absorption energies for the three proteins. Overall, the interaction of the PSB retinal with the counterion groups contributes to the large blue-shift upon the binding of the chromophore into the opsin through not only the reduction of the interaction upon the electronic excitation discussed above but also the change in the bond alternation pattern.

The results clearly explain strong correlation between vibrational frequency of the ethylenic mode and the absorption energy revealed by Ebrey et al.[50] The resonance Raman experiments revealed that the retinal proteins with the larger absorption energies possess the higher ethylenic frequencies of the chromophore. The higher ethylenic frequency is attributed to the stronger bond alternation due to the stronger interaction of PSB with the counterion groups, which results in the larger absorption energy as discussed above.

There is a large difference in the bond alternation between the X-ray crystallographic structural model (1L9H) and the refined QM/MM one optimized at the DFT/B3LYP level of theory (B3LYP/MM). Figure 5 depicts the bond alternation patterns, showing much stronger bond alternation in the chromophore of 1L9H. We also carried out the SAC-CI absorption energy calculation for the chromophore of 1L9H, where positions of the missing hydrogen atoms in the X-ray crystallographic structures were determined with AMBER99 force field. Table 5 gives a comparison of the absorption energy between 1L9H and B3LYP/MM. The computed SAC-

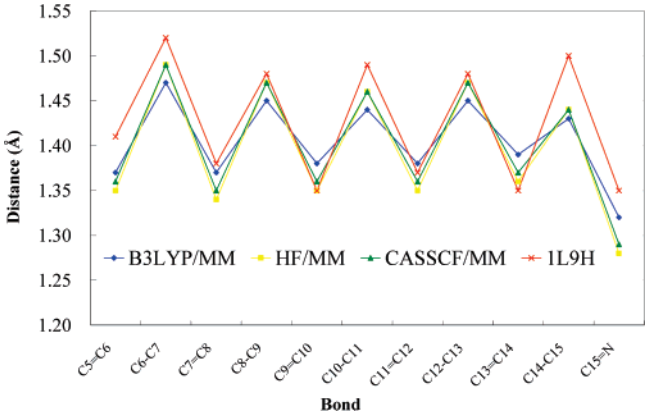

Figure 5. Bond length on π-chain of PSB in Rh (Å).

Table 5. SAC-CI Absorption Energies and Oscillator Strength for the Isolated Chromophores

optimize	main config(|C|>0.3)	E_{ex} (eV)	f (au)
B3LYP/MM	0.91 (H → L)	1.36	0.63
HF/MM[a]	0.89 (H → L)	1.72	0.53
CASSCF/MM[b]	0.89 (H → L)	1.70	0.52
1L9H[c]	0.72 (H → L)	1.72	0.42

[a] Reference 14. [b] Reference 13. [c] Reference 7.

CI absorption energy for 1L9H is 1.72 eV and significantly deviates from the B3LYP/MM one (1.36 eV), indicating again great impact of the bond alternation on the absorption energy. Note that the differences in the bond length responsible for the bond alternation between 1L9H and B3LYP/MM are less than 0.1 Å and much smaller than the resolution of the X-ray crystallographic model (2.6 Å).[7] Hence the X-ray crystallographic model is not "chemically precise" enough, and the structure refinement by the reliable methods is definitively requisite for the accurate prediction of the absorption energy. The bond alternation in the HF optimized structure is much stronger than that of the B3LYP one. The SAC-CI absorption energy of the bare chromophore using the HF/MM optimized structure was estimated to be 1.72 eV and is greatly overestimated compared to that for the B3LYP/MM optimized structure, as shown in Table 5.

Torsion of the polyene chain of the chromophore in the binding pocket also alters the absorption energy as mentioned above. The differences in the absorption energies of the bare chromophore mainly represent the contributions of the torsional distortion of the chromophore. The contributions of the torsion play an important role in the fine-tuning of the absorption energies among the retinal proteins as seen below.

Table 6 summarizes dihedral angles of the polyene chain of the chromophore in bR, sRII, and Rh. The chromophores in bR and sRII are both in all-trans conformation, and their structures are similar to each other. The polyene chains are almost planar except for the dihedral angles of C_{13}═C_{14}. Deviations of the C_{13}═C_{14} dihedral angles from planarity are by 21 and 15 degrees in bR and sRII, respectively, indicating crucial pretorsions for the selective photoisomerizations for those functions.[42] In Rh, the chromophore assumes a C_{11}-cis and C_6-syn conformation and is strongly twisted in the protein confinement. Again, torsion around the C_{11}═C_{12} bond by 14 degrees could play a role in the selective photoisomerization. In addition, a dihedral angle around the C_6−C_7 single bond largely deviates from planarity by −47 degrees. This value is larger than the one by Sugihara (−42 degrees)[26] and smaller than by Andruniów (−54 degrees).[13]

The single bond rotation around C_6−C_7 in Rh greatly contributes to the spectral shift. In order to illustrate this, we performed SAC-CI absorption energy calculations for the chromophore with torsions around the C_6−C_7 bond by 30 and 60 degrees. The absorption energies were computed for the chromophore in vacuo, and the geometries were optimized with constraints for the C_6−C_7 dihedral angles. Table 7 lists the dependence of the SAC-CI absorption energy on the bond rotation. The absorption energy increases as the torsion around C_6−C_7 is larger, consistent with calculations by Wanko et al. using other high level methods.[12] The blue-shift induced by the torsion around the single bond is due to shortened π conjugation of the polyene chain. The higher absorption energy of the bare chromophore of Rh is therefore suggested to be mainly attributed to the large torsion around C_6−C_7 by −47 degree.

4. Discussion

4.1. Mechanism of the Color Tuning. Based on the successful prediction of the absorption energies for those retinal proteins, we analyzed the spectral shifts in order to identify the molecular factors determining the color tuning. We consider the absorption energy of bR as a reference and analyzed a similar amount of the blue-shifts (0.31 eV) in the absorption energy of Rh and sRII from that of bR. As mentioned above, the apparently similar shifts are expected to involve different mechanisms of the color tuning, since the protein structures of Rh and sRII are quite different. The structure of sRII is found to be very similar to that of bR,[8,9] whereas the structure of Rh[7] is incompatible with those of bR and sRII.

The spectral shifts can be decomposed into three contributions. The first one is the structural distortion of the chromophore due to the protein confinement (structural effect). The second one is the ES interaction of the chromophore with the surrounding proteins (ES effect). The last one is the quantum effect of the counterion and a water molecule in the vicinity of PSB. This is a correction by the quantum-mechanical effect which cannot be described by the ES calculations (counterion QM correction). Those contributions can be deduced from the absorption energies listed in Table 2. The structural distortion effect is the difference of the absorption energies of the bare chromophores. The ES effect is the difference of the spectral shift due to the electrostatic environment modeled by the point-charges. The counterion QM correction is the difference in the absorption energy between the AS and RET systems.

As shown in Figure 6, the underlying mechanisms are completely different between Rh and sRII. In Rh, the ES contribution mainly contributes to the blue-shift (0.12 eV), which is more than the half of the total computed shift of 0.22 eV. This is because the interaction between PSB and the counterion group in Rh is stronger than that in bR. As

Table 6. Dihedral Angles (in deg) of the Retinal Chromophore in the QM/MM Refined Structures of Rh, bR, and sRII

	Rh				bR			sRII	
	B3LYP/MM[a]	HF/MM[b]	SCC-DFTB/MM[c]	CASSCF/MM[d]	B3LYP/MM[a]	HF/MM[e]	SCC-DFTB/MM[f]	B3LYP/MM[a]	HF/MM[e]
$N_\zeta-C_\epsilon$	93.12	94.48	100.76	109.36	−114.84	−112.87	−100.66	−95.56	−99.17
$C_{15}=N_\zeta$	179.36	−179.87	169.94	173.39	−166.73	−166.82	−168.91	−173.06	−169.23
$C_{14}-C_{15}$	172.78	172.91	178.59	177.74	−179.24	−178.15	174.36	177.85	179.12
$C_{13}=C_{14}$	−178.52	−178.94	173.29	178.79	−158.95	−161.34	−153.17	−164.99	−164.93
$C_{12}-C_{13}$	170.07	166.07	171.71	166.81	175.74	173.72	174.37	177.78	176.83
$C_{11}=C_{12}$	−13.93	−9.67	−18.25	−8.19	−172.60	−172.70	−171.71	−176.86	−175.56
$C_{10}-C_{11}$	175.00	168.85	175.34	166.24	−176.11	−174.98	−172.71	176.98	178.55
$C_9=C_{10}$	166.77	168.88	170.43	169.63	−179.23	−178.80	174.69	−174.28	−174.57
C_8-C_9	176.73	−176.46	171.30	−170.31	−173.57	−172.59	−164.50	177.36	178.89
$C_7=C_8$	175.81	174.96	−179.52	175.95	175.50	176.86	171.00	−173.90	−174.41
C_6-C_7	−46.73	−51.78	−42.09	−53.95	172.90	169.73	167.10	174.95	173.50

[a] Present geometry. [b] Reference 14. [c] Reference 26. [d] Reference 13. [e] Reference 10. [f] Reference 12.

Table 7. Dependence of Absorption Energies on C6−C7 Rotation of all-trans PSB

	TD-B3LYP			SAC-CI		
angle (deg)	main config |C|>0.3	*f* (au)	E_{ex}(eV)	main config |C|>0.3	*f* (au)	E_{ex}(eV)
≅0[a]	0.55(H → L)	2.03	2.41	0.92(H → L)	1.03	1.18
30.0[b]	0.56(H → L)	1.87	2.39	0.92(H → L)	0.98	1.21
60.0[b]	0.56(H → L)	0.81	2.17	0.92(H → L)	0.95	1.42

[a] Fully optimized geometry with B3LYP in the gas phase. [b] B3LYP geometry in the gas phase only fixed C6−C7 dihedral angle.

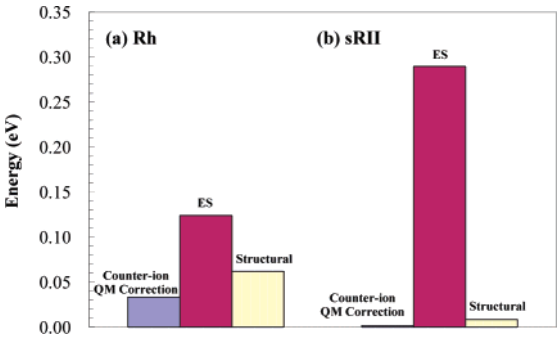

Figure 6. Decomposition of absorption energy difference from bR.

seen in Figure 2, the PSB in Rh forms a salt bridge directly to the counterion group, Glu113, whereas bR has a water molecule (Wat402) intervening between PSB and the counterion groups. The structural distortion and the QM correction contribute to the blue-shift by 0.06 and 0.04 eV, respectively. The structural contribution is mainly attributed to the torsion around the C_6-C_7 bond as described above.

In contrast to the case of Rh, the structural distortion is only a minor contribution of 0.01 eV to the overall blue-shift (0.30 eV). As seen above, the structures of the chromophore in bR and sRII are quite similar, resulting in the small contribution of the geometric structural distortion. Instead, the ES contribution (0.29 eV) dominates the overall blue-shift, as suggested previously.[10,14,20] In contrast to the shift of Rh, the counterions (0.00 eV) do not compensate the ES one.

On the large blue-shift of sRII from bR, a mechanism suggested in a previous study[51] is different in interaction of the chromophore with the positively charged guanidinium group of a proximal arginine between bR and sRII. The X-ray crystallographic studies revealed that orientation of the side chain of the proximal arginine, Arg72, in sRII is opposite to that of the corresponding arginine, Arg82, in bR.[8,9] The mechanism, however, has been questioned by other studies.[10,14] We therefore again analyzed the mechanism with the improved computational strategy used in the present study.

In order to examine the mechanism, we carried out theoretical mutation: the positive ES potentials produced by Arg82 of bR and Arg72 of sRII were replaced by that by alanine, a neutral residue, and the B3LYP/MM geometry optimizations were performed for the mutant proteins. As shown in Table 2, the SAC-CI absorption energies of the mutated bR and sRII are 2.34 and 2.58 eV, respectively, and undergo small changes from those of the native ones, 2.23 and 2.53 eV, respectively. This is consistent with results of the mutation experiments.[48,49] It is therefore concluded from the analysis that the ES interaction with the arginines plays a minor role in the spectral shift. This is because those arginines create the positive ES potential almost equally over the entire polyene chain of the chromophore, and the migration of the positive charge of PSB upon the excitation does not alter the ES interaction with the arginines.[10,14] It should be noted that the present analysis does not rule out non-negligible contribution of the electronic interaction with arginines through a hydrogen-bond network spanned between PSB and the arginines, as shown in Figure 2(d,e).

Recently, Hoffmann et al. proposed the two main and equally important factors about the color tuning mechanism between bR and sRII: the difference of neutral amino acids in the binding pocket and the difference in the extended hydrogen-bond network at the extracellular side of the proteins (counterion region).[52] We are also working on this subject, and the contribution of the electronic interaction has to be quantitatively determined in the near future.

The use of HF optimized structures overestimates the effect of the geometric distortion in the spectral blue-shift of Rh from bR. The contribution was estimated to be 0.21 eV for the HF/MM optimized structures, which is much larger than 0.06 eV for the B3LYP ones of the present study. The HF

Table 8. SAC-CI Absorption Energies, Oscillator Strength, and Difference Dipole Moment between the First Excited and Ground State in the Gas Phase (eV)

model		main config (|C|>0.3)	E_{ex}(eV)	f(au)	$\Delta\mu_{ex-gr}$ (debye)
PSB	bare[a]	0.92(H → L)	1.30	0.91	−13.29
	in vacuo[b]	0.92(H → L)	1.18	1.01	−9.36
SB	in vacuo[b]	0.94(H → L)	3.06	1.73	7.79

[a] B3LYP/MM geometry in bR. [b] B3LYP geometry in the gas phase.

method gives too strong bond alternation of the polyene structure, which overemphasizes torsions around the single and double bonds, respectively, although the overall distortion of the chromophore in the biding pocket is not greatly altered. Table 6 compares the torsions of the HF optimized structure with those of the B3LYP one, clearly indicating the errors of torsion. These errors lead to the overestimation of the contribution of geometric distortion. Note that the π bonding distributions of HOMO and LUMO locate mainly on the double and single bonds, respectively. Thus the overestimation of torsions around the single bonds lifts the energy level of LUMO. Similarly, the underestimation of torsions around the double bonds lowers that of HOMO. Both of the errors therefore give rise to the overestimated contribution of geometric distortion.

It is noteworthy that the absorption energy for the SB retinal in vacuo is computed to be 3.06 eV by the SAC-CI method as shown in Table 8. This result indicates that deprotonation at the Schiff base causes significant blue-shift in the absorption energy. In this sense, the present result supports the deprotonated form to be the UV pigments as suggested by Blatz et al. and Dukkipati et al.[53] We are now investigating the excited states for the UV pigments, and these will be reported in the future.

4.2. Comparison with Previous Theoretical Studies. The present results are compared with the previous studies using the modern electron-correlation methods.

In our previous SAC-CI study,[14] the QM region included only the chromophore. The counterion and water molecules in the vicinity of PSB were treated by the point-charge model. Furthermore, the HF method was employed for the QM/MM optimization.[14] The present study shows that these two inaccurate treatments introduce sizable errors in the absorption energy. As suggested previously[11,12,20] and confirmed in the present study, excluding the counterion groups from the QM segment significantly underestimates the absorption energy. As discussed in section 3.5, the HF optimized structure causes large overestimation of the absorption energy due to too strong bond alternation[12,25] and the torsion of the $\pi-$chain. The agreement in the previous study therefore is the result of cancellation between the underestimation due to the neglect of the electronic interaction and the overestimation due to the use of HF method for the geometry optimization.

Hayashi and Ohmine computed the absorption energy of bR at the MRMP level of theory.[20] With the HF/MM optimized structure and the QM segment including only the chromophore, they obtained the absorption energy of 2.75 eV, which rather overestimated the experimental data. The active space of the MRMP calculation was restricted to 12 electrons in 9 π orbitals instead of the full valence 12 electrons in 12 π orbitals. This would be another source of the overestimation in the absorption energy.

Recently, Andruniów et al. have revised their CASPT2 calculation by including the entire chromophore in the QM region.[13] The absorption energy was improved to be 2.59 eV, which nicely agrees with the experimental one (2.49 eV). However, in their mechanism proposed for the color tuning, the interaction of the chromophore with the protein surroundings provides a red shift in the absorption energy of the chromophore compared with the gas-phase result, which is significantly different from the present one: the absorption energy of the bare chromophore (2.72 eV) they obtained was larger than the absorption energy of the chromophore in the opsin (2.49 eV by the experiment[1] and 2.59 eV by CASPT2[13]). This gas-phase result is also much larger than those by Schreiber and Buss[11] using CASPT2 (1.88 eV) and Wanko et al.[12] using SORCI (1.93 eV). This discrepancy would originate from the CASSCF optimized structure used for the CASPT2 calculation. As shown in Figure 5, the CASSCF bond-alternation pattern of the chromophore[13] is very close to the HF one. We also note that the QM segment of their CASPT2 calculation did not include the counterion group and a water molecule proximal to PSB.

The QM/MM optimization procedure is also different. In the study of Andruniów et al.,[13] they used a X-ray crystallographic structure determined by Teller et al.[54] (PDB code: 1HZX), whereas the present study employed that determined by Okada et al. (PDB code: 1L9H).[7,29] The chromophore structure determined by Teller et al. is strongly distorted in the Schiff base region, showing remarkable deviation from the other three X-ray crystallographic structures by Palczewski et al.[55] and Okada et al.[7,29] In the QM/MM optimization, the protein atoms other than the Lys296 side chain and proximal water molecules were kept frozen, whereas the whole protein including the chromophore is fully and consistently relaxed in the present QM/MM optimization. Consequently, Andruniów et al. obtained the N_ζ(PSB)-O(Glu113) distance of 3.7 Å, which is much longer than that obtained in the present study (2.66 Å). The present structure is in accord with other QM/MM[26,29,56] and MD[57] studies. The large increase in the N_ζ-O distance breaks the salt bridge structure and can greatly lower the absorption energy as previously suggested.[10]

It should be noted, however, that the distance between PSB and the counterion group is still controversial in the experiment. NMR studies have suggested a longer distance and presence of a water molecule intervening between PSB and the counterion group.[58] In contrast, FTIR experiments have given the stronger hydrogen bonding between PSB and the counterion (Glu113) in Rh compared with that in bR.[59] In addition, low-temperature FTIR experiments of an internal water molecule in Rh has revealed a weaker hydrogen bonding of a water molecule (Wat2b) to the counterion (Glu113) compared with that of bR and sRII.[60] These data would indicate a shorter distance between PSB and the counterion group. In the latest X-ray crystallographic structure model by Okada et al.,[29] the distance is observed to be

~3.28 Å, although the resolution of the measurement (2.2 Å) does not seem to be high enough to detect the distance precisely. In this regard, further examination is required.

Wanko et al. clarified important methodological aspects in calculating the absorption energy of retinal proteins, such as the bond alternation and the distortion around C_6-C_7 bond of the chromophore.[12] They performed QM(SCC-DFTB[26,61])/MM calculations for the structural optimization and the SORCI calculations for the excited states. The computed absorption energy is 2.34 eV for bR, which is close to the experimental value of 2.18 eV. However, the calculation did not include the counterion groups in the QM segment. The computed absorption energy may become higher if the electronic interaction is taken into account.

Gascon and Batista computed the absorption energy of Rh with the TD-DFT method.[23] The absorption energy obtained (2.57 eV) is in good agreement with the experimental one (2.49 eV). However, the TD-DFT method is not appropriate for calculating the excited states of the PSB retinal chromophore. On twisting the C_6-C_7 bond, the TD-DFT gives even wrong trends as shown by Wanko et al.[12] and also in Table 7. In addition, TD-DFT failed to reproduce the absorption energy of Rh, bR, and sRII as shown in Table 2.

Schreiber et al. performed CASPT2 calculations in the gas phase for the PSB retinal chromophore and some counterion group (Glu113 and Wat2b).[11] They pointed out the remarkable influence of the counterion groups to the absorption energy. Sugihara et al. performed CASPT2//QM (SCC-DFTB)/MM calculations.[26] However, the chromophore was simplified, and the calculated excitation energy overestimated the experimental data.[11,26]

Recently, Andersen et al. measured the absorption spectrum of the C_6-syn PSB retinal chromophore in vacuo by using the ELISA, and the absorption maximum was observed at 2.03[39] and 2.00 eV.[40] As expected, this gas-phase absorption energy shows the red-shift of 0.79 eV compared with that in methanol (2.82 eV).[62] So far we have not performed calculations comparable to this experiment. With the geometry reported by Cembran et al.,[63] we obtained 1.87 eV by the SAC-CI method. However, this structure was optimized by the CASSCF method and is expected to be insufficient due to the overestimation in the bond alternation. The absorption energy also strongly depends on the torsional angle around the polyene chain. Therefore, it would be necessary for further investigation to simulate the ELISA experiment.

5. Conclusion

The excited states of Rh, bR, and sRII were studied by the SAC-CI calculations with the QM(B3LYP)/MM(AMBER99) optimized structures. This is the first study which successfully reproduced the major three retinal proteins, Rh, bR, and sRII. Based on the SAC-CI results, the mechanism of the color tuning in Rh, bR, and sRII was analyzed. The result shows that the spectral differences in these proteins are mainly due to the opsin ES interaction, and the chromophore conformation have only a minor effect on the color tuning in these proteins especially for bR and sRII. Our result also shows that the electronic interactions (quantum effect) between the chromophore and the counterion group are minor contributions to the color tuning. However the interactions are indispensable for predicting absolute absorption energies. There is so far no previous theoretical study reporting the electronic interaction due to the computational restriction, and thus current information would be much useful in a future study.

The effect of retinal bond alternation to the absorption energy was examined by comparison of several geometries both in vacuo and bare environment. The results clearly indicate that the retinal structure should be described precisely since the absorption spectrum is highly sensitive to retinal bond alternation. In addition, it is also shown that the dynamic electron-correlation being included in the wave function is necessary to expect the accurate absorption energy.

The present results indicate that the SAC-CI calculation with the QM/MM optimized structure is a promising approach to study the absorption spectrum of the retinal protein. This approach would be useful for studying the color tuning of the retinal proteins led by the theoretical prediction. Future studies of the color tuning are to account for the polarization effect in the protein. In the present study, though the polarization of the QM region (PSB, counteranion, water) by the protein environment is included, that of the protein environment by the chromophore is neglected. Houjou et al. suggested the large polarization effect on the absorption energies by the polarizable continuum calculations.[19,27] It is furthermore necessary to clarify the effect of protein polarization in an atomic detail. A study in this regard is ongoing.

Acknowledgment. This study was supported by a Grant-in-Aid for Creative Scientific Research and for Young Researcher from the Ministry of Education, Culture, Sports, Sciences, and Technology of Japan and also by Japan Science and Technology Agency. One of the authors (K. F.) gratefully acknowledges the Research Fellowship for Young Scientists from the Japan Society for the Promotion of Science. A part of the computation was carried out in RCCS in Okazaki.

Supporting Information Available: The SAC-CI inputs on Gaussian03. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Kandori, H.; Schichida, Y.; Yoshisawa, T. *Biochemistry (Moscow*) **2001**, *66*, 1197. Mathies, R. A.; Lugtenburg, J. In *Handbook of Biological Physics*; Stavenga, D. G., Grip, W. J. d., Pugh, E. N., Eds.; Elsevier Science B. V.: Amsterdam, 2000.

(2) Stavenga, D. G.; Grip, W. J. d.; Pugh, E. N. In *Molecular Mechanisms in Viral Transduction*; Elsevier Science: New York, 2000.

(3) Nakanishi, K.; Crouch, R. *Isr. J. Chem.* **1995**, *35*, 253.

(4) Kleinschmidt, J.; Harosi, F. I. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 9181.

(5) Irving, C. S.; Byers, G. W.; Leermakers, P. A. *Biochemistry* **1970**, *9*, 858. Blatz, P. E.; Liebman, P. A. *Exp. Eye Res.* **1973**, *17*, 573. Honig, B.; Kinur, U.; Nakanishi, K.; Balogh-Nair, V.; Gawinowicz, M. A.; Arnaboldi, M.; Motto, M. G. *J. Am. Chem. Soc.* **1979**, *101*, 7084. Schulten, K.; Dinur, U.; Honig, B. *J. Chem. Phys.* **1980**, *73*, 3927. Béjá, O.; Spudich, E. N.; Spudich, J. L.; Leclerc, M.; DeLong, E. F. *Nature* **2001**, *411*, 786.

(6) Lin, S. W.; Kochendoerfer, G. G.; Carroll, K. S.; Wang, D.; Mathies, R.; Sakmar, T. P. *J. Biol. Chem.* **1998**, *273*, 24583.

(7) Okada, T.; Fujiyoshi, Y.; Silow, M.; Navarro, J.; Landau, E. M.; Shichida, Y. *Proc. Natl. Acad. Sci. U.S.A* **2002**, *99*, 5982.

(8) Luecke, H.; Schobert, B.; Richter, H. T.; Cartailler, J. P.; Lanyi, J. K. *J. Mol. Biol.* **1999**, *291*, 899.

(9) Royant, A.; Nollert, P.; Edman, K.; Neutze, R.; Landau, E. M.; Pebay-Peyroula, E. Navarro, J. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 10131.

(10) Hayashi, S.; Tajkhorshid, E.; Pebay-Peyroula, E.; Royant, A.; Landau, E. M.; Navarro, J.; Schulten, K. *J. Phys. Chem. B* **2001**, *105*, 10124.

(11) Schreiber, M.; Buss, V.; Sugihara, M. *J. Chem. Phys.* **2003**, *119*, 12045.

(12) Wanko, M.; Hoffmann, M.; Strodel, P.; Koslowski, A.; Thiel, W.; Neese, F.; Frauenheim, T.; Elstner, M. *J. Phys. Chem. B* **2005**, *109*, 3606.

(13) Andruniów, T.; Ferré, N.; Olivucci, M. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 17908.

(14) Fujimoto, K.; Hasegawa, J.; Hayashi, S.; Kato, S.; Nakatsuji, H. *Chem. Phys. Lett.* **2005**, *414*, 239.

(15) Kandori, H. *Chem. Ind.* **1995**, *18*, 735.

(16) Mathies, R.; Stryer, L. *Proc. Natl. Acad. Sci. U.S.A.* **1976**, *73*, 2169.

(17) Warshel, A. *J. Phys. Chem.* **1979**, *83*, 1640. Sakurai, M.; Sakata, K.; Saito, S.; Nakajima, S.; Inoue, Y. *J. Am. Chem. Soc.* **2003**, *125*, 3108.

(18) Warshel, A.; Chu, Z. T. *J. Phys. Chem. B* **2001**, *105*, 9857.

(19) Houjou, H.; Inoue, Y.; Sakurai, M. *J. Phys. Chem. B* **2001**, *105*, 867.

(20) Hayashi, S.; Ohmine, I. *J. Phys. Chem. B* **2000**, *104*, 10678.

(21) Vreven, T.; Morokuma, K. *Theor. Chem. Acc.* **2003**, *109*, 125.

(22) Ferré, N.; Olivucci, M. *J. Am. Chem. Soc.* **2003**, *125*, 6868.

(23) Gascon, J. A.; Batista, V. S. *Biophys. J.* **2004**, *87*, 2931.

(24) Nakayama, K.; Nakano, H.; Hirao, K. *Int. J. Quantum Chem.* **1998**, *66*, 157.

(25) Hufen, J.; Sugihara, M.; Buss, V. *J. Phys. Chem. B* **2004**, *108*, 20419.

(26) Sugihara, M.; Hufen, J.; Buss, V. *Biochemistry* **2006**, *45*, 801.

(27) Houjou, H.; Koyama, K.; Wada, M.; Sameshima, K.; Inoue, Y. Sakurai, M. *Chem. Phys. Lett.* **1998**, *294*, 162.

(28) Sugihara, M.; Buss, V.; Entel, P.; Elstner, M.; Frauenheim, T. *Biochemistry* **2002**, *41*, 15259.

(29) Okada, T.; Sugihara, M.; Bondar, A.; Elstner, M.; Entel, P.; Buss, V. *J. Mol. Biol.* **2004**, *342*, 571.

(30) Chizhov, I.; Schmies, G.; Seidel, R.; Sydor, J. R.; Lüttenberg, B.; Engelhard, M. *Biophys. J.* **1998**, *75*, 999.

(31) Birge, R. R.; Zhang, C. F. *J. Chem. Phys.* **1990**, *92*, 7178.

(32) Nakatsuji, H.; Hirao, K. *J. Chem. Phys.* **1978**, *68*, 2053.

(33) Nakatsuji, H. *Chem. Phys. Lett.* **1978**, *59*, 362. Nakatsuji, H. *Chem. Phys. Lett.* **1979**, *67*, 329. Nakatsuji, H. *Chem. Phys. Lett.* **1979**, *67*, 334.

(34) Nakatsuji, H. In *Computational Chemistry - Reviews of Current Trends*; Leszcynski, J., Ed.; World Scientific: Singapore, 1997.

(35) Ehara, M.; Hasegawa, J.; Nakatsuji, H. In *Theory and Applications of Computational Chemistry*; Dykstra, C., Frenking, G., Kim, K., Scuseria, G., Eds.; Elsevier Science: New York, 2006.

(36) Nakatsuji, H. *Chem. Phys.* **1983**, *75*, 425.

(37) Neese, F. *J. Chem. Phys.* **2003**, *119*, 9428.

(38) Lee, H. M.; Kim, J.; Kim, C. J.; Kim, K. S. *J. Chem. Phys.* **2002**, *116*, 6549.

(39) Andersen, L. H.; Nielsen, I. B.; Kristensen, M. B.; Ghazaly, M. O. A. E.; Haacke, S.; Nielsen, M. B.; Petersen, M. Å. *J. Am. Chem. Soc.* **2005**, *127*, 12347.

(40) Nielsen, I. B.; Lammich, L.; Andersen, L. H. *Phys. Rev. Lett.* **2006**, *96*, 018304.

(41) Gao, J.; Truhlar, D. G. *Annu. Rev. Phys. Chem.* **2002**, *53*, 467. Gogonea, V.; Suárez, D.; van der Vaart, A.; Merz, K. W. J. *Curr. Opin. Struct. Biol.* **2001**, *11*, 217. Cui, Q.; Karplus, M. *J. Am. Chem. Soc.* **2002**, *124*, 3093. Field, M. J.; Bash, P. A.; Karplus, M. *J. Comput. Chem.* **1990**, *11*, 700. Warshel, A.; Levitt, M. *J. Mol. Biol.* **1976**, *103*, 227. Warshel, A.; Chu, Z. T.; Hwang, J.-K. *Chem. Phys.* **1991**, *158*, 303.

(42) Hayashi, S.; Tajkhorshid, E.; Schulten, K. *Biophys. J.* **2002**, *83*, 1281.

(43) Dunning, T. H., Jr.; Hey, P. J. In *Method of Electronic Structure Theory*; Shaefer, H. F., III, Ed.; Plenum Press: New York, 1977; p 1.

(44) Wang, J.; Cieplak, P.; Kollman, P. A. *J. Comput. Chem.* **2000**, *21*, 1049.

(45) Jorgensen, W. L.; Chandreskhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1982**, *79*, 926.

(46) Schmidt, M. W.; Baldridge, K. K.; Boatz, J. A.; Elbert, S. T.; Gordon, M. S.; Jensen, J. H.; Koseki, S.; Matsunaga, N.; Nguyen, K. A.; Su, S. J.; Windus, T. L.; Dupuis, M.; Montgomery, J. A. *J. Comput. Chem.* **1993**, *14*, 1347.

(47) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R. J. A.; Montgomery, J.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J.; Ayala, P. Y.; Morokuma, K.; Voth, G.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.;

Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. Gaussian, Inc.: Pittsburgh, PA, 2003.

(48) Balashov, S. P.; Govindjee, R.; Kono, M.; Imasheva, E.; Lukashev, E.; Ebrey, T. G.; Crouch, R. K.; Menick, D. R.; Feng, Y. *Biochemistry* **1993**, *32*, 10331.

(49) Ikeura, Y.; Shimono, K.; Iwamoto, M.; Sudo, Y.; Kamo, N. *Photochem. Photobiol.* **2003**, *77*, 96.

(50) Aton, B.; Doukas, A. G.; Callender, R. H.; Becher, B.; Ebrey, T. G. *Biochemistry* **1977**, *16*, 2995.

(51) Ren, L.; Martin, C. H.; Wise, K. J.; Gillespie, N. B.; Luecke, H.; Lanyi, J. K.; Spudich, J. L.; Birge, R. R. *Biochemistry* **2001**, *40*, 13906.

(52) Hoffmann, M.; Wanko, M.; Strodel, P.; König, P. H.; Frauenheim, T.; Schulten, K.; Thiel, W.; Tajkhorshid, E.; Elstner, M. *J. Am. Chem. Sci.* **2006**, *128*.

(53) Blatz, P. E.; Mohler, J. H.; Navangul, H. V. *Biochemistry* **1972**, *11*, 848. Dukkipati, A.; Kusnetzow, A.; Babu, K. R.; Ramos, L.; Singh, D.; Knox, B. E.; Birge, R. R. *Biochemistry* **2002**, *41*, 9842.

(54) Teller, D. C.; Okada, T.; Behnke, C. A.; Palczewski, K.; Stenkamp, R. E. *Biochemistry* **2001**, *40*, 7761.

(55) Palczewski, K.; Kumasaka, T.; Hori, T.; Behnke, C. A.; Motoshima, H.; Fox, B. A.; Trong, I. L.; Teller, D. C.; Okada, T.; Stenkamp, R. E.; Yamamoto, M.; Miyano, M. *Science* **2000**, *289*.

(56) Röhrig, U. F.; Guidoni, L.; Laio, A.; Frank, I.; Rothlisberger, U. *J. Am. Chem. Soc.* **2004**, *126*, 15328.

(57) Röhrig, U. F.; Guidoni, L.; Rothlisberger, U. *Biochemistry* **2002**, *41*, 10799. Saam, J.; Tajkhorshid, E.; Hayashi, S.; Schulten, K. *Biophys. J.* **2002**, *83*, 3097.

(58) Eilers, M.; Reeves, P. J.; Ying, W.; Khorana, H. G.; Smith, S. O. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 487; Verhoeven, M. A.; Creemers, A. F. L.; Bovee-Geurts, P. H. M.; Grip, W. J. d.; Lugtenburg, J.; Groot, H. J. M. d. *Biochemistry* **2001**, *40*, 3283.

(59) Baasov, T.; Friedman, N.; Sheves, M. *Biochemistry* **1987**, *26*, 3210.

(60) Furutani, Y.; Shichida, Y.; Kandori, H. *Biochemistry* **2003**, *42*, 9619.

(61) Bondar, A.; Fischer, S.; Smith, J. C.; Elstner, M.; Suhai, S. *J. Am. Chem. Soc.* **2004**, *126*, 14668.

(62) Freedman, K. A.; Becker, R. S. *J. Am. Chem. Soc.* **1986**, *108*, 1245.

(63) Cembran, A.; González-Luque, R.; Altoè, P.; Merchán, M.; Bernardi, F.; Olivucci, M.; Garavelli, M. *J. Phys. Chem. A* **2005**, *109*, 6597.

CT6002687