Proton Transfer and Associated Molecular Rearrangements in the Photocycle of Photoactive Yellow Protein: Role of Water Molecular Migration on the Proton Transfer Reaction

Motoshi Kamiya,[†] Shinji Saito,[‡] and Iwao Ohmine*[,†]

Chemistry Department, Nagoya University, 1 Furocho, Chikusaku, Nagoya 464-8602, Japan, and Department of Computational Molecular Science, Institute for Molecular Science, Myodaiji, Okazaki, Aichi 444-8585, Japan

Received: September 25, 2006; In Final Form: December 28, 2006

The mechanism of the proton transfer and the concomitant molecular structural and hydrogen bond rearrangements after the photoisomerization of the chromophore in the photocycle of photoactive yellow protein are theoretically investigated by using the QM/MM method and molecular dynamics calculations. The free energy surface along this proton-transfer process is determined. This work suggests the important role of the water molecular migration into the moiety of chromophore, which facilitates proton transfer by the hydrogen bond rearrangement and the hydration of the pB′ state.

1. Introduction

Photoactive yellow protein (PYP)[1] from purple bacteria *Halorhodospira halophila*[2] is considered to be the photosensor for the negative phototaxis.[3] PYP is a small water-soluble protein and has a structure motif called PAS,[4,5] named after the three proteins Per, Arnt, and Sim. PYP contains a *p*-coumaric acid as a chromophore which covalently binds to Cys69. In the ground state, pG, the chromophore is deprotonated. The photocycle of PYP, shown in Figure 1, is initiated by the visible light absorption with absorption maximum 446 nm. Upon this excitation, the vinyl C═C bond of the chromophore isomerizes from the trans to cis form.[6,7] Then, the first long-lived intermediate pR state (also known as I_1 and PYP_L) with the absorption maximum at 465 nm appears. The secondary and tertiary structures of the pR state are very similar to those of the pG state. In both states, the chromophore makes hydrogen bonds to the surrounding residues, Tyr42 and Glu46.[7,8] The next step of the photocycle involves proton transfer from Glu46 to the chromophore.[9,10] The intermediate state after this proton transfer is called pB′.[10,11] The protonated chromophore then moves to the protein surface. These structural changes, the proton-transfer reaction and the surface exposure of the chromophore, induce the unfolding of the N-terminal region of the protein.[12,13] There exists the partially unfolded structure called pB (also known as I_2 and PYP_M) after the pB′ state. The existence of the pR_2 state is also suggested between the pR and pB′ states.[14−16]

In this work, we focus on the molecular structural rearrangement associated with the proton transfer from Glu46 to the chromophore after photoisomerization. As this proton-transfer reaction takes several hundred microseconds for completion,[10] it is conceivable that a large structural rearrangement around the chromophore associated with the proton transfer takes place. We here investigate what kind of the rearrangement takes place by using molecular dynamics (MD) and quantum mechanical/molecular mechanical (QM/MM) calculations.[17−21] We find that

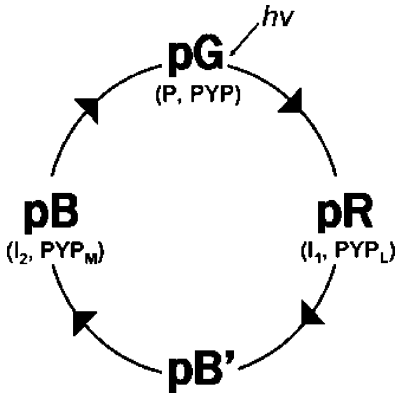

Figure 1. Photocycle of PYP.

migration and coordination of water molecules are involved in this structural rearrangement.[22] These water molecules play an essential role in the displacement of Arg52, which is the critical structural change for the surface exposure of the chromophore after the proton transfer.[23]

2. Method

The ground-state X-ray structure (PDB: 2PHY)[24] was used as an initial structure of the pG state. For the pR state, we employed the cryotrapped crystal structure (PDB: 3PYP)[25] deposited at the Protein Data Bank, which is known as the PYP_{BL}[26] or PYP_B[27] state, as an initial configuration. After the equilibration of this pR structure (see below), we obtain the structure similar to that relaxed from the recently proposed pR structure (PDB: 1UWP).[28] We added 9568 and 9652 TIP3P water molecules[29] spherically surrounding the protein for the pG and pR states, respectively. In these states, Glu46 was assumed to be protonated. The AMBER99 force field[30,31] was used for the parameters of amino acids. Atomic charges in the chromophore were estimated from the electrostatic potential (ESP) of the optimized structure with the HF/6-31G* level calculation.[32,33,34] Here, we used Gaussian98[35] for the ESP and the quantum chemical energy contribution of QM/MM calculations. In the QM/MM calculation, we calculated the quantum mechanical contribution with B3LYP/6-31G**.[36]

* Corresponding author. E-mail: ohmine@chem.nagoya-u.ac.jp.
† Nagoya University. E-mail for M.K.: kamo@chem.nagoya-u.ac.jp.
‡ Institute for Molecular Science. E-mail for S.S.: shinji@ims.ac.jp.

10.1021/jp066256u CCC: $37.00 © 2007 American Chemical Society
Published on Web 02/27/2007

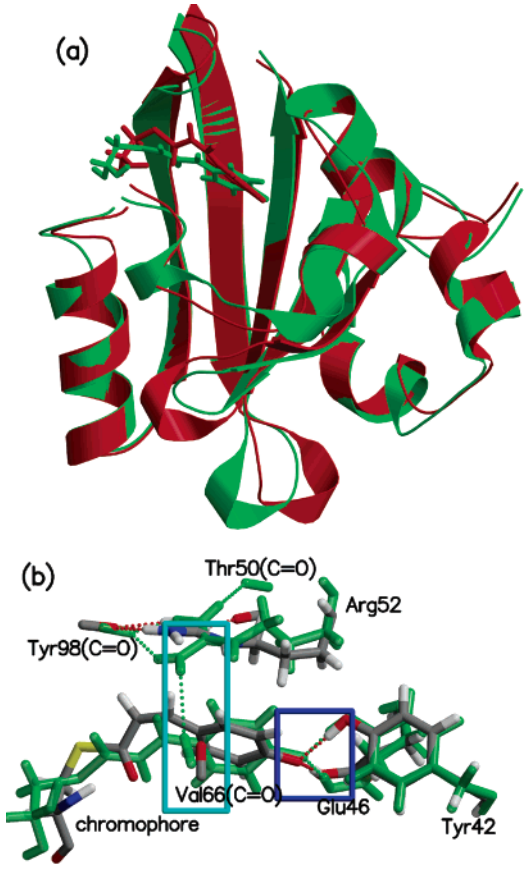

Figure 2. Average structures of pG (green) and pR (red), which are superimposed. (a) Shown is the overall structure of PYP. The chromophore is shown in sticks. (b) Shown is the expanded structure around Arg52, in which hydrogen bonds are shown in dashed lines, hydrogen bonds of the chromophore are enclosed with the blue line, and the hydrogen bond between Arg52 and Val66 in pG is enclosed with the cyan line. Atoms of pR are colored in a standard way (i.e., carbon atoms with gray, hydrogen atoms with white, oxygen atoms with red, nitrogen atoms with blue, and sulfur atoms with yellow), while all atoms of pG are colored with semitransparent green. The Fortran subroutine written by Seok et al.[45] is used for the structural superimposition. All figures of the protein structure are made with MOLSCRIPT[46] and Raster3D.[47]

Molecular dynamics (MD) calculations were performed using a molecular mechanics (MM) method described below (MM-MD) at constant-temperature condition with Nose-Hoover's thermostat.[37] Velocity Verlet method with 1 fs time step was used for the time evolution of the system. In all MD simulations, covalent bond lengths involving hydrogen atoms were constrained with RATTLE.[38] The cell multipole method[39] was used for the computation of the electrostatic interactions. It was shown that a small energy drift sometimes appears when the cell multipole method or fast multipole method is employed for the calculation of the electrostatic interaction.[40] However, the drift is found to be almost negligible in the present calculation where the thermostat is employed. To keep all water molecules in the hydration sphere, we applied a harmonic force with 1.5 kcal/mol Å^2 force constant on water molecules if they are tend to go out from the sphere. The effects of the boundary condition and the system size might be considered.[41] The center of mass of the protein was also constrained by the harmonic potential with the same force constant of the water constraint above to avoid escaping from the sphere. The whole system was equilibrated at 150 K for 50 ps. The system was then equilibrated at 300 K for 1 ns. After the initialization, all MD calculations were performed in the NVT ensemble at room temperature (300 K).

The free energy profiles for structural changes except for the proton transfer were calculated by using the umbrella sampling method,[42] in which a harmonic potential was used for the constraint potential. The weighted histogram analysis method (WHAM)[43,44] was used for combining the unbiased free energies for different windows. Configuration samplings are repeated until the free energy profile is converged. For example, 600 000 configurations from 24 ns trajectory (as total, not counting those for the equilibration) were needed to generate the free energy profile shown in Figure 4. Free energy calculation was performed on the basis of the perturbation method. The free energy difference between states i and j is expressed by

$$\Delta F_{ij} \equiv F_j - F_i = -k_B T \ln \frac{Z_j}{Z_i}$$

where Z_i is a partition function of state i and k_B is the Boltzmann constant. The ratio of the partition functions is rewritten as

$$\frac{Z_j}{Z_i} = \frac{\int e^{-\beta E_j}\, dr^N}{\int e^{-\beta E_i}\, dr^N} = \frac{\int e^{-\beta(E_j - E_i)} e^{-\beta E_i}\, dr^N}{\int e^{-\beta E_i}\, dr^N} \qquad (2\text{-}1)$$

where E_i is the QM/MM energy of state i. Since the configuration sampling with the QM/MM calculation is very time-consuming, we used the following equations to reduce the computational time:

$$\begin{aligned}
\frac{\int e^{-\beta(E_j - E_i)} e^{-\beta E_i} dr^N}{\int e^{-\beta E_i}\, dr^N} &= \\
&\frac{\int e^{-\beta(E_j - E_i)} e^{-\beta(E_j - E_{i,MM})} e^{-\beta E_{i,MM}} dr^N}{\int e^{-\beta E_{i,MM}} dr^N} \frac{\int e^{-\beta E_{i,MM}}\, dr^N}{\int e^{-\beta E_i}\, dr^N} \\
&= \langle e^{-\beta(E_j - E_i)} e^{-\beta(E_j - E_{i,MM})} \rangle_{i,MM} \\
&\qquad \frac{\int e^{-\beta E_{i,MM}}\, dr^N}{\int e^{-\beta(E_i - E_{i,MM})} e^{-\beta E_{i,MM}} dr^N} \\
&= \langle e^{-\beta(E_j - E_i)} e^{-\beta(E_i - E_{i,MM})} \rangle_{i,MM} / \langle e^{-\beta(E_i - E_{i,MM})} \rangle_{i,MM}
\end{aligned} \qquad (2\text{-}2)$$

Here $E_{i,MM}$ is the MM energy of state i. In eq 2-2, the average is taken over MM-generated configurational space for the state i and the i and j state energies, E_i and E_j, are calculated by QM/MM calculation for those configurations. These representations, eqs 2-1 and 2-2, are identical if the QM/MM-generated configurations are the same as the MM-generated ones. It is found that there is a good correlation between $E_{i,MM}$ and E for MM-MD-generated configurations (not shown in figures), and thus we may expect that the QM/MM-MD (if performed) generated configurations are not much different from the MM-MD-generated configurations.

The free energy profile of the proton-transfer reaction between Glu46 and the chromophore was evaluated with the QM/MM method as described above, in which i denotes the position of the transferring proton. A proton-transfer reaction involves the cleavage and the formation of covalent bonds. The parameter set to evaluate the energy of the proton-transfer segment, that is, D (donor)−H (transferring proton)−A (acceptor) segment, does not exist in the present MM method (AMBER99), and

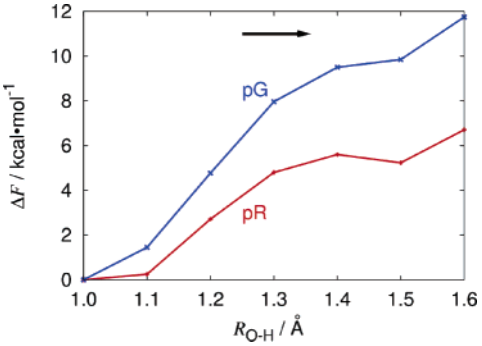

Figure 3. Free energy profiles of the proton-transfer reaction in pG (blue) and pR (red). The abscissa axis is the distance between one of the oxygen atoms of Glu46 forming the hydrogen bond with the chromophore and the transferring proton. The black arrow indicates the direction of the proton transfer from Glu46 to the chromophore. The O−O distance between the oxygen atom of Glu46 and the phenolic oxygen of the chromophore is fixed at 2.6 Å, and all other degrees of freedom except for the length of covalent bonds involving the hydrogen atom fixed using RATTLE are not constrained. In these calculations, only Tyr42, Glu46, and the chromophore are treated as the QM region. See the Method section for the details of the free energy calculations.

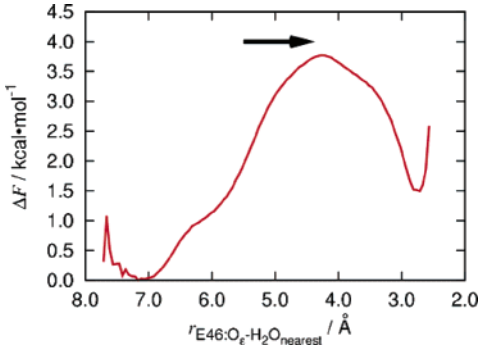

Figure 4. Free energy profile of the coordination of a water molecule to Glu46. The abscissa axis is the distance between the side-chain carbonyl oxygen atom of Glu46, not forming the hydrogen bond with the chromophore, and the nearest water molecule from that oxygen atom. The black arrow indicates the direction of the water molecular coordination to Glu46.

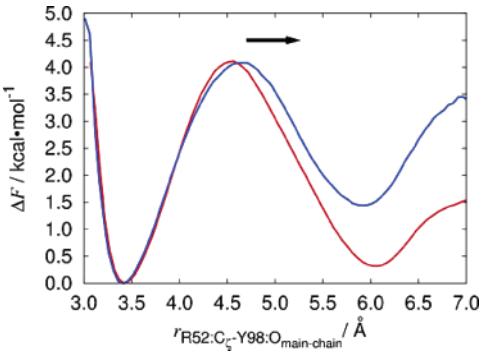

Figure 5. Free energy profile of the displacement of Arg52 without (blue) and with (red) a coordinated water molecule to Glu46. The abscissa axis is the distance between the backbone carbonyl oxygen atom of Tyr98 and the central carbon atom of the guanidinium group (C_ζ) of Arg52. The black arrow indicates the direction of the breaking of the hydrogen bond between Arg52 and Tyr98.

thus, $E_{i,MM}$ cannot be obtained. But it is not of the matter. The energy of this D (donor)−H (transferring proton)−A (acceptor) segment is constant if its geometry is fixed; the rigid D−H−A structure is assumed when we generate the MM configurations for each *i*. The constant contributions of $E_{i,MM}$ to the numerator cancels out to those of the denominator in eq 2-2. The distribution of the atomic charges in this moiety is evaluated from the QM/MM calculation. The MM-MD calculation is then performed using these atomic charges to generate the surrounding MM configurations (see below). The QM region consists of Tyr42, Glu46, and the chromophore Cys69. In addition to these residues, if hydrophilic residues or water molecules locate near the chromophore, they are also included in the QM region. We adopt the conventional treatment for the interaction between QM and MM moieties,[18] and inserted hydrogen atoms between the α- and *β*-carbons of Tyr42, Glu46, and Cys69, at the QM-MM interfaces as link atoms. All other atoms, amino acid residues and solvent molecules, are treated as the MM part and assumed to interact with QM region via Lennard-Jones and electrostatic interactions. Charges of atoms in those residues are recalculated at given *i* using the same method for the determination of atomic charges in the chromophore described above.

These recalculated charges are used for classical samplings. A trajectory calculation is performed for each position of the proton, *i*, for longer than 1 ns, and its initial 500 ps is discarded as the equilibration. For each *i*, evenly separated 1000−4000 configurations are sampled from the trajectory, and then QM/MM energies are evaluated for these configurations. The free energies are calculated in both forward and backward directions of the proton transfer. Computational software for the MM part of QM/MM calculation, MD calculation, free energy calculation, etc., used in the work has been developed in the laboratory of the present authors.

3. Results and Discussion

3.1. Structural Changes Caused by Photoisomerization. Figure 2a superimposes the average structures of pG (green) and pR (red) obtained using the MD calculation. As clearly seen in Figure 2a, there are no significant differences in the secondary and the tertiary structures between the pG and pR states. In the pG state, the chromophore makes the hydrogen bonds to Tyr42 and Glu46.[7] During the photoisomerization of the chromophore, leading to the pR state, these hydrogen bonds are intact, while a small structural change is induced around Arg52, as seen in Figure 2b. Arg52 has three hydrogen bonds with carbonyl oxygen atoms of Thr50, Val66, and Tyr98 in the pG state. The isomerization of the vinyl C═C bond of the chromophore pushes away Arg52 and results in the weakening of the hydrogen bond between Arg52 and Val66 in the pR state.

The photoisomerization is then followed by the proton transfer from Glu46 to the chromophore.[9,10,27,48,49] Figure 3 plots the QM/MM free energy profile of the proton transfer in the pR state (red), where the abscissa axis is taken to be the O−H distance between the carboxyl oxygen atom of Glu46 and the proton transferring to the chromophore, and the O−O distance between the phenolic oxygen atom of the chromophore and the carboxyl oxygen atom of Glu46 is fixed at 2.6 Å. In the configuration samplings, all other degrees of freedom are not constrained, except for the lengths of the covalent bonds involving hydrogen atoms fixed using RATTLE method. No significant structural rearrangement of the surrounding is seen during this proton transfer. We can see that the proton transferred state is unstable by about 5 kcal/mol compared with the pR state. Figure 3 also plots the free energy profile along the proton transfer assumed to take place in the pG state (blue). We can see in the figure that the free energy increase with the proton transfer in the pR state is smaller than that in the pG state. This is attributed to the decreased *π*-conjugation of the chromophore in the pR state because of the decreased planarity of the

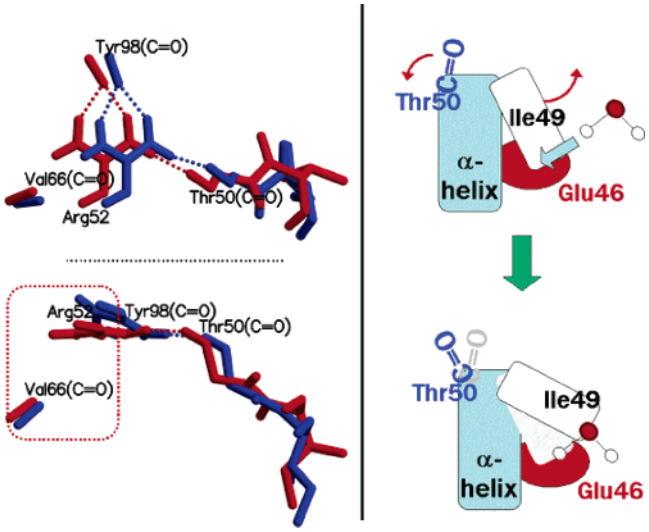

Figure 6. Schematic views of the displacement of Arg52: (right) scheme of the structural rearrangement caused by the coordination of a water molecule to Glu46; (left) superimposed average structure of before (blue) and after (red) the coordination of a water molecule to Glu46. The region enclosed in the red line shows that the distance between Arg52 and Val66 shortens as a result of the coordination of a water molecule to Glu46.

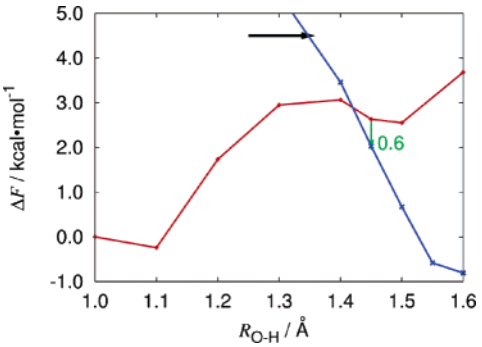

Figure 7. Free energy profile of the proton-transfer reaction for the water coordinated (red) and water migrated (blue) structures. The abscissa axis is taken to be the O−H distance between the carboxyl oxygen atom of Glu46 and the proton transferring to the chromophore, and the O−O distance between the phenolic oxygen atom of the chromophore and the carboxyl oxygen of Glu46 is fixed at 2.6 Å, as the same as those in Figure 3. The black arrow indicates the direction of the proton transfer from Glu46 to the chromophore. The distance between phenolic oxygen atoms of the chromophore and Tyr42 is also fixed. The water molecule coordinated to Glu46 is included in the QM region of the QM/MM calculation of the red line profile. For the blue line profile, six closest water molecules from C_δ of Glu46, the carbon atom of the side-chain carboxyl group, and two arginine residues, Arg52 and Arg124, are also included in the QM region. The difference in structure and energy between the red and blue curves is discussed in Figures 8 and 9.

chromophore caused by the steric repulsion between the thioester oxygen and the aromatic ring of the chromophore and to the weakening/break of hydrogen bonds among the thioester oxygen atom of the chromophore and the surrounding. The thioester oxygen atom makes hydrogen bonds with the amide hydrogen atoms in the protein backbone of Cys69, Asp97, and Tyr98 in the pG state. Among these three hydrogen bonds, only the first one is strong and stable, while the other two hydrogen bonds are weak, and thus repeats breaking and re-forming in the pG state. These hydrogen bonds become much weaker in the pR state.

It is noted here that, to obtain the more accurate free energy profile of the proton transfer, the long-range charge-transfer effect on the proton moiety should be taken into account.[21] The further improvement may be obtained in the present method if the constant coefficients are introduced to reduce the interactions between the QM and the MM parts to reproduce pK_a values.[21,50] The MD calculations including the electronic effects, such as empirical valence bond (EVB) method,[51−58] might be employed, instead of the MM-MD calculation.

3.2. Coordination of a Water Molecule to Glu46 and the Rearrangement of the Hydrogen Bond Network around Arg52. Since Glu46 locates near the protein surface, water molecules may coordinate to Glu46. Figure 4 shows, for example, the free energy change when a water molecule approaches the side chain of Glu46. Here, the abscissa axis is the distance between one of the side-chain oxygen atoms of Glu46, not forming hydrogen bond with the chromophore, and the nearest water molecule from this oxygen atom. The minimum at the distance around 7 Å corresponds to the pR state before the proton transfer without the coordination of a water molecule, while the other minimum at distance around 3 Å corresponds to the pR state with a water molecule coordinating to Glu46. Since the free energy difference between these two minima is, although positive, not too large (1.5 kcal/mol), it is conceivable that the coordination of a water molecule to Glu46 takes place.

We found in the MD trajectories that subsequent structural changes follow this coordination; Arg52 changes its hydrogen bond partner from Tyr98 to Val66. Along this process, the nature of the hydrogen bond between Thr50 and Arg52 is found to be almost unchanged. The free energy profile of this process is plotted in Figure 5 (red). The distance between the central carbon atom of the guanidinium group of Arg52 and the backbone carbonyl oxygen atom of Tyr98 is taken as the abscissa axis. Arg52 forms two hydrogen bonds, to Thr50 and Tyr98, respectively, around 3.5 Å. On the other hand, Arg52 forms two hydrogen bonds, to Thr50 and Val66, around at 6.0 Å. As two hydrogen bonds are formed at these two minima, their energies are about equal.

The blue line plots the free energy profile of the same process but when a water molecule is not coordinated to Glu46. In this case, Arg52 makes two hydrogen bonds, to Thr50 and Tyr98, respectively, around 3.5 Å, while it makes one hydrogen bond only to Thr50 around at 6.0 Å. Thus, the free energy at 6.0 Å in the blue line becomes higher than that in the red line, by about 1.5 kcal/mol. The displacement of the guanidinium group of Arg52 after the photon absorption could be attributed to the weakening of the hydrogen bond between Arg52 and Val66.[59]

Since the water molecule which coordinates to Glu46 is very far, about 10 Å apart from Arg52, there is a long-range effect disturbing the hydrogen bond network around Arg52. This may be a domino-like mechanism. The coordination of a water molecule to Glu46 displaces Ile49, which locates between Glu46 and solvent molecules (bulk water). The displaced Ile49 then pushes Thr50 and Arg52. As a result, Arg52 approaches Val66, as shown in the red-colored structure of Figure 6.

3.3. Proton-Transfer Reaction and the Migration of Water Molecules. Figure 7 shows the free energy profile of the proton transfer from Glu46 to the chromophore, with the water coordination to Glu46. In this calculation, the water molecule coordinating to Glu46 is added to the QM region, and the distance between the phenolic oxygen atoms of the chromophore and Tyr42 is constrained to be fixed to avoid the large rearrangement in the hydrogen bond network. We can see in the figure that the proton-transfer reaction with a water molecule coordinated to Glu46 (red) is unstable by about 2 kcal/mol at

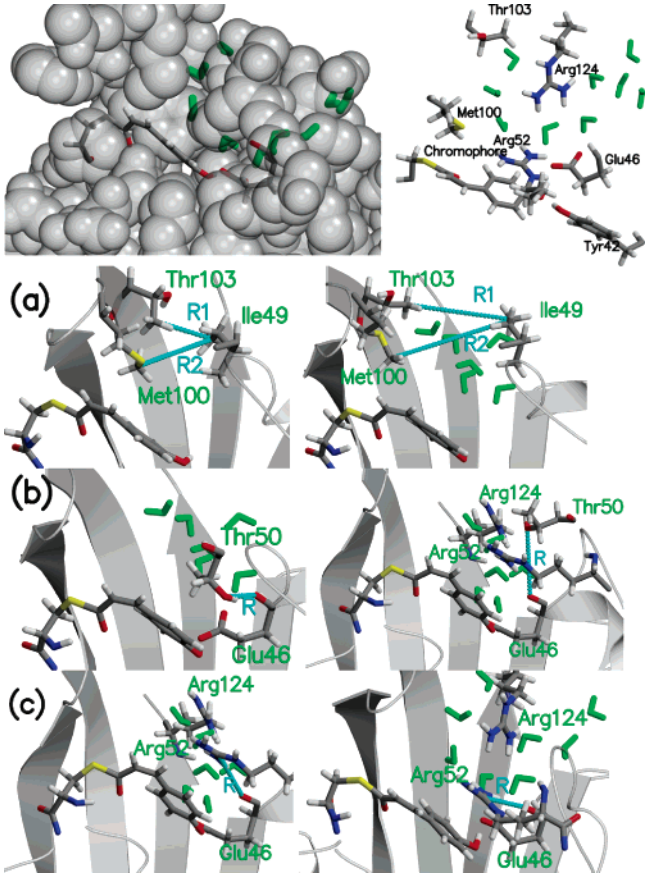

Figure 8. Instantaneous structure after the migration of water molecules. In the upper left figure, Glu46, the chromphore, and migrated water molecules are shown in sticks, and all other atoms are shown in spacefill. Migrated water molecules are colored with green. (a) Shown is the the structure of before (left) and after (right) the opening up of the protein. The sum of distances of R_1 and R_2 which are shown in cyan is used as the reaction coordinate for the free energy calculation in Figure 9a. (b) Shown is the breaking of the hydrogen bond between Glu46 and Thr50. This hydrogen bond forms as a result of the migration of water molecules, and thus it does not exist in the left structure of (a). The distance R shown in cyan is used as the reaction coordinate for the free energy calculation in Figure 9b. (c) Shown is the formation of the hydrogen bond between Glu46 and Arg52. The distance between backbone carbonyl oxygen of Glu46 and the central carbon atom of the guanidinium group of Arg52, R, is used as the reaction coordinate for the free energy calculation in Figure 9c.

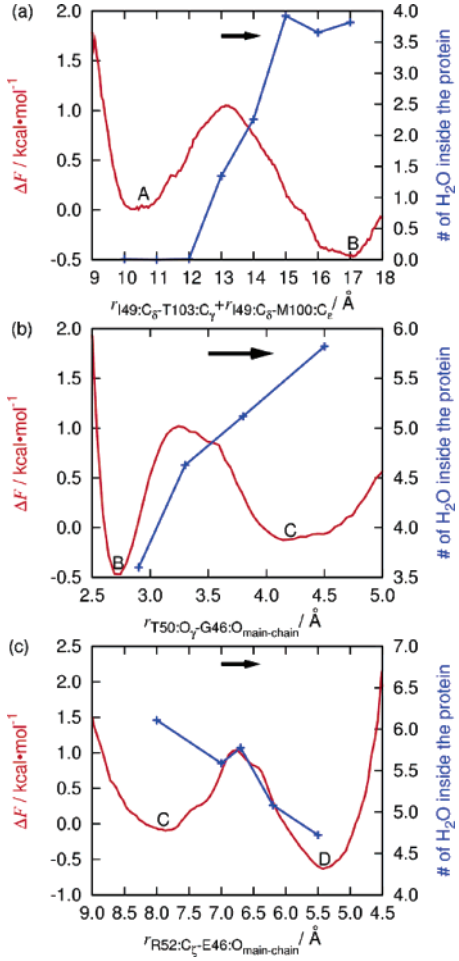

Figure 9. Free energy profile of the migration of water molecules at $R_{O-H} = 1.45$ Å (see Figure 7). Each plot shows the free energy profile of (a) the opening up of the protein and migration of water molecules, (b) the breaking of the hydrogen bond between Glu46 and Thr50, and (c) the formation of the hydrogen bond between Glu46 and Arg52. The reaction proceeds from "A" to "D" via "B" and "C". The abscissa axis of (a)−(c) is the sum of the distance between C_δ of Ile49 and C_γ of Thr103 and the distance between C_δ of Ile49 and C_ϵ of Met100, the distance between the backbone oxygen atom of Glu46 and side-chain hydroxyl oxygen atom of Thr50, and the distance between the Glu46 oxygen atom and the central carbon atom of the guanidinium group of Arg52, respectively (see Figure 8). The average number of water molecules located inside the polyhedron, which vertices the carboxyl carbon C_δ of Glu46, the vinyl carbon atom adjacent to the phenol ring of the chromophore, and C_α atoms of Ile49, Phe96, Thr103, Lys106, Val122, and Arg124 are also shown. The black arrows in (a)−(c) indicate the direction of the opening up of the protein, breaking of the hydrogen bond between Glu46 and Thr50, and the formation of the hydrogen bond between Glu46 and Arg52, respectively.

$R_{O-H} = 1.5$ Å. Only one water molecule thus cannot sufficiently stabilize the negative charge on Glu46.

In MD trajectories employed for the free energy calculation, we have found another kind of the protein structure for the proton transferred state, as shown in Figure 8, where several water molecules, colored with green, migrate into the interior of the protein. The bridging between Ile49 and the segment of Met100 and Thr103 is collapsed by the breaking of the hydrogen bond between Arg52 and Tyr98, which is caused with the coordination of a water molecule to Glu46 (as described above) and the increase of the Glu46 negative charge associated with the proton transfer. These migrated water molecules hydrate the negatively charged Glu46.[60] During this hydration process, two arginine residues of Arg52 and Arg124 with plus charges are pulled toward Glu46 with the minus charge (see the right side of Figure 8c) and then the hydrogen bonds of Arg52 to Thr50 and Val66 are broken, inducing the destabilization of the helix involving Glu46, Ile49, and Thr50. The free energy profile of the proton transfer on the water-migrated structure is shown in the blue line of Figure 7, in which six closest water molecules from Glu46 and the two arginine residues, Arg52 and Arg124, are also included in the QM region. The distance between the phenolic oxygen atoms of the chromophore and Tyr42 here is again constrained to be fixed to avoid the large rearrangement in the hydrogen bond network, as used in the calculation of the red line profile of the figure. In this water-migrated structure, the proton-transferred structure is quite stable due to the stabilization of the negative charge on Glu46 with the migrated water molecules. We use the term pB_w to refer the water-migrated proton transferred state, which corresponds to the state at $R_{O-H} = 1.60$ Å on the blue line of Figure 7. Such structural change was not found in the calculation before. There were several energy barriers (Figures 4 and 5) to accomplish the full migration of water molecules, and the system

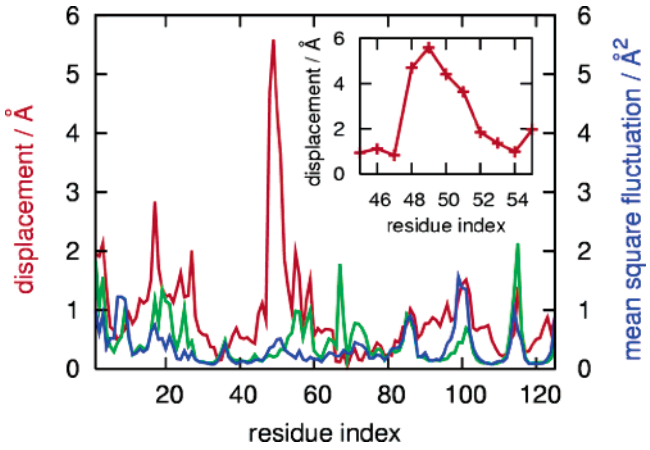

Figure 10. Difference in the average position of individual C_αs between pR and pB_w. The differences of residues around Glu46 are expanded in the inset. Mean square fluctuations of each of the C_αs in pR (green) and pB_w (blue) are also shown.

cannot readily go over these barriers in MD calculation within a few nanoseconds.

We evaluate the free energy difference between the water-coordinated structure (red) and the water-migrated structure (blue) at the fixed R_{O-H} of 1.45 Å, O−O distance between the chromophore and Glu46 of 2.6 Å, and O−O distance between the chromophore and Tyr42 of 2.73 Å. To do so, we have divided their structural difference into three parts: (a) the collapse of the surface exposed region and the migration of the water molecules, (b) the breaking of the hydrogen bond between the backbone carbonyl oxygen of Glu46 and the hydroxyl group of Thr50, and (c) the orientational change of Arg52, as shown in Figure 8a−c, respectively. The free energy profiles of individual divisions are calculated by using the classical umbrella sampling. The reaction coordinate of part a is chosen as the sum of the distance between C_δ of Ile49 and C_γ of Thr103 and the distance between C_δ of Ile49 and C_ϵ of Met100. These distances largely change with the opening up of the protein and the migration of water molecules. The distance between the backbone oxygen atom of Glu46 and the side-chain hydroxyl oxygen atom of Thr50 is chosen for the reaction coordinate of part b, and the distance between the Glu46 oxygen atom and the central carbon atom of the guanidinium group of Arg52 is chosen for the reaction coordinate of part c. The free energy profiles obtained for parts a−c are shown in Figure 9a−c, respectively, in which the number of water molecules inside the protein is also plotted. The total free energy difference between two structures, namely "A" in Figure 9a and "D" in Figure 9c, at $R_{OH} = 1.45$ Å, is −0.6 kcal/mol, as indicated with the green line in Figure 7. Since the relative position of these two free energy curves is thus determined, we can compare the two profiles, the red and the blue curves in the figure, and construct a "one-dimensional" picture of water migration; that is, the sequential or the simultaneous migration of water molecules takes place along the proton transfer. The exact mechanism of the water molecular migration can be obtained only when the multidimensional free energy surfaces could be determined for different sets of the water coordination and the different degrees of hydration.

The number of water molecules increases associated both with the collapsing of the surface exposed region from "A" to "B" in Figure 9a and with the breaking of the hydrogen bond between backbone carbonyl oxygen of Glu46 and Thr50, from "B" to "C" in Figure 9b. The number of migrated water molecules decreases due to their replacement with Arg52 (Figure 9c). As a result, about 5 water molecules migrate into the protein in the final structure, "D". These water molecules and two arginine residues significantly stabilize the negative charge on Glu46.

The difference in the average position of individual C_αs between pR and pB_w is seen in Figure 10. Interestingly, the displacements of Glu46 (1.1 Å) and Arg52 (1.8 Å) are not so large. On the other hand, residues from Asp48 to Gly51 show large displacements (inset of Figure 10). Although the position of Gly47 itself yields the small displacement, 0.83 Å, the torsion angle of the protein backbone changes significantly at Gly47. Asp48 thus shows the extended displacement of 4.7 Å. It should be noted that the N-terminal region, residue positions 1−28, fluctuates over the extended space, that is, its mean-deviation from the average structure is large in both pR (green) and pB_w (blue) states, and so the direct comparison between their average structures is not meaningful for this region.

The free energy diagram of the pathway from pR to pB_w is summarized in Figure 11: (i) the coordination of a water molecule to Glu46 (corresponding to Figure 4); (ii) the breaking of the hydrogen bond between Arg52 and Tyr98 (corresponding

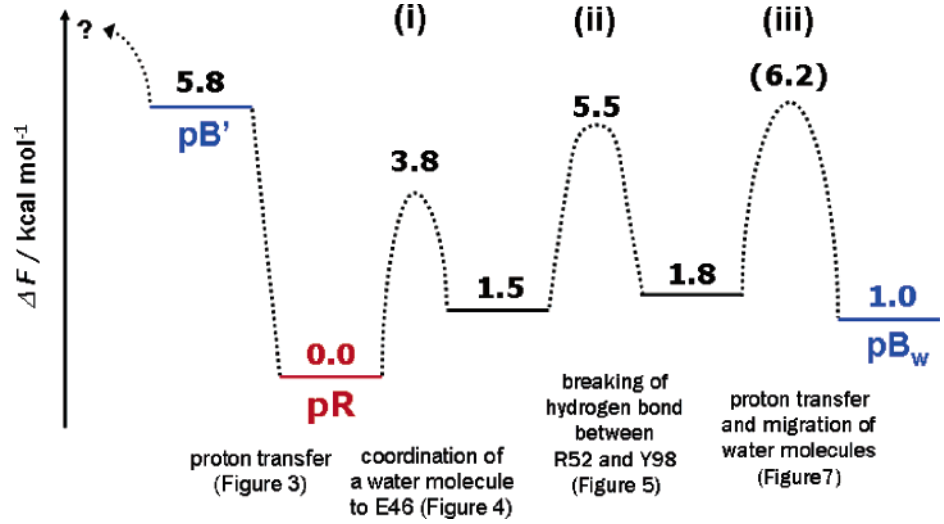

Figure 11. Free energy diagram, in which i−iii represents a water molecule coordinates to Glu46 (corresponding to Figure 4), the breaking of the hydrogen bond between Arg52 and Tyr98 (corresponding to Figure 5) and the proton transfer from Glu46 to the chromophore and the migration of the water molecules, and the orientational change of two arginines (corresponding to Figure 7 which involves the steps of Figure 9a−c), respectively.

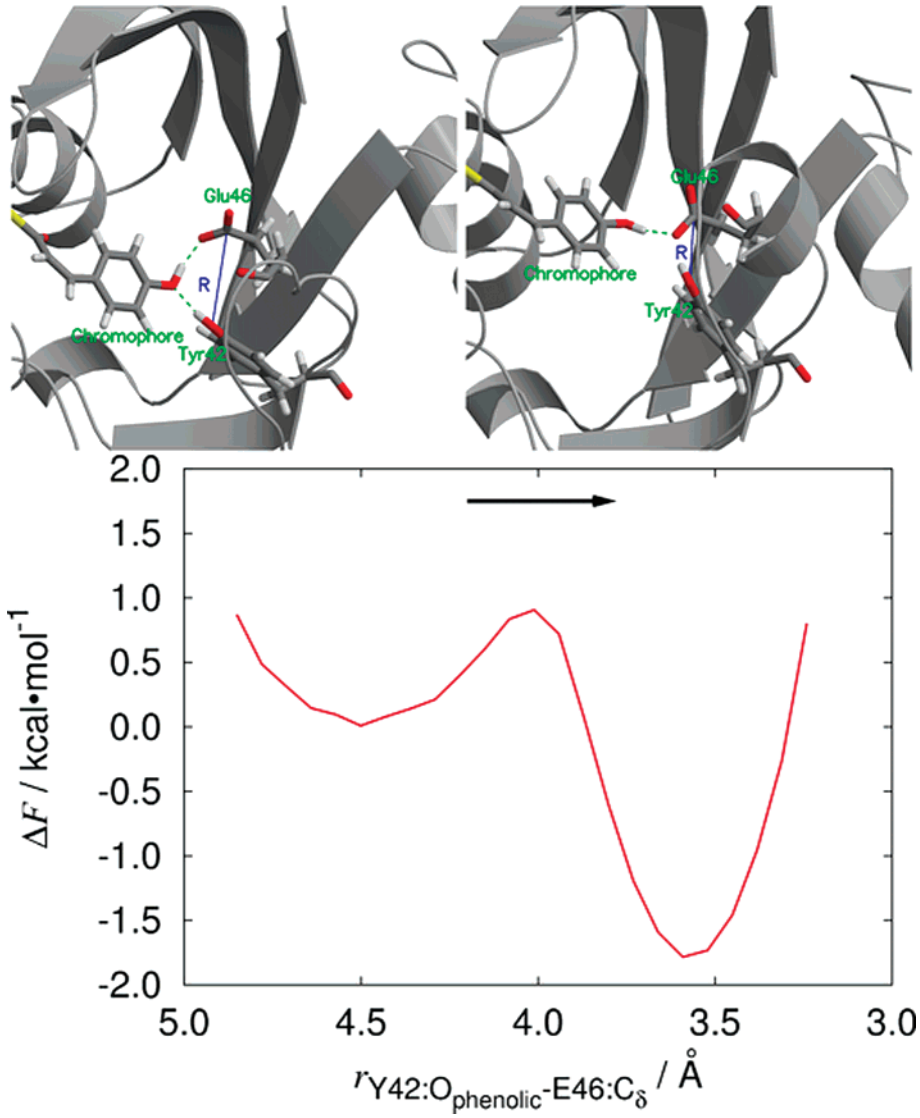

Figure 12. Free energy profile of the local hydrogen bond network relaxation around the chromophore after the proton transfer and the migration of water molecules. The abscissa axis is the distance between the phenolic oxygen atom of Tyr42 and the carboxyl carbon atom of Glu46, and the black arrow indicates the formation of the hydrogen bond between Tyr42 and Glu46 and the breaking of the hydrogen bond between Tyr42 and the chromophore. The left (right) structure corresponds to the average structure of the left (right) minimum of the free energy profile.

to Figure 5); (iii) the proton transfer from Glu46 to the chromophore and the migration of the water molecules and the orientational change of two arginines (corresponding to Figure 7 which involves the steps of Figure 9a−c), where the pB_w state has the protonated chromophore and has several water molecules inside the protein moiety. Of course, the sequence of these structural changes may be altered; for example, a sequence in which the step ii appears first and then steps i and iii follow is possible. Apparently, the pR state is more stable than the pB′ and pB_w states. In any case, the barrier heights along the free energy diagram, relative to pR, are not too large, and furthermore the pB_w state is much more stable than the pB′ state, showing that the water molecular migration associated with the proton-transfer reaction is energetically feasible and thus it may actually take place in the PYP photocycle.

In addition, it can explain that the experimental result[22] suggesting that the extent of the hydration of the protein affects the process from the pR to pB′ states. Moreover, the migrated water molecules may reduce the energy barrier for the surface exposure of the chromophore and the breaking of the hydrogen bond between Glu46 and the chromophore after the proton transfer by means of hydrogen bond exchanges. The position of Arg52 in the pR state locates on the pathway of the surface exposure of chromophore after the proton-transfer reaction. The displacement of Arg52 (Figure 5) is thus required for this exposure.

The analysis here is limited to only deal with a path of the local water migration associated with the proton transfer. There might be many other paths of HB and molecular rearrangements, which take place after the photoisomerization of chromophore. To find such possibilities, kinetic Monte Carlo,[61] hyperdynamics,[62] and other methods[63] should be applied.

The pB_w state is still higher in free energy by 1.0 kcal/mol in comparison with the pR state. This difference decreases and even their ordering in the free energy alters, when we take account of the relaxation of the hydrogen bond network around the chromophore in pB_w. The full consideration of the hydrogen bond network relaxation in the pB_w state, however, is very difficult since its entire configurational space is so vast. We have thus evaluated it partially by calculating the free energy change associated with only a small rearrangement in the moiety of the chromophore in the pB_w state. Figure 12 shows the free energy profile of the relaxation of the hydrogen bond network around the chromophore, in which the reaction coordinate is

the distance between the phenolic oxygen atom of Tyr42 and the carboxyl carbon atom of Glu46. Note that the constraints of fixing the O−O distance between chromophore and Glu46 and that between chromophore and Tyr42, which were used in the free energy calculation of Figure 7, are released in this calculation. The left (right) minimum corresponds to the pB_w state before (after) the relaxation of the hydrogen bond network. The stabilization by about 1.8 kcal/mol is obtained. By this stabilization, pR and pB_w have almost the same free energy. Other numerous relaxations, for example, the surface exposure of the chromophore, will stabilize the pB_w state further.

It is known that the N-terminal region of PYP collapses after the proton-transfer reaction.[12] Since the chromophore and its surrounding residues, which involve Glu46 and Arg52, should not interact with the residues of the N-terminal region directly, the key molecular movement triggering the collapse of the N-terminal region is not clear and needs to be investigated. In the structure of the pB_w state, we find that water molecules which approach Gly29 during the coordination and the migration of water molecules disturb the hydrogen bond between two residues in the β-scaffold, Gly29 and Val122. If these water molecules break the hydrogen bond between Gly29 and Val122, the binding pocket of Phe6 which consists of these five residues, Phe28, Lys106, His108, Phe121, and Lys123, could also be a site of migration by the water molecules. This may lead to the release of Phe6 from the pocket and the unfolding of the N-terminal region. Such a mechanism should be elucidated by further investigations.

4. Conclusion

The molecular structural and the hydrogen bond rearrangements associated with the proton transfer, after the photoisomerization of the chromophore, in the photocycle of PYP were investigated by evaluating the free energy and examining the role of water molecular coordination and migration. It was found that the photoisomerization of the chromophore largely affects the hydrogen bond network around Arg52 and Val66. The coordination of a water molecule to Glu46 reduces the energy barrier of the proton transfer and also induces further hydrogen bond network rearrangement around Arg52, resulting in the collapse of the surface-exposed region involving Ile49 and thus inducing the migration of water molecules to the interior of the protein. These migrated water molecules play essential roles in stabilizing the proton-transfer process, causing the structural change after this transfer reaction, and assisting the large displacement of the chromophore to the protein surface by the hydrogen bond exchange and also provide enough space for the phenolic ring of the chromophore to flip. The steric repulsion between Ile49 and a coordinated water molecule, which weakens the hydrogen bond network around Arg52, causes the facile displacement of Arg52 and then causes the large structural rearrangement. The well-packed native structure of the protein moiety cannot release the steric frustration brought by the water molecular coordination.

The crucial role of the water molecular migration and coordination in PYP photocycle is suggested in this work. Future intensive investigations will be required to find how the free energy of the pB_w state is stabilized with further structural rearrangements and thus to elucidate the key trigger motion to cause the unfolding of the N-terminal region. Other possible roles of the water molecular migration and the hydration in the PYP photocycle might be found by exploring the larger configurational space.

Acknowledgment. The calculations have been carried out by using the supercomputers at the Information Technology Center in Nagoya University and at the Research Center for Computational Science in Okazaki, Japan. This work was supported by the Grant-in Aid for Specially Promoted Research (No. 14100100), Grants-in Aid for Scientific Research (Nos. 16350008 and 18066018), Molecular-Based New Computational Science Program, NINS, and the Next Generation Super Computing Project, Nanoscience Program, MEXT, of Japan. M.K. thanks the financial support from the 21C-COE program of the Nagoya University Chemistry Department.

References and Notes

(1) Meyer, T. E. *Biochim. Biophys. Acta* **1985**, *806*, 175.
(2) Imhoff, J. F.; Süling, J. *Arch. Microbiol.* **1996**, *165*, 106.
(3) Sprenger, W. W.; Hoff, W. D.; Armitage, J. P.; Hellingwerf, K. J. *J. Bacteriol.* **1993**, *175*, 3096.
(4) Pellequer, J. L.; Wager-Smith, K. A., Getzoff, E. D. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 5884.
(5) Ponting, C. P.; Aravind, L. *Curr. Biol.* **1997**, *7*, 674.
(6) Hoff, W. D.; Düx, P.; Hård, P.; Devreese, B.; Mugteren-Roodzant, I. M.; Crielaard, W.; Boelens, R.; Kaptein, R.; van Beeumen, J.; Hellingwerf, K. J. *Biochemistry* **1994**, *33*, 13959.
(7) Brudler, R.; Rammelsberg, R.; Woo, T. T.; Getzoff, E. D.; Gerwert, K. *Nat. Struct. Biol.* **2001**, *8*, 265.
(8) Unno, M.; Kumauchi, M.; Sasaki, J.; Tokunaga, F.; Yamauchi, S. *Biochemistry* **2002**, *41*, 5668.
(9) Hendriks, J.; Hoff, W. D.; Crielaard, W.; Hellingwerf, K. J. *J. Biol. Chem.* **1999**, *274*, 17655.
(10) Xie, A.; Kelemen, L.; Hendriks, J.; White, B. J.; Hellingwerf, K. J.; Hoff, W. D. *Biochemistry* **2001**, *40*, 1510.
(11) Pan, D.; Philip, A.; Hoff, W. D.; Mathies, R. A. *Biophys. J.* **2004**, *86*, 2374.
(12) van der Horst, M. A.; van Stokkum, I. H.; Crielaard, W.; Hellingwerf, K. J. *FEBS Lett.* **2001**, *497*, 26.
(13) Imamoto, Y.; Kamikubo, H.; Harigai, M.; Shimizu, N.; Kataoka, M. *Biochemistry* **2002**, *41*, 13595.
(14) Takeshita, K.; Imamoto, Y.; Kataoka, M.; Mihara, K.; Tokunaga, F.; Terazima, M. *Biophys. J.* **2002**, *83*, 1567.
(15) Chen, E.; Gensch, T.; Gross, A. B.; Hendriks, J.; Hellingwerf, K. J.; Kilger, D. S. *Biochemistry* **2003**, *42*, 2062.
(16) Unno, M.; Kumauchi, M.; Hamada, N.; Tokunaga, F.; Yamauchi, S. *J. Biol. Chem.* **2004**, *279*, 23855.
(17) Warshel, A.; Levitt, M. *J. Mol. Biol.* **1976**, *103*, 227.
(18) Field, M. J.; Bash, P. A.; Karplus, M. *J. Comput. Chem.* **1990**, *11*, 700.
(19) Gao, J. *Acc. Chem. Res.* **1996**, *29*, 298.
(20) Hayashi, S.; Ohmine, I. *J. Phys. Chem. B* **2000**, *104*, 10678.
(21) Yagasaki, T.; Saito, S.; Ohmine, I. *J. Chem. Phys.* **2005**, *122*, 144504.
(22) van der Horst, M. A.; van Stokkum, I. H. M.; Dencher, N. A.; Hellingwerf, K. J. *Biochemistry* **2005**, *44*, 9160.
(23) Genick, U. K.; Borgstahl, G. E. O.; Ng, K.; Ren, Z.; Pradervand, C.; Burke, P. M.; Šrajer, V.; Teng, T.-Y.; Schildkamp, W.; McRee, D. E.; Moffat, K.; Getzoff, E. D. *Science* **1997**, *275*, 1471.
(24) Baca, M.; Borgstahl, G. E. O.; Boissinot, M.; Burke, P. M.; Williams, D. R.; Slater, K. A.; Getzoff, E. D. *Biochemistry* **1994**, *33*, 14369.
(25) Genick, U. K.; Soltis, S. M.; Kuhn, P.; Canestrelli, I. L.; Getzoff, E. D. *Nature* **1998**, *392*, 206.
(26) Imamoto, Y.; Kataoka, M.; Tokunaga, F. *Biochemistry* **1996**, *35*, 14047.
(27) Thompson, M. J.; Bashford, D.; Noodleman, L.; Getzoff, E. D. *J. Am. Chem. Soc.* **2003**, *125*, 8186.
(28) Kort, R.; Hellingwerf, K. J.; Ravelli, B. G. *J. Biol. Chem.* **2004**, *279*, 26417.
(29) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926.
(30) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179.
(31) Wang, J.; Cieplak, P.; Kollman, P. A. *J. Comput. Chem.* **2000**, *21*, 1049.
(32) Bayly, C. I.; Cieplak, P.; Cornell, W.; Kollman, P. A. *J. Phys. Chem.* **1993**, *97*, 10269.
(33) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Kollman, P. A. *J. Am. Chem. Soc.* **1993**, *115*, 9620.
(34) Cieplak, P.; Cornell, W. D.; Bayly, C.; Kollman, P. A. *J. Comput. Chem.* **1993**, *16*, 1357.

(35) Frisch, M. J.; Trucks, G.W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, V. G.; Montgomery, J. A., Jr.; Stratmann, R. E.; Burant, J. C.; Dapprich, S.; Millam, J. M.; Daniels, A. D.; Kudin, K. N.; Strain, M. C.; Farkas, O.; Tomasi, J.; Barone, V.; Cossi, M.; Cammi, R.; Mennucci, B.; Pomelli, C.; Adamo, C.; Clifford, S.; Ochterski, J.; Petersson, G. A.; Ayala, P. Y.; Cui, Q.; Morokuma, K.; Rega, N.; Salvador, P.; Dannenberg, J. J.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Cioslowski, J.; Ortiz, J. V.; Baboul, A. G.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Gomperts, R.; 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.; Andres, J. L.; Gonzalez, C.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A. *Gaussian98*, revision A.11.4; Gaussian, Inc.: Pittsburgh, PA, 2002.

(36) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648.

(37) Martyna, G. J.; Tuckerman, M. E.; Tobias, D. J.; Klein, M. L. *Mol. Phys.* **1996**, *87*, 1117.

(38) Andersen, H. C. *J. Comput. Phys.* **1983**, *52*, 24.

(39) Ding, H.-Q.; Karasawa, N.; Goddard, W. A., III. *J. Chem. Phys.* **1992**, *97*, 4309.

(40) Bishop, T. C.; Skeel, R. D.; Schulten, K. *J. Comput. Chem.* **1997**, *18*, 1785.

(41) Sham, Y. Y.; Warshel, A. *J. Chem. Phys.* **1998**, *109*, 7940.

(42) Frenkel, D.; Smit, B. *Understanding Molecular Simulation*, 2nd ed.; Academic Press: New York, 2002.

(43) Kumar, S.; Rosenberg, J. M.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A. *J. Comput. Chem.* **1992**, *13*, 1011.

(44) Kumar, S.; Rosenberg, J. M.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A. *J. Comput. Chem.* **1995**, *16*, 1339.

(45) Coutsias, E. A.; Seok, C.; Dill, K. A. *J. Comput. Chem.* **2004**, *25*, 1849.

(46) Kraulis, P. J. *J. Appl. Crystallogr.* **1991**, *24*, 946.

(47) Merrit, E. A.; Bacon, D. J. *Methods Enzymol.* **1997**, *277*, 505.

(48) Groenhof, G.; Lensink, M. F.; Berendsen, H. J. C.; Mark, A. E. *Proteins: Struct., Funct., Bioinf.* **2002**, *48*, 212.

(49) Groenhof, G.; Bouxin-Cademartory, M.; Hess, B.; de Visser, S. P.; Berendsen, H. J. C.; Olivucci, M.; Mark, A. E.; Robb, M. A. *J. Am. Chem. Soc.* **2004**, *126*, 4228.

(50) Tu, Y.; Laaksonen, A. *J. Chem. Phys.* **2000**, *113*, 11264.

(51) Warshel, A.; Weiss, R. M. *J. Am. Chem. Soc.* **1980**, *102*, 6218.

(52) Kato, M.; Pisliakov, A. V.; Warshel, A. *Proteins: Struct., Funct., Bioinf.* **2006**, *64*, 829.

(53) Schmitt, U. W.; Voth, G. A. *J. Chem. Phys.* **1999**, *111*, 9361.

(54) Braun-Sand, S.; Strajbl, M.; Warshel, A. *Biophys. J.* **2004**, *87*, 2221.

(55) Olsson, M. H. M.; Sharma, P. K.; Warshel, A. *FEBS Lett.* **2005**, *579*, 2026.

(56) Olsson, M. H. M.; Warshel, A. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 6500.

(57) Braun-Sand, S.; Burykin, A.; Chu, Z. T.; Warshel, A. *J. Phys. Chem. B* **2005**, *109*, 583.

(58) Štrajbl, M.; Hong, G.; Warshel, A. *J. Phys. Chem. B* **2002**, *106*, 13333.

(59) Ren, Z.; Perman, B.; Šrajer, V.; Teng, T.-Y.; Pradervand, C.; Bourgeois, D.; Schotte, F.; Ursby, T.; Kort, R.; Wulff, M.; Moffat, K. *Biochemistry* **2001**, *40*, 13788.

(60) Shiozawa, M.; Yoda, M.; Kamiya, N.; Asakawa, N.; Higo, J.; Inoue, Y.; Sakurai, M. *J. Am. Chem. Soc.* **2001**, *123*, 7445.

(61) Fichthorn, K. A.; Weinberg, W. H. *J. Chem. Phys.* **1991**, *95*, 1090.

(62) Voter, A. F. *J. Chem. Phys.* **1997**, *106*, 4665.

(63) Voter, A. F.; Montalenti, F.; Germann, T. C. *Annu. Rev. Mater. Res.* **2002**, *32*, 321.