A Molecular Dynamics Study for the Structure Determination of the Signaling State in the Photocycle of Photoactive Yellow Protein

Motoshi Kamiya[†] **and Iwao Ohmine***

Chemistry Department, Nagoya University, 1, Furocho, Chikusa-ku, Nagoya, Japan 464-8602, and Fukui Institute for Fundamental Chemistry, Kyoto University, 34-4, Takano Nishihiraku-machi, Sakyo-ku, Kyoto, Japan 606-8103

Received: October 19, 2009; Revised Manuscript Received: March 31, 2010

The structure of the putative signaling state in the photocycle of photoactive yellow protein, pB, is investigated by molecular dynamics (MD), steered MD, and free energy calculations. The present work suggests that the structural change toward pB involves a translational displacement of the N-terminal region. The structural rearrangement around α3, which is induced by the water molecular migration to the interior part of the protein during the formation of pB′, makes the joint region between β1 and the N-terminal domain flexible and induces the structural displacement of β1 and the N-terminal domain.

1. Introduction

Photoactive yellow protein (PYP)[1] from purple bacterium *Halorhodospira halophila*[2] is considered to be the photosensor for the negative phototaxis.[3] PYP absorbs the blue light and undergoes a photocycle, involving four steps (Figure 1). Starting from the initial ground state pG, PYP undergoes the trans−cis photoisomerization of the *p*-coumaric acid[4,5] to the state called pR (also known as PYP_L or I_1). In the next step, the protonation of the isomerized chromophore and the deprotonation of Glu46, which is directly hydrogen-bonded to the chromophore, take place.[6,7] This protonated chromophore state is called pB′ (also known as $PYP_{M'}$, PYP_M^{acid}, or I_2). Then the unfolding of the N-terminal domain, which is about 10 Å apart from the chromophore, is induced.[8,9] The structure of the chromophore binding pocket is expected to be almost unchanged during this unfolding process, according to the time-resolved X-ray experiments in the crystalline phase.[10] This partially unfolded state is called pB (also known as PYP_M, PYP_M^{alkali}, or I_2') and is considered to be the active state for the signal transduction. The signal transduction is induced by the binding of PYP in the pB state to other molecules such as DNAs or proteins. It is thus important to understand the structure and properties of pB state to understand the function of PYP.

Although many investigations have been performed,[11,12] the structure and properties of the pB state are not well-disclosed yet. For example, it is difficult to examine the partially unfolded protein structure by NMR, in which the large fluctuation of the N-terminal domain and the short lifetime of wild-type pB make the structural determination difficult,[11] and by the X-ray crystallography, in which the large structural change on the N-terminal domain does not fully take place in the crystalline phase.[13] The unfolding mechanism of the N-terminal region is well-understood neither. Only models such as the protein quake,[7] the hydrophobic collapse,[14] and the helix capping[10] have been proposed for its mechanism. Furthermore, computationally, simulations of millisecond-order and protein-wide structural changes are still beyond the scope of the current computational capability even though various extended ensemble molecular dynamics (MD) methods have been employed.[15]

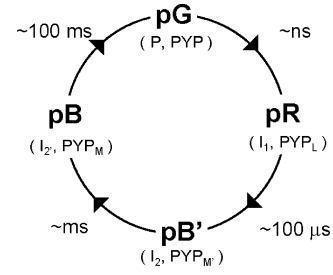

Figure 1. Photocycle of PYP. pG is the initial ground state. Upon the photon excitation of the chromophore, the isomerization of the chromophore is induced (the pR state). To relieve the frustration brought by the isomerization, the protonation of the chromophore and the structural change of the surroundings take place (the pB′ state). After pB′, the system undergoes the structural change on the N-terminal region, yielding the putative signaling state pB. To explore this structural change during the formation of the pB state is the main task of the present work. Finally, the reisomerization of the chromophore and the refolding of the protein take place to recover the pG state.

In the present work, we investigate the structure and physical properties of the pB state and the mechanism of how large structural changes around N-terminal are induced, by using the all-atom classical MD, steered MD, and free energy calculations. The mechanism of these sequential reaction processes from pG to pB is an example of the typical biological functions, amplifying small changes, here the isomerization of the chromophore, to large structural reorganizations, the structural rearrangement on the N-terminal domain.

2. Method

MD, steered MD, and free energy calculations are performed to search and validate the structure of the putative signaling pB. The steered MD technique enforcing the structural rearrangement on the N-terminal domain is used to obtain the candidate structures of the pB state. The free energy profile along the structural rearrangement leading to the pB state is then determined. As the initial state searching for pB, the structure obtained from our previous study for pB′ is used. It was found

* Corresponding author. E-mail: ohmine@fukui.kyoto-u.ac.jp.
† E-mail: kamo@chem.nagoya-u.ac.jp.

10.1021/jp910018d
Published on Web 04/22/2010

in the pB′ state that the chromophore is protonated, Glu46 is deprotonated, and several water molecules are migrated into the interior part of the protein.[16]

The distance from a specific configuration is used as the reaction coordinate for MD and free energy calculations, where the reaction coordinate λ is defined as

$$\lambda = \frac{1}{N}\sum_{W} m_i|\mathbf{r}_i - \mathbf{r}_{i,0}|^2$$

Here, m_i is the mass of atom i, $\mathbf{r}_i$ is the position of atom i, $\mathbf{r}_{i,0}$ is the position of atom i in the reference coordinate, and the translation and the rotation of the protein are removed.[17] W consists of residues in the N-terminal region and β1, 6−33 in the residue index. The residues 1−5 exhibit large fluctuations, but their change does not, causing the substantial inherent structural rearrangement; these residues are excluded from W. A configuration with a large λ value has a very different coordinate from the reference coordinate, and the configuration with $\lambda = 0$ is identical to the reference coordinate. Since the structural change in the formation of pB is not known, we choose this general reaction coordinate, λ.

To obtain the putative structure of the pB state, the steered MD technique is employed. The steered MD calculation is performed by adding the term expressed by $H_{add}(\lambda) = k(\lambda - \lambda_0)^2$ to the Hamiltonian of the system, where 0.1 kcal·mol^{-1}·Da^{2-}·Å^{-4} is used for the value of k in all of the calculations with this additional Hamiltonian. In the present work, a configuration is selected from those in the equilibrium trajectory without this additional term of the Hamiltonian as the reference coordinate $\mathbf{r}_0$. The system is equilibrated at $\lambda_0 = 30$, and the initial coordinates for the steered MD runs are generated at first. We then perform a few hundred steered MD runs by increasing λ_0 as $\lambda_0(t) = 30 + t\Delta\lambda$, where the increase rate of λ_0, $\Delta\lambda$, during the steered MD calculation was set to 10^{-4} Da·Å^2·fs^{-1}. We chose the configuration with $\lambda_0 = 30$ as the initial configuration of the steered MD calculations instead the one with $\lambda_0 = 0$, since the configuration with $\lambda = 0$ is at the unique position of $\mathbf{r}_0$, and the system configuration yielding the fluctuation hardly coincides with it. At $\lambda = 30$, atoms in the region W are displaced by 2.1 Å on average, which is in the range of the fluctuation, from the reference coordinate. Three distinct molecular rearrangements, called "Rearrangements 1−3", are found along λ in the steered MD calculations (see Results Section).

A new crude method is used for the free energy calculation. The free energy profile is expected to be a smooth function of λ and thus the free energy difference be written as $\Delta F(\lambda) = b(\lambda - \lambda_0)$ around λ_0. To evaluate the free energy gradient, b, we performed MD calculations with Hamiltonian $H + H_{add}$ at the selected values of λ_0, that is, at $\lambda_0 = 30$ and every hundred, 100, 200, 300, and so on, and then obtained the distribution of states (configurations) as a function of λ at each λ_0. Each distribution has a maximum at λ_{eq} near λ_0; λ_{eq} should be at the minimum of the function $H_{add}(\lambda) + b(\lambda - \lambda_0)$. Using this λ_{eq}, the gradient b is determined by $-2k(\lambda_{eq} - \lambda_0)$. The free energy at an arbitrary value of λ is interpolated by a quadratic function from these gradients. To obtain λ_{eq} at a given λ_0 value, a MD calculation with Hamiltonian $H + H_{add}$ is performed by starting from the structure which is obtained from a steered MD trajectory representing each "Rearrangement". To calculate the gradient, we performed an at least 5 ns length equilibration first and then performed a 10−20 ns MD calculation to obtain the probability distribution along the reaction coordinate around λ_0. The scheme of this free energy estimation is summarized in Figure 2.

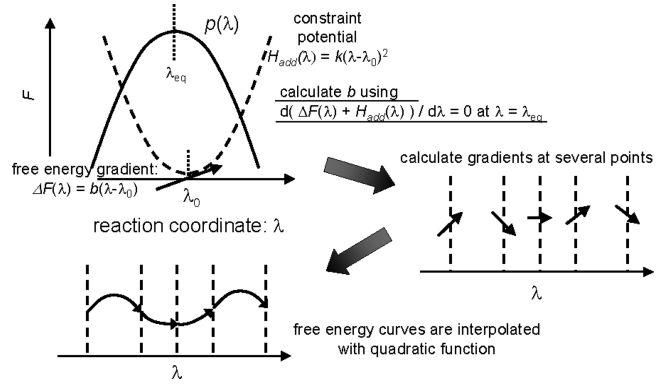

Figure 2. Free energy calculation scheme is summarized. (top) The gradient of the free energy at λ_0, b, is calculated using the position of the average (or the most probable) λ value (λ_{eq}) obtained from the MD calculations. (right) The gradients are calculated at several λ_0 values. (bottom) The free energy curve between two gradients are interpolated using a quadratic function $F(\lambda) = a\lambda^2 + b\lambda + c$, where the coefficients are uniquely determined by the two gradients.

Although this method does not provide very accurate free energies, it can provide the rough estimation of the free energy profiles over an extended region at least qualitatively by MD trajectory calculations within available computer power. There exist many free energy evaluation methods, such as the umbrella integration[18] and the methods based on Jarzynski's equality, which are more accurate than the present one.[19−21] The structural changes focused in this study are, however, much extended, and the free energy curves of these changes are hardly evaluated by using these accurate methods.

The MD parameters are the same as those used in the previous study.[16] The AMBER99 force field[22,23] and the TIP3P water model[24] were used for the interaction parameters of amino acids and water molecules, respectively. The protein molecule is spherically surrounded by water molecules with the radius of 42 Å. The system consists of 30 885 atoms (protein: 1929, water: 28 956), and the density of the sphere is about 1.0 g/mL. Water molecules which go out the sphere are constrained by a harmonic potential with 1.5 kcal·mol^{-1}·Å^{-2} of the force constant. MD calculations were performed at the constant temperature with a Nose−Hoover's thermostat.[25] A velocity Verlet method with 2 fs of time for each step was employed for the time-integration. In all MD simulations, covalent bond lengths involving hydrogen atoms are constrained with RATTLE.[26] The cell multipole method[27] was used for the calculation of the long-range electrostatic interaction. His108 in the structures is assumed to be electrically neutral in these calculations, although the experiment by Fisher et al. reports that this histidine in pG is protonated.[28] Since the pB state has been observed in nonacidic conditions,[29,30] the histidine in pB is considered to be exposed to the solvent and thus expected to be neutral. The system employed in the steered MD and the free energy calculations is spherical and consists of 30 885 atoms in MD and free energy calculations. All of the MD, steered MD, and free energy calculations are performed at 300 K.

CRYSOL[31] and STRIDE[32] are used for the calculation of the radius of gyration, R_g, and the secondary structure assignment, respectively.

3. Results and Discussions

Structure of pB′. pB′ is the state before the pB state in the photocycle of PYP (see Figure 1). In the previous study,[16] it

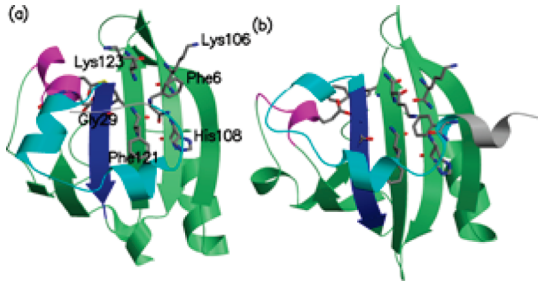

Figure 3. Instantaneous structures of pB′ are shown in ribbons and sticks, where the colors are cyan for N-terminal domain except for residues 1−5, blue for β1, magenta for α3, and green for the others. Heavy atoms of Phe6, Gly29, Glu46, Lys106, His108, Phe121, Lys123, and the chromophore are shown in sticks: (a) a dominant structure of pB′, where the chromophore is exposed to the solvent, α3 is stable and N-terminal domain is intact. (b) A minor structure of pB′ having disturbed α3 and slightly elongated β1. The pictures of the structures are prepared with MOLSCRIPT.[43]

was found that the water molecular migration to the chromophore binding pocket is essential for the formation of pB′ which has the protonated chromophore and the deprotonated Glu46.[16,33] In this pB′ structure, the chromophore and Glu46 are thus hydrated but still buried in the protein.

In the present work, additional extended MD calculations are performed on this pB′ state obtained in the previous study to investigate its further structural relaxation, and two main structures shown in Figure 3 are obtained as the relaxed pB′ state (just called pB′ state, hereafter). The first one, Figure 3a, is similar to the structure found in the previous study,[16] where Glu46 is hydrated and the β-sheet is intact. The second one, shown in Figure 3b, is similar to the first one but has the slightly displaced α3 (helix involving Glu46) and extended β1, associated with the orientational change of Gly29. The experimental observation of the enhancement on the H/D exchange rate of the residues in α3 and β1 by the continuous light illumination might be attributed to the fluctuation between these two structures.[34] The second structure, however, appears less frequently in the MD calculation and thus must be slightly more unstable in comparison with the first one. In these two structures (Figures 3a,b), the chromophore is slowly fluctuating between the buried and the solvent exposed form, although the chromophore favors to be exposed to the solvent. In addition to this fluctuation of the chromophore, the β4−β5 loop involving Met100 also largely fluctuates, since the hydrogen bond between Arg52 and Tyr98, which is the anchor of the loop, is already broken. The structure and its fluctuation of pB′ shown here are in good agreement with experimental results. For example, the high-angle X-ray scattering study suggests that the PYP_M intermediate of the N-terminally truncated mutant, having the similar structure to wild-type pB′, exhibits the fluctuation of the β4−β5 loop.[35] The time-resolved X-ray crystallography experiment has observed the displacement of Arg52 and the structural shift of α3 during the formation of pB′.[36]

It was also reported with the transient-grating experiments that the diffusion coefficient of the pB′ is about 15% smaller than that of pG.[37] We have performed the preliminary estimation of the diffusion coefficients of pG and pB′ by using MD calculations and found about a 5% decrease of the diffusion coefficient in pB′.[38] The decrease of the diffusion coefficient with the formation of pB′ can be attributed to the increase of the solvent accessible surface area, which is caused by the water migration associated with the proton transfer reaction.

Structure of pB and Free Energy Calculations. Employing these two pB′ structures, one with stable α3 (Figure 3a) and the other with unstable α3 (Figure 3b), as initial structures, we search for the structural change of the N-terminal region in the pB state by using the methods described in Section 2.

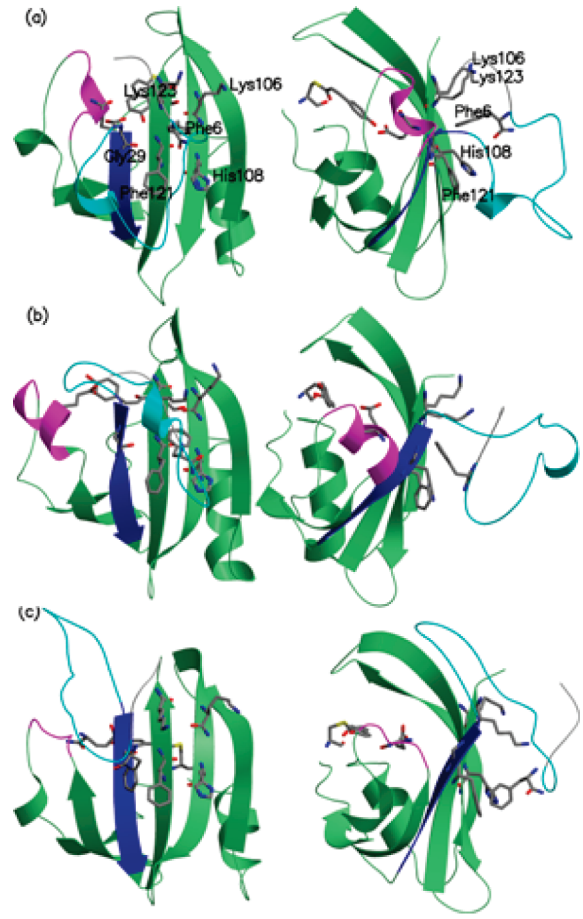

Figure 4. Structures with the disturbed N-terminal domain for the states at λ = 600 obtained by starting from (a) pB′ with stable α3 ("Rearrangement 1"; black dashed curve in Figure 6), (b) pB′ with disturbed α3 ("Rearrangement 2"; black solid curve), and (c) another structure for pB′ with disturbed α3 ("Rearrangement 3"; blue curve), respectively, are shown in ribbons and sticks. The structure (c) is a candidate for the equilibrium state of pB, denoted as "pB structure" (see text).

Steered MD calculations starting from pB′ structures having the stable α3 (Figure 3a) results in the unfolding and extension of the N-terminal domain (this path is denoted as "Rearrangement 1", hereafter). The structure is shown in Figure 4a, where the salt bridge between Lys110 in the β-sheet and Glu9/Glu12 in the N-terminal domain is completely broken and a part of the β-sheet close to the β5−β6 loop involving Lys110 is thus exposed to the solvent by this rearrangement. On the other hand, the steered MD calculations starting from pB′ structures with unstable α3 (Figure 3b) yield two distinct structural change pathways; the first one (denoted as "Rearrangement 2"; Figure 4b) is the dominant pathway and is similar to the "Rearrangement 1" except that several residues in β6 such as Phe121 are exposed to the solvent. The second one ("Rearrangement 3"; Figure 4c) involves not just an unfolding but also a significant translation, sliding up the β-sheet, of the N-terminal domain. Most of the residues in β5 and β6 are exposed to solvent in this structure, while the N-terminal domain excluding residues 1−5 has a slightly more compact structure than that in "Rearrangements 1 and 2". "Rearrangement 3" pathway bifurcates from the "Rearrangement 2" pathway around λ =

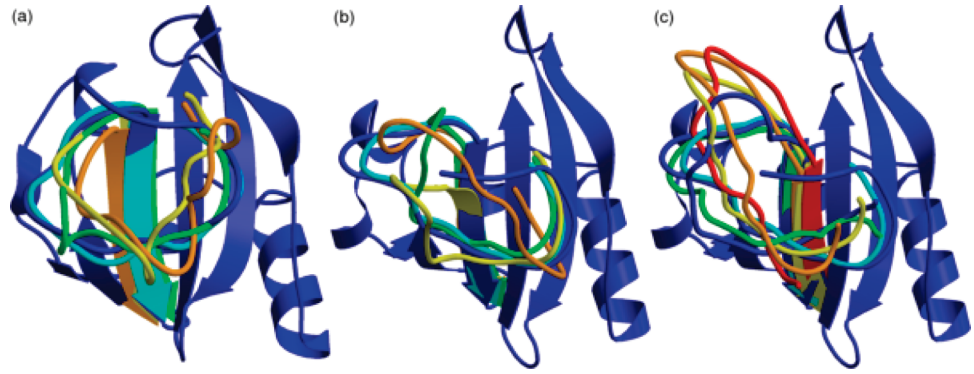

Figure 5. Intermediate structures of the N-terminal domain in the steered MD calculations for (a) "Rearrangement 1", (b) "Rearrangement 2", and (c) "Rearrangement 3" are shown. The structures colored with cyan, green, yellow, and orange correspond to the structure at $\lambda = 30$, 200, 400, and 600, respectively. For "Rearrangement 3", the structure at $\lambda = 800$ is also plotted with the red color. The whole structure of the reference coordinate $\mathbf{r}_0$ is plotted with the dark blue color.

100−200 in the steered MD trajectories, but the yielding of such bifurcation is small.

Instantaneous structures selected from those in "Rearrangements 1, 2, and 3" are shown in Figure 5, parts a, b, and c, respectively. One can see that β1 is kept intact in all of the intermediate structures. The structures of "Rearrangements 2 and 3", being almost identical at $\lambda = 30$, become distinct for $\lambda \geq 200$. The bifurcation of the pathways arises from the directionality of the structural displacements in "Rearrangements". The N-terminal domain moves to the direction leaving the β-sheet in "Rearrangement 2" (Figures 5b) while approaching the β-sheet in "Rearrangements 3" (Figure 5c). Because of this directionality difference of the displacements, the structures of these "Rearrangements" are mutually largely separated when λ is increased, and thus the branching from "Rearrangement 2" to "Rearrangement 3" becomes kinetically difficult. It should be noted that the branching rate must significantly increase if a better order parameter (reaction coordinate) describing the more precise global structural change, rather than λ employed in the present work, could be defined. Around $\lambda = 400$ in "Rearrangement 3" (see Figure 5c), the N-terminal domain is sliding over the β-sheet to the up direction of the figure. Associated with this sliding motion, the joint region of the N-terminal domain and β1 is twisted toward the same direction, to the up direction in the figure. The structures of "Rearrangement 3", used as the initial structures to calculate the free energy calculation, for the smaller λ values (< 200) are obtained by performing the steered MD by decreasing λ with starting from the configuration at $\lambda = 200$ on "Rearrangement 3".

Using the intermediate structures from the steered MD trajectories and the subsequent MD calculations, the free energy gradients are determined for various λ values. Figure 6 shows the obtained free energy curves. The free energy curve of "Rearrangement 1" (indicated with black dashed line in the figure) yields a local minimum around $\lambda = 250$ and then increases rapidly with λ. The free energy profile of "Rearrangement 2" (black solid line) is also monotonically increasing with the increase of λ, while that of "Rearrangement 3" (blue color line) first decreases up to $\lambda = 220$ and then yields a small barrier before reaching another minimum around $\lambda = 600$. To evaluate these free energy curves, one MD trajectory is used, starting with the structure from a randomly selected steered MD calculation for each λ_0 value, except for $\lambda_0 = 400$, 500, and 600 in "Rearrangement 3", where the structural variation is very large. In the latter case, multiple MD calculations are performed starting from the structures of different steered MD trajectories at each of these λ_0 values. A set of MD trajectories is then

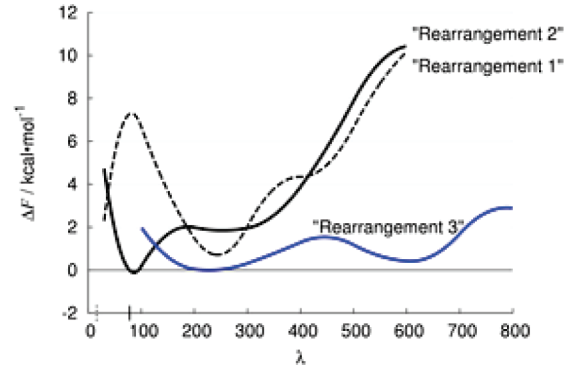

Figure 6. Free energy curves for the structural change of the N-terminal domain are shown. The black dashed curve shows the free energy change, corresponding to "Rearrangement 1", obtained from the steered MD trajectories staring from the pB′ structures of Figure 3a type (see text). The solid black and blue free energy curves corresponding to "Rearrangements 2 and 3", respectively, that is, those obtained by starting from the pB′ structures of Figure 3b type (see text). Small black ticks on the abscissa axis represent the equilibrium position without constraint of $H_{add}(\lambda)$. The line styles of ticks correspond to those of the free energy curves, respectively. MD calculations without constraint of λ starting from any configurations at small λ values on the blue curve (with "Rearrangement 3") yield almost identical equilibrium position with the solid tick, and thus the blue tick is not shown. The solid black tick locates at $\lambda = 78$ and the dashed at $\lambda = 16$. The free energies at these equilibrium λ values ($\lambda = 16$ for "Rearrangement 1" and $\lambda = 78$ for "Rearrangement 2") obtained by the MD calculations without $H_{add}(\lambda)$ are taken to be zero. The free energy value at $\lambda = 16$ (i.e., at its minimum) for "Rearrangement 1" is extrapolated from the value at $\lambda = 30$ using its gradient. The free energy curve for "Rearrangement 3" is shifted relative to the free energy curve for "Rearrangement 2" in the value so that the blue curve crosses with the black solid one around $\lambda = 125$ (see text).

chosen, one from those multiple MD trajectories at each λ_0, in a way that the average structures of a chosen set are mutually most resembled; a set of trajectories which minimize the total sum of the root-square displacements of the average structure of the adjacent λ grids is selected. This is done to define one path representing a "reaction coordinate" of "Rearrangement 3" and to calculate its free energy curve but is still denoted by λ. Even if full multiple MD configurations starting from the structures in different steered MD trajectories are used to evaluate λ_{eq} value, the free energy obtained is very close (within 1−2 kcal·mol^{-1} at most for "Rearrangement 3" and almost identical for other "Rearrangements") to that obtained with a single trajectory.

In Figure 6 small colored ticks are marked on the abscissa axis to represent the λ values of the equilibrium positions

obtained from the MD trajectory calculations (MD without H_{add}) starting from the structures selected from MD trajectories with $H + H_{add}$ at $\lambda_0 = 30$ for each "Rearrangement". The equilibrium positions of the ticks for "Rearrangements 1 and 2" nearly coincide with the minima of their free energy curves. This clearly shows the validity of the present method for the free energy estimation. The λ value at the equilibrium staring from the "Rearrangement 3" configuration at $\lambda_0 = 30$ is found to be near equal to that of "Rearrangement 2" (the blue tick coincides with the black tick and thus is not shown in the figure). The equilibrium structure of "Rearrangement 3" is also found to be similar to the equilibrium structure in "Rearrangement 2". This means that the system cannot follow the pathway corresponding to "Rearrangement 3" but does trace the one corresponding to "Rearrangement 2" for small λ.

The exact energy difference between the free energy curves of "Rearrangement 2" and that of "Rearrangement 3" is extremely difficult to be determined by any kinds of methods such as umbrella samplings, since their structural difference and fluctuations are so large. We must obtain, for example, the free energy value, not a relative one, for the minimum for "Rearrangement 2" (the black solid curve in Figure 6), corresponding to the pB′ structure shown in Figure 3b, and that for the right side minimum for "Rearrangement 3" (blue curve in Figure 6), corresponding to the candidate of the pB structure. This estimation is, however, also computationally extremely difficult, since the degree of freedom of the system is so large. Instead, in this study, we simply shift the free energy curve for "Rearrangement 3" relative to that for "Rearrangement 2" in the value so that these two curves are crossed around $\lambda = 125$ (see Figure 6), where most of the branching from "Rearrangement 2" to "Rearrangement 3" takes place in steered MD calculations. The branching from "Rearrangement 2" to "Rearrangement 3" could happen thermodynamically at any larger λ value, but it must be very difficult kinetically, since the structural rearrangements associated with such branching are too large to occur in the MD calculations (see Figure 5b,c).

It has been shown by the early NMR study that the dynamical fluctuation of the N-terminal residues in the pB state is very large.[39] In Figure 6, only "Rearrangement 3" has a near flat profile and thus can yield the largest fluctuation, among the free energy curves. This curve has a shallow minimum around $\lambda = 600$, and this structure at $\lambda = 600$, shown in Figure 4c, thus can be the representative structure for the putative active state, pB; this structure is denoted as "pB-structure", hereafter. Notice here that the pB state is expected to exhibit a large structural fluctuation over the extended range of λ (e.g., $\lambda = 200-700$) on the flat free energy curve. One of the reasons why the free energy curve for "Rearrangement 3" has a near flat profile can be attributed to the fact that the structural rearrangement along this curve, although involves large displacements of the residues, does not increase the total area of the solute (protein)−solvent surface much, thus causing only the minimal change in the protein−solvent interaction. The formation of the stable hydrogen bonds by the further elongation of β1 would also contribute to the stability along the pathway for "Rearrangement 3". In addition to the free energy curves of "Rearrangements 1−3", we have estimated the free energy curve of the pG state along λ. This curve is found (data not shown in figures) to monotonically increase with λ.

The displacements of the α-carbons during the structural changes are shown in Figure 7. One can see that the structural change on β1 is very small, although β1 is included in the region W. There is the hydrophobic interaction between the N-terminal and the β-sheet. The side chains of Lys106, His108, Phe121, and Lys123 in the β-sheet form a "pin-holder" sticking into the N-terminal cap (see Figure 3a or 4a). The side chain of Phe6 in the N-terminal domain is tightly bounded with this "pin-holder". This small hydrophobic interaction region is kept intact in all structures except when λ becomes very large ($\lambda \geq 600$) in "Rearrangement 3". Around $\lambda = 600$ in "Rearrangement 3", this hydrophobic region is in either of the broken or the intact

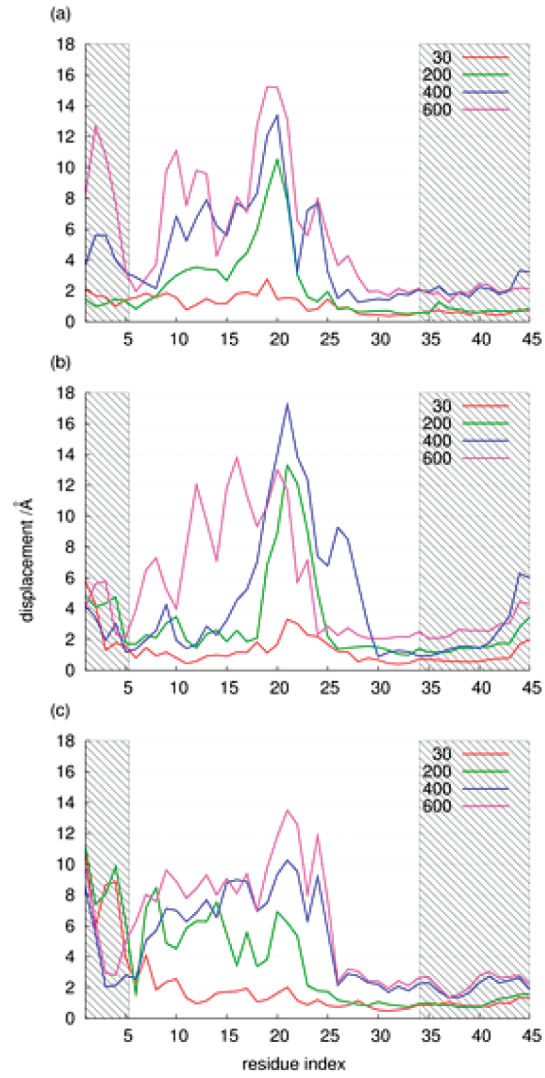

Figure 7. Average displacements of α-carbons in the N-terminal domain and β1 from their reference coordinate $\mathbf{r}_{i,0}$ at $\lambda = 30$, 200, 400, 600 during (a) "Rearrangement 1" (starting from pB′ with stable α3), (b) "Rearrangement 2" (starting pB′ with distorted α3), and (c) "Rearrangement 3" (branched pathway from b) are shown. Note that the residues in the shaded/unshown region (residues of 1−5 and 34−125) are not included in the region W of H_{add} and thus are free from the constraint by λ (see Section 2).

TABLE 1: Radius of Gyration R_g in Each State

state	calcd R_g[a] (expt. value[29])/Å
pG	15.3 (15.2)
pB′	15.4 (15.7[b])
	α3-disturbed: 15.5
pB	$\lambda = 600$ in "Rearrangement 1": 16.3
	"pB-structure": 16.1 (16.2[c])

[a] Average value; calculated using CRYSOL.[31] [b] At pH = 5.0. [c] At pH = 8.0.

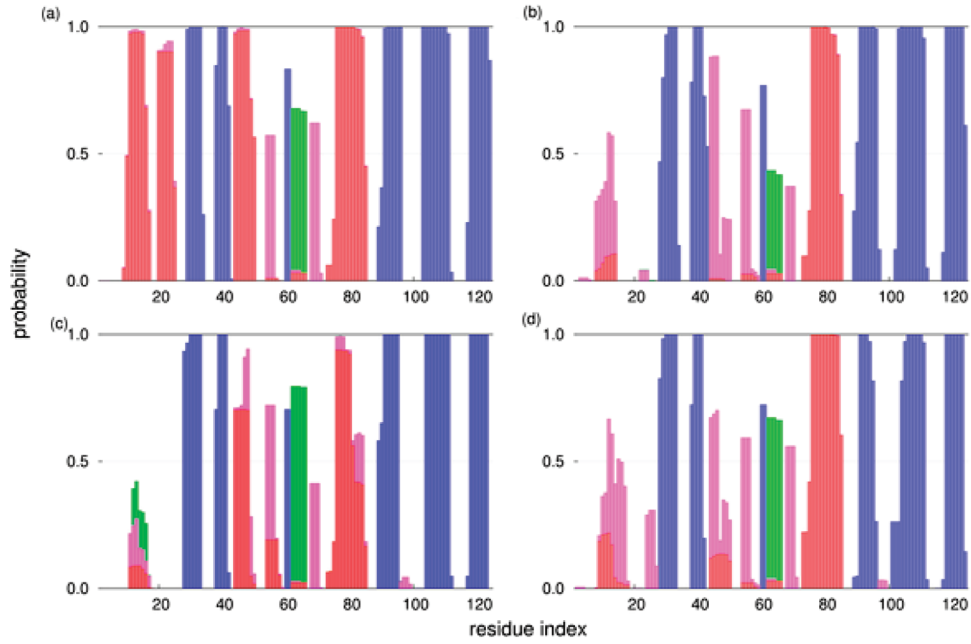

Figure 8. Average secondary structure abundances in (a) pG at (b) $\lambda = 600$ on the free energy curve for "Rearrangement 1" (see Figure 4a for the structure), (c) $\lambda = 600$ on the free energy curve for "Rearrangement 2" (see Figure 4b for the structure), and (d) $\lambda = 600$ on the free energy curve for "Rearrangement 3" ("pB-structure"; see Figure 4c for the structure), depicted with stacked bars. The abundances in pG are shown for the comparison. The red, magenta, green, and blue bars represent α-helices, 3_{10}-helices, π-helices, and β-strands, respectively. The empty bars represent other secondary structures, for example, turns and coils. STRIDE[32] is used for the secondary structure assignment.

state, and thus Phe6 yields a large displacement by fluctuating between these states, as seen in the Figure 7c.

This can explain the mutational experimental results. It was found that the mutation of Phe6 or Lys123 significantly slows down the kinetics of the photocycle.[40,41] The hydrophobic interaction just described is acting to restrict the translational displacement of the N-terminal region and thus make the formation of pB difficult. When this interaction is removed by the mutation or by the partial truncation of the N-terminal cap, the system can easily undergo the translation, sliding up the β-sheet, of the N-terminal domain, and its pB state becomes too stable, resulting in the reduction of photocycle kinetics.

The radii of gyration (R_g) of the obtained structures are calculated with CRYSOL,[31] and the results are listed in Table 1. The table shows the increase of R_g with the advance of PYP photocycle, from pG through pB′ to pB. This is in the good agreement with the small-angle X-ray experiment.[29] Since "pB-structure" (Figure 4c) has a slightly more compact structure in comparison with the other structures, the increased R_g of pB is not only attributed to the formation of the extended, partially unfolded structure of the N-terminal domain, but also the increase of the structural asphericity associated with the displacement of the N-terminal domain. We have also estimated the diffusion coefficient of the "pB-structure". In agreement with the experiment,[37] the diffusion coefficient of the "pB-structure" is found to be very close to that of pB′, suggesting that the diffusion coefficient of the protein molecules is not significantly influenced by the structural change on the N-terminal region.

It is indicated by the CD experiment that the helical segments of the N-terminal domain, namely, α1 and α2, diminish, and the other helices, α3 and/or α5, destabilize in the progress of the photocycle.[42] The Fourier transform infrared spectroscopy (FTIR) study[42] further suggests certain structural changes in the β-sheet region. The calculated abundances of the average secondary structure are shown for pG state and pB state in Figure 8, parts a and d, respectively. We can indeed see in Figure 8d that the "pB-structure" has the more destabilized α1−3 helices and the more elongated β1 strand in comparison with those in pG, in agreement with these experimental results.

4. Summary

The structural change on the N-terminal region of PYP was investigated by MD and the free energy calculations. It was found that the N-terminal region undergoes not just an unfolding but also a translational rearrangement toward the chromophore side, which is induced by the destabilization of α3 and the orientational change of Gly29. The reaction model from this and the previous studies can be summarized as follows. The isomerization of the chromophore brings a structural frustration on the protein and an increase of the chromophore proton affinity. The chromophore-protonated state pB′, which has deprotonated Glu46 and the water molecules inside the protein, becomes stable due to the frustration generated by the isomerization. The hydrogen bond bridge between Arg52 and Tyr98 and the small hydrophobic region consisting of Ile49, Met100, and Thr103 break during the formation of pB′ to incorporate water molecules. These structural changes make α3 unstable and flexible. The orientation of Gly29 becomes also flexible, since the steric repulsion with α3 is weakened. The orientational change of Gly29 induces the sliding up of the N-terminal domain over the β-sheet, finally resulting in the formation of the putative active state pB. It should be noted that the overall picture of the unfolding mechanism of the N-terminal presented here has a certain resemblance to the helix-capping model.[10]

There are several limitations in this study. The most important one is concerned with the method. In our method, we assume that the structural change takes place mainly on the N-terminal domain, and we perform only a rough estimation of the free energy change. To obtain a more precise description, further investigations are still necessary to understand this slow, large, and complicated structural change.

To explore the whole photocycle of PYP, we need to know its last step, that is, one from pB to pG, where the counter to the photoisomerization, namely, the nonphotoactivated cis−trans isomerization bringing back the chromophore to the trans-form, and the concomitant slow refolding process of the protein take place. The quantum chemical calculation is being performed to elucidate the mechanism of this isomerization process by the present authors, and the result will be published elsewhere.

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. The present work was supported by the Grant-in Aid for Specially Promoted Research (No. 14100100) and the Grant-in Aid for Scientific Research (No. 19350009). M.K. thanks the financial support from 21C-COE program of 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) Hoff, W. D.; Dux, P.; Hard, P.; Devreese, B.; Nugteren-Roodzant, I. M.; Crielaard, W.; Boelens, R.; Kaptein, R.; van Beeumen, J.; Hellingwerf, K. J. *Biochemistry* **1994**, *33*, 13959.
(5) Brudler, R.; Rammelsberg, R.; Woo, T. T.; Getzoff, E. D.; Gerwert, K. *Nat. Struct. Mol. Biol.* **2001**, *8*, 265.
(6) Hendriks, J.; Hoff, W. D.; Crielaard, W.; Hellingwerf, K. J. *J. Biol. Chem.* **1999**, *274*, 17655.
(7) Xie, A.; Kelemen, L.; Hendriks, J.; White, B. J.; Heliingwerf, K. J.; Hoff, W. D. *Biochemistry* **2001**, *40*, 1510.
(8) Pan, D.; Philip, A.; Hoff, W. D.; Mathies, R. A. *Biophys. J.* **2004**, *86*, 2374.
(9) van der Horst, M. A.; van Stokkum, I. H.; Crielaard, W.; Hellingwerf, K. J. *FEBS Lett.* **2001**, *497*, 26.
(10) Rajagopal, S.; Anderson, S.; Šrajer, V.; Schmidt, M.; Pahl, R.; Moffat, K. *Structure* **2005**, *13*, 55.
(11) Fuentes, G.; Nederveen, A. J.; Kaptein, R.; Boelens, R.; Bonvin, A. M. J. J. *J. Biomol. NMR* **2005**, *33*, 175.
(12) Bernard, C.; Houben, K.; Derix, N. M.; Marks, D.; van der Horst, M. A.; Hellingwerf, K. J.; Boelens, R.; Kaptein, R.; van Nuland, N. A. J. *Structure* **2005**, *13*, 953.
(13) Kort, R.; Ravelli, R. B.; Schotte, F.; Bourgeois, D.; Crielaard, W.; Hellingwerf, K. J.; Wulff, M. *Photochem. Photobiol.* **2003**, *78*, 131.
(14) Cusanovich, M. A.; Meyer, T. E. *Biochemistry* **2003**, *42*, 4759.
(15) Vreede, J.; Crielaard, W.; Hellingwerf, K. J.; Bolhuis, P. G. *Biophys. J.* **2005**, *88*, 3525.
(16) Kamiya, M.; Saito, S.; Ohmine, I. *J. Phys. Chem. B* **2007**, *111*, 2948.
(17) Coutsias, E. A.; Seok, C.; Dill, K. A. *J. Comput. Chem.* **2004**, *25*, 1849.
(18) Kästner, J.; Thiel, W. *J. Chem. Phys.* **2006**, *123*, 234106.
(19) Jarzynski, C. *Phys. Rev. Lett.* **1996**, *78*, 2690.
(20) Park, S.; Khalili-Araghi, F.; Tajkhorshid, E.; Schulten, K. *J. Chem. Phys.* **2003**, *119*, 3559.
(21) Maragakis, P.; Ritort, F.; Bustamante, C.; Karplus, M.; Crooks, G. E. *J. Chem. Phys.* **2008**, *129*, 024102.
(22) 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.
(23) Wang, J.; Cieplak, P.; Kollman, P. A. *J. Comput. Chem.* **2000**, *21*, 1049.
(24) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926.
(25) Martyna, G. J.; Tuckerman, M. E.; Tobias, D. J.; Klein, M. L. *Mol. Phys.* **1996**, *87*, 1117.
(26) Andersen, H. C. *J. Comput. Phys.* **1983**, *52*, 24.
(27) Ding, H.-Q.; Karasawa, N.; Goddard, W. A., III. *J. Chem. Phys.* **1992**, *97*, 4309.
(28) Fisher, S. Z.; Anderson, S.; Henning, R.; Moffat, K.; Langan, P.; Thiyagarajan, P.; Shultz, A. J. *Acta Crystallogr.* **2007**, *D63*, 1178.
(29) Shimizu, N.; Imamoto, Y.; Harigai, M.; Kamikubo, H.; Yamazaki, Y.; Kataoka, M. *J. Biol. Chem.* **2006**, *281*, 4318.
(30) Otto, H.; Hoersch, D.; Meyer, T. E.; Cusanovich, M. A.; Heyn, M. P. *Biochemistry* **2005**, *44*, 16804.
(31) Svergun, D.; Barberato, C.; Koch, M. H. J. *J. Appl. Crystallogr.* **1995**, *28*, 768.
(32) Frishman, D.; Argos, P. *Proteins: Struct., Funct., Bioinf.* **1995**, *23*, 566.
(33) Leenders, E. J. M.; Vande Vondele, J. I.; Bolhuis, P. G.; Meijer, E. J. *J. Phys. Chem. B* **2007**, *111*, 13591.
(34) Brudler, R.; Gessner, C. R.; Li, S.; Tyndall, S.; Getzoff, E. D.; Woods, V. L., Jr. *J. Mol. Biol.* **2006**, *363*, 148.
(35) Hironari, K.; Shimizu, N.; Harigai, M.; Yamazaki, Y.; Imamoto, Y.; Kataoka, M. *Biophys. J.* **2007**, *92*, 3633.
(36) Ihee, H.; Rajagopal, S.; Šrajer, V.; Pahl, R.; Anderson, S.; Schmidt, M.; Schotte, F.; Anfinrud, P. A.; Wulff, M.; Moffat, K. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 7145.
(37) Hoshihara, Y.; Imamoto, Y.; Kataoka, M.; Tokunaga, F.; Terazima, M. *Biophys. J.* **2008**, *94*, 2187.
(38) This calculation is performed in the cubic periodic system consisting of about 30 000 atoms. The diffusion coefficient is calculated from the mean-square displacement of the protein center of mass using the relation $6Dt = \langle|\mathbf{r}(t) - \mathbf{r}(0)|^2\rangle$, where D is the diffusion coefficient. The SANDER module of AMBER7 is used for the MD calculation. The particle-mesh Ewald method is used for the calculations of the electrostatic interaction. The force field parameters used are same as those in the calculations on the water sphere system.
(39) Rubinstenn, G.; Vuister, G. W.; Mulder, F. A. A.; Düx, P. E.; Boelens, R.; Hellingwerf, K. J.; Kaptein, R. *Nat. Struct. Mol. Biol.* **1998**, *5*, 568.
(40) Harigai, M.; Kataoka, M.; Imamoto, Y. *J. Am. Chem. Soc.* **2006**, *128*, 10646.
(41) Harigai, M.; Kataoka, M.; Imamoto, Y. *Photochem. Photobiol.* **2008**, *84*, 1031.
(42) Harigai, M.; Imamoto, Y.; Kamikubo, H.; Yamazaki, Y.; Kataoka, M. *Biochemistry* **2003**, *42*, 13893.
(43) Kraulis, P. J. *J. Appl. Crystallogr.* **1991**, *24*, 946.

JP910018D