JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

Article

pubs.acs.org/JACS

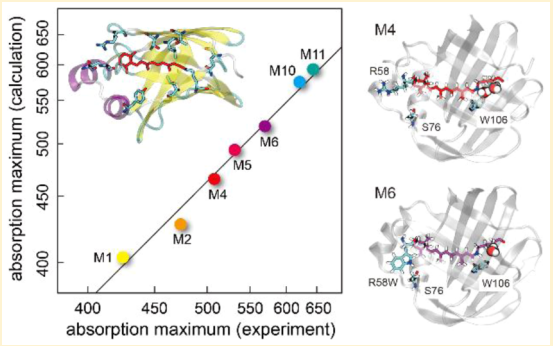
Molecular Mechanism of Wide Photoabsorption Spectral Shifts of Color Variants of Human Cellular Retinol Binding Protein II

Cheng Cheng, Motoshi Kamiya, Yoshihiro Uchida, and Shigehiko Hayashi*

Department of Chemistry, Graduate School of Science, Kyoto University, Kyoto 606-8502, Japan

Supporting Information

ABSTRACT: Color variants of human cellular retinol binding protein II (hCRBPII) created by protein engineering were recently shown to exhibit anomalously wide photoabsorption spectral shifts over ~200 nm across the visible region. The remarkable phenomenon provides a unique opportunity to gain insight into the molecular basis of the color tuning of retinal binding proteins for understanding of color vision as well as for engineering of novel color variants of retinal binding photoreceptor proteins employed in optogenetics. Here, we report a theoretical investigation of the molecular mechanism underlying the anomalously wide spectral shifts of the color variants of hCRBPII. Computational modeling of the color variants with hybrid molecular simulations of free energy geometry optimization succeeded in reproducing the experimentally observed wide spectral shifts, and revealed that protein flexibility, through which the active site structure of the protein and bound water molecules is altered by remote mutations, plays a significant role in inducing the large spectral shifts.



Photoabsorption maxima of rhodopsin photoreceptor membrane proteins are known to be widely distributed depending on their functions such as color vision, phototaxis, and photosynthesis.[1,2] The proteins called opsins bind a common chromophore, retinal, to their lysine residues through a Schiff base linkage. Retinal possesses a linear polyene structure with six conjugated double bonds, and protonation of the Schiff base brings its absorption maxima to the visible region. Absorption maximum (λ_{max}) of retinal protonated Schiff base (PSB) is highly sensitive to the surrounding environment, and thus undergoes large shifts from that in polar solvent (440 nm in methanol solution) upon binding to the proteins, called the opsin shift. Molecular mechanism of the color regulation of the opsin shift has extensively been investigated to obtain a molecular understanding of color vision as well as guiding principles for engineering of color variants of microbial-opsin based optogenetics tools to augment controllability of neuronal activity.[3,4]

Recently, Wang et al. engineered a soluble protein, human cellular retinol binding protein II (hCRBPII), and created mutants that can bind retinal with PSB linkage through Q108K mutation (Figure 1) and exhibit anomalously large spectral shifts.[5] Absorption maxima of the color variants made by point mutations at 10 positions at most widely range over ~200 nm from 425 to 644 nm. Especially, the opsin shift of the most red-shifted variant with 9 point mutations (λ_{max} = 644 nm) exceeds any of those of rhodopsin proteins ever found (λ_{max} < 590 nm), and even that of retinal PSB in vacuo (λ_{max} = 610 nm).[6] The remarkably wide spectral shifts over 200 nm and the "super-red" phenomenon obtained by engineering only with point mutations at no more than 10 positions provide one with a unique opportunity to uncover precise physicochemical mechanism of the color regulation of retinal binding proteins which opens the way to design of engineered photoreceptor proteins for optical visualization and manipulation of biological activity working with light in a wide spectral range including near-infrared optical window in biological tissues.

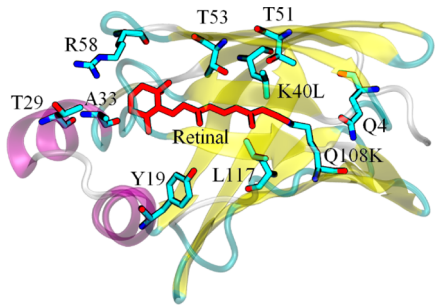


Figure 1. Protein structure of M4 variant of hCRBPII and positions of mutations introduced in the color variants.

Fortunately, steady accumulation of knowledge on molecular mechanism of the color tuning of rhodopsins has been achieved both theoretically and experimentally.[3,7] One of the major mechanisms is the electrostatic tuning. Since the chromophore with the PSB group exhibits positive charge redistribution along the polyene chain upon excitation, the absorption maximum is sensitive to the electrostatic field at the chromophore generated

Received: August 6, 2015
Published: September 29, 2015

DOI: 10.1021/jacs.5b08316
J. Am. Chem. Soc. 2015, 137, 13362−13370

by the surrounding protein environment. The other major mechanism is the conformational change of the chromophore exerted by the confinement in the protein binding pocket.

The chromophore−protein interaction leading to the spectral shift can theoretically be analyzed by hybrid quantum mechanical/molecular mechanical (QM/MM) methods.[8−17] One can calculate electrostatic contributions of the individual surrounding groups of protein, water molecules, and ions to the spectral shift to identify the molecular origins of the spectral shift. However, a difficulty arises in understanding anomalous features of the spectral shifts upon mutations experimentally observed for hCRBPII[5] through the conventional QM/MM analyses. First, large synergetic effects of mutations on the spectral shifts were often observed, indicating that the introduction of mutations causes significant structural changes of the protein. Furthermore, spectral shifts upon some mutations are hardly explained by the simple electrostatic mechanism. For example, all of mutations that replace a positively charged Arg58 in the vicinity of the β-ionone ring of retinal PSB to neutral amino acids (leucine, alanine, and glutamine) and negatively charged acidic ones (aspartic acid and glutamic acid) give rise to similar and moderate blue shifts by 8−9 nm, while mutations to aromatic amino acids (tryptophan, phenylalanine, and tyrosine) lead to considerably large red shifts by 16−27 nm. The observation of the hydrophobic and nonelectrostatic contributions suggested involvement of water molecules in the spectral shifts.[5] To elucidate the mechanism underlying the remarkable phenomena, therefore, a methodology capable of properly treating structural changes of the protein upon mutations in thermally fluctuating water environment yet with keeping high accuracy of the description of the electronic and geometric structures of the chemically complex chromophore is required.

Here, we report a theoretical investigation of the wide spectral shifts of the color variants of hCRBPII by means of hybrid molecular simulations of ab initio QM/MM free energy geometry optimization with QM/MM reweighting free energy-self-consistent field (RWFE-SCF) method developed recently.[18] In the method, the electronic wave function and the molecular geometry of the QM part are optimized on a mean field free energy surface constructed with statistical samples of MM configurations obtained by classical MD simulations. Accordingly, the direct coupling of the quantum mechanically treated electronic and geometric changes of the chromophore with slow and/or large conformational changes of the protein upon mutations treated with MM force field can be taken into account. Although the method neglects thermal fluctuation of the QM part, the approximation enables complete separation between the very time-consuming QM/MM calculation and the classical MD trajectory calculation for the statistical sampling of the protein conformations. The method, therefore, allows one to refine the modeled mutant structures through sufficient thermal relaxation by the classical MD simulations for the protein environment while maintaining the high accuracy of the structural determination of the electronically complex chromophore with ab initio QM treatment. The method was successfully applied to theoretical design of color variants of microbial rhodopsins, which exhibit spectral shifts by maximally 100 nm.[4]

In this study, we modeled seven color variants including mutations at 2−9 positions, M1, M2, M4, M5, M6, M10, and M11, experimentally created by Wang et al.[5] (see Figure 2a for the labeling), from the X-ray crystallographic structure of M4,

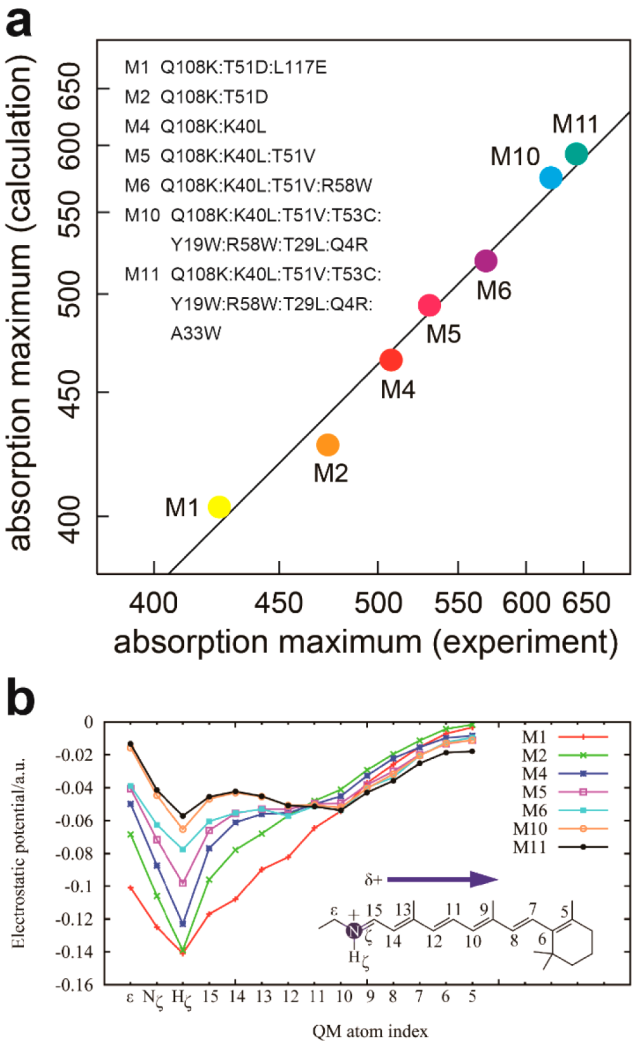


Figure 2. Calculated absorption maxima and electrostatic potentials on the chromophore of the color variants. (a) Comparison of calculated absorption maxima with experimental ones. The x and y axes were expressed in the energy scale. A line with a slope of 1 fitted to the data points is indicated. The translational shift of the line by a constant of 4.6 kcal/mol represents the systematic overestimation of the computed absorption excitation energies. Mutations in the color variants are indicated. (b) Electrostatic potentials along the polyene chain. Positive charge in PSB region delocalizes along the polyene chain upon photoexcitation.

which includes the fewest mutations at 2 positions. Excitation energies calculated for the computationally modeled structures of the mutants were shown to well reproduce the experimentally observed wide spectral shifts ranging over ~200 nm. Theoretical analyses for the mutant models revealed significant role of protein flexibility, through which the structure of the active site including bound water molecules is affected by remote mutations, in inducing the large spectral shifts.

COMPUTATIONAL METHODS

The initial structure of the protein (M4) was taken from an X-ray crystallographic structure[5] (PDB ID: 4EXZ, molecule A). The protein with the crystalline water molecules is immersed in a water box in the periodic boundary condition (Figure S1). We assigned the standard protonation state at the neutral pH to the titratable groups, except for Glu72 which was set to be protonated (see Supporting Information for

DOI: 10.1021/jacs.5b08316
J. Am. Chem. Soc. 2015, 137, 13362−13370

details and Figure S2). Sodium ions were put in the box with LEaP to neutralize the system. When the total charge of the protein changes upon the mutation, the number of sodium ions was changed to neutralize the total charge of the simulation system. The sodium ions in the system were not observed to play any significant role in the spectral shift (see Supporting Information for details).

The retinal PSB moiety was taken as the QM part. QM/MM RWFE-SCF free energy geometry optimizations incorporated in GAMESS program package[19] were carried out for all the color variants described above. The QM part is treated by density functional theory with M06-2X functional[20] and 6-31G** basis functions. As described below, accurate description of the electronic and geometric structure of the chemically complex chromophore treated quantum mechanically is a key factor for the modeling of the color variants. M06-2X functional, which is known to possess well-balanced functional forms and well-optimized parameters,[20] is therefore considered to be one of the best choices of DFT functional for the present purpose. All classical MD simulations for the sampling of MM conformations were performed with AMBER12 program package.[21] The MM region is described with ff99SB/TIP3P parameter set. For each cycle of the sequential sampling of QM/MM RWFE-SCF calculation,[18] a MD trajectory calculation of the MM region for 10 ns was carried out, and 50 000 conformations from the last half (5 ns) of the trajectory were used for the QM/MM RWFE-SCF calculation in each cycle of the sequential sampling.

Excitation energies were evaluated by XMCQDPT2(12,12) method[22] with 6-31G** basis functions under mean electrostatic field of MM conformational samples. The XMCQDPT method is one of the complete active space (CAS)-SCF based multiconfigurational perturbation theories, and thus, one of those providing physically straightforward descriptions of the excited state electronic state including both of effects of the multiconfiguration and the dynamics correlation with less empirical approximation at an affordable computational cost. The XMCQDPT method is therefore preferable to the present purpose to examine the spectral shifts by physical perturbations of the surrounding environment upon mutations, rather than to reproduce the absolute excitation energies. For analysis of electrostatic interaction between the chromophore and the MM environment, restrained electrostatic potential derived charges of the QM part with CAS-SCF(12,12) electron density were used. The computed electrostatic contributions by the analysis are therefore expected to be somewhat overestimated due to overpolarization of the CAS-SCF wave function.

Details of the simulation systems and the computational procedures are given in Supporting Information.

RESULTS

Modeling of Structures of Color Variants. The free energy geometry optimization was first performed for M4 color variant. The optimization required 50 cycles of the sequential sampling,[18] i.e., MD simulation for 500 ns in total (Figure S3c). During the simulations of M4 variant, two water molecules occupied a cavity in the vicinity of PSB in 15-syn form and one of them formed a hydrogen-bond with PSB (Figures 3a, S3, and S4). Although such water molecules occupying the cavity near PSB were absent in the X-ray crystallographic model employed for the initial structure in the present simulation (PDB ID: 4EXZ, molecule A), the water occupation in the simulations was found to be stable; even one of the water molecules was removed, and it returned to the cavity within 20 ns. The water occupation is attributed to a vacant cavity near PSB in 15-syn form existing in the X-ray crystallographic structure (Figure 3b); it is natural to observe the water occupation in the cavity in the MD simulation of the protein in bulk water molecules (Figure S1). The water occupation observed in the simulation is in line with the experimental observation that one water molecule was found in the corresponding cavity in the other protein models (PDB ID, 4EXZ, molecule B, and PDB ID, 4RUU), respectively.

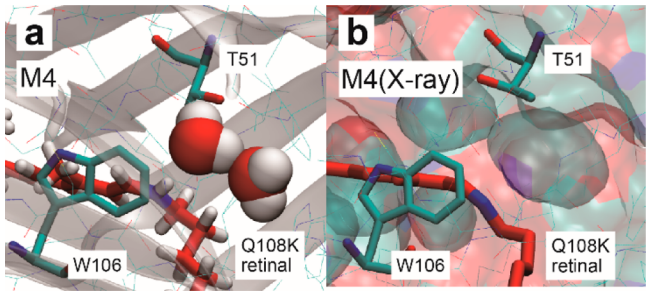


Figure 3. Bound water molecules around PSB in M4. (a) Bound water molecules around PSB in M4. PSB forms a direct hydrogen bond with one of them. (b) Cavities in an X-ray crystallographic structure (4EXZ).

The optimized structure of M4 variant was used for the starting structure for modeling of other color variants, M1, M2, M5, M6, M10, and M11. We first introduced mutations in M4 variant to create M2 and M5 ones and refined them by QM/MM RWFE-SCF calculations. The modeling was then extended to M1 and M6 to M11 one by one. For the modeling of M6 and M10, which involved large conformational changes, more careful procedures were employed. To minimize artifacts of manual replacements of the mutated residues, we first modeled M6:T29L mutant, which is one of the mutations to create M10 from M6 (see Figure 2a), for the modeling of M10. Moreover, for the modeling of M6 and M6:T29L, we carried out QM/MM RWFE-SCF calculations for ∼180 ns with fixed QM regions before QM/MM RWFE-SCF geometry optimizations (Figure S3e,f) to remove structural artifacts at the manually mutated residues through the sufficient structural relaxation by the preparatory QM/MM RWFE-SCF calculations. The numbers of the cycles of the sequential samplings required for the modeling of the color variants from M4 were 5−12, i.e., MD trajectory calculations for 50−120 ns in total (Figure S3), which are much shorter than the simulation for the free energy optimization of M4 from the X-ray crystallographic structure for more than 500 ns, because the starting structure of M4 has already been refined thoroughly by the free energy optimization. The present scheme of the one by one modeling of the various color variants from M4 is therefore considered to be computationally more efficient than the modeling of each color variant from each X-ray crystallographic structure of the corresponding color variant, which may require long simulations comparable to the modeling of M4, despite the fact that some of the X-ray crystallographic structures of the mutants are available. To confirm stability of the structure of M6, QM/MM RWFE-SCF geometry optimization was continued until 1 μs (Figure S5). Comparison of the free energetically optimized structures of the color variants with the X-ray crystallographic ones is given in Figure S6. Although the geometries of the chromophore underwent significant changes by the quantum chemical refinement, the overall mutual alignment of the chromophore and the surrounding groups was qualitatively kept. Details of the structural features of the modeled color variants other than M4 are described later.

Shifts of Absorption Maxima. Figure 2 and Table S1 compare the computed absorption maxima of the color variants with the experimental ones. The present calculations succeeded in reproducing the experimentally observed spectral shifts in the wide range of 219 nm (22.9 kcal/mol in energy). The

DOI: 10.1021/jacs.5b08316
J. Am. Chem. Soc. 2015, 137, 13362−13370

Table 1. Difference of Decomposed Electrostatic Contributions of Individual Groups and a Contribution of Conformational Change of the Chromophore to Exciting Energies of Color Variants from Those of M4 (kcal/mol)

	M1	M2	M4[f]	M5	M6	M10	M11
Q4R	-	-	0 (2.5)	-	-	−9.9	−7.5
Y19W	-	-	0 (−0.3)	-	-	0.2	0.2
T29L	-	-	0 (−0.2)	-	-	−0.2	−0.4
A33W	-	-	0 (0.2)	-	-	-	−0.9
(-K40L)[a]	−14.2	−11.8	0 (−0.7)	-	-	-	-
T51D/V	20.5[g]	20.2[g]	0 (1.3)	−1.5[h]	−1.6[h]	−2.6[h]	−2.5[h]
T53C	-	-	0 (0.6)	-	-	0.7	0.6
R58W	-	-	0 (8.7)	-	−8.9	−7.2	−7.1
L117E	12.6	-	0 (−1.4)	-	-	-	-
protein	16.6	8.1	0 (32.4)	4.3	−3.8	−12.5	−12.2
H_2O near PSB[b]	−6.4	0.1	0 (12.9)	−1.0	−6.9	−7.9	−12.6
H_2O (bulk)[c]	0.7	−1.2	0 (−28.5)	−6.2	5.9	10.0	12.1
QM/isolated[d]	1.1	1.1	0 (50.7)	−0.6	−2.1	−4.8	−5.4
Total shift[e]	9.4	5.5	0 (61.4)	−3.5	−6.3	−11.6	−13.2

[a]Removal of K40L mutation, which is introduced to create M4 variant from the wild type hCRBPII and is absent in M1 and M2. [b]Contributions of water molecules near PSB, i.e., those within 6 Å from both N_ζ and C_δ of Q108K. [c]Contributions of all of the water molecules in the simulation box except for the water molecules near PSB. Sodium ions in the bulk water provided contributions of −0.1 to ∼0.4 kcal/mol; the effect of the sodium ions is negligible. [d]Excitation energies computed with 3SA-XMCQDPT(12,12) method for the QM molecule without the surrounding environment. [e]Total spectral shifts from M4 variant (Figure 2 and Table S1). The total shifts do not coincide with the sums of the electrostatic contributions and the contributions of QM/isolated listed in the table (see Supporting Information for details). [f]Values in parentheses are contributions of M4 variant. [g]T51D mutant. [h]T51V mutant.

standard deviation of the computed spectral shifts of the color variants from M4 with respect to the experimental ones is 1.1 kcal/mol, which is much smaller than the overall shift (22.9 kcal/mol). The computed absolute values are, on the other hand, systematically overestimated by 4.6 kcal/mol (see below for detailed discussions). Nevertheless, the agreement of the relative spectral shifts in the wide range indicates that the present calculations successfully captured the molecular origins of the anomalously large spectral shifts.

As mentioned above, the primary mechanism that determines the spectral shifts is the electrostatic interaction of the chromophore with the surrounding environment. Figure 2b shows electrostatic potentials generated by the surrounding environment along the polyene chain of the chromophore. One can clearly discern that slopes of the electrostatic potentials of the color variants become more gradual as the absorption spectra are more red-shifted. Because the positive charge of the chromophore is relatively localized in the PSB region in the electronically ground state and it delocalizes along the polyene chain upon excitation, the slope of electrostatic potentials well correlates with the spectral shift caused by the electrostatic interaction.[3,7] The differences in the profile of electrostatic potential originate from those in polarity of the mutated residues as well as structural changes of the surrounding environment.

Color Variants with Blue Shifts: M1 and M2. The blue shifts of M1 and M2 from M4 are well explained by changes of direct electrostatic interactions by introductions of negatively charged carboxylate at the positions of 51 and 117 and removal of positively charged amine at the position 40 which locate in the PSB side of the chromophore (Figure 1). Decomposition of electrostatic contributions of the mutated residues to the excitation energy indicates electrostatic contributions by the mutations of T51D and K40L well explain the blue shift of M2 from M4 (Table 1 and Figure S7). In the case of M1, further introduction of the negatively charged group by the mutation of L117E gives a contribution to the further blue-shift (Table 1 and Figure S7), which is partly compensated by screening by water molecules brought in the vicinity of PSB by the introduction of the negatively charge group (Table 1 and Figure S3a).

Color Variants with Red Shifts: M5 and M6. In the case of the red-shifted color variants, indirect effects due to structural changes of the surrounding environment upon the mutations give significant contributions to the spectral shifts in addition to the contributions of the direct electrostatic interaction with the mutated residues. Especially, behavior of the bound water molecules near PSB found in the M4 model described above was observed to remarkably change upon the mutations.

The water molecules near PSB are perturbed by T51V mutation which is introduced in M5 and is persistently present in the further red-shifted color variants (Figures 2 and S3). In the case of M5, one of the two water molecules is absent (Figures S3d and S4), presumably due to the removal of the hydroxyl group of Thr51 which reduces polarity of the cavity near PSB (Figure 3). However, the hydrogen-bond between PSB and the remaining water molecule is still almost intact in M5 (Figure S4), and thus, the displacement of the water molecules near PSB gives a minor impact on the spectral shift (Table 1).

Drastic changes then appeared upon R58W mutation introduced in M6 and the further red-shifted variants. Arg58 is located on the surface of the protein in the vicinity of the β-ionone ring, which is the other end of the polyene chain of the chromophore opposite to PSB (Figure 1). Arg58 carrying a positive charge is exposed into bulk water solvent in M4 (Figure 4a and Movie S1). On the other hand, the indole ring of R58W firmly attaches to the protein surface due to hydrophobic interaction and forms a hydrogen-bond with the main-chain carbonyl group of Ser76 (Figure 4b and Movie S2). The formation of the hydrophobic interaction and the hydrogen-bond compresses the chromophore, and induces the structural changes around the remote PSB region, resulting

DOI: 10.1021/jacs.5b08316
J. Am. Chem. Soc. 2015, 137, 13362−13370

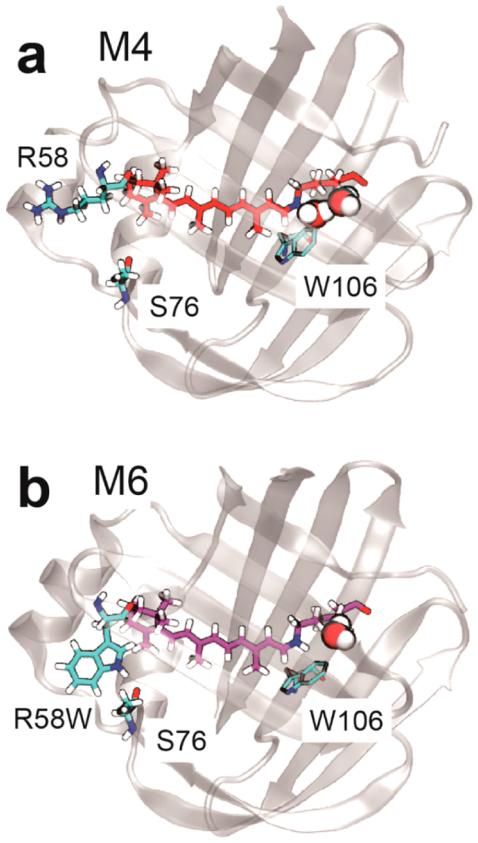


Figure 4. Conformations of Arg58 in M4 (a) and R58W in M6 (b).

in exclusion of one water molecule from the cavity near PSB (Figures 4a,b, 5a,b, and S3e). Furthermore, unlike M5, the exclusion of the water molecule is accompanied by complete dissociation of the hydrogen-bond between PSB and the remaining water molecule (Figure S4), and thus gives a significant electrostatic contribution (−6.9 kcal/mol) to the spectral shift (Table 1).

Figure 5c−e compares the free energetically optimized structures of the chromophore along with relative positions of Trp106 in the vicinity of PSB in M4 and M6. During the free energy optimization of M6, the chromophore underwent slight bend expressed by displacements of its both ends due to the compression of the chromophore (Figure 5d). The structural change of the chromophore accompanies the mutual positional shift of PSB and the indole ring of Trp106 by ∼0.4 Å. Consequently, the ring comes to cover H_ζ of PSB (Figures 5d and S8) and sterically hinders the formation of hydrogen-bond with the nearby water molecule (Figures 5b,c and S4). The dissociation of the hydrogen-bond was coupled with relaxation of torsion of the chromophore around PSB induced by the hydrogen bond (Figures 5e and S3e), and was observed to be stable in a QM/MM RWFE-SCF simulation for ∼800 ns (Figure S5). It is noteworthy that the displacement of the water molecules in the cavity near PSB did not occur during the simulation with the fixed geometry of the QM part at that of M5 for the first 180 ns (Figure S3e). This clearly indicates that the observed slight but significant changes of the geometry of the chromophore in the QM part are indispensable for the water displacement leading to the large spectral red shift, and the present method of the free energy geometry optimization capable of treating the direct coupling between the QM geometry change described at the ab initio level of theory and slow and global MM conformational changes on tens-to-hundreds of nanosecond time scale successfully uncovered the complex phenomena caused by the remote mutation.

The compression of the chromophore by R58W also significantly alters the electrostatic interaction of the chromophore with the surrounding protein groups correlated to the spectral shift. The removal of a positive charge by R58W mutation gives a large electrostatic contribution (−8.9 kcal/mol) to the spectral shift (Table 1 and Figure S9). However, the contribution of the direct interaction is largely reduced by screening by bulk water molecules which is estimated to be 5.9 kcal/mol in total (Table 1). The large dielectric compensation from bulk water molecules was also found in a previous study by Scott and Callis.[23] The observed screening of bulk water molecule is consistent with the experimentally observed insensitivity of polarity of residues replacing arginine at the position of 58 described above. Moreover, the contribution is largely compensated by an accumulated contribution of many amino acids in the C-terminal half of the protein, which form a β-sheet secondary structure next to PSB half of the chromophore (Figure S10), and consequently, the overall contribution of the protein comes to be −3.8 kcal/mol, which is smaller than that of the water molecules near PSB, −6.9 kcal/mol (Table 1). The feature of the accumulation of the small contributions clearly indicates that the contributions originate from a geometric shift of the chromophore in the binding pocket. The significant compensation may represent a counterbalance to the destabilization of PSB by the dissociation of the direct hydrogen-bond with the nearby water molecule.

As described above, the indole ring of R58W was observed to form a hydrophobic contact at the β-ionone ring side, inducing the compression of the chromophore. Given the experimental observation[5] that mutations with other aromatic amino acids, phenylalanine and tyrosine, give rise to the red shifts similar to that of R58W as described above, the similar mechanism is suggested for the hydrophobic spectral tuning. Furthermore, larger synergetic effects of R58W on the red shifts caused by additional mutations than that of R58Y were experimentally observed.[5] The formation of the hydrogen-bond of R58W with the main-chain carbonyl group of Ser76 observed in the simulation, which was also observed in the computational models of the further red-shifted color variants of M10 and M11, is naturally expected to be absent in R58Y, and thus explains the experimentally observed synergetic effect through better encapsulation of the chromophore in the binding pocket as suggested by Wang et al.,[5] by the supplementary hydrogen-bond formation. The effect may also be explained by a stronger hydrophobic interaction of tryptophan with a larger aromatic planar ring which cooperatively assists effects of the other mutations. Especially, as seen below, mutation of T29L, which is included in the further red-shifted color variants, M8-M11[5] (Figure 2), also induces remarkable conformational changes. Since the position of 29 is close to that of 58 (Figure 1) and the replacement of threonine by leucine increases hydrophobic environment as well, the establishment of strong hydrophobic interactions in the β-ionone ring side is also suggested to be a key factor for the conformational changes leading to the large red shifts.

Color Variants with Large Red Shifts: M10 and M11. The further red shifts of M10 and M11 are primary caused by the displacement of a water molecule. The free energy geometry optimization for T29L mutant introduced in M6, which is also included in M10 and M11 variants, showed

DOI: 10.1021/jacs.5b08316
J. Am. Chem. Soc. 2015, 137, 13362−13370

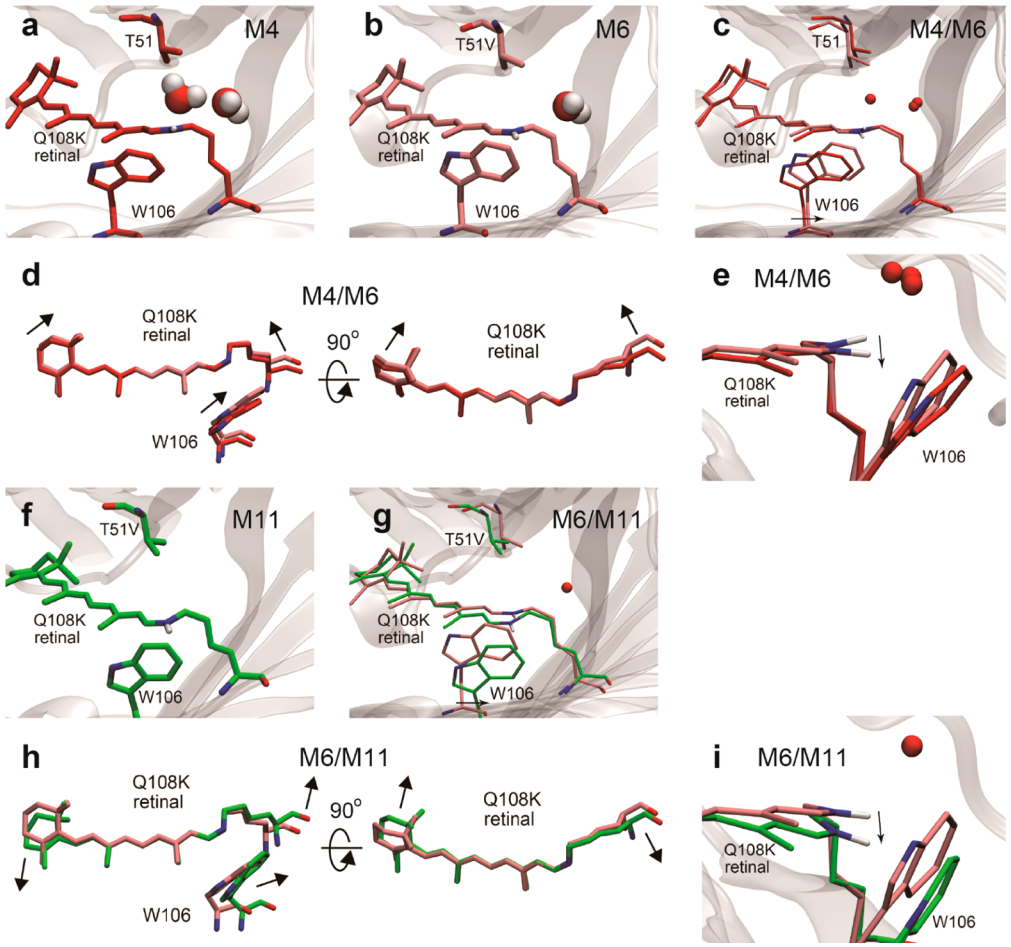


Figure 5. Conformational changes around the chromophore in the color variants. (a) Structure of M4. (b) Structure of of M6. A water molecule is excluded from the cavity near PSB. (c) Structural comparison between M4 (red) and M6 (pink). Trp106 is displaced with respect to PSB. (d) Comparison of the structures of the chromophore and Trp106 between M4 (red) and M6 (pink). Heavy atoms of the polyene chain of the chromophore (C_7 to N_ζ) are fitted. Bent of the chromophore accompanied by the positional shift of Trp106 is observed. (e) Torsional relaxation around PSB in M6 (pink) from M4 (red). (f) Structure of M11. (g) Structural comparison between M6 (pink) and M11 (green). Trp106 is largely displaced with respect to PSB. (h) Comparison of the structures of the chromophore and Trp106 between M6 (pink) and M11 (green). Heavy atoms of the polyene chain of the chromophore (C_7 to N_ζ) are fitted. Bend and torsion of the chromophore accompanied by the positional shift of Trp106 is observed. (i) Displacement of N_ζ-H_ζ bond of PSB and the indole ring of Trp106 in M6 (green) from M4 (pink).

exclusion of the water molecule near PSB (Figure S3f). Because, similar to R58W mutation, T29L mutation is located in the vicinity of the β-ionone ring (Figure 1) and increases hydrophobicity around it, a mechanism of the geometric shift of the chromophore in the binding pocket similar to that seen for R58W is considered to exclude the water molecule near PSB (see below). It is again noted that, similar to the case of M6, the exclusion of the water molecule did not occur during the simulation with the fixed geometry of the QM part at that of M6 for the first 180 ns (Figure S3f), indicating that the direct coupling between the geometric change of the QM part and the conformational changes of the protein upon the mutations is requisite for the displacement of the water molecule.

In the case of M10, however, Q4R mutation relatively close to PSB attracts water molecules in a cleft between PSB and Q4R (Figure S11), which compensate the red-shift. On the other hand, mutation of A33W added in M11 variant, which introduces a hydrophobic bulky group near the β-ionone ring (Figure 1), induces the further geometric shift of the chromophore in the binding pocket, completely excluding water molecules around PSB (Figures 5f and S3h). Figure 5g–i compares the free energetically optimized structures of the chromophore along with relative positions of Trp106 in M6 and M11. In M11, the chromophore undergoes significant bend and torsion, which are accompanied by the further developed mutual positional shift of PSB and the indole ring of Trp106 by ~1.1 Å with respect to M4 (Figures 5h and S8). The N_ζ–H_ζ bond of PSB is now completely covered by the indole ring (Figure 5i). The complete exclusion of the water molecules from the cavity near PSB gives a large electrostatic contribution of −12.6 kcal/mol (Table 1).

Changes of direct electrostatic interactions upon Q4R mutation may also contribute to the red shifts of M10 and M11 (Table 1, and Figures S12 and S13). Since the position of 4 locates in the PSB side of the chromophore (Figure 1), the introduction of the positive charge leads to a more gradual slope of electrostatic potential along the polyene chain of the chromophore (Figure 2b). However, contributions of screening by bulk water molecules compensating the electrostatic contribution of the direct interactions with the protein groups including Q4R are comparable (Table 1), and thus the direct electrostatic contribution of Q4R is considered to be less significant.

DOI: 10.1021/jacs.5b08316
J. Am. Chem. Soc. 2015, 137, 13362–13370

Experimental measurement of mutations of glutamine at the position of 4 in a red-shifted variant M8 (λ_{max} = 591 nm) showed that all the replacements with various residues including aromatic, hydrophobic, and polar ones gave large red shifts by more than 20 nm.[5] This remarkable phenomena can be explained by the observation in the simulations of M4 and M6 that the water molecule kept in the cavity near PSB, which plays a role in the spectral shifts as described above, forms a hydrogen-bond with the carbonyl group of the side chain of Gln4 (Figure S14). The replacement of the side chain of glutamine with any other side chains removing the hydrogen-bond is therefore expected to cause the displacement of the water molecule leading to the red shifts. The larger red shift of Q4R mutation may also include the electrostatic contribution described above.

Loss of an ordered bound water near PSB was also observed in an X-ray crystallographic model of a mutant almost identical to M10 (PDB ID: 4EEJ).[5] Since 15-syn conformation of PSB originally found in M4 variant is altered to be 15-anti one in the mutant, the displacement of the water molecule may be correlated with the syn−anti conformational change of PSB. We tested the hypothesis by modeling M10 with 15-anti form based on the X-ray crystallographic structure (4EEJ). The shift of the excitation energy from that of M4 calculated for the 15-anti model of M10 is −3.9 kcal/mol, which is much smaller than the experimental one (−10.3 kcal/mol) and the calculated one for the 15-syn model of M10 (−11.6 kcal/mol) described above. The large blue shift of the 15-anti model from that of the 15-syn one originates from formation of a direct hydrogen-bond of PSB with a water molecule occupying space opposite to Trp106; PSB in 15-syn form is free from screening of water molecules by Trp106, allowing it to attract water molecules and bind them (Figure S15).

Contributions of Chromophore's Structural Changes. Structural changes of the chromophore itself induced by the mutations also contribute to the spectral shifts to some extent as shown by the calculations of excitation energy in the absence of the interaction with the surrounding environment (Table 1). The main contributions come from changes in bond-alternation of the polyene chain (Figure S16) caused by (de)localization of the positive charge of PSB linking to the changes in electrostatic potentials (Figure 2b). The other contributions originate from torsion around C_6−C_7 bond.[4] The torsion by 154° in M4 is relaxed to be ∼168° in M10 and M11, giving contributions of ∼0.5 kcal/mol to the red shifts (Figure S17).

Electronic Contributions to the Spectral Shifts. The QM-MM interaction lacks various electronic interactions between the QM region and the MM one that affect the spectral shifts such as charge transfer,[24,25] excitonic coupling,[24] the MM polarization,[10,24−27] and the MM dispersion.[10] Contribution of the electronic interactions of the chromophore with the surrounding environment other than the electrostatic interaction to the spectral shift of M6 from M4 was estimated with the conformational samples of the free energetically optimized structures with the enlarged QM part including the side chains of the mutated residues, those of Glu38, Ser55, Thr60, and Trp106, which are close to the chromophore, and the water molecules near PSB in addition to the original QM part (see Supporting Information for the computational details). One of the water molecules included in the enlarged QM part is the one forming a direct hydrogen-bond with PSB in M4 and undergoes dissociation of the hydrogen-bond in M6. Since the excitation energy calculation with XMCQDPT method for the enlarged QM systems is computationally too demanding, we estimated the spectral shifts by the QM/MM calculation with configuration interaction single (CIS) method. The CIS calculations can take into account the electronic interactions of charge transfer and excitonic coupling between the chromophore and the surrounding groups upon excitation.[24]

Table S2 summarizes the spectral shifts computed with CIS for the QM/MM systems with the enlarged QM parts as well as the original QM parts. The spectral shifts for the enlarged QM system are comparable to those for the original systems (Table S2), indicating minor effect of the electronic interactions of charge transfer and excitonic coupling on the spectral shifts. Unlike other microbial rhodopsins where the positively charged PSB strongly interacts with negatively charged carboxylates (Asp85 and Asp212 in the case of bacteriorhodopsin) through hydrogen-bond network, and where the conjugated π systems of the polyene chain of the chromophore strongly interacts with an aromatic conjugated π system of tyrosine (Trp185 in the case of bacteriorhodopsin), hCRBP variants investigated in the present study do not include such strong interactions with the mutated residues. The minor effects of the electronic interactions of charge transfer and exitonic coupling on the spectral shift of hCRBP variants are therefore attributed to the small contributions of the individual mutated residues. Similarly, the polarization interaction, which can be regarded as a higher order term than the electrostatic one, and the dispersion interaction, the magnitude of which is usually much smaller than the electrostatic one, are expected to give only small contributions to the shift upon the point mutations. Overall, a minor contribution of the electronic interactions other than the electrostatic one to the spectral shifts is expected.

The result of CIS calculations also clearly indicates that the interaction between PSB and the water molecule forming the hydrogen-bond with PSB is described reasonably accurately by the QM-MM interaction. Since the hydrogen-bond of PSB with the water molecule is one of the key determinants of the anomalous spectral shift from M4 to M6, inclusion of the water molecule in addition to the chromophore in the QM region could give a better treatment of the interaction. However, in this case, one would need to perform a QM/MM MD simulation with the extended QM region for tens of nanoseconds to describe the large translocation of the water molecule upon the mutations of M6 from M4 as observed in the present study. Unfortunately, such a QM/MM MD simulation for tens of nanoseconds is computationally not feasible at presently available computational resources. The extended QM region for the QM/MM RWFE-SCF calculations is also not considered to be a good choice, because conformational changes of the QM region are described by geometry optimization without thermal fluctuation in the methodology, and thus, the large translocations of the water molecule induced by thermal relaxation upon the mutations revealed in the present study are not expected to be well captured by the simulations with the extended QM region. Given the reasonably accurate description of the QM-MM interaction between PSB and the water molecule described above, therefore, we consider that the choice of the QM region in the present study is the best one for the purpose of the present study, i.e., elucidation of the molecular mechanism underlying the anomalous large spectral shift upon the mutations.

DOI: 10.1021/jacs.5b08316
J. Am. Chem. Soc. 2015, 137, 13362−13370

Systematic Overestimation of Calculated Excitation Energies. The excitation energies calculated in the present study are systematically overestimated by 4.6 kcal/mol on average from the experimental ones. A part of the systematic error is considered to be associated with quality of the QM calculations such as that of basis functions utilized for the excitation energy calculation and methodology for the geometry optimization calculation[28] (see Supporting Information for details). It should be noted that the systematic deviations associated with the basis functions and the QM methodologies are not expected to largely affect the spectral shifts upon the point mutations focused on in the present study, and thus, the excitation energy calculations of the color variants with the basis functions and the methodologies employed in the present study are sufficiently accurate for the present purpose.

The systematic overestimation of the excitation energies also possibly come from poor description of the QM-MM interaction which lacks the electronic interactions other than the electrostatic one. The electronic interactions can give rise to red-shifts of the absorption maximum.[10,24−27] As described above, the electronic interactions are expected to give only small contributions to the spectral shift upon the point mutations. On the other hand, the accumulation of those contributions from the entire protein and the surrounding bulk water environment can give rise to a significant overall shift of the absolute excitation energy, even though each contribution from the individual group in the system is small. The lack of the accumulated electronic contributions from the entire groups of the system is presumably responsible for a part of the systematic overestimation of the excitation energies. Especially, unlike rhodopsins composed of transmembrane α-helixes, hCRBPII variants possess a β-sheet in the C-terminal half of the protein which is located next to PSB half of the chromophore. The β-sheet consists of many π bonds of the amide groups of the main chain aligning parallel to the π conjugated plane of the chromophore. Since the π bonds are electronically more active than σ ones, the electronic interactions of the chromophore upon the excitation with the π bonds are larger than those of σ bonds. The proximity of the many π bonds of the β-sheet and the chromophore and their parallel alignment are therefore expected to give rise to the accumulated shifts of the excitation energy.

DISCUSSIONS AND CONCLUDING REMARKS

The present study theoretically elucidated molecular mechanisms of the anomalously wide spectral shifts of the color variants of hCRBPII. Increase of hydrophobicity through exclusion of water molecules from interior of the protein by mutations has been suggested to account for the large spectral shifts previously.[5,29] The present study revealed in atomistic detail that the hydrophobic mutations around the β-ionone ring lead to the displacements of the water molecules in a remote cavity near PSB, which greatly contribute to the large spectral shifts. This hardly predictable remote effect of the mutations is realized by characteristic nature of the globally correlated flexibility of protein. The present approach of the free energy geometry optimization with QM/MM RWFE-SCF method succeeded in taking into account the complex protein flexibility underlying the spectral shifts through combining the highly accurate QM method at the ab initio level of theory to treat the electronic and geometric structure of the chromophore with long time MD simulations to describe sufficient structural relaxation of the protein upon the mutations.

The present analysis provides a clue to design of color variants with further red-shifted absorption maxima toward the optical window. In the most red-shifted variant, M11, electrostatic potentials in PSB half of the chromophore are greatly increased and flattened (Figure 2b), although the potentials are still negative and thus keep the Schiff base protonated. Further increase of the electrostatic potentials in PSB half could therefore lead to deprotonation of the Schiff base. On the other hand, negative electrostatic potentials in the β-ionone ring half are weak (Figure 2b), suggesting mutations that generates negative electrostatic potentials in the β-ionone ring producing a further red-shift. However, mutations introducing polar groups in the hydrophobic environment around the β-ionone ring may evoke complex structural response to the mutations. The theoretical modeling will be a powerful approach to decipher such complex structural changes upon mutations and to provide physicochemical basis for development of functionally novel variants of photoreceptor and luminescent proteins including rhodopsins.[4]

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/jacs.5b08316.

Additional methods and data (PDF)
Movie S1 (MPG)
Movie S2 (MPG)

AUTHOR INFORMATION

Corresponding Author

*hayashig@kuchem.kyoto-u.ac.jp

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work was financially supported by grants from the Japanese Ministry of Education, Culture, Sports, Science and Technology (25104004 and 25291034), and the Programme for Promotion of Basic and Applied Researches for Innovations in Bio-oriented Industry and Science and Technology Research Promotion Program for Agriculture, Forestry, Fisheries and Food Industry. Some computations were performed at the Research Center for Computational Science, Okazaki, Japan. Molecular images are created with VMD.[30]

REFERENCES

(1) Palczewski, K. *J. Biol. Chem.* **2012**, *287*, 1612.
(2) Spudich, J. L.; Yang, C. S.; Jung, K. H.; Spudich, E. N. *Annu. Rev. Cell Dev. Biol.* **2000**, *16*, 365.
(3) Ernst, O. P.; Lodowski, D. T.; Elstner, M.; Hegemann, P.; Brown, L. S.; Kandori, H. *Chem. Rev.* **2014**, *114*, 126.
(4) Kato, H. E.; Kamiya, M.; Sugo, S.; Ito, J.; Taniguchi, R.; Orito, A.; Hirata, K.; Inutsuka, A.; Yamanaka, A.; Maturana, A. D.; Ishitani, R.; Sudo, Y.; Hayashi, S.; Nureki, O. *Nat. Commun.* **2015**, *6*, 7177.
(5) Wang, W.; Nossoni, Z.; Berbasova, T.; Watson, C. T.; Yapici, I.; Lee, K. S. S.; Vasileiou, C.; Geiger, J. H.; Borhan, B. *Science* **2012**, *338*, 1340.
(6) Andersen, L. H.; Nielsen, I. B.; Kristensen, M. B.; El Ghazaly, M. O. A.; Haacke, S.; Nielsen, M. B.; Petersen, M. A. *J. Am. Chem. Soc.* **2005**, *127*, 12347.

DOI: 10.1021/jacs.5b08316
J. Am. Chem. Soc. 2015, 137, 13362−13370

(7) Schulten, K.; Hayashi, S. *Quantum Effects in Biology*; Cambridge University Press: Cambridge, United Kingdom, 2014 Chapter 11.
(8) Hayashi, S.; Ohmine, I. J. *J. Phys. Chem. B* **2000**, *104*, 10678.
(9) Hayashi, S.; Tajkhorshid, E.; Pebay-Peyroula, E.; Royant, A.; Landau, E. M.; Navarro, J.; Schulten, K. *J. Phys. Chem. B* **2001**, *105*, 10124.
(10) 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.
(11) Ferré, N.; Olivucci, M. *J. Am. Chem. Soc.* **2003**, *125*, 6868.
(12) Hoffmann, M.; Wanko, M.; Strodel, P.; König, P. H.; Frauenheim, T.; Schulten, K.; Thiel, W.; Tajkhorshid, E.; Elstner, M. *J. Am. Chem. Soc.* **2006**, *128*, 10808.
(13) Fujimoto, K.; Hayashi, S.; Hasegawa, J.; Nakatsuji, H. *J. Chem. Theory Comput.* **2007**, *3*, 605.
(14) Rajamani, R.; Lin, Y.-L.; Gao, J. *J. Comput. Chem.* **2011**, *32*, 854.
(15) Sekharan, S.; Katayama, K.; Kandori, H.; Morokuma, K. *J. Am. Chem. Soc.* **2012**, *134*, 10706.
(16) Pal, R.; Sekharan, S.; Wei, J. N.; Batista, V. S. *J. Am. Chem. Soc.* **2013**, *135*, 9624.
(17) Huntress, M. M.; Gozem, S.; Malley, K. R.; Jailaubekov, A. E.; Vasileiou, C.; Vengris, M.; Geiger, J. H.; Borhan, B.; Schapiro, I.; Larsen, D. S.; Olivucci, M. *J. Phys. Chem. B* **2013**, *117*, 10053.
(18) Kosugi, T.; Hayashi, S. *J. Chem. Theory Comput.* **2012**, *8*, 322.
(19) 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.; Windus, T. L.; Dupuis, M.; Montgomery, J. A. *J. Comput. Chem.* **1993**, *14*, 1347.
(20) Zhao, Y.; Truhlar, D. G. *Theor. Chem. Acc.* **2008**, *120*, 215.
(21) Case, D. A. et al. *AMBER 12*, University of California, San Francisco, CA, 2012 (see Supporting Information for the complete list of the developers).
(22) Granovsky, A. A. *J. Chem. Phys.* **2011**, *134*, 214113.
(23) Scott, J. N.; Callis, P. R. *J. Phys. Chem. B* **2013**, *117*, 9598.
(24) Hasegawa, J. *Chem. Phys. Lett.* **2013**, *571*, 77.
(25) Wanko, M.; Hoffmann, M.; Frauenheim, T.; Elstner, M. *J. Phys. Chem. B* **2008**, *112*, 11462.
(26) Houjou, H.; Inoue, Y.; Sakurai, M. *J. Phys. Chem. B* **2001**, *105*, 867.
(27) Wanko, M.; Hoffmann, M.; Frähmcke, J.; Frauenheim, T.; Elstner, M. *J. Phys. Chem. B* **2008**, *112*, 11468.
(28) Valsson, O.; Angeli, C.; Filippi, C. *Phys. Chem. Chem. Phys.* **2012**, *14*, 11015.
(29) Wang, W.; Geiger, J. H.; Borhan, B. *BioEssays* **2014**, *36*, 65.
(30) Humphrey, W.; Dalke, A.; Schulten, K. *J. Mol. Graphics* **1996**, *14*, 33.

DOI: 10.1021/jacs.5b08316
J. Am. Chem. Soc. 2015, 137, 13362−13370