Supplementary Information

Combined probes of X-ray scattering and optical spectroscopy Reveal How Global Conformational Change Is Temporally and Spatially Linked to Local Structural Perturbation in Photoactive Yellow Protein

Tae Wu Kim,[1,2,†] Cheolhee Yang,[1,2,†] Youngmin Kim,[1,2] Jong Goo Kim,[1,2] Jeongho Kim,[3] Yang Ouk Jung,[1,2] Sunhong Jun,[1,2] Sang Jin Lee,[1,2] Sungjun Park,[1,2] Irina Kosheleva,[4] Robert Henning,[4] Jasper J. van Thor,[5] & Hyotcherl Ihee[1,2,]*

[1] Department of Chemistry, Korea Advanced Institute of Science and Technology (KAIST), Daejeon 305-701, Korea

[2] Center for Nanomaterials and Chemical Reactions, Institute for Basic Science (IBS), Daejeon 305-701, Korea

[3] Department of Chemistry, Inha University, Incheon, 402-751, Korea

[4] Center for Advanced Radiation Sources, The University of Chicago, Chicago, IL60637, USA

[5] Division of Molecular Biosciences, Imperial College London, South Kensington Campus, SW7 2AZ London, United Kingdom

[*]Corresponding author: hyotcherl.ihee@kaist.ac.kr

[†]These authors contributed equally to this work.

(1) Elimination of solvent heating contribution from time-resolved difference scattering curves

Time-resolved difference scattering curves contain the contribution of solvent heating induced by the energy dissipated from the excited proteins into the solvent (see Figure S1). To extract only the contributions from the structural change of solute (protein) molecules, the solvent heating contribution needs to be eliminated properly from the difference scattering curves. To do so, we implemented time-resolved X-ray scattering experiment on reduced equine heart cytochrome *c* (Cyt *c*) that is in a folded conformation. Upon photoexcitation, Cyt *c* undergoes an exothermic reaction starting from the excited state and dissipates the excitation energy into the solvent without much change in the global conformation of the protein.[1] Therefore, the scattering signal at 1 ms time delay can be approximated as the solvent heating response of the water solvent. As shown in Figure S2, the solvent heating response is dominant in the wide-angle region and is small and featureless in the small-angle region (q < 0.6 Å^{-1}). The solvent heating response of Cyt *c* in the wide-angle region (1.0 – 2.3 Å^{-1}) can be reconstructed by a linear combination of the temperature change at constant density, $(\partial S(q)/\partial T)_\rho$, and the density change at constant temperature, $(\partial S(q)/\partial \rho)_T$.[2] The experimental condition for Cyt *c* was almost the same as the one for PYP except for the excitation wavelength (532 nm for Cyt *c* and 460 nm for PYP): 293 K temperature, nanosecond laser pulses of 1 mJ/mm^2 fluence, and X-ray pulses of ~1.5 μs duration. Therefore, the solvent heating response determined for Cyt *c* can be used to eliminate the solvent heating contribution from time-resolved X-ray scattering curves of E46Q-PYP. The heating-free difference scattering curves were obtained by subtracting the scaled solvent heating response from the difference scattering curves of E46Q-PYP at various positive time delays.

(2) Kinetic analysis for the time-resolved X-ray scattering data

We applied singular value decomposition (SVD)[3-6] to the solvent heating-free experimental scattering curves to determine the kinetics of structural transitions of the protein. By the SVD analysis, the data matrix of time-resolved difference scattering curves, **A**, can be decomposed into left- and right-singular vectors using the relationship of $\mathbf{A} = \mathbf{USV}^{T}$, where **U** consists of the left-singular vectors (lSVs) corresponding to time-independent $\boldsymbol{q}$ spectra, **V** consists of the right-singular vectors (rSVs) corresponding to time-dependent amplitude change of **U**, and **S** contains the weights of the singular vectors. The lSVs contains the information on the scattering curves of structurally distinct transient intermediates while the rSVs represent the population dynamics of those transient intermediates. Thus, the SVD analysis allows us to obtain the number of structurally distinct species and the dynamics of each species regardless of the kinetic model.

From the SVD analysis in the *q*-range of 0.05 – 1.0 $Å^{-1}$ and the time range from 3.16 μs to 100 ms, we identified three significant singular vectors judged from the singular values (**S**) and the autocorrelation factors of the corresponding singular vectors. Figures S3a shows the first three lSVs while Figure S3b shows the first three rSVs scaled with singular values obtained from the SVD analysis. By globally fitting the three rSVs by a sum of multiple exponentials sharing common relaxation times, the time constants of 3.1 (± 2.2) μs, 143 (± 4) μs, and 6.7 (± 1.0) ms were determined. In fact, the difference scattering curve at 100 ms still has apparent oscillatory features in the entire *q*-range with slightly smaller amplitude than the difference scattering curve at 1 ms, indicating that the global conformational change of E46Q-PYP might be still in progress at 100 ms toward the ground state. Therefore, in addition to the three exponentials described above, an extra exponential with the time constant of ~1 sec was needed to describe the recovery to the ground state in the time range

of hundreds of ms to seconds. However, we were not able to determine the recovery time of the ground state precisely from the exponential fitting of rSVs due to the limited time window (up to 100 ms) of our experimental data. In addition, to check whether any other intermediate is formed beyond the time range of our experimental data, we performed the SVD analysis in several reduced time ranges. For example, from the SVD analysis in the time range from 3.16 μs to 1 ms (instead of 100 ms) as shown in Figure S4, we identified two significant lSVs and rSVs, suggesting that at least two intermediates are formed in the time range from 3.16 μs to 1 ms. However, an early intermediate formed within the time resolution (~1.5 μs) of our experiment can be obscured in the SVD analysis due to the following reasons; (1) our experimental data measured from 3.16 μs are not enough to resolve the dynamics earlier than 3.1 μs and (2) the difference scattering curve at 3.16 μs already has significant amplitudes. According to previous time-resolved studies on E46Q-PYP,[7-11] there exists an intermediate in the early μs time range. Therefore, by adding an intermediate formed before 3.1 μs, we can make a sequential kinetic model based on four structurally distinct intermediates and four time constants (3.1 μs, 143 μs, 6.7 ms, and ~1 second).

$$\mathbf{A} \xrightarrow{\mathbf{3.1}\ \boldsymbol{\mu s}} \mathbf{B} \xrightarrow{\mathbf{143}\ \boldsymbol{\mu s}} \mathbf{C} \xrightarrow{\mathbf{6.7}\ \boldsymbol{ms}} \mathbf{D} \xrightarrow{\mathbf{1}\ \boldsymbol{s}} \mathbf{pG}$$

Based on the above sequential kinetic model and the time constants estimated from the SVD analysis, we implemented principal component analysis (PCA) whereby the experimental scattering curves were decomposed into four species-associated scattering curves corresponding to the four intermediates as follows:

$$\Delta \boldsymbol{S}_{\boldsymbol{theory}}(\boldsymbol{q}_{\boldsymbol{i}}, \boldsymbol{t}_{\boldsymbol{j}}) = \sum_{\boldsymbol{k}=\mathbf{1}}^{\mathbf{4}} \left[\boldsymbol{C}_{\boldsymbol{k}}(\boldsymbol{t}_{\boldsymbol{j}})\right] \Delta \boldsymbol{S}_{\boldsymbol{C}_{\boldsymbol{k}}}(\boldsymbol{q}_{\boldsymbol{i}})$$

where $\Delta \boldsymbol{S}_{\boldsymbol{theory}}(\boldsymbol{q}_{\boldsymbol{i}}, \boldsymbol{t}_{\boldsymbol{j}})$ is the theoretical difference scattering curve at given q and t values, $\Delta \boldsymbol{S}_{\boldsymbol{C}_{\boldsymbol{k}}}(\boldsymbol{q}_{\boldsymbol{i}})$ is the species-associated difference scattering curve corresponding to the k-th

intermediate species at a given q value, $[\boldsymbol{C_k}(\boldsymbol{t_j})]$ is the instantaneous population of the k-th intermediate at a given t value and can be easily calculated using the time constants obtained from the SVD analysis. Then, we minimized the discrepancy between the theoretical and experimental scattering curves using the Minuit packages.

(3) Reconstruction of three-dimensional molecular shape of intermediates from species-associated difference scattering curves

The species-associated difference scattering curves determined from the kinetic analysis contain the information on the global conformation of the protein intermediates. We employed ab-initio shape reconstruction to extract the molecular shape of the protein from the species-associated difference scattering curves. This approach consists of two steps: (i) Species-associated "static" scattering curves are obtained by adding the scattering curve of the ground state to the scaled species-associated difference scattering curves of the intermediates. (ii) For the species-associated static scattering curves, the three-dimensional shape reconstruction is implemented by DAMMIN package[12] based on the dummy atom modelling.[13,14]

The species associated "static" scattering curve for each intermediate species is obtained by adding a static scattering curve of the ground state (pG) to the scaled species-associated difference scattering curve. A scaling factor to appropriately scale the species-associated difference scattering curve is calculated by considering two parameters: (1) the relative ratio between experimental and theoretical X-ray scattering curves and (2) the effective photoreaction yield that depends on the primary quantum yield of a protein. The detailed procedure of reconstructing the species-associated "static" scattering curve is described in a previous study of wild-type PYP (wt-PYP).[13]

For both E46Q-PYP and wild-type PYP, we used the same experimental condition for the laser fluence (1 mJ/mm^2 at a focal spot of 0.50 × 0.13 mm^2) and sample concentration (4.4 mM). From the laser fluence and sample concentration, we calculated the number of photons and the number of sample molecules at the focal spot, respectively, and found that the photoexcitation condition used in our measurement was in the saturation regime, guaranteeing highly efficient conversion to the signaling state. According to a previous study on E46Q-PYP,[11] E46Q-PYP has the primary quantum yield of a similar value as wild-type PYP. To compensate for the difference of extinction coefficient at 460 nm between the wild type (35.4 mM^{-1}cm^{-1}) and E46Q-PYP (53.8 mM^{-1}cm^{-1}), the laser path-length through the sample capillary was set to be 100 μm and 150 μm for E46Q-PYP and the wild type, respectively. As a result, the optical densities of the solution samples of E46Q-PYP and wild-type PYP were equivalent to each other:

$$OD_{wt-PYP} = 4.4\ mM \times 35.4\ mM^{-1}cm^{-1} \times 0.015\ cm = 2.35$$
$$OD_{E46Q-PYP} = 4.4\ mM \times 53.8\ mM^{-1}cm^{-1} \times 0.010\ cm = 2.36$$

As shown in Figure S5, we compared the difference scattering curves of E46Q-PYP and wild-type PYP. Each scattering curve was normalized by the flux of incident X-rays. In the small-angle scattering region (q < 0.3 Å^{-1}), the amplitude of difference scattering curve for E46Q-PYP is much smaller than that for wt-PYP, implying that the protein conformational change is smaller in E46Q-PYP than in wt-PYP.

To acquire the static scattering curve of the ground state of E46Q-PYP, we used a high-resolution crystal structure of E46Q-PYP (PDB: 1OTA[15]). To check whether this crystal structure is compatible with the structure of E46Q-PYP in solution, we compared the theoretical X-ray scattering curve calculated from the PDB structure with the experimental static scattering curve of E46Q-PYP in the ground state (4.4 mM E46Q-PYP in 20 mM Na-phosphate, 20 mM NaCl, and pH 7.0), measured at the BioCARS 14-ID-B beamline at

Advanced Photon Source. The theoretical scattering curve well matches the experimental one within the error range (see Figure S6), indicating that the crystal structure of E46Q-PYP can be used as the ground-state structure in solution. To consider the thermal fluctuation of the crystal structure in solution, we implemented a MD simulation with the crystal structure as a starting point. The MD simulation was performed with the GROMACS 4.5.5 package,[16] employing Gromos96 43a1 force field[17] in combination with the SPC water model and the Berendsen thermostat.[18] The simulation box was equilibrated for 10 ns at 293 K. From the analysis of the MD trajectory, the thermal fluctuation of the protein was less than 0.2 nm in terms of the root mean square fluctuation (RMSF) of the atomic positions. By clustering the protein structures in the trajectory, a series of representative protein structures (40 structures) were determined and used to calculate the static X-ray scattering curve. As shown in Figure S7a, the averaged theoretical X-ray scattering curve of the 40 structures extracted from the clustering analysis for MD snapshots has a similar shape as the scattering curve calculated from the crystal structure in the entire q range. To take into account the thermal fluctuation of the protein at 293 K, the averaged theoretical scattering curve from the simulated structures was used as the static X-ray scattering curve of E46Q-PYP in the ground state.

To identify the three-dimensional shape of a protein intermediate, it is required to determine the optimal pair distribution function. From the reconstructed species-associated "static" scattering curve of each intermediate, the pair distribution function, $P(r)$, and the radius of gyration, R_g, of the intermediate can be obtained by the following Fourier transform relationship:

$$S(q) = \int_0^{D_{max}} dr\, P(r) \frac{\sin(qr)}{qr}$$

where D_{max} is the maximum intramolecular distance. As the pair distribution function is a real-space representation of the X-ray scattering profile, it allows us to obtain the direct

information on the protein conformation. To determine the pair distribution function from the reconstructed species-associated static scattering curves, we implemented the indirect Fourier transform using the GNOM software.[19] The resultant pair distribution function of each intermediate was used as a basis of the ab-initio shape reconstruction. Using the optimal pair distribution function and the species-associated static scattering curve, the ab-initio dummy-atom modeling was implemented to reconstruct the molecular shape of a protein intermediate using the DAMMIN software.[12] In the shape reconstruction, a protein conformation is described as a collection of densely packed bead-atoms inside a constrained volume with its maximum diameter determined from the indirect Fourier transform of the species-associated static scattering curve. The configuration of randomly oriented bead-atoms was allowed to change until the discrepancy between the species-associated static scattering curve and the theoretical scattering curve becomes minimized. By this procedure, we obtained an optimized arrangement of bead-atoms that reflects the global shape of the protein (see Figure S8a). For each species-associated scattering curve, multiple runs of shape reconstruction were performed to make a pool of protein structures reflecting the structural fluctuation of the protein. From this structure pool containing ~50 reconstructed structures, a representative reconstructed shape of the protein was generated to determine an average structure of each intermediate using DAMAVER and DAMFILT[20] (see Figure S8a). To verify the effectiveness of our approach for the protein shape reconstruction, we performed the shape reconstruction from the ground-state structure of E46Q-PYP obtained from the MD simulation. From the averaged X-ray scattering curve calculated from the MD-simulated structures, we reconstructed the molecular shape of the ground state. As shown in Figure S7b, the representative molecular shape well matches the 40 protein structures obtained from the MD simulation. We applied the ab-initio shape reconstruction to the species-associated static scattering curves of all the intermediates (pR_{α}, pR_{β}, pB_{α}, and pB_{β}). The final representative

shapes of the intermediates clearly show conformational changes of the protein accompanying the protrusion in a local region during the signaling process of E46Q-PYP (see Figure 4).

(4) Effect of different experimental conditions between TRXSS and TA measurements

The experimental conditions (for example, sample concentration and laser fluence) used for the TRXSS and TA measurements were different from each other. Unless both TRXSS and TA experiments are conducted under the exactly same experimental condition, which is not practically possible with current technology, we cannot completely rule out the possibility that the different experimental conditions for the two types of measurements may affect the protein dynamics measured by the two techniques. However, for the following reasons, we believe that the difference in experimental conditions between the two methods did not affect our conclusion on the transition dynamics from pR to pB.

(a) Sample concentration

Among various experimental parameters, a major factor that may influence the measured protein dynamics is the concentration of a protein solution. According to a comprehensive study on the influence of the protein environment (i.e. dilute solution vs. crystalline sample vs. macromolecular crowding) on photoinduced dynamics of wt-PYP and E46Q-PYP,[10] the crowded solution of wt-PYP containing a high concentration (~10 mM) of bovine serum albumin as the co-solute macromolecules exhibits the same TA dynamics as the diluted solution (~25 μM) of wt-PYP while the crystalline sample shows much faster ground state recovery than the dilute solution. Since the concentrations of solution samples used for the TRXSS (4.4 mM) and TA (70 μM) measurements lie within the concentration range between the dilute and crowded solution samples examined in the previous work and is much lower

than that of the crystalline sample, the difference between TRXSS and TA dynamics is not likely to result from the difference in concentration of the protein solutions.

(b) Laser fluence

Another experimental parameter that may affect the measured protein dynamics is the laser fluence. In this work, the laser fluence was much higher in the TRXSS measurement (1 mJ/mm^2) than in the TA measurement (76 $\mu J/mm^2$), but we should consider that the sample concentration was also much higher in the TRXSS measurement. For more valid comparison of excitation condition, we calculated the excitation photon density per sample molecule using the laser fluence and the sample concentration and found that the photon density of the TRXSS measurement (6.7 photons/molecule) was actually lower than that of the TA measurement (50 photons/molecule) but still lies in the saturation regime. Meanwhile, we used nanosecond laser pulses as an excitation source for both measurements and therefore coherent multi-photon excitation should be minimal. In fact, previous TA studies employing nanosecond excitation pulses[10,21] reported the same pR-to-pB transition dynamics as the present work. Thus, the excitation condition for the TRXSS and TA measurements can be regarded equivalent to each other.

References

1. T. W. Kim, J. G. Kim, C. Yang, H. Ki, J. Jo and H. Ihee, *Bull. Korean Chem. Soc.*, 2014, **35**, 697-698.
2. M. Cammarata, M. Lorenc, T. K. Kim, J. H. Lee, Q. Y. Kong, E. Pontecorvo, M. Lo Russo, G. Schiro, A. Cupane, M. Wulff and H. Ihee, *J. Chem. Phys.*, 2006, **124**.
3. E. R. Henry and J. Hofrichter, *Methods Enzymol.*, 1992, **210**, 129-192.
4. M. Schmidt, R. Pahl, V. Srajer, S. Anderson, Z. Ren, H. Ihee, S. Rajagopal and K. Moffat, *Proc. Natl. Acad. Sci. U.S.A.*, 2004, **101**, 4799-4804.
5. H. Ihee, S. Rajagopal, V. Srajer, R. Pahl, S. Anderson, M. Schmidt, F. Schotte, P. A. Anfinrud, M. Wulff and K. Moffat, *Proc. Natl. Acad. Sci. U.S.A.*, 2005, **102**, 7145-7150.
6. K. Y. Oang, J. G. Kim, C. Yang, T. W. Kim, Y. Kim, K. H. Kim, J. Kim and H. Ihee, *J. Phys. Chem. Lett.*, 2014, **5**, 804-808.
7. S. Rajagopal, S. Anderson, V. Srajer, M. Schmidt, R. Pahl and K. Moffat, *Structure*, 2005, **13**, 55-63.
8. S. Devanathan, S. Lin, M. A. Cusanovich, N. Woodbury and G. Tollin, *Biophys. J.*, 2000, **79**, 2132-2137.
9. Y. O. Jung, J. H. Lee, J. Kim, M. Schmidt, K. Moffat, V. Srajer and H. Ihee, *Nat. Chem.*, 2013, **5**, 212-220.
10. S. Yeremenko, I. H. M. van Stokkum, K. Moffat and K. J. Hellingwerf, *Biophys. J.*, 2006, **90**, 4224-4235.
11. L. J. G. W. van Wilderen, M. A. Van der Horst, I. H. M. van Stokkum, K. J. Hellingwerf, R. van Grondelle and M. L. Groot, *Proc. Natl. Acad. Sci. U.S.A.*, 2006, **103**, 15050-15055.
12. D. I. Svergun, *Biophys. J.*, 1999, **77**, 2896-2896.
13. T. W. Kim, J. H. Lee, J. Choi, K. H. Kim, L. J. van Wilderen, L. Guerin, Y. Kim, Y. O. Jung, C. Yang, J. Kim, M. Wulff, J. J. van Thor and H. Ihee, *J. Am. Chem. Soc.*, 2012, **134**, 3145-3153.
14. J. G. Kim, T. W. Kim, J. Kim and H. Ihee, *Acc. Chem. Res.*, 2015, **48**, 2200-2208.
15. S. Anderson, S. Crosson and K. Moffat, *Acta Crystallogr. Sect. D*, 2004, **60**, 1008-1016.
16. D. Van der Spoel, E. Lindahl, B. Hess, G. Groenhof, A. E. Mark and H. J. C. Berendsen, *J. Comput. Chem.*, 2005, **26**, 1701-1718.
17. X. Daura, A. E. Mark and W. F. van Gunsteren, *J. Comput. Chem.*, 1998, **19**, 535-547.
18. H. J. C. Berendsen, J. P. M. Postma, W. F. Vangunsteren, A. Dinola and J. R. Haak, *J. Chem. Phys.*, 1984, **81**, 3684-3690.
19. D. I. Svergun, *J. Appl. Crystallogr.*, 1992, **25**, 495-503.
20. V. V. Volkov and D. I. Svergun, *J. Appl. Crystallogr.*, 2003, **36**, 860-864.
21. U. K. Genick, S. Devanathan, T. E. Meyer, I. L. Canestrelli, E. Williams, M. A. Cusanovich, G. Tollin and E. D. Getzoff, *Biochemistry*, 1997, **36**, 8-14.

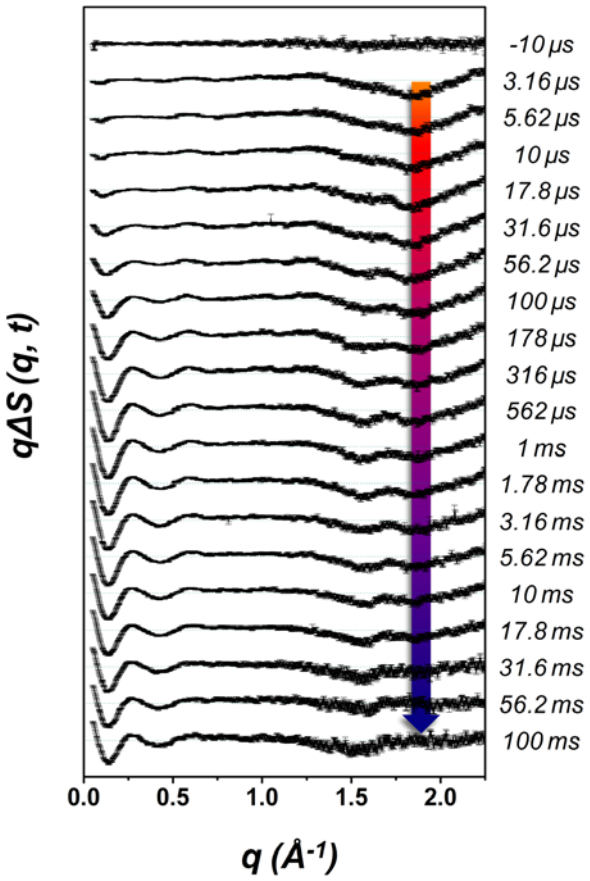

Figure S1. Time-resolved X-ray solution scattering curves of E46Q-PYP in the *q*-range from 0.05 to 2.25 $Å^{-1}$. The coloured arrow indicates the contribution of solvent heating induced by the dissipation of laser excitation energy.

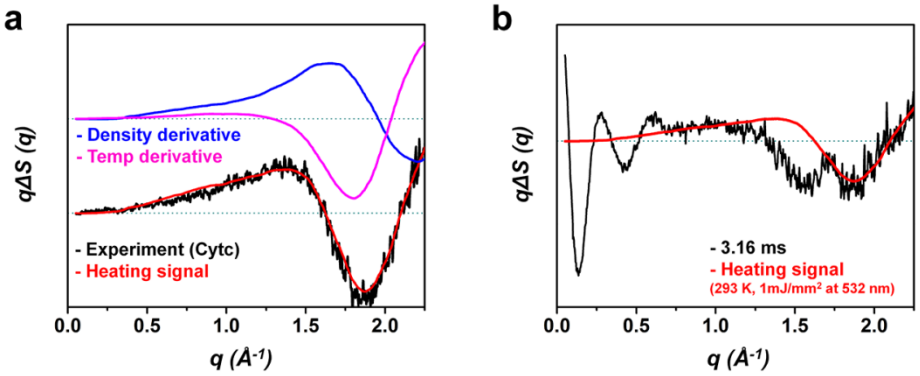

Figure S2. (a) Time-resolved X-ray solution scattering curve of reduced equine heart cytochrome *c* (Cyt *c*) at 1 ms time delay (black line). The difference scattering curve at 1 ms time delay is dominated by the solvent heating response (red line), which can be described by a linear combination of the temperature change at constant density $(\partial S(q)/\partial T)_\rho$ (magenta line) and the density change at constant temperature $(\partial S(q)/\partial \rho)_T$ (blue line). (b) Comparison of the solvent heating response of Cyt *c* (red line) and the difference scattering curve of E46Q-PYP at 3.16 ms time delay (black line). It can be seen that the small-angle region of $q < 0.6$ Å is not much affected by the solvent heating.

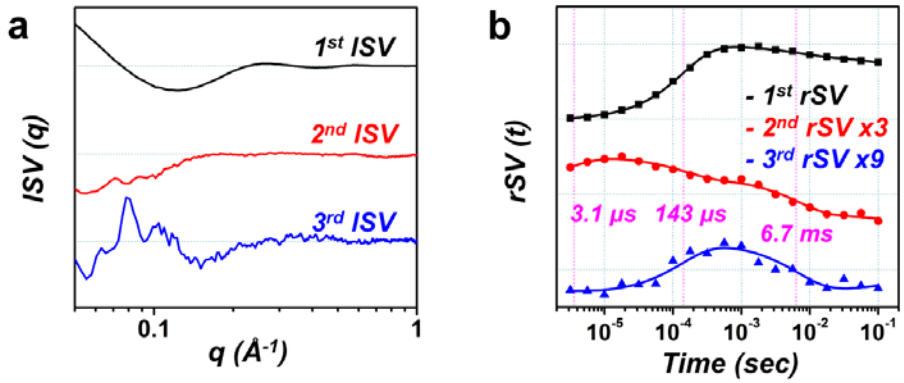

Figure S3. Singular value decomposition (SVD) analysis of the difference scattering curves. (a) The first three left-singular vectors (lSVs). (b) The most significant right-singular vectors (rSVs) multiplied by their corresponding singular values (S). The rSVs were globally fitted by multiple exponential functions (solid lines), yielding the time constants of 3.1 µs, 143 µs, 6.7 ms, and ~1 s (fixed value).

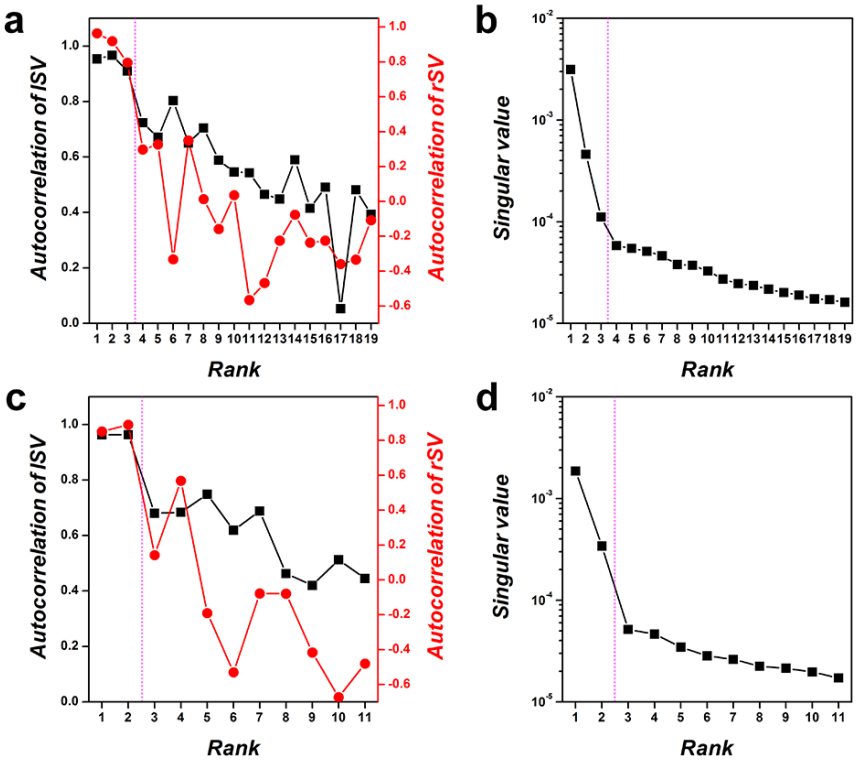

Figure S4. (a) Autocorrelation values of left singular vectors (lSVs, black squares) and right singular vectors (rSVs, red circles) and (b) singular values obtained from the time-resolved X-ray scattering data of E46Q-PYP in the time range from 3.16 μs to 100 ms. (c) Autocorrelation values of lSVs and RSVs and (d) singular values in the time range from 3.16 μs to 1 ms. For the data from 3.16 μs to 1 ms, the first two lSVs and rSVs are significant while the first three lSVs and rSVs are significant for the data from 3.16 μs to 100 ms.

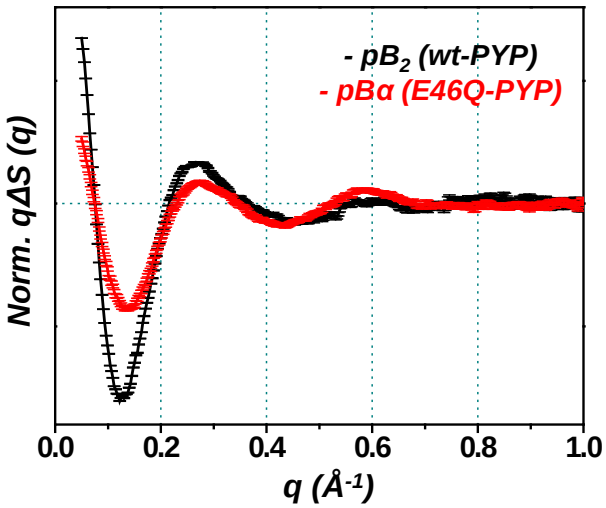

Figure S5. Comparison of the difference scattering signals of wt-PYP (black line) and E46Q-PYP (red line). The scattering curves were normalized by the photon flux of incident X-rays. Considering that the excitation condition, primary quantum yield, and optical density of sample solutions are equivalent for wt-PYP and E46Q-PYP, the smaller amplitude of the E46Q-PYP signal indicates that E46Q-PYP undergoes the signaling transition with smaller structural change than wt-PYP.

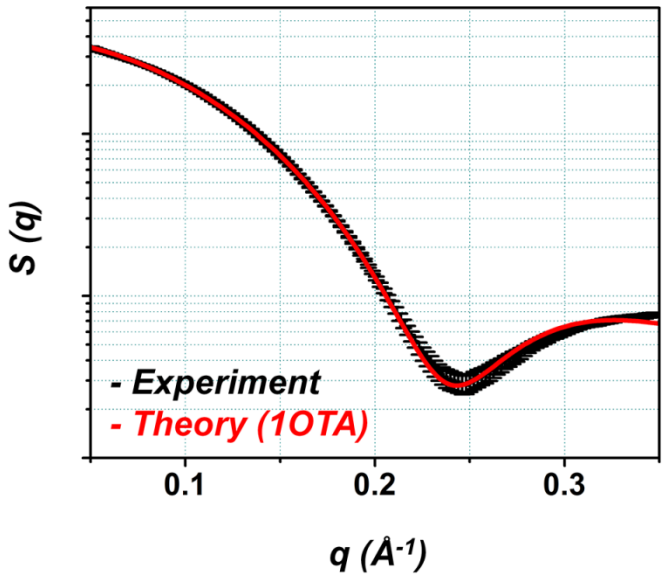

Figure S6. Comparison of experimental (black line) and theoretical (red line) static scattering curves for the ground state (pG). In the small-angle region (0.05 – 0.3 $Å^{-1}$), the theoretical scattering curve calculated from the crystal structure (PDB: 1OTA) well matches the experimental scattering curve (4.4 mM E46Q-PYP in 20 mM Na-phosphate buffer, 20 mM NaCl, and pH 7.0) within the error range.

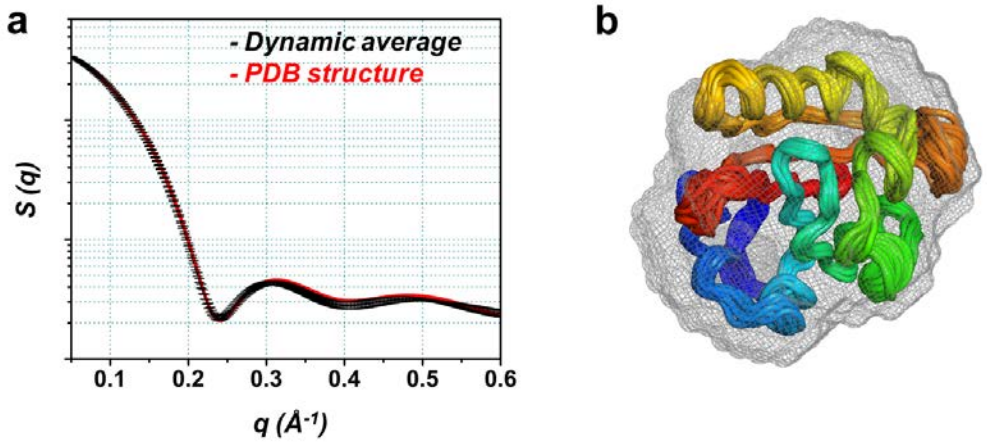

Figure S7. Test for the effectiveness of the structural refinement. The molecular dynamics (MD) simulation was implemented for the crystal structure of E46Q-PYP (PDB: 1OTA[15]) to consider the thermal fluctuation of protein backbone at 293 K. (a) The averaged theoretical X-ray scattering curve (black line) calculated from the representative structures (40 structures) obtained from the MD simulation well matches the scattering curve (red line) calculated for the crystal structure in the small-angle region (q < 0.6 Å^{-1}). (b) The molecular shape of the pG state (grey mesh) reconstructed from the averaged theoretical scattering curve (black line in Figure S7a) well matches the representative structures obtained from the MD simulation (rainbow ribbon).

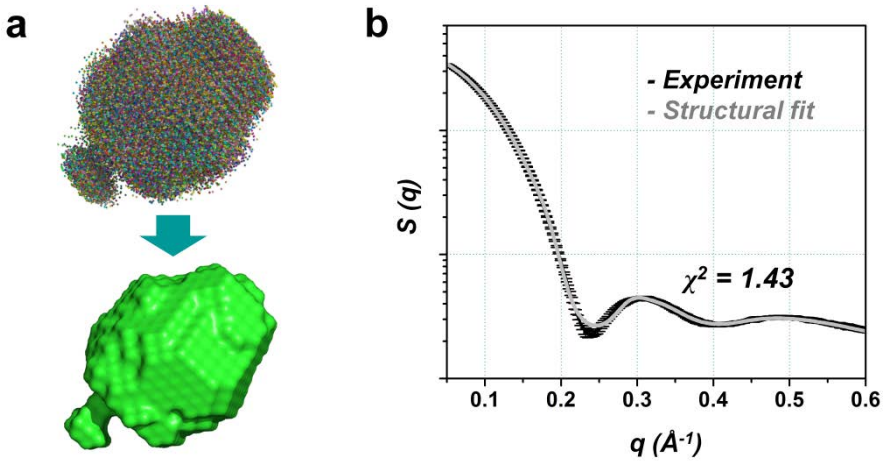

Figure S8. Shape reconstruction of protein intermediate structures from species-associated static scattering curves. (a) The optimized molecular shapes of pB_{α} reconstructed from the pair distribution function. The probability map (that is, conformational space) represented by 50 reconstructed protein shapes (rainbow color-coded dots). The most probable structure (green surface) is determined by filtering the probability map. (b) Comparison of the experimental species-associated static curve (black line) and the theoretical scattering curves from 50 independent reconstructions (gray line). The theoretical scattering curve from the shape reconstruction well fits the experimental curve within the error bars. The good agreement between the two curves demonstrates the reliability of our result.

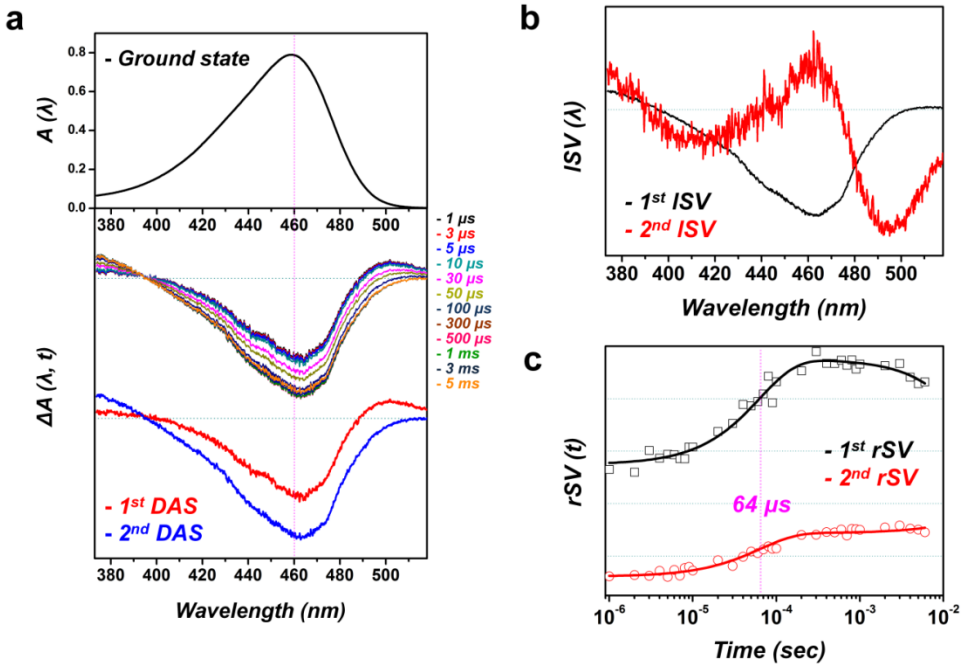

Figure S9. Kinetic analysis of the transient absorption spectra of E46Q-PYP. (a) Transient absorption spectra at representative time delays (three points per decade) are compared with the absorption spectrum of E46Q-PYP in the ground state. (b) The first two left-singular vectors (lSVs) from the SVD analysis of the transient absorption spectra. (c) The most significant right-singular vectors (rSVs) multiplied by their corresponding singular values (*S*). The rSVs were globally fitted by multiple exponential functions, yielding the time constants of 64 (± 4) μs and 100 ms (fixed value). With the transition from pR to pB, the 2nd decay associated spectra (DAS), corresponds to the species-associated difference scattering curve on TRXSS, shows a positive feature (that is, transient absorption) around 380 nm, which is related to the formation of the spectrally blue-shifted pB species, and a negative feature (that is, bleaching) around 500 nm, which is associated with the disappearance of the spectrally red-shifted pR species.

Table S1. Comparison of time constants determined from time-resolved X-ray solution scattering (TRXSS) and transient absorption (TA) spectroscopy.

	TRXSS	**TA**[a)]
$\mathbf{pR_\alpha \rightarrow pR_\beta}$	3.1 (± 2.2) μs	
$\mathbf{pR_\beta \rightarrow pB_\alpha}$	143 (± 4) μs	64 (± 4) μs
$\mathbf{pB_\alpha \rightarrow pB_\beta}$	6.7 (± 1.0) ms	

a) The time constant determined by TA spectroscopy was assigned to the transition from pR to pB that corresponds to the $pR_\beta \rightarrow pB_\alpha$ transition in the sequential kinetic model applied to the data from the TRXSS measurement.