

Length and Charge of the N-terminus Regulate the Lifetime of the Signaling State of Photoactive Yellow Protein

Cheolhee Yang,[§] Youngmin Kim,[§] Seong Ok Kim, Sang Jin Lee, Jungkweon Choi, and Hyotcherl Ihee*

Cite This: J. Phys. Chem. B 2023, 127, 9001–9013



III Metrics & More

Article Recommendations

ABSTRACT: Photoactive yellow protein (PYP) is one of the most extensively studied photoreceptors. Nevertheless, the role of the N-terminus in the photocycle and structural transitions is still elusive. Here, we attached additional amino acids to the N-terminus of PYP and investigated the effect of the length and charge of additional N-terminal residues using circular dichroism, two-dimensional nuclear magnetic resonance (2D-NMR), transient absorption (TA), and transient grating (TG) spectroscopic techniques. TA experiments showed that, except for negatively charged residues (5D-PYP), additional N-terminal residues of PYP generally enable faster dark recovery from the putative signaling state (pB₂) to the ground state (pG). TG data showed that although the degree of structural changes can be controlled by adjusting specific amino acid residues in the extended N-terminus of N-terminal extended PYPs (NE-PYPs), the dark recovery times of wt-PYP and NE-PYPs, except for 5D-PYP, are independent of the structural differences between pG and pB₂ states. These results demonstrate that the recovery time and the degree of structural



change can be regulated by controlling the length and sequence of N-terminal residues of PYP. The findings in this study emphasize the need for careful attention to the remaining amino acid residues when designing recombinant proteins for genetic engineering purposes.

INTRODUCTION

Photoactive yellow protein (PYP) is a small, globular protein with 125 amino acids (14 kDa) and has *p*-coumaric acid (*p*CA) as the chromophore. The photocycle reaction of PYP initiated with the photoisomerization of pCA has been extensively investigated as a representative model system of photosensing proteins using various time-resolved techniques.¹⁻²³ Briefly, upon the blue light irradiation, the ground state of PYP, pG (λ_{max} = 446 nm), undergoes the photocycle as shown in Figure 1a. The red-shifted intermediates, pR_1 (λ_{max} = 465 nm) and pR_2 , decay within a microsecond and hundred microsecond time scales, respectively, to form a blue-shifted intermediate ($\lambda_{max} = 355$ nm), pB_1 . The pB_1 intermediate mostly transforms to another blue-shifted intermediate, pB_2 , which returns to pG. It has been reported that the transition from pB₂ to pG occurs through the interaction between N-terminal amino acids and the β -scaffold. The pB₂ species shows the largest conformational change associated with the partial unfolding of the N-terminus and is presumed as the signaling state of PYP.^{1,2}

To understand the role of the N-terminus in the photocycle of PYP, N-terminal truncated PYPs (NT-PYPs) have been investigated with various techniques.^{3–7,24,25,44,45} Studies using small-angle X-ray scattering (SAXS),²⁴ X-ray crystallography,²⁵ and nuclear magnetic resonance (NMR)⁶ demonstrated that the ground state of NT-PYPs has a similar globular structure to that of wild-type PYP (wt-PYP). In transient absorption (TA) spectroscopy studies, NT-PYPs showed significantly slower dark recovery time from pB₂ to pG than wt-PYP.^{3,4} It was proposed

that the deceleration of the dark recovery observed in NT-PYPs is due to the absence of the interaction between N-terminal amino acids and the β -scaffold surrounding the chromophorebinding domain.³ In the pG state, the N-terminal loop interacts with the β -scaffold through hydrogen bonds and ionic bonds, stabilizing the pG state. It was also suggested that the interaction between Phe6 and Lys123 is important among the interactions between the N-terminus and the β -sheet.⁴⁸ When the Nterminal loop is shortened, the interaction between the loop and the β -scaffold is weakened, leading to the destabilization of the pG state and a slower return to the intact state. In other words, the N-terminal loop acts as a "molecular switch" that regulates the structural changes and stability of PYP. Indeed, Khan et al. showed that the NT-PYPs exhibited a smaller difference than wt-PYP in diffusion coefficients between pG species and pB species, implying that the conformational change in the pB species of NT-PYPs is less significant than that observed in wt-PYP.⁷ They also suggested that the conformational change in the N-terminal group is the main cause of the diffusion coefficient change. A study using Fourier-transform infrared (FTIR) spectroscopy to

Received:June 7, 2023Revised:September 14, 2023Published:October 11, 2023







Figure 1. Photocycle and structure of PYP and various constructs investigated in this work. (a) Schematic illustration for the photocycle of PYP. Appearances of pR_1 and pR_2 are according to the parallel model. The process corresponding to the recovery from pB_2 to pG was too slow to be detected in TG spectroscopy. (b) Molecular structure of wt-PYP (PDB ID: 2PHY). In this work, additional amino acid residues with varying charges and lengths were added to the N-terminus to investigate the effect of the extension of the N-terminus on the photocycle kinetics and structures of intermediates. (c) Schematic representations of PYP constructs, Saa-PYP and 17aa-PYP, each with an extended N-terminal region by additional amino acids. The N-terminal of Saa-PYP is extended by 5 amino acids (GPLGS), which are highlighted in orange, while the N-terminal of 17aa-PYP is extended by 17 amino acids (MRGSHHHHHHGSDDDDK), which are highlighted in blue. (d) Schematic representation of PYP constructs, SX-PYP, with N-terminal extension by 5 repeating amino acids, which are highlighted in purple. Using Gly(G), Leu(L), Ser(S), Lys(K), and Asp(D) as the repeating amino acids, a total of five SX-PYP constructs, SG-PYP, SL-PYP, SS-PYP, SK-PYP, and SD-PYP, were generated. (e) Sequence representations of PYP constructs considered in this work.

investigate NT-PYPs found that the blue light irradiation induces smaller changes of the absorption bands corresponding to the amide mode of β -sheet in NT-PYPs, compared to those in wt-PYP.⁵ Additionally, the amide modes of the β -sheet in NT-PYPs were observed in downshifted regions compared to wt-PYP. These bands are also linked to hydrogen bonding between the N-terminus and β -sheet through water molecules. In this regard, Harigai et al. suggested that the interaction between the N-terminus and the β -scaffold is related to the conformational change during the photocycle reaction of PYP.⁵ Additionally, they suggested that the electrostatic interaction plays an important role in the recovery of PYP.

While the role of the N-terminus of PYP in its biological functions has been extensively studied through investigations of NT-PYPs, here, we aimed to explore the potential impact of extending, rather than removing, the N-terminus of PYP on its biological functions. If the length of the N-terminus regulates PYP functions, modifying it can offer a means of modulating PYP activity. It is also worth noting that the extension of a protein terminus occurs frequently during the preparation of recombinant proteins. Several research groups reported that the His-tag sequence used for protein purification frequently could induce changes in intrinsic properties of a protein, such as solubility^{26–29} and structure.^{30–35} For example, the His-tagged sensory rhodopsin II from *Natronobacterium pharaonis* showed a slight difference in the volume change and the activation energy for the production of K510 intermediate compared to wild-type sensory rhodopsin II, indicating a structural difference between

the K510s of two constructs.³⁵ The studies imply that the incorporation of additional N-terminal residues can alter the inherent properties of a protein. This notion may hold true for PYP as well, where the addition of extra N-terminal residues has the potential to influence its photocycle and signal transduction.

Therefore, in this study, we aimed to investigate the effect of the additional N-terminal residues on the photocycle of PYP using various spectroscopic techniques, including circular dichroism (CD), NMR, TA spectroscopy, and transient grating (TG) techniques. For this purpose, we first designed two Nterminal extended PYPs (NE-PYPs), 5aa-PYP and 17aa-PYP, where frequently observed sequences after affinity columns were attached to the N-terminus (see Methods). One (5aa-PYP) is with additional five amino acids (GPLGS), and the other (17aa-PYP) is with 17 additional residues (MRGSHHHHHHGSDDDDK). The GPL amino acid residual sequence is frequently observed after HRV C3 protease cleavage of the His-tag. The amino acid sequences MRGS and GS frequently appear for the purpose of subcloning genes into vectors. The six-histidine residue sequence (HHHHHH) is known as the His-tag. The DDDDK site is a typical enterokinase cleavage site for cutting the His-tag. To elucidate the effect of the electrostatic interaction on the photocycle of PYP, we also prepared five NE-PYPs with five additional repetitive sequences: DDDDD-PYP (5D-PYP) with a negatively charged side chain, GGGGG-PYP (5G-PYP) which has hydrogen as its side chain, KKKKK-PYP (5K-PYP) with a positively charged side chain, LLLLL-PYP (5L-PYP) with a hydrophobic side chain, and SSSSS-PYP (5S-PYP) with an uncharged polar side chain at pH 7.0. The TG technique can provide the global structural information on a protein by monitoring the diffusion coefficients of a reactant and a product, whereas the TA technique gives information related to local protein structural changes occurring around the chromophore. The results presented herein show that the pB₂ state can be regulated by the length and charge of the amino acid residues in the extended N-terminus of PYP. PYP has been used as a photoresponsive module for developing optogenetic tools that can manipulate the activity of target proteins through linear fusion with PYP and the target protein, such as DNA transcription factors (i.e., bZIP, GCN4)⁴¹ and hemolytic activity (staphylococcal α -hemolysin).⁴² Therefore, our findings may have implications for the development of optogenetic tools that can be controlled using photoactive proteins via adjusting the length and sequence of the N-terminus to regulate the magnitude and duration of the signaling state of the optogenetic tools.

METHODS

Reconstruction of pQE80L-His Tag-HRV 3C Protease **Recognition site-LGS-PYP.** The PYP gene with the HRV 3C protease recognition site and Leu–Gly–Ser was amplified by using PCR and was inserted into the pQE80L vector. The primers for the PYP gene containing *Bam*HI restriction enzyme sites are as follows (5' to 3'):

Forward: TCACCATCACGGATCCCTG-GAAGTGCTGTTTCAGGGCCCGCTGGGCAGCATG-GAACACGTAGCCTTC

Reverse: GCTCGCATGCGGATCCCTAGACGCGCTT-GACGAAGAC

The amplified HRV 3C protease-LGS-PYP gene was subcloned into the pQE80L vector by using an EZ-Cloning kit (Enzynomics).

Purification of wt-PYP and 17aa-PYP. The pQE80L-EK-PYP plasmid containing the enterokinase protease site in front of the PYP gene was transformed into Escherichia coli BL21 (DE3). Then, a single colony from the transformed plate was used to inoculate 50 mL of LB broth as the seed culture. The overnightgrown seed culture was used to inoculate 12 L of LB broth for a large culture at 37 °C. When the optical density (O.D.) at 600 nm of the large culture reached 0.6, 1 mM IPTG was added, and the culture was further incubated overnight at 18 °C. After incubation, the culture was centrifuged at 6000 rpm for 15 min. The pellet cells were resuspended in a 20 mM Tris-HCl buffer (pH 7.0) containing 150 mM NaCl and 1 mM PMSF. To disrupt the cells, 30 min of sonication was applied. To incorporate the chromophore, the *p*-coumaric anhydride was added to the cell lysate, and the color change was inspected. Then, the mixture was further centrifuged at 17,000 rpm for an hour to separate the cell debris. The supernatant was subjected to nickel affinity chromatography and purified with a gradient elution method. The purified PYP solution was dialyzed against an enterokinase reaction buffer (20 mM Tris-HCl (pH 7.4) containing 50 mM NaCl and 2 mM CaCl₂). The dialyzed protein sample was treated with enterokinase and incubated overnight to cleave the His-tag. The cleaved protein sample was dialyzed against 20 mM Tris-HCl (pH 7.0) buffer. The dialyzed protein solution was applied to an ion exchange column (HiTrap, GE Healthcare Life Sciences). Then, the protein was eluted in a gradient of slowly increasing concentrations of NaCl. The purity index of each fraction was determined using the absorbance ratio of 280 to 460 nm. The fractions that have a purity index lower than 0.5 were collected. Finally, the wt-PYP solution was dialyzed against 20 mM Tris-HCl (pH 7.0) containing 20 mM NaCl.

To obtain 17aa-PYP instead of wt-PYP, the digestion step with enterokinase was bypassed, and the same purification procedure was applied. The size of each construct was confirmed using SDS-PAGE gel electrophoresis and MALDI-TOF mass spectrometry.

Purification of 5aa-PYP. The cell culture and harvest purification procedure were the same as those used in wt-PYP preparation except for the following. We used the pQE80L-His Tag-HRV 3C protease recognition site-LGS-PYP plasmid for 5aa-PYP. After purification using nickel affinity chromatography, the prepared PYP has a His-tag with an HRV 3C protease recognition site. The purified solution was dialyzed against HRV 3C protease reaction buffer (50 mM Tris-HCl (pH 7.5) containing 0.15 M NaCl). The dialyzed sample was treated with HRV 3C protease and incubated overnight to cleave His-tag. The HRV 3C protease recognizes the Leu-Glu-Val-Leu-Phe-Gln-Pro amino acid sequence and cleaves between Gln and Gly. After cleavage, two amino acids, Gly-Pro, remained. Because we inserted Leu-Gly-Ser amino acids, 5aa-PYP has five additional amino acids (GPLGS). After the removal of the His-tag, the remaining purification steps were the same as those used for wt-PYP.

UV–Visible (UV–vis) and Circular Dichroism (CD) Spectroscopy Measurements. After the purification of the proteins, the UV–vis spectra were collected at 0.5 nm intervals using a UV–vis spectrometer (UV-2550, Shimadzu Corp., Japan) with a 1 cm path length quartz cuvette. The baseline was measured with the final buffer (20 mM Tris–HCl (pH 7.0) and 20 mM NaCl) in the same cuvette and was used for correcting the baseline via subtraction. To confirm the secondary structures of the proteins, the CD spectra were collected using a CD spectrometer (Jasco-815, JASCO Inc., Japan) with a 1 mm quartz cuvette at room temperature. The spectral window was from 190 to 260 at 1 nm intervals. The baseline was measured with the same buffer in the same cuvette and was used for subtraction.

¹⁵N-Labeled Sample Preparation and NMR Experiments. Uniformly ¹⁵N-labeled wt-PYP, 5aa-PYP, and 17aa-PYP were expressed in *E. coli* BL21(DE3) using minimal media with ¹⁵NH₄Cl as the nitrogen source for the isotope-labeled proteins. The expressed proteins were purified as described previously. The purified proteins were dialyzed against 5 mM potassium phosphate (pH 5.8) and concentrated using an Amicon Ultra-15 centrifugal filter (MWCO 10 kDa) to a final concentration of 750 μ M with 10% of D₂O.

NMR experiments were performed on a 600 MHz Varian Unity Inova (KAIST, Daejeon, Korea) spectrometer equipped with a 5 mm triple-resonance probe with *xyz* gradient coils at 298.15 K. Two-dimensional ¹H, ¹⁵N-heteronuclear single quantum coherence (HSQC) spectra of wt-PYP, Saa-PYP, and 17aa-PYP were acquired. Obtained NMR data were processed with NMRPipe¹ and were analyzed and visualized by SPARKY software.²

Transient Absorption (TA) Measurement. The experimental setup of TA measurement was similar to the one used in our previous studies.³⁻⁶ A third harmonic (355 nm) of a Nd:YAG laser (NL 301G, EKSPLA) was used to pump an optical parametric oscillator (LT-2214-PC, LOTIS), producing a 460 nm pump beam. A continuous-wave 250 W Xe lamp (66902, Newport) was used as a probe source. The probe light was directed to the sample position at an acute angle against the pump laser beam to maximize the overlap between the pump and probe. The resulting TA signal was passed through a spectrometer (SpectraPro 2300i, Princeton Instrument) and detected by an ICCD (iStar, Andor) or a photomultiplier tube (PMT). The observed wavelength region using the ICCD was from 350 to 490 nm. The time delays were from 100 ns to 10 ms. To improve the signal-to-noise ratio, TA spectra were averaged 200 times. The TA signal measured by using a PMT was digitized by a digital oscilloscope (TDS 3052B, Tektronix). The signal was averaged 320 times to achieve a sufficient signal-tonoise ratio. The concentrations of PYPs were adjusted to be about 70 μ M. The sample solution was flown by using a syringe pump and a flow cell (2 mm path length) to avoid the unnecessary photosaturation of the sample solution. The repetition rate of the pump beam was set at 1 Hz when the sample was flowing. To collect the data with a longer time delay than 1 s in the PMT measurement, the flow was stopped, and the repetition rate of the pump beam was adjusted to 0.5 Hz. Even though the entire Δ O.D. at three wavelengths decreased by the stopping of flow, the temporal profiles could be overlapped by scaling, which means the stopping of flow only affected the signal level, not the kinetics (data are not shown). For the analysis of TA spectra measured using ICCD, singular value decomposition (SVD) was used. The $m \times n$ matrix, A, was composed of the measured TA spectra, where *m* is the number of the wavelength point for each absorption spectra and *n* is the number of the time delay. This matrix A can be decomposed into the separate matrices U, S, and V, which have the relationship of $A = USV^{T}$. U is the left singular matrix of $m \times n$ dimension, and each column vector of U is a time-independent spectrum. S is the singular value matrix of the $n \times n$ dimension. **S** is a diagonal matrix, and the diagonal values of S are called the singular values, which are related to the contribution of each left singular vector (RSV) and

right singular vector (LSV) corresponding to the singular value. **V** is the right singular matrix $n \times n$, and each column of the **V** matrix is a temporal change corresponding to the time-independent spectrum.

Transient Grating (TG) Measurement. The TG measurement setup and procedures were similar to those employed in previous studies.³⁻⁶ The 460 nm laser pulse was used as a pump beam, which was produced from an optical parametric oscillator (LT-2214-PC, LOTIS) pumped by a third harmonic (355 nm) of an Nd:YAG laser (NL 301G, EKSPLA). A cw photodiode laser (835 nm, Thorlabs) was used as a probe beam. The pump beam was split into two equivalent beams, and these excitation beams and the probe beam were intersected with each other inside the sample. A signal was diffracted in the direction of the vector sum of the three beams. The diffracted signal (TG signal) was separated from the scattered excitation beams by using a long-pass filter (>780 nm) and detected by a photomultiplier tube module (H7732-10, Hamamatsu). The detected signal was digitized by a digital oscilloscope (TDS 3052B, Tektronix). To obtain a sufficient signal-to-noise ratio, the signal was averaged 160 times. The repetition rate of the pump beam was 0.5 Hz. The total energy of the excitation beams was adjusted under ~15 μ J to prevent multiphoton absorption. The size of the excitation beam at the sample position was focused to $\sim 1 \text{ mm}$ diameter. The irradiated volume by the excitation laser was small compared to the total volume of the sample. The TG signals were measured at various grating wavenumbers, q, which varied by changing the angle between the two pump beams. The rate constants of the thermal decay process in each setup at various q values were determined from the TG signal of the reference sample, bromocresol purple (BCP), which releases the absorbed photon energy only as heat without any photoreactions. Using the thermal decay rate and well-known thermal diffusion coefficient for water, $D_{\rm th}$, the q value at each alignment was determined accurately. The concentrations of PYP constructs were adjusted to be about 340 μ M. The sample solution was contained by using a quartz cell (2 mm path length).

Molecular Dynamics (MD) Simulation. MD simulations for structures under the following four conditions were conducted utilizing MD simulations: (i) pG state of wt-PYP, (ii) pB₂ state of wt-PYP, (iii) pG state of 5D-PYP, and (iv) pB₂ state of 5D-PYP. For (i), the MD simulation used an X-ray crystallographic structure with the Protein Data Bank (PDB) entry of 2PHY as the starting structure to generate the equilibrium structures. For simulations, the GROMACS 2020 package with the Charmm36 force field was used, in combination with the tip4p water model. The system was equilibrated under NVT conditions for 100 ps with a velocity rescaling thermostat ($\tau_{\rm T}$ = 0.1 ps, T = 300 K) and was then subsequently equilibrated under NPT conditions for 500 ps with the velocity rescaling thermostat ($\tau_{\rm T}$ = 0.1 ps, *T* = 300 K) and a Parrinello–Rahman barostat ($\tau_{\rm T}$ = 0.5 ps, P = 1 bar). Then, a 300 ns long production simulation was conducted by using the equilibrated structure. For the MD simulation of (ii), we used the NMR structure (PDB ID: 2KX6) as the initial structure, employing the same approach as that used for (i). In the case of (iii) and (iv), the initial structure for each state was generated by elongating the crystal (or NMR) structure with five aspartic acid residues and was applied to the MD simulation employing the same approach used for (i) and (ii). Subsequently, for each simulation, the MD snapshots were sampled at 100 ps intervals from the MD trajectory. From the MD snapshots, we calculated the root-mean-square fluctuation (RMSF) of the protein moiety



Figure 2. (a) UV–vis spectra PYP constructs. Black, red, blue, green, cyan, magenta, dark yellow, and purple curves indicate UV–vis spectra of wt-PYP, Saa-PYP, 5D-PYP, SG-PYP, SK-PYP, SL-PYP, SS-PYP, and 17aa-PYP, respectively. (b) CD spectra of wt-PYP (black), 5aa-PYP (red), and 17aa-PYP (blue). For the comparison, CD spectra of Saa-PYP and 17aa-PYP were scaled using the intensity at 220 nm as a scaling point. (c) ¹H, ¹⁵N-HSQC spectrum of wt-PYP. (d) Overlaid ¹H, ¹⁵N-HSQC spectra of wt-PYP and Saa-PYP. (e) Overlaid ¹H, ¹⁵N-HSQC spectra of Saa-PYP and 17aa-PYP. (n) panels (c–e), orange, green, and blue peaks correspond to wt-PYP, Saa-PYP, and 17aa-PYP, respectively.

and the center of mass (COM) distance between the N-terminus and the β -sheets for each condition.

RESULTS AND DISCUSSION

Structures of 5aa-PYP and 17aa-PYP in the Ground State. To characterize NE-PYPs, UV-visible (UV-vis) absorption and CD spectra of wt-PYP, 5aa-PYP, and 17aa-PYP were measured. As shown in Figure 2a, the absorption spectra of these constructs were found to superimpose in the spectral region from 350 to 550 nm, indicating that the chromophore-site environment is consistently preserved, regardless of the presence of the extra N-terminal residues. As shown in the CD spectra (Figure 2b), wt-PYP, 5aa-PYP, and 17aa-PYP exhibit similar molar ellipticities at 220 nm, whereas the shoulders around 207 nm of 5aa-PYP and 17aa-PYP are slightly more pronounced than that of wt-PYP. Given that random coil content typically exhibits a negative band in the 190 to 210 nm range, we suggest that the variation observed in the CD spectra of wt-PYP, 5aa-PYP, and 17aa-PYP likely arises from the additional amino acids rather than a significant secondary structural change of the protein. Furthermore, the two negative peaks (at approximately 207 and 223 nm) in the CD spectra still resemble each other, indicating that the secondary structures of wt-PYP, 5aa-PYP, and 17aa-PYP are relatively similar. In summary, the CD spectra of 5aa-PYP and 17aa-PYP are similar to those of wt-PYP, indicating that the compositions of the secondary structures of 5aa-PYP and 17aa-PYP are unaffected despite the additional residues.

UV-vis and CD spectra suggest that both 5aa-PYP and 17aa-PYP share common structural features with wt-PYP. To further elucidate the ground state structures of 5aa-PYP and 17aa-PYP, we performed a heteronuclear single quantum coherence (HSQC) experiment of uniformly ¹⁵N-labeled PYP constructs. As previously reported, the chemical shifts of amide protons of

residues in wt-PYP, especially the structured region, are well spread and easy to be manually assigned based on the BMRB entry 18122.³⁶ Comparison of the HSQC spectra of the three constructs shows that the peak patterns and chemical shifts in the structured region (residues 26-125 of wt-PYP) of 5aa-PYP and 17aa-PYP are almost the same as those of wt-PYP (see Figure 2c-e), strongly suggesting that the overall structures of the three constructs are identical. However, it is noticeable that the remaining peaks, which are related to the N-terminal part of PYP (residue 1-25 of wt-PYP), show slight differences when the N-terminus is extended. Moreover, the extra peaks in 5aa-PYP and 17aa-PYP, which correspond to the additional residues, are located in a cluster of severely overlapped regions (~8.3 ppm), indicating that the additional residues of 5aa-PYP and 17aa-PYP seem to be significantly unstructured.⁴⁶ Based on the results of UV-vis, CD, and HSQC spectra, we conclude that the secondary and tertiary structures of PYP constructs used in this study remain intact and are unaffected by the addition of extra N-terminal residues.

Recovery Times of 5aa-PYP and 17aa-PYP. We conducted two distinct TA experiments to gain a comprehensive understanding of the PYP photocycle's kinetics and spectral changes. The first experiment involved a time-focused measurement, where we measured the TA signal at three wavelengths to analyze the detailed kinetics over time. In contrast, the second experiment employed a spectra-focused measurement, where we measured TA spectra at a relatively coarse number of time points to identify spectral changes for each PYP construct. For the time-focused measurement, we investigated the effect of additional N-terminal residues on the photocycle kinetics by studying the detailed kinetics through temporal profiles of TA signals for wt-PYP, 5aa-PYP, and 17aa-PYP at 380, 465, and 494 nm (Figure 3). The absorption of pB₁ and pB₂ intermediates (blue-shifted intermediates) is evident at 380 nm, while at 494



Figure 3. TA temporal profiles at 380 nm (orange), 465 nm (pale blue), and 494 nm (green) of (a) wt-PYP, (b) 5aa-PYP, (c) 5D-PYP, (d) 5G-PYP, (e) 5K-PYP, (f) 5L-PYP, (g) 5S-PYP, and (h) 17aa-PYP. Black dotted lines on the graph represent the fitting lines for each profile. The experimental curves are well reproduced by the sum of four exponentials. The time constants obtained from the fits are shown in Table 1. The τ_4 corresponding to the dark recovery time from pB₂ to pG, which is indicated in each panel with a vertical blue dotted line, shows noticeable dependence on the length of N-terminal residues.

nm, the TA signal is primarily influenced by the absorption of pR_1 and pR_2 intermediates (red-shifted intermediates). Although there may be some influence from the red-shifted intermediates at 465 nm, this wavelength can also be affected by bleaching, resulting from the decrease of the ground state, pG. By employing these three wavelengths, we monitored the timeresolved changes of the ground state as well as the red-shifted and blue-shifted intermediates. For each construct, the temporal profiles at three wavelengths were simultaneously fit with a sum of four exponential functions (tetra-exponential functions) sharing the common time constants. In previous studies, both the sequential^{37,38} (Figure S1) and the parallel^{1,18,19,39,40} appearances (Figure 1a) of pR_1 and pR_2 were suggested. In this work, we used the parallel model, but we note that both models require the same number of time constants in our time window, which is four, thereby requiring tetra-exponential functions (Supporting Note 1). Therefore, the fitting result does not depend on the kinetic model, and our data can be explained by both models equally well. Here, we focus on the effect of extra N-terminal residues on the time constants. All decay profiles for wt-PYP, 5aa-PYP, and 17aa-PYP are satisfactorily reproduced by tetra-exponential functions, as shown in Figure 3. The time constants determined from the global fitting are summarized in Table 1. It is well established that four intermediates (pR_1, pR_2)

Table 1. Time Constants (τ_1 , τ_2 , τ_3 , and τ_4) Determined v	ia
the Global Fitting of TA Signals at 380, 465, and 494 nm	by
Using the Tetra-Exponential Function ^a	

time constants	$\tau_1 \ (\mu s)$	$\tau_2 \ (\mu s)$	$\tau_3 (ms)$	$ au_4 ({ m ms})$
wt-PYP	6.2 ± 6.4	280 ± 110	1.3 ± 0.6	110 ± 3
5aa-PYP	6.1 ± 6.5	260 ± 58	1.4 ± 0.2	81 ± 2
17aa-PYP	3.1 ± 4.5	210 ± 67	1.5 ± 0.2	55 ± 3
5G-PYP	2.8 ± 1.6	150 ± 15	0.7 ± 0.1	89 ± 2
5L-PYP	3.7 ± 1.5	210 ± 15	1.0 ± 0.1	100 ± 1
5S-PYP	3.8 ± 1.8	210 ± 15	0.9 ± 0.1	95 ± 1
5K-PYP	2.6 ± 1.5	120 ± 14	0.6 ± 0.1	110 ± 2
5D-PYP	4.3 ± 1.8	220 ± 17	1.4 ± 0.1	350 ± 5
<i>a</i> ₁₇₇₁ 1		1 1		

"The errors here account for only random errors, not systematic errors.

pB₁, pB₂) exist within this time interval (after hundreds of nanoseconds), as confirmed by previous experimental studies, including time-resolved spectroscopies, ^{12,17} X-ray Laue crystallography,¹⁹ and X-ray solution scattering (liquidography).¹ The first intermediate, pR₁, is often not well resolved in some experiments, and it can be discerned only when the signal-to-noise ratio is sufficiently high. The decay components of τ_1 , τ_2 , and τ_3 are assigned to the decay process of pR₁, the decay of pR₂,



Figure 4. TA spectra of PYP constructs. (a–h) TA spectra of (a) wt-PYP, (b) 5aa-PYP, (c) 5D-PYP, (d) 5G-PYP, (e) 5K-PYP, (f) 5L-PYP, (g) 5S-PYP, and (h) 17aa-PYP. The contour plots display the changes over time in TA spectra. (i) TA spectra at 100 ns and 1 ms, and 10 ms. Black, red, blue, green, cyan, magenta, dark yellow, and orange indicate TA spectra of wt-PYP, 5aa-PYP, 5D-PYP, 5G-PYP, 5K-PYP, 5L-PYP, 5S-PYP, and 17aa-PYP, respectively.

and the $pB_1 \rightarrow pB_2$ transition, respectively. The decay component of τ_4 corresponds to the dark recovery time from pB_2 to pG. As shown in Table 1, the dark recovery times (τ_4) show a noticeable dependence on the length of N-terminal residues. Especially, 17aa-PYP exhibits a faster dark recovery time (55 ± 3 ms) compared with that of 5aa-PYP (81 ± 2 ms), whose recovery time is also slightly faster than that of wt-PYP (110 ± 3 ms). The variations in τ_4 are significantly greater than the associated errors. On the other hand, those in τ_1 , τ_2 , and τ_3 fall within their respective error margins, and in this work, we did not attempt to find meaningful trends in these three time constants.

For the spectra-focused measurement, we examined TA spectra for wt-PYP, Saa-PYP, and 17aa-PYP to assess the impact of external N-terminal residues on the TA spectral features. The TA spectra for 17aa-PYP at early ($\Delta t = 100 \text{ ns}$) and late delay (>1 ms) times show a notable difference compared to those for wt-PYP (Figure 4). To quantify the differences in spectral shape and kinetics, we employed singular value decomposition (SVD),^{53–55} which decomposes the original data into time-invariant features (left singular vectors, LSVs), their relative

contributions (singular values), and their time profiles (right singular vectors, RSVs). The first two singular values are meaningful for all three proteins (Figure S3a). We emphasize that obtaining only two species in the spectral-focused measurement, in contrast to the four species assigned in the time-focused measurement (as described earlier), is not unexpected. As depicted in Figure 1a and in the time-focused measurement, the photocycle of PYP encompasses four pathways within our time window: $pR_1 \rightarrow pB_1$ (τ_1), $pR_2 \rightarrow$ $pB_1(\tau_2), pB_1 \rightarrow pB_2(\tau_3)$, and $pB_2 \rightarrow pG(\tau_4)$. In the spectralfocused measurement, the $pR_1 \rightarrow pB_1$ process was not well resolved due to the coarse time interval, and the $pB_2 \rightarrow pG$ process was absent due to the limited data available, which did not extend beyond 10 ms. Nevertheless, this measurement allowed us to compare the spectral shapes of the red- and blueshifted intermediates of the PYP constructs. The similarity of two RSVs found in all three proteins suggests that there is little difference in kinetics from 1 ns to 10 ms, which is the time range barely affected by τ_4 , even when the N-terminus is modified. This observation aligns with the kinetic analysis of the temporal profiles, which indicates that τ_1 , τ_2 , and τ_3 are unaffected

Article



Figure 5. (a–h) TG signals of (a) wt-PYP, (b) Saa-PYP, (c) SD-PYP, (d) SG-PYP, (e) SK-PYP, (f) SL-PYP, (g) SS-PYP, and (h) 17aa-PYP. TG signals were measured at various q^2 values for each protein: for wt-PYP and Saa-PYP, $q^2 = 0.7$, 1.2, 2.5, 3.9, and 5.6 × 10¹² m⁻²; for 17aa-PYP, $q^2 = 0.7$, 1.3, 2.7, 4.1, and 4.9 × 10¹² m⁻²; for SD-PYP and SL-PYP, $q^2 = 0.5$, 1.1, 2.1, 3.1, and 4.6 × 10¹² m⁻², respectively; for SG-PYP and SK-PYP, $q^2 = 0.4$, 1.1, 2.0, 4.5, and 7.4 × 10¹² m⁻²; for SS-PYP, $q^2 = 0.4$, 1.0, 1.8, 4.1, and 7.0 × 10¹² m⁻². The experimental curves for each q^2 are color-coded in ascending order of q as follows: red, orange, green, blue, and pale purple. The black dotted lines in all panels indicate the fitting lines obtained using eq 1.

significantly. On the other hand, the LSVs for 17aa-PYP show spectral differences compared with those for wt-PYP (Figure S3b,c). Those spectral differences are evidence that the environmental change near the chromophore during the photocycle of PYP can be affected by extra N-terminal residues.

Recovery Times of 5X-PYPs. The TA data of wt-PYP, 5aa-PYP, and 17aa-PYP revealed that increasing the length of the extra N-terminal residues results in a faster dark recovery. Apparently, the length of the extra N-terminal residues contributes to faster dark recovery. Nevertheless, one cannot rule out the effect of the charge of a specific residue of the extra N-terminal residues. To systematically investigate the charge effect, we prepared additional constructs with five extra residues and conducted the same TA experiments performed for wt-PYP, 5aa-PYP, and 17aa-PYP. Five additional constructs (5D-PYP, 5K-PYP, 5S-PYP, 5L-PYP, and 5G-PYP) were prepared. 5D-PYP has five additional negatively charged aspartic acids, 5K-PYP has five additional positively charged lysines, 5S-PYP has five additional neutral polar serines, 5L-PYP has five additional neutral nonpolar leucines, and 5G-PYP has five additional glycines, which are the smallest neutral amino acids. If the length of the N-terminus was the sole factor determining the dark recovery time, we would expect all 5X-PYPs, which stand for 5D-PYP, 5K-PYP, 5L-PYP, 5S-PYP, and 5G-PYP, to exhibit

similarly faster dark recovery times. The TA time profiles are shown in Figure 3, and the fitted time constants are listed in Table 1.

As shown in Table 1, 5X-PYPs, except for 5D-PYP, show a pattern of slightly faster dark recovery times than that of wt-PYP. This result suggests that additional N-terminal residues can induce a faster dark recovery. In contrast, the presence of negatively charged additional residues in 5D-PYP leads to a markedly slower dark recovery time $(350 \pm 5 \text{ ms})$ compared to the other PYP constructs. On the other hand, 5K-PYP does not show any noticeable difference. These results suggest that the negatively charged residues in the additional N-terminus part may interfere with the dark recovery, whereas the neutral residues play a significantly smaller role, and the positively charged residues have negligible impact. To find the origin of the charge effect, we considered the reported correlation between the dark recovery time and the interaction between the Nterminus and the PYP body. It was suggested that the slower dark recovery time for NT-PYPs is due to the lack of electrostatic interactions between Gly7, Glu9, and Glu12 of the N-terminus and β -scaffold (β 5 and β 6).⁵ In particular, in the crystal structure (PDB 2PHY) of the dark state of wt-PYP, the amide nitrogen of Gly7 and the amide oxygen of Glu12 are bound to His108 and Trp119 via water-mediated hydrogen bonding. The amide

pubs.acs.org/JPCB

fitting parameters	$\tau'_{1}(\mu s)$	$\tau'_{2}(\mu s)$	τ'_{3} (ms)	$D_{\rm pG} \left(10^{-10} \ {\rm m}^2 \ {\rm s}^{-1} \right)$	$D_{\rm pB}~(10^{-10}~{ m m^2~s^{-1}})$	$\Delta f (10^{-12} \text{ kg} \cdot \text{s}^{-1})$
wt-PYP	6.4 ± 6.0	280 ± 49	1.7 ± 0.3	1.29 ± 0.01	1.09 ± 0.01	5.9
5aa-PYP	6.0 ± 16	250 ± 19	0.92 ± 0.15	1.16 ± 0.03	1.06 ± 0.02	3.3
17aa-PYP	1.2 ± 0.4	240 ± 24	0.91 ± 0.04	1.16 ± 0.03	1.05 ± 0.07	3.7
5G-PYP	2.4 ± 0.4	220 ± 22	1.2 ± 0.1	1.26 ± 0.01	1.20 ± 0.01	1.6
5L-PYP	1.2 ± 0.2	180 ± 12	1.1 ± 0.1	1.34 ± 0.02	1.16 ± 0.04	4.7
5S-PYP	1.6 ± 0.1	220 ± 2	1.0 ± 0.1	1.32 ± 0.01	1.06 ± 0.02	7.5
5K-PYP	2.0 ± 0.2	190 ± 7	1.1 ± 0.1	1.12 ± 0.01	1.07 ± 0.01	1.7
5D-PYP	1.3 ± 0.2	190 ± 10	0.79 ± 0.27	1.25 ± 0.03	1.04 ± 0.02	6.5
For each PYP constru	ict, the TG signa	ls at various a^2 (a	· orating wavenum	per) are fitted simultaneo	usly, τ'_1 , τ'_2 , τ'_2 , D_{-C} , D_{-C}	$_{\rm p}$, and Δf values were

For each PYP construct, the TG signals at various q^2 (q: grating wavenumber) are fitted simultaneously. τ'_1 , τ'_2 , τ'_3 , D_{pG} , D_{pB} , and Δf values were determined from the fitting of TG signals using eq 1. The errors here account for only random errors, not systematic errors.

oxygen of Glu9 is directly hydrogen-bonded to Lys110. Additionally, the side chains of Glu12 are bound to Lys110 by a salt bridge.⁵ It was suggested that in some species, Glu12 is substituted by Asp, which maintains its electrostatic properties.⁵ Since the aspartic acid in the additional N-terminal part of 5D-PYP is also negatively charged as the glutamic acid, the interaction between the additional sequence (DDDDD) in the N-terminus and the β -scaffold may weaken or compete with the electrostatic interaction between the original N-terminus and β scaffold during the dark recovery process. The interaction between the additional 5D residue and Lys123 can also weaken the interaction between Phe6 and Lys123. Any disruption or competition resulting from the addition of the negatively charged residues may ultimately cause a slower dark recovery process in 5D-PYP compared to that in wt-PYP. We observe that the extended residues in 17aa-PYP, associated with the fastest dark recovery, include four repetitive D residues, one K residue, and six repetitive H residues. The expedited recovery in 17aa-PYP, despite the presence of four negatively charged residues (D) and seven positively charged residues (K and H), suggests the involvement of more intricate interactions in 17aa-PYP, possibly influenced by the relatively long N-terminus with high flexibility in the pB₂ state.

MD Simulation. Regarding the interaction between the additional sequence (DDDDD) in the N-terminus and the β sheets, we calculated the surface charge distribution for wt-PYP in both the pG and pB₂ states. Furthermore, we performed molecular dynamics (MD) simulations for the following four conditions and calculated root-mean-square fluctuation (RMSF) of protein moiety and the center of mass (COM) distance between the N-terminus and the β -sheets from the MD snapshots of these simulations: (i) pG state of wt-PYP, (ii) pB₂ state of wt-PYP, (iii) pG state of 5D-PYP, and (iv) pB₂ state of 5D-PYP (details are described in the Methods Secton). The surface charge distributions of the pG and pB2 states exhibit positive charge within the β -sheets while showing negative charge in the rest of the structures (Figure S4). For the pG state, the RMSF plot of 5D-PYP shows decreased values compared to those of wt-PYP in the N-terminal region, while the distributions of the distances between the N-terminus and the β -sheets are similar for wt-PYP and 5D-PYP. For the pB₂ state, the RMSF plot of 5D-PYP displays decreased values compared to those of wt-PYP in the N-terminal region, similar to the case in the pG state (Figure S5c). Unlike the similar distributions of the distances between the N-terminus and the β -sheets observed for the pG state, the distance distributions of pB_2 show a significant decrease in 5D-PYP compared to wt-PYP (Figure S5d). The results from MD simulations demonstrate that in the pB₂ state of 5D-PYP, the N-terminus can form additional interaction with β - sheets, whereas in the pG state, it does not. Considering that the side chains of the five aspartic acid residues are negatively charged under our experimental condition, the positive charge within the β -sheets of the pB₂ state indicates that the interaction might be driven by the electrostatic properties of the N-terminus and β -sheets (details are described in the Supporting Note 4).

TG Signals of 5aa-PYP, 5X-PYPs, and 17aa-PYP. To obtain information about the structures of the pG and pB₂ species of NE-PYPs, we measured the TG signals of NE-PYPs. The diffusion coefficient (D) of a chemical species determined from the TG experiment can provide information about the size and shape of a transient species formed during the photocycle of PYP. Figure 5 shows TG signals of wt-PYP, 5aa-PYP, 5X-PYPs, and 17aa-PYP at various grating wavenumbers, *q*. The features of TG signals of 5aa-PYP, *SX*-PYPs, and 17aa-PYP as well as wt-PYP are almost the same as those of wt-PYP reported in several previous studies.^{7,14–17} To compare the TG result in this study with the result in previous studies on the photocycle reaction of PYP, all TG signals were fitted using the following function

$$\begin{split} I_{\rm TG} &= \left[\delta n_1 \exp\left(-\frac{t}{\tau_1}\right) + \delta n_{\rm th} \exp(-D_{\rm th}q^2 t) \\ &+ \delta n_2 \exp\left(-\left(\frac{1}{\tau_2} + D_{\rm pG}q^2\right)t\right) \\ &+ \delta n_3 \exp\left(-\left(\frac{1}{\tau_3} + D_{\rm pB}q^2\right)t\right) \\ &- \delta n_{\rm pG} \exp(-D_{\rm pG}q^2 t) + \delta n_{\rm pB} \exp(-D_{\rm pB}q^2 t) \right]^2 \end{split}$$
(1)

where δn_1 , δn_2 , δn_3 , $\delta n_{\rm th}$, $\delta n_{\rm pG}$, and $\delta n_{\rm pB}$ are refractive index changes corresponding to exponential decays, τ'_1 , τ'_2 , and τ'_3 are time constants of each exponential decay, $D_{\rm th}$ is a thermal diffusivity of the solvent, and $D_{\rm pG}$ and $D_{\rm pB}$ are diffusion coefficients of pG and pB₂ states of PYP, respectively. Since only k_1 (1/ τ'_1) is independent of q values, it is set as a shared parameter for every q value. More details about TG techniques are explained in Supporting Note 2.

In eq 1, the contribution due to the slow dark recovery from pB₂ to pG in the diffusion peak was not considered because the dark recovery times for wt-PYP, 5aa-PYP, 5X-PYPs, and 17aa-PYP are significantly longer than the diffusion times of pB₂ and pG. Indeed, as shown in Figure 5, all TG signals of each construct at various q^2 are well-fitted by eq 1. Table 2 shows the fit results for all of the TG signals. The three time constants (τ'_1 , τ'_2 , and τ'_3) are assigned to the decay of pR₁, the decay of pR₂,

and the $pB_1 \rightarrow pB_2$ transition, respectively. The predominant time constants closely match those measured by TA experiments with some exceptions noted. For instance, τ'_3 for 5D-PYP, as determined with TG (Table 2), is twice as fast as compared to wt-PYP, whereas τ_3 for 5D-PYP, as determined with TA (Table 1), is comparable to that for wt-PYP. These discrepancies between TA and TG results likely stem from the distinct observables used in each method, which are the absorbance changes of the chromophore and the structural and diffusion changes of the protein, respectively. The time constants for the $pR_2 \rightarrow pB_1$ transition are not significantly affected by the Nterminal extension, whereas it was reported that they became smaller by N-terminal truncation.⁷ This difference between NE-PYPs and NT-PYPs suggests that the extension of the Nterminus mostly affects the pB₂ state, while truncation of the Nterminus influences the whole photocycle of PYP.

On the other hand, the determined D_{pG} values for Saa-PYP, SK-PYP, and 17aa-PYP are slightly smaller than those for wt-PYP, whereas 5D-PYP, 5G-PYP, 5L-PYP, and 5S-PYP have almost the same D_{pG} values as wt-PYP within experimental errors. These results suggest that 5aa-PYP, SK-PYP, and 17aa-PYP have a large molecular volume and/or different shapes compared to wt-PYP, while SD-PYP, 5G-PYP, SL-PYP, and SS-PYP have a similar molecular volume and/or shape compared to wt-PYP. According to the Stokes–Einstein relationship, the diffusion coefficient is related to the molecular weight of a molecule (eq 2, details in Supporting Note 3)

$$\frac{D_{\rm wt}}{D_{\rm NE}} = \frac{M_{\rm NE}^{1/3}}{M_{\rm wt}^{1/3}} \tag{2}$$

Here, $D_{\rm wt}$ and $D_{\rm NE}$ are the diffusion coefficients corresponding to a certain state of wt-PYP and NE-PYP, respectively, and $M_{\rm wt}$ and $M_{\rm NE}$ are the molecular weights of wt-PYP and NE-PYP, respectively. The D_{pG} values of 5aa-PYP, 5K-PYP, and 17aa-PYP calculated using eq 2 are similar to those of wt-PYP (1.27, 1.26, 1.22, and 1.29×10^{-10} m² s⁻¹ for 5aa-PYP, 5K-PYP, 17aa-PYP, and wt-PYP respectively), meaning that the extension of Nterminus in the dark state does not significantly affect a diffusion coefficient of a protein. In contrast to the prediction based on the molecular weight, the TG experiments for 5aa-PYP, 5K-PYP, and 17aa-PYP show that the measured D_{pG} values of these constructs are slightly smaller than those of wt-PYP (Table 2). This observation suggests that the molecular volumes of 5aa-PYP, 5K-PYP, and 17aa-PYP are slightly larger than those of wt-PYP. In wt-PYP, the N-terminal domain interacts with the β scaffold (which is part of the chromophore-containing core domain of the protein) via electrostatic interaction.5,43,48 Therefore, if the extension of the N-terminus induces a change in the interaction between the N-terminus and β -scaffold linked to the chromophore-containing domain of the protein, this can lead to a change in the molecular volume and/or shape of PYP. In this regard, the smaller D_{pG} values of 5aa-PYP, 5K-PYP, and 17aa-PYP are probably caused by the interaction between the extended N-terminus and the β -scaffold. Especially, 5K-PYP exhibits a much smaller D_{pG} value than other 5X-PYPs, which suggests that water-mediated salt bridges may have formed between the five lysines and the PYP body in the pG state. Lysine (K) is a positively charged amino acid that often forms watermediated salt bridges with negatively charged amino acids in a protein.⁴⁷ The surface charge distribution of PYP calculated using pdb2pqr and APBS program⁵² shows a relatively high negative charged surface (Figure S4). The additional five lysine

residues in 5K-PYP are likely to form water-mediated salt bridges with the negatively charged amino acid residues on the surface of PYP, ultimately leading to an increase in the volume of 5K-PYP and a decrease in D_{pG} . While typical hydrophobic interactions have distances of 3.3-4.0 Å,⁴⁹ water-mediated salt bridges are known to interact at distances of up to 2 nm.⁵⁰ Moreover, water molecules tightly incorporated into the protein framework are considered to be part of the protein structure.⁵¹ As a result, water-mediated salt bridges can increase the apparent volume of the protein. In the case of 17aa-PYP, there are a total of 7 positively charged amino acid residues (six histidines and one lysine). Even if there are four negatively charged amino acid residues, the seven positively charged amino acid residues would still be able to form salt bridges with the surface of 17aa-PYP, thereby resulting in an increased volume of the protein. The relatively small D_{pG} of 5aa-PYP can be attributed to the unique proline structure contained in the additional five amino acid residues, which can restrict the free movement of these residues and lead to an irregular shape, ultimately causing increased friction. Other 5X-PYPs that contain neutral extra amino acids may not have any significant interactions in the pG state, which results in experimentally determined D_{pG} values similar to the value obtained using eq 2.

Upon examination of the diffusion coefficient of $pB_2(D_{pB})$ of NE-PYPs, the extra amino acid residues with neutral uncharged side chains in 5G-PYP and 5L-PYP have D_{pB} values that are consistently slightly larger than that of wt-PYP. This suggests that their volumes may be smaller than that of pB₂ of wt-PYP, or their shape is more spherical. In a previous study of PYP that used double electron-electron resonance spectroscopy (DEER), NMR, and time-resolved pump-probe X-ray solution scattering (TR-SAXS/WAXS), it was suggested that the I₂' state, which corresponds to the pB2 state, has a relatively well-ordered structure, and in this state, part of the N-terminal 20 residues contacts with the chromophore-binding cleft.² The additional N-terminal residues of 5G-PYP and 5L-PYP might affect the interaction between the N-terminal residues and the chromophore-binding cleft and make a new interaction between the extended neutral N-terminus and the other hydrophobic core side in the pB_2 state, thereby making the pB_2 states of 5G-PYP and 5L-PYP compact. In contrast, the rest of the constructs have smaller D_{pB} compared to the wt-PYP, indicating that they have become larger and/or more irregular in shape. Previous work has shown that slow diffusion in NT-PYPs is caused by the unfolding of two α -helixes in the N-terminal region, increasing intermolecular interactions due to hydrogen bonding with water molecules.⁵ In line with this, a plausible explanation is that highly charged amino acids facilitate the formation of hydrogen bonds with water molecules. The additional aspartic acid residues in 5D-PYP have the potential to form water-mediated salt bridges with positively charged amino acid residues, which are hidden within the protein in the pG state but may become exposed to the bulk solvent in the partially unfolded pB₂ state. Such salt bridges may have the effect of increasing the overall volume. A similar process may be operational also in 5S-PYP, as the additional serine residues of 5S-PYP also can form watermediated salt bridges like 5D-PYP. 5K-PYP is likely to possess salt bridges between five lysines and the PYP body in the pG state, and these linkages are expected to be retained in pB_{2} , leading to the smaller D_{pB} of 5K-PYP. The relatively small D_{pB} of 5aa-PYP can be explained with the same logic used to explain its relatively small D_{pG} , that is, the unique proline structure restricting the free movement of these residues.

Khan et al. showed that the partial unfolding of α -helices of the N-terminal region could be interpreted by the change of the friction coefficients $(\Delta f = k_{\rm B}T(1/D_{\rm pB} - 1/D_{\rm pG}))$. The change in Δf can serve as an indicator of changes in friction resulting from the transition of the protein from its ground state with D_{pG} to a partially unfolded putative signaling state with $D_{\rm pB}$. Based on the obtained D_{pG} and D_{pB} , the Δf values for wt-PYP, 5aa-PYP, 5X-PYPs, and 17aa-PYP were calculated (Table 2). 5G-PYP (1.6) and 5K-PYP (1.7) have much smaller Δf values than wt-PYP (5.9), implying that the pB_2 state of 5G-PYP and 5K-PYP undergoes less structural change than that of wt-PYP. On the other hand, 5D-PYP and 5S-PYP have much larger Δf values than wt-PYP, suggesting that the pB₂ state of 5D-PYP and 5S-PYP undergoes larger structural changes than wt-PYP. Other NE-PYPs have similar Δf values. These results indicate that the degree of structural changes can be controlled by adjusting specific amino acid residues in the extended N-terminus of NE-PYPs. The plot of the dark recovery time versus Δf (Figure 6),



Figure 6. Dark recovery time vs Δf for various PYP constructs studied in this work. The plot suggests that the dark recovery times of wt-PYP and NE-PYPs, except for SD-PYP, are independent of the structural differences between pG and pB₂. The red dashed line represents the average τ_4 (~92 ms) value of the constructs, excluding SD-PYP.

furthermore, shows that the dark recovery time of PYP does not correlate with structural differences between pG and pB₂. This result suggests that the dark recovery times of wt-PYP and NE-PYPs, except for 5D-PYP, are independent of the structural differences between pG and pB₂.

CONCLUSIONS

This study demonstrates that the extension of the N-terminus of PYP, except for negatively charged residues, shortens the dark recovery time in proportion to the length of the extension, contrary to the N-terminal truncated PYPs (NT-PYPs), showing slower dark recovery time than wt-PYP.^{3,4} The extension of the N-terminus appears to have a relatively smaller influence on the dark recovery time than that of the truncation of the N-terminus. The acceleration of dark recovery for 5aa-PYP, 5G-PYP, SL-PYP, 5S-PYP, and 17aa-PYP may be due to the increased interaction between the extended N-terminus and β -scaffold in the pG state. Notably, 5D-PYP with additional negatively charged residues exhibits several times slower dark recovery than

wt-PYP, probably because the extra N-terminal five aspartic acid residues may interfere with the recovery from pB₂ to pG state through interaction with β -scaffold. This indicates that the additional negatively charged residues have a greater effect on kinetics than the length of the extension has. Additionally, the TG data show that 5G-PYP and 5K-PYP have a much smaller Δf value than that of wt-PYP, implying that the conformational change upon the formation of pB₂ in these NE-PYPs is much smaller than that in wt-PYP. 5D-PYP and 5S-PYP have a larger Δf value than wt-PYP, implying that the conformational change upon the formation of pB₂ in these NE-PYPs is much larger than that in wt-PYP. These findings help us better understand the function of proteins that have PYP modules or domains. Our results for NE-PYPs and the previous results for NT-PYPs together show that both the lifetime of the pB₂ state and the magnitude of the structural change can be regulated by controlling the length and the sequence of N-terminal residues. This provides new insights into controlling the photocycle of photoactive proteins, whose photocycles are influenced by interactions between the N-terminal and the protein core. Moreover, this study demonstrates that leaving certain residues uncut during the production of recombinant PYP can impact its photocycle and how extending the N-terminus affects the function and properties of PYP. This study suggests that the existence of residual amino acid residues during recombinant protein production can affect their functionality, emphasizing the need for careful consideration of such residues in the general production of recombinant proteins.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jpcb.3c03841.

Details about theoretical backgrounds and MD simulation; schematic figure of the PYP sequential photocycle, detailed assignment of ¹H, ¹⁵N-HSQC spectra, SVD results of TA spectra, the calculated surface charge distribution of PYP, and the result from MD simulations; and calculated contents of the secondary structure element (PDF)

AUTHOR INFORMATION

Corresponding Author

Hyotcherl Ihee – Center for Advanced Reaction Dynamics, Institute for Basic Science (IBS), Daejeon 34141, Republic of Korea; Department of Chemistry and KI for the BioCentury, KAIST, Daejeon 34141, Republic of Korea; orcid.org/ 0000-0003-0397-5965; Email: hyotcherl.ihee@kaist.ac.kr

Authors

- **Cheolhee Yang** Center for Advanced Reaction Dynamics, Institute for Basic Science (IBS), Daejeon 34141, Republic of Korea
- Youngmin Kim Center for Advanced Reaction Dynamics, Institute for Basic Science (IBS), Daejeon 34141, Republic of Korea
- Seong Ok Kim Center for Advanced Reaction Dynamics, Institute for Basic Science (IBS), Daejeon 34141, Republic of Korea; Department of Chemistry and KI for the BioCentury, KAIST, Daejeon 34141, Republic of Korea
- Sang Jin Lee Center for Advanced Reaction Dynamics, Institute for Basic Science (IBS), Daejeon 34141, Republic of

Korea; Department of Chemistry and KI for the BioCentury, KAIST, Daejeon 34141, Republic of Korea

Jungkweon Choi – Center for Advanced Reaction Dynamics, Institute for Basic Science (IBS), Daejeon 34141, Republic of Korea; Department of Chemistry and KI for the BioCentury, KAIST, Daejeon 34141, Republic of Korea

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jpcb.3c03841

Author Contributions

[§]C.Y. and Y.K. contributed equally to this work.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work was supported by the Institute for Basic Science (IBS-R033).

REFERENCES

(1) Kim, T. W.; Lee, J. H.; Choi, J.; Kim, K. H.; van Wilderen, L. J.; Guerin, L.; Kim, Y.; Jung, Y. O.; Yang, C.; Kim, J.; et al. Protein Structural Dynamics of Photoactive Yellow Protein in Solution Revealed by Pump-Probe X-ray Solution Scattering. *J. Am. Chem. Soc.* **2012**, *134* (6), 3145–3153.

(2) Ramachandran, P. L.; Lovett, J. E.; Carl, P. J.; Cammarata, M.; Lee, J. H.; Jung, Y. O.; Ihee, H.; Timmel, C. R.; van Thor, J. J. The Short-Lived Signaling State of the Photoactive Yellow Protein Photoreceptor Revealed by Combined Structural Probes. *J. Am. Chem. Soc.* **2011**, *133* (24), 9395–9404.

(3) Harigai, M.; Yasuda, S.; Imamoto, Y.; Yoshihara, F.; Tokunaga, F.; Kataoka, M. Amino Acids in the N-Terminal Region Regulate the Photocycle of Photoactive Yellow Protein. *J. Biochem.* **2001**, *130* (1), 51–56.

(4) van der Horst, M. A.; van Stokkum, I. H.; Crielaard, W.; Hellingwerf, K. J. The Role of the N-Terminal Domain of Photoactive Yellow Protein in the Transient Partial Unfolding during Signalling State Formation. *FEBS Lett.* **2001**, 497 (1), 26–30.

(5) Harigai, M.; Imamoto, Y.; Kamikubo, H.; Yamazaki, Y.; Kataoka, M. Role of an N-Terminal Loop in the Secondary Structural Change of Photoactive Yellow Protein. *Biochemistry* **2003**, *42* (47), 13893–13900.

(6) 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. The Solution Structure of a Transient Photoreceptor Intermediate: $\Delta 25$ Photoactive Yellow Protein. *Structure* **2005**, *13* (7), 953–962.

(7) Khan, J. S.; Imamoto, Y.; Harigai, M.; Kataoka, M.; Terazima, M. Conformational Changes of PYP Monitored by Diffusion Coefficient: Effect of N-Terminal α -Helices. *Biophys. J.* **2006**, *90* (10), 3686–3693.

(8) Meyer, T. E.; Yakali, E.; Cusanovich, M. A.; Tollin, G. Properties of a Water-Soluble, Yellow Protein Isolated from a Halophilic Phototrophic Bacterium That Has Photochemical Activity Analogous to Sensory Rhodopsin. *Biochemistry* **1987**, *26* (2), 418–423.

(9) Hoff, W. D.; Vanstokkum, I. H. M.; Vanramesdonk, H. J.; Vanbrederode, M. E.; Brouwer, A. M.; Fitch, J. C.; Meyer, T. E.; Vangrondelle, R.; Hellingwerf, K. J. Measurement and Global Analysis of the Absorbance Changes in the Photocycle of the Photoactive Yellow Protein from Ectothiorhodospira halophila. *Biophys. J.* **1994**, *67* (4), 1691–1705.

(10) Meyer, T. E.; Tollin, G.; Hazzard, J. H.; Cusanovich, M. A. Photoactive Yellow Protein from the Purple Phototrophic Bacterium, Ectothiorhodospira halophila - Quantum Yield of Photobleaching and Effects of Temperature, Alcohols, Glycerol, and Sucrose on Kinetics of Photobleaching and Recovery. *Biophys. J.* **1989**, *56* (3), 559–564.

(11) Hendriks, J.; van Stokkum, I. H. M.; Hellingwerf, K. J. Deuterium Isotope Effects in the Photocycle Transitions of the Photoactive Yellow Protein. *Biophys. J.* **2003**, *84* (2), 1180–1191. (12) Yeremenko, S.; van Stokkum, I. H. M.; Moffat, K.; Hellingwerf, K. J. Influence of the Crystalline State on Photoinduced Dynamics of Photoactive Yellow Protein Studied by Ultraviolet-Visible Transient Absorption Spectroscopy. *Biophys. J.* **2006**, *90* (11), 4224–4235.

(13) Borucki, B.; Otto, H.; Joshi, C. P.; Gasperi, C.; Cusanovich, M. A.; Devanathan, S.; Tollin, G.; Heyn, M. P. pH Dependence of the Photocycle Kinetics of the E46Q Mutant of Photoactive Yellow Protein: Protonation Equilibrium between I_1 and I_2 Intermediates, Chromophore Deprotonation by Hydroxyl Uptake, and Protonation Relaxation of the Dark State. *Biochemistry* **2003**, *42* (29), 8780–8790.

(14) Takeshita, K.; Imamoto, Y.; Kataoka, M.; Tokunaga, F.; Terazima, M. Themodynamic and Transport Properties of Intermediate States of the Photocyclic Reaction of Photoactive Yellow Protein. *Biochemistry* **2002**, *41* (9), 3037–3048.

(15) Takeshita, K.; Imamoto, Y.; Kataoka, M.; Mihara, K.; Tokunaga, F.; Terazima, M. Structural Change of Site-Directed Mutants of PYP: New Dynamics during pR State. *Biophys. J.* **2002**, *83* (3), 1567–1577.

(16) Khan, J. S.; Imamoto, Y.; Yamazaki, Y.; Kataoka, M.; Tokunaga, F.; Terazima, M. A Biosensor in the Time Domain Based on the Diffusion Coefficient Measurement: Intermolecular Interaction of an Intermediate of Photoactive Yellow Protein. *Anal. Chem.* **2005**, 77 (20), 6625–6629.

(17) Hoshihara, Y.; Imamoto, Y.; Kataoka, M.; Tokunaga, F.; Terazima, M. Conformational Changes in the N-Terminal Region of Photoactive Yellow Protein: A Time-Resolved Diffusion Study. *Biophys. J.* **2008**, *94* (6), 2187–2193.

(18) Jung, Y. O.; Lee, J. H.; Kim, J.; Schmidt, M.; Moffat, K.; Srajer, V.; Ihee, H. Volume-Conserving Trans-Cis Isomerization Pathways in Photoactive Yellow Protein Visualized by Picosecond X-Ray Crystallography. *Nat. Chem.* **2013**, *5* (3), 212–220.

(19) Ihee, H.; Rajagopal, S.; Srajer, V.; Pahl, R.; Anderson, S.; Schmidt, M.; Schotte, F.; Anfinrud, P. A.; Wulff, M.; Moffat, K. Visualizing Reaction Pathways in Photoactive Yellow Protein from Nanoseconds to Seconds. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102* (20), 7145–7150.

(20) Genick, U. K.; Borgstahl, G. E. O.; Ng, K.; Ren, Z.; Pradervand, C.; Burke, P. M.; Srajer, V.; Teng, T. Y.; Schildkamp, W.; McRee, D. E.; et al. Structure of a Protein Photocycle Intermediate by Millisecond Time-Resolved Crystallography. *Science* **1997**, *275* (5305), 1471–1475.

(21) Xie, A. H.; Kelemen, L.; Hendriks, J.; White, B. J.; Hellingwerf, K. J.; Hoff, W. D. Formation of a New Buried Charge Drives a Large-Amplitude Protein Quake in Photoreceptor Activation. *Biochemistry* **2001**, *40* (6), 1510–1517.

(22) Brudler, R.; Rammelsberg, R.; Woo, T. T.; Getzoff, E. D.; Gerwert, K. Structure of the I₁ Early Intermediate of Photoactive Yellow Protein by FTIR Spectroscopy. *Nat. Struct. Biol.* **2001**, *8* (3), 265–270.

(23) Mizuno, M.; Kamikubo, H.; Kataoka, M.; Mizutani, Y. Changes in the Hydrogen-Bond Network around the Chromophore of Photoactive Yellow Protein in the Ground and Excited States. *J. Phys. Chem. B* **2011**, *115* (29), 9306–9310.

(24) Imamoto, Y.; Kamikubo, H.; Harigai, M.; Shimizu, N.; Kataoka, M. Light-Induced Global Conformational Change of Photoactive Yellow Protein in Solution. *Biochemistry* **2002**, *41* (46), 13595–13601.

(25) Vreede, J.; van der Horst, M. A.; Hellingwerf, K. J.; Crielaard, W.; van Aalten, D. M. F. PAS Domains - Common Structure and Common Flexibility. *J. Biol. Chem.* **2003**, 278 (20), 18434–18439.

(26) Woestenenk, E. A.; Hammarstrom, M.; van den Berg, S.; Hard, T.; Berglund, H. His Tag Effect on Solubility of Human Proteins Produced in Escherichia coli: a Comparison between Four Expression Vectors. *J. Struct. Funct. Genomics* **2004**, *5* (3), 217–229.

(27) Xu, C. G.; Fan, X. J.; Fu, Y. J.; Liang, A. H. Effect of Location of the His-Tag on the Production of Soluble and Functional Buthus martensii Karsch Insect Toxin. *Protein Expression Purif.* **2008**, *59* (1), 103–109.

(28) Calidas, D.; Lyon, H.; Culver, G. M. The N-Terminal Extension of S12 Influences Small Ribosomal Subunit Assembly in Escherichia coli. RNA **2014**, *20* (3), 321–330.

(29) Song, L.; Dumon, C.; Siguier, B.; Andre, I.; Eneyskaya, E.; Kulminskaya, A.; Bozonnet, S.; O'Donohue, M. J. Impact of an N-Terminal Extension on the Stability and Activity of the GH11 Xylanase from Thermobacillus xylanilyticus. *J. Biotechnol.* **2014**, *174*, 64–72.

(30) Chant, A.; Kraemer-Pecore, C. M.; Watkin, R.; Kneale, G. G. Attachment of a Histidine Tag to the Minimal Zinc Finger Protein of the Aspergillus nidulans Gene Regulatory Protein AreA Causes a Conformational Change at the DNA-Binding Site. *Protein Expression Purif.* **2005**, 39 (2), 152–159.

(31) Tajika, Y.; Sakai, N.; Tamura, T.; Yao, M.; Watanabe, N.; Tanaka, I. Crystal Structure of Hypothetical Protein PH0828 from Pyrococcus horikoshii. *Proteins* **2004**, *57* (4), 862–865.

(32) McGonigle, R.; Yap, W. B.; Ong, S. T.; Gatherer, D.; Bakker, S. E.; Tan, W. S.; Bhella, D. An N-Terminal Extension to the Hepatitis B Virus Core Protein Forms a Poorly Ordered Trimeric Spike in Assembled Virus-Like Particles. *J. Struct. Biol.* **2015**, *189* (2), 73–80.

(33) Raz, Y.; Adler, J.; Vogel, Á.; Scheidt, H. A.; Haupl, T.; Abel, B.; Huster, D.; Miller, Y. The Influence of the Δ K280 Mutation and N- or C-Terminal Extensions on the Structure, Dynamics, and Fibril Morphology of the Tau R2 Repeat. *Phys. Chem. Chem. Phys.* **2014**, *16* (17), 7710–7717.

(34) Korepanova, A.; Douglas, C.; Leyngold, I.; Logan, T. M. N-Terminal Extension Changes the Folding Mechanism of the FK506-Binding Protein. *Protein Sci.* **2001**, *10* (9), 1905–1910.

(35) Losi, A.; Wegener, A. A.; Engelhard, M.; Gartner, W.; Braslavsky, S. E. Time-Resolved Absorption and Photothermal Measurements with Recombinant Sensory Rhodopsin II from Natronobacterium pharaonis. *Biophys. J.* **1999**, 77 (6), 3277–3286.

(36) Pool, T. J.; Oktaviani, N. A.; Kamikubo, H.; Kataoka, M.; Mulder, F. A. A. ¹H, ¹³C, and ¹⁵N Resonance Assignment of Photoactive Yellow Protein. *Biomol. NMR Assignments* **2013**, *7* (1), 97–100.

(37) Chen, E. F.; Gensch, T.; Gross, A. B.; Hendriks, J.; Hellingwerf, K. J.; Kliger, D. S. Dynamics of Protein and Chromophore Structural Changes in the Photocycle of Photoactive Yellow Protein Monitored by Time-Resolved Optical Rotatory Dispersion. *Biochemistry* **2003**, *42* (7), 2062–2071.

(38) Losi, A.; Gensch, T.; van der Horst, M. A.; Hellingwerf, K. J.; Braslavsky, S. E. Hydrogen-Bond Network Probed by Time-Resolved Optoacoustic Spectroscopy: Photoactive Yellow Protein and the Effect of E46Q and E46A Mutations. *Phys. Chem. Chem. Phys.* **2005**, 7 (10), 2229–2236.

(39) Anderson, S.; Srajer, V.; Moffat, K. Structural Heterogeneity of Cryotrapped Intermediates in the Bacterial Blue Light Photoreceptor, Photoactive Yellow Protein. *Photochem. Photobiol.* **2004**, *80* (1), 7–14.

(40) Rajagopal, S.; Anderson, S.; Srajer, V.; Schmidt, M.; Pahl, R.; Moffat, K. A Structural Pathway for Signaling in the E46Q Mutant of Photoactive Yellow Protein. *Structure* **2005**, *13* (1), 55–63.

(41) Morgan, S. A.; Al-Abdul-Wahid, S.; Woolley, G. A. Structure-Based Design of a Photocontrolled DNA Binding Protein. *J. Mol. Biol.* **2010**, *399* (1), 94–112.

(42) Ui, M.; Tanaka, Y.; Araki, Y.; Wada, T.; Takei, T.; Tsumoto, K.; Endo, S.; Kinbara, K. Application of Photoactive Yellow Protein as a Photoresponsive Module for Controlling Hemolytic Activity of Staphylococcal α -Hemolysin. *Chem. Commun.* **2012**, 48 (39), 4737– 4739.

(43) Borgstahl, G. E. O.; Williams, D. R.; Getzoff, E. D. 1.4 Å Structure of Photoactive Yellow Protein, a Cytosolic Photoreceptor - Unusual Fold, Active-Site, and Chromophore. *Biochemistry* **1995**, 34 (19), 6278–6287.

(44) Kim, Y.; Yang, C.; Kim, T. W.; Thamilselvan, K.; Kim, Y.; Ihee, H. Regulation of Protein Structural Changes by Incorporation of a Small-Molecule Linker. *Int. J. Mol. Sci.* **2018**, *19*, No. 3714, DOI: 10.3390/ ijms19123714.

(45) Lee, S. J.; Kim, Y.; Kim, T. W.; Yang, C.; Thamilselvan, K.; Jeong, H.; Hyun, J.; Ihee, H. Reversible molecular motional switch based on circular photoactive protein oligomers exhibits unexpected photo-induced contraction. *Cell Rep. Phys. Sci.* **2021**, *2*, No. 100512.

(46) Wishart, D. S.; Bigam, C. G.; Holm, A.; Hodges, R. S.; Sykes, B. D. ¹H, ¹³C and random coil NMR chemical shifts of the common amino acids. *J. Biomol. NMR* **1995**, *5* (1), 67–81.

(47) Sabarinathan, R.; Aishwarya, K.; Sarani, R.; Vaishnavi, M. K.; Sekar, K. Water-mediated ionic interactions in protein structures. *J. Biosci.* **2011**, *36* (2), 253–263.

(48) Harigai, M.; Kataoka, M.; Imamoto, Y. Interaction between Nterminal loop and beta-scaffold of photoactive yellow protein. *Photochem. Photobiol.* **2008**, *84* (4), 1031–1037.

(49) Onofrio, A.; Parisi, G.; Punzi, G.; Todisco, S.; Di Noia, M. A.; Bossis, F.; Turi, A.; De Grassi, A.; Pierri, C. L. Distance-depending hydrophobic-hydrophobic contacts in protein folding simulations. *Phys. Chem. Chem. Phys.* **2014**, *16*, 18907–18917.

(50) Kuffel, A.; Zielkiewicz, J. Water-mediated long-range interactions between the internal vibrations of remote proteins. *Phys. Chem. Chem. Phys.* **2015**, *17*, 6728–6733.

(51) Papoian, G. A.; Ulander, J.; Wolynes, P. G. Role of water mediated interactions in protein-protein recognition landscapes. *J. Am. Chem. Soc.* **2003**, *125* (30), 9170–9178.

(52) Jurrus, E.; Engel, D.; Star, K.; Monson, K.; Brandi, J.; Felberg, L. E.; Brookes, D. H.; Wilson, L.; Chen, J.; Liles, K.; et al. Improvements to the APBS biomolecular solvation software suite. *Protein Sci.* **2018**, 27 (1), 112–128.

(53) Golub, G. H.; Reinsch, C. Singular value decomposition and least squares solutions. *Numer. Math.* **1970**, *14*, 403–420.

(54) Schmidt, M.; Rajagopal, S.; Ren, Z.; Moffat, K. Application of singular value decomposition to the analysis of time-resolved macromolecular x-ray data. *Biophys. J.* **2003**, *84* (3), 2112–2129.

(55) Jung, Y. O.; Lee, J. H.; Kim, J.; Schmidt, M.; Moffat, K.; Srajer, V.; Ihee, H. Volume-conserving trans-cis isomerization pathways in photoactive yellow protein visualized by picosecond X-ray crystallography. *Nat. Chem.* **2013**, *5* (3), 212–220.