Supporting Information Appendix

Dynamics and mechanism of ultrafast water-protein interactions

Department of Physics, Department of Chemistry and Biochemistry, and Programs of Biophysics, Chemical Physics, and Biochemistry, The Ohio State University, Columbus, Ohio, 43210

Yangzhong Qin, Lijuan Wang, Dongping Zhong*

SI Note 1. Data analyses

The methodology of detailed data analyses can be found elsewhere (16). Briefly, we measure the solvation dynamics by defining the solvation correlation function as the following:

$$C(t) = \frac{\nu_s(t) - \nu_l(t)}{\nu_s(0) - \nu_l(0)} \qquad \text{(S1)}$$

where ν_s is the overall emission peak including the solvation dynamics and population decay contributions. ν_l is the total lifetime emission peak due to the multiple lifetimes with different emission peaks. In one way, the fluorescence transients at certain wavelength $I_\lambda(t)$ will be taken and can be well fitted with multiple exponential decays after deconvolution from the instrument response function (IRF). Therefore, we can separate the contributions from solvation and lifetimes by defining the following formula:

$$I_\lambda(t) = I_\lambda^{solv}(t) + I_\lambda^{popul}(t) = \sum_i a_i e^{-t/\tau_i} + \sum_j b_j e^{-t/\tau_j} \qquad \text{(S2)}$$

where the first term is for solvation and the second term is for lifetime emission. The coefficient a_i is positive (decay dynamics) at the blue side of the emission peak and is negative (initial rise) at the red side, which is a typical signature of solvation. The coefficient b_j is always positive and represents the contribution from multiple lifetime decays. With the steady-state emission intensity I_λ^{ss} at λ, we can easily construct the 3D femtosecond-resolved emission spectrum (FRES) for the overall emission spectrum from eq. (S3) and the lifetime emission spectrum from eq. (S4) below.

$$I(\lambda,t) = \frac{I_\lambda^{ss} I_\lambda(t)}{\sum_i a_i\tau_i + \sum_j b_j\tau_j} \qquad \text{(S3)}$$

$$I^{popul}(\lambda,t) = \frac{I_\lambda^{ss} I_\lambda^{popul}(t)}{\sum_i a_i\tau_i + \sum_j b_j\tau_j} \qquad \text{(S4)}$$

We can construct the FRES and derive the solvation correlation function by following eq. (S1). The solvation correlation function usually exhibits multiple exponential decays, indicating the different hydration water motions.

In another way, we can also directly scan the emission spectra at different delay times and can immediately construct 3D FRES after deconvolution. Thus, we can directly read out the emission peaks $\nu_s(t)$. However, we still need to fit a series of transients from the 3D FRES following eq. (S2) and obtain $\nu_l(t)$. Similarly, using eq. (S1), we then obtain the final solvation function c(t). Both methods of either taking the fluorescence transients or directly scanning the spectra will give the similar results of the solvation dynamics.

The time-resolved fluorescence anisotropy directly measures the tryptophan relaxation dynamics. With the two femtosecond-resolved fluorescence intensity of $I_{//}$ and $I_{\perp}$, we can construct the fluorescence anisotropy $r(t)=(I_{//}-I_{\perp})/(I_{//}+2I_{\perp})$, which follows eq. (S5) below.

$$r(t) = r_{IC}(t) + r_{2W}(t) + r_{3W}(t) + r_T(t) \qquad \text{(S5)}$$

The initial ultrafast decay (r_{IC}) results from the internal conversion between two electronically excited states of ^1L_a and ^1L_b and happens in less than 100 fs[12]. The following two components (r_{2W} and r_{3W}) indicate the local wobbling motions of the tryptophan sidechain. r_T represents the contribution from the overall protein tumbling motion. Following a simple axially symmetric oscillation model, we can estimate the wobbling semiangle (θ) using the following expression:

$$1 - \frac{r_{iW}(0)}{r_{iW}(0) + r_T(0)} = \left[\frac{3\cos^2\theta - 1}{2}\right]^2, \ \mathrm{i} = 2, 3 \qquad \text{(S6)}$$

Each fluorescence anisotropy component from the wobbling and tumbling motions follows an exponential decay and their decay times can be determined from an exponential fit.

SI Note 2. Molecular dynamics (MD) simulations

The amber package 11 was used for the MD simulation with the ff99SB force field. The crystal structure was taken from the protein data bank (2RDI). Mutation was performed in the Pymol software and the most probable configuration was chosen as the initial structure. Then, the whole protein was solvated by explicit water TIP3P with a cutoff of 12Å in an octahedral box. Cl^{-1} and Na^{+} ions were added to neutralize the system. A minimization process was used to remove the bad contacts. First, fix the solute and run 1000 steps of minimization to relax the water and ion molecules with the constant volume periodic boundaries; second, relax the entire system without any constraints for 1500 steps. And then, while adding a weak restraint on the protein molecules, the system was heated up from 0 to 300 K in 10000 steps with a step size of 2 fs by using the Langevin dynamics to control temperature. SHAKE algorithm was used to constrain bonds involving hydrogen. Finally, a 2-ns MD simulation trajectory was obtained after a 0.5-ns equilibration process at 300 K under a constant pressure of 1 atm to mimic the typical experiment condition.

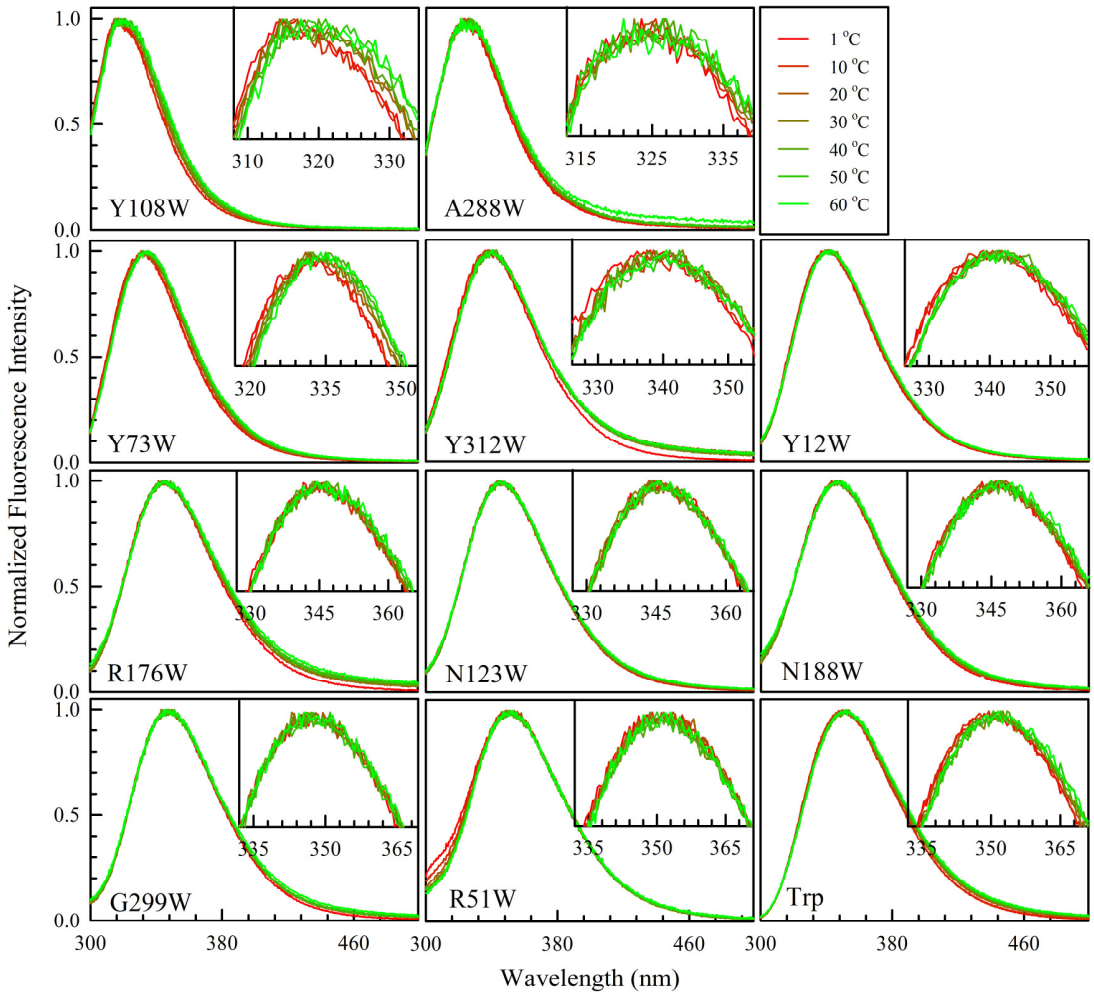

Fig. S1. The steady-state emission spectra of 10 mutants and tryptophan in buffer. The steady-state emission spectra only slightly red shifted when temperature increases from 1 °C to 60 °C, indicating the overall stable structure in the temperature range. The insets show close-up views around the emission peaks.

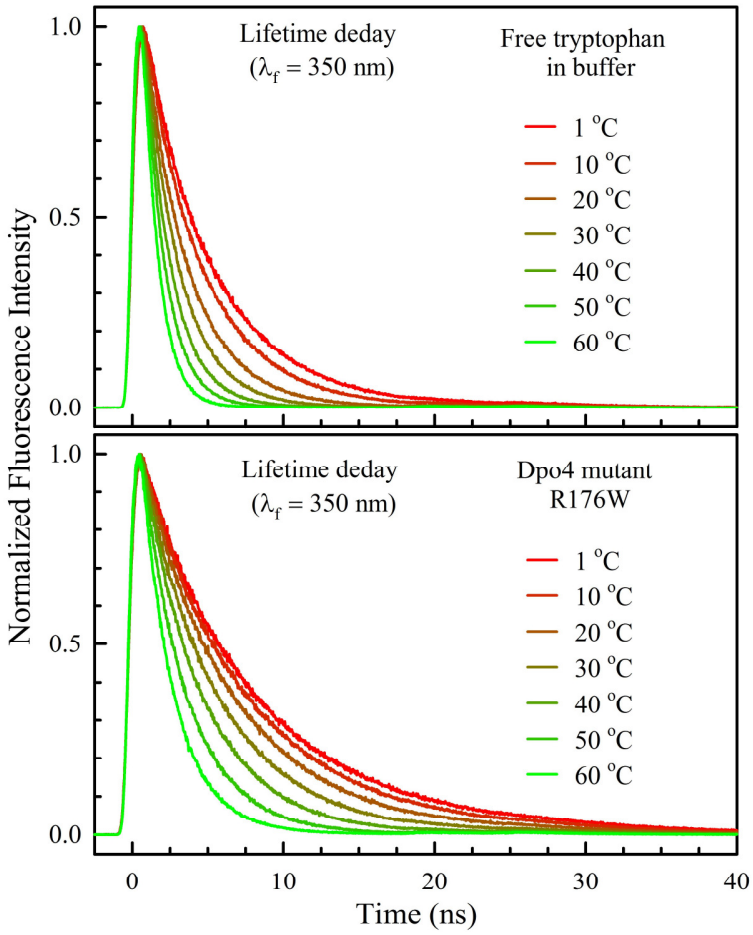

Fig. S2. Lifetime decays of free tryptophan and an example of Dpo4 mutant R176W. Lifetime decays of free tryptophan in buffer and Dpo4 mutant R176W become faster when temperature increases. Time-resolved fluorescence transients were taken at 350 nm by TCSPC.

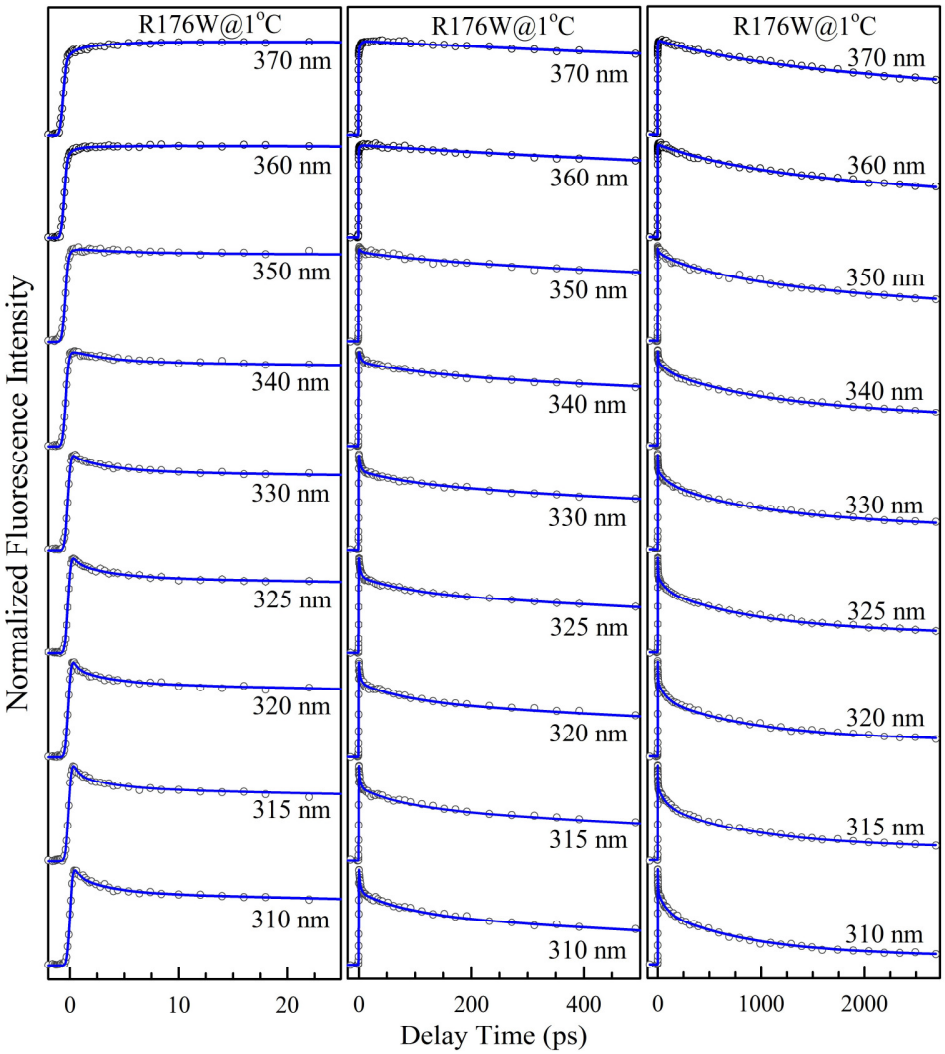

Fig. S3A. Fluorescence transients of R176W at 1 °C. Fluorescence transients were taken at different wavelengths to cover the overall emission spectrum. The transients were taken up to 2.75 ns and are shown in three panels in three time windows.

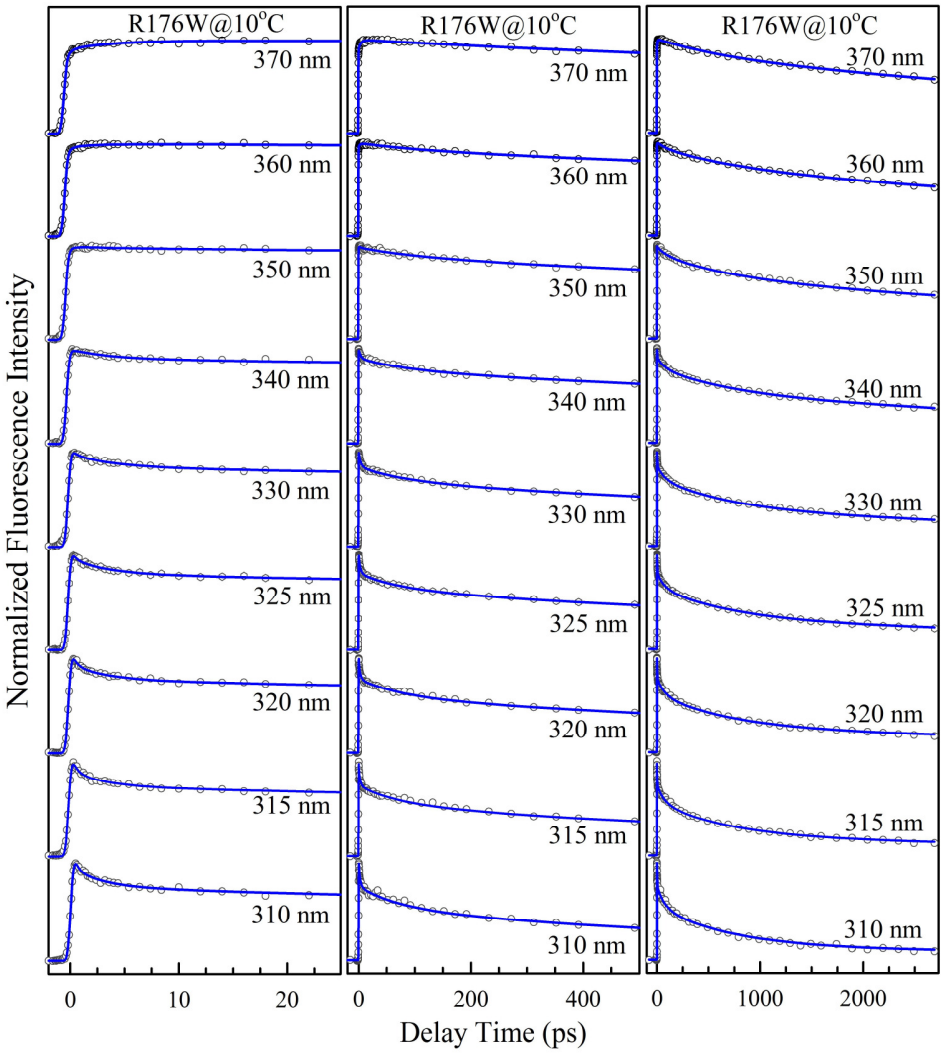

Fig. S3B. Fluorescence transients of R176W at 10 °C. Fluorescence transients were taken at different wavelengths to cover the overall emission spectrum. The transients were taken up to 2.75 ns and are shown in three panels in three time windows.

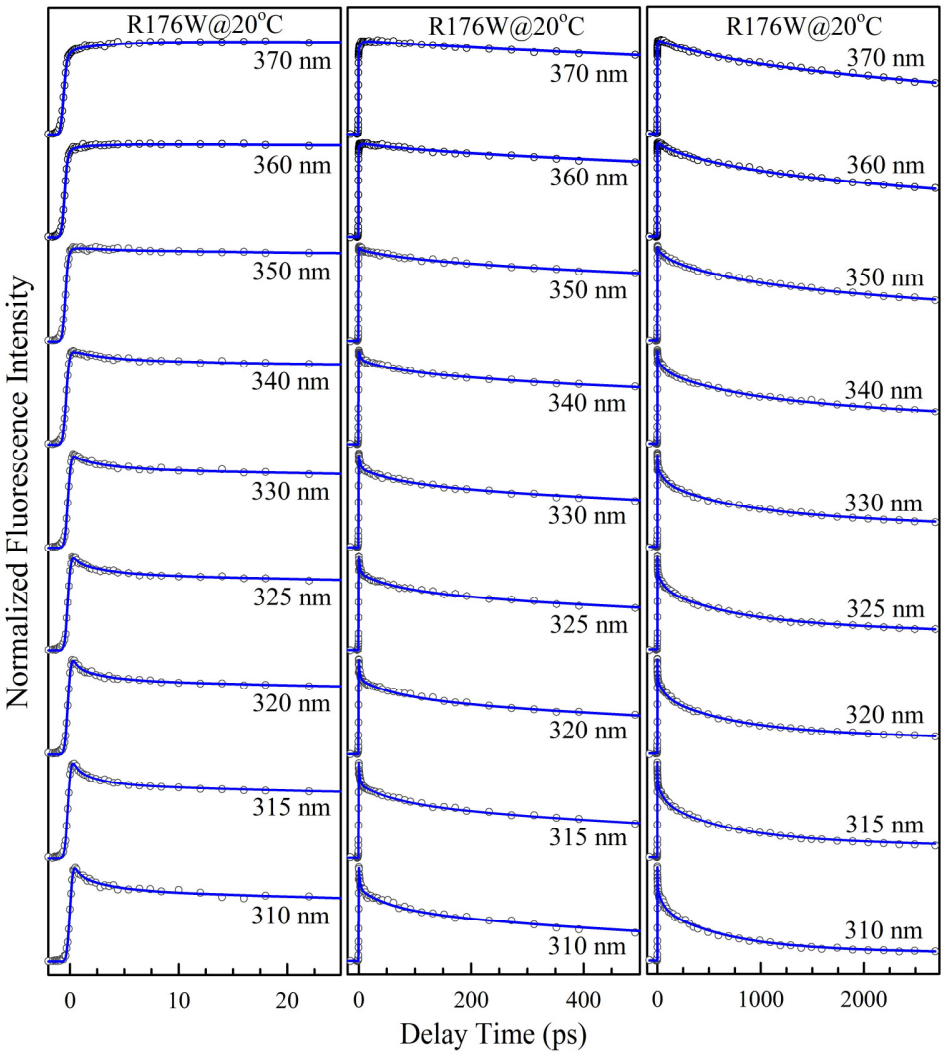

Fig. S3C. Fluorescence transients of R176W at 20 °C. Fluorescence transients were taken at different wavelengths to cover the overall emission spectrum. The transients were taken up to 2.75 ns and are shown in three panels in three time windows.

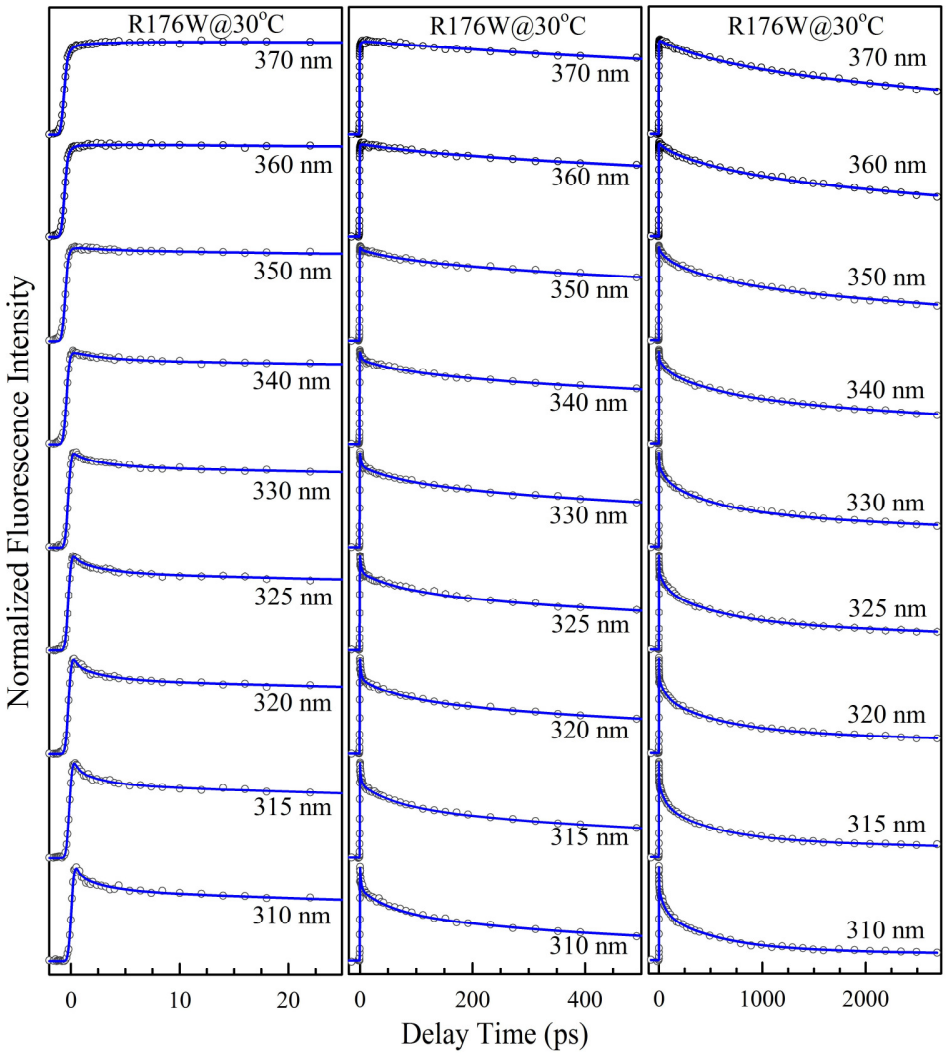

Fig. S3D. Fluorescence transients of R176W at 30 °C. Fluorescence transients were taken at different wavelengths to cover the overall emission spectrum. The transients were taken up to 2.75 ns and are shown in three panels in three time windows.

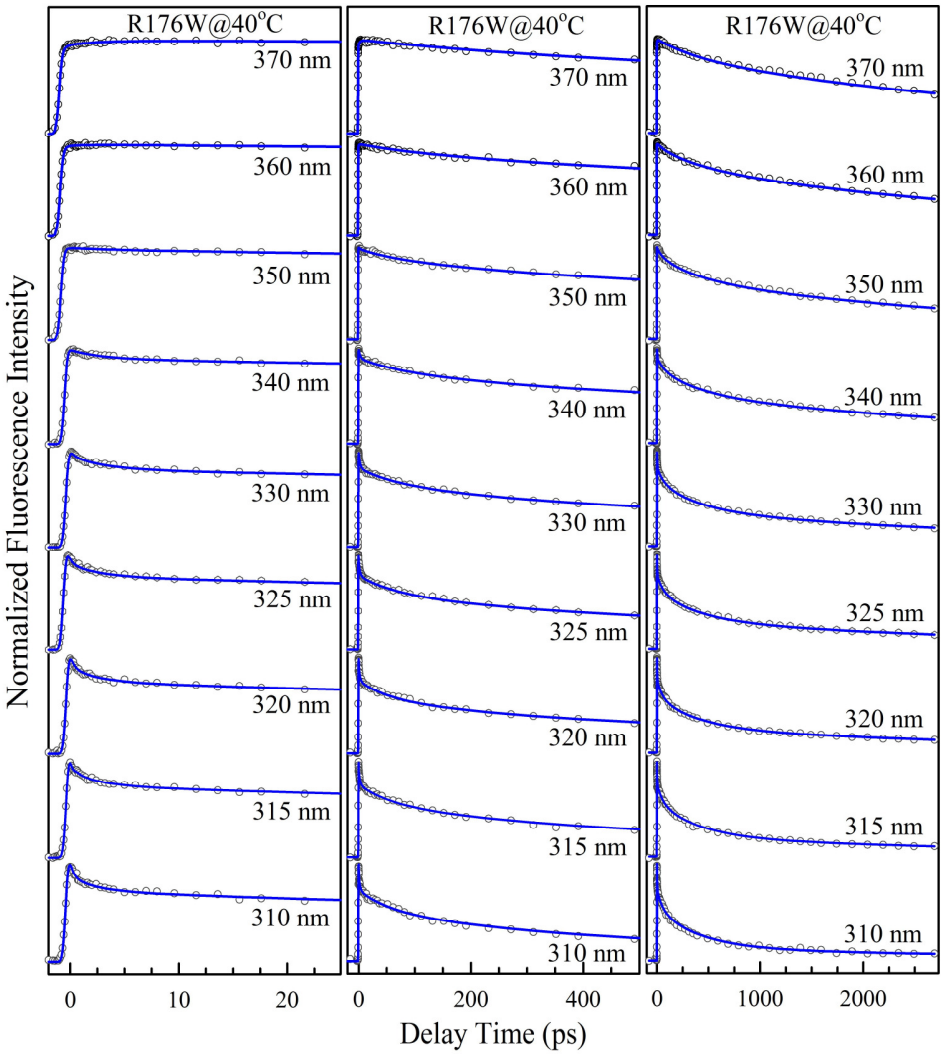

Fig. S3E. Fluorescence transients of R176W at 40 °C. Fluorescence transients were taken at different wavelengths to cover the overall emission spectrum. The transients were taken up to 2.75 ns and are shown in three panels in three time windows.

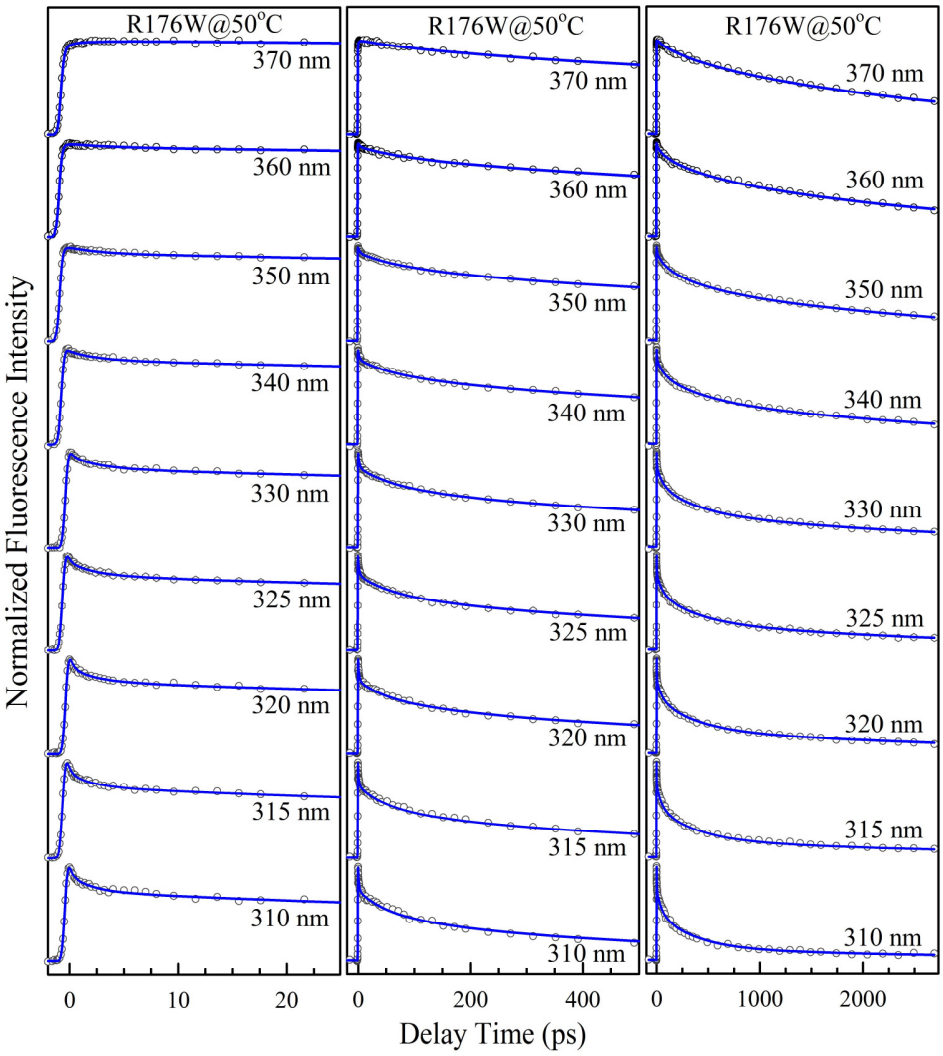

Fig. S3F. Fluorescence transients of R176W at 50 °C. Fluorescence transients were taken at different wavelengths to cover the overall emission spectrum. The transients were taken up to 2.75 ns and are shown in three panels in three time windows.

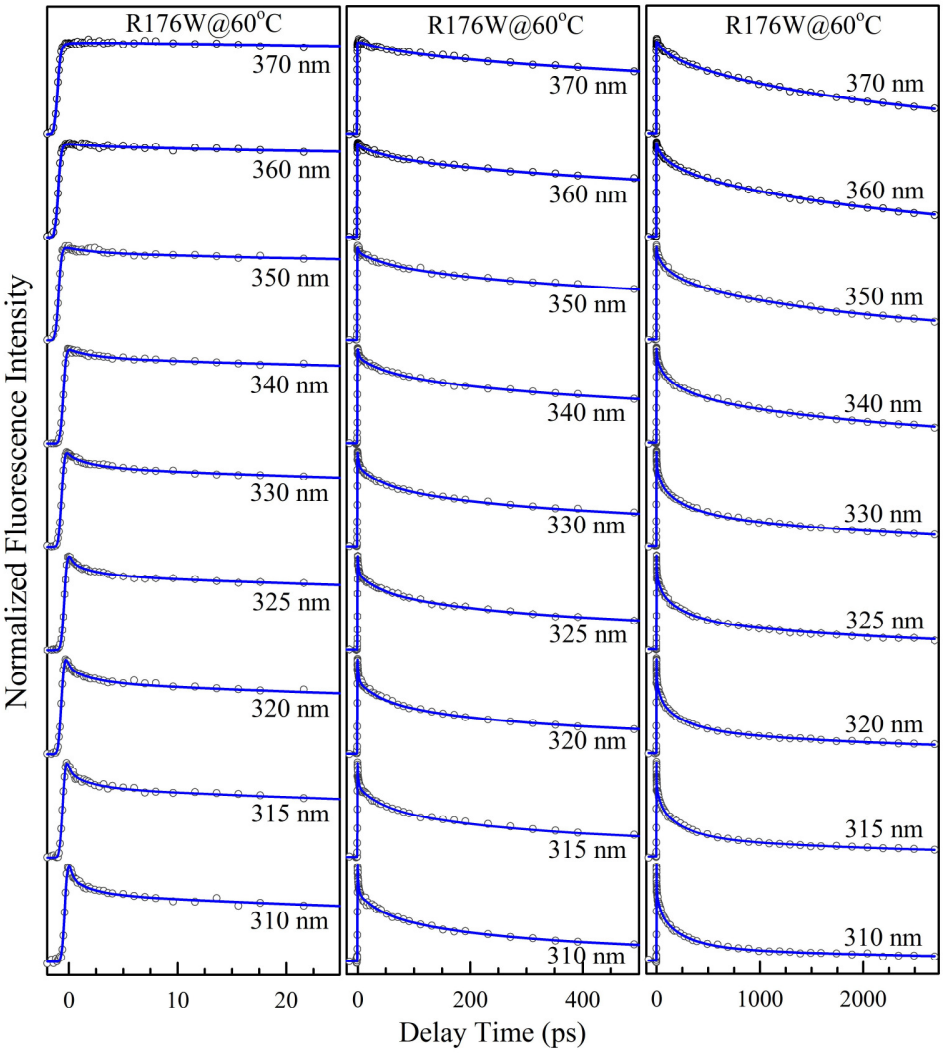

Fig. S3G. Fluorescence transients of R176W at 60 °C. Fluorescence transients were taken at different wavelengths to cover the overall emission spectrum. The transients were taken up to 2.75 ns and are shown in three panels in three time windows.

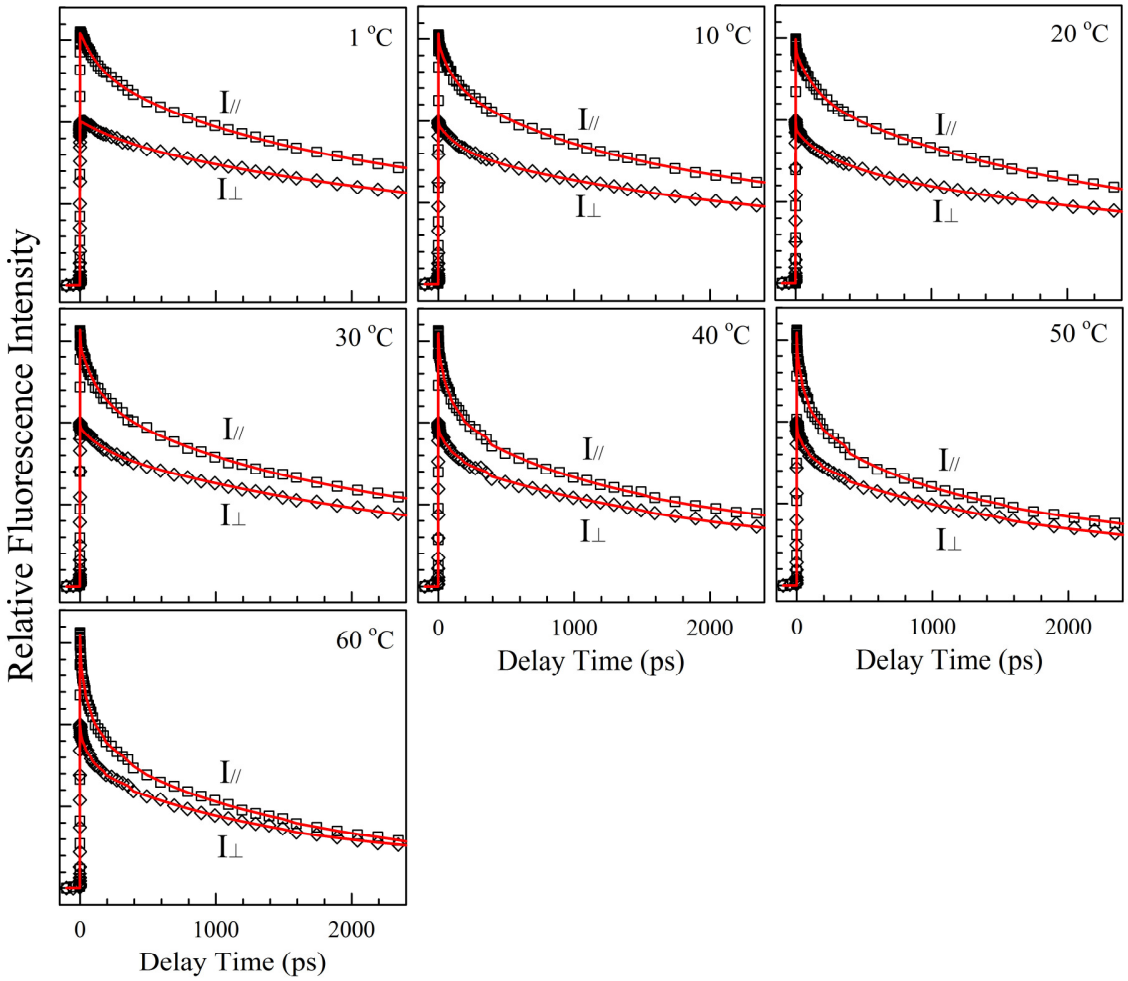

Fig. S4. Parallel and perpendicular fluorescence transients of mutant R176W taken at different temperatures for the fluorescence anisotropy dynamics. The parallel ($I_{//}$) and perpendicular ($I_{\perp}$) fluorescence transients were taken at 350 nm for different temperatures.

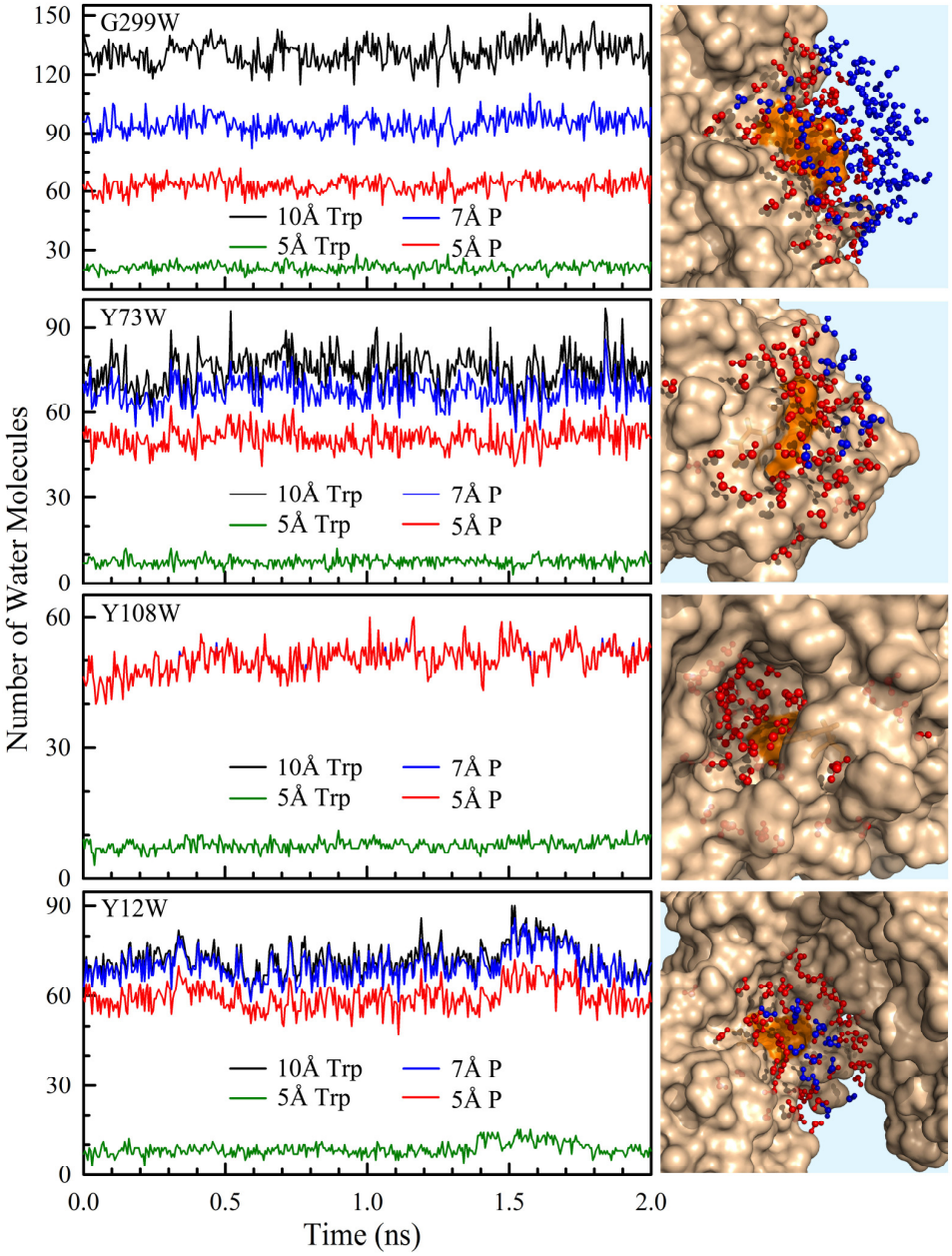

Fig. S5. MD simulations of four typical mutants. Left column shows the number of water molecules within 5 Å and 10 Å from the tryptophan indole ring for four mutants G299W, Y73W, Y108W and Y12W. Among all water molecules within 10 Å from the indole ring, water molecules within 5 Å and 7 Å to the protein surface are also shown. Right column displays one snapshot of MD simulations for the four mutants, showing that all water molecules within 10 Å from the indole ring can be separated into two parts: 5 Å from protein surface (red) and 5-10 Å from the protein surface (blue). The orange patches on the protein surface are the tryptophan probes. Note that all the water molecules are within 5 Å from the protein surface for the buried mutant Y108W.

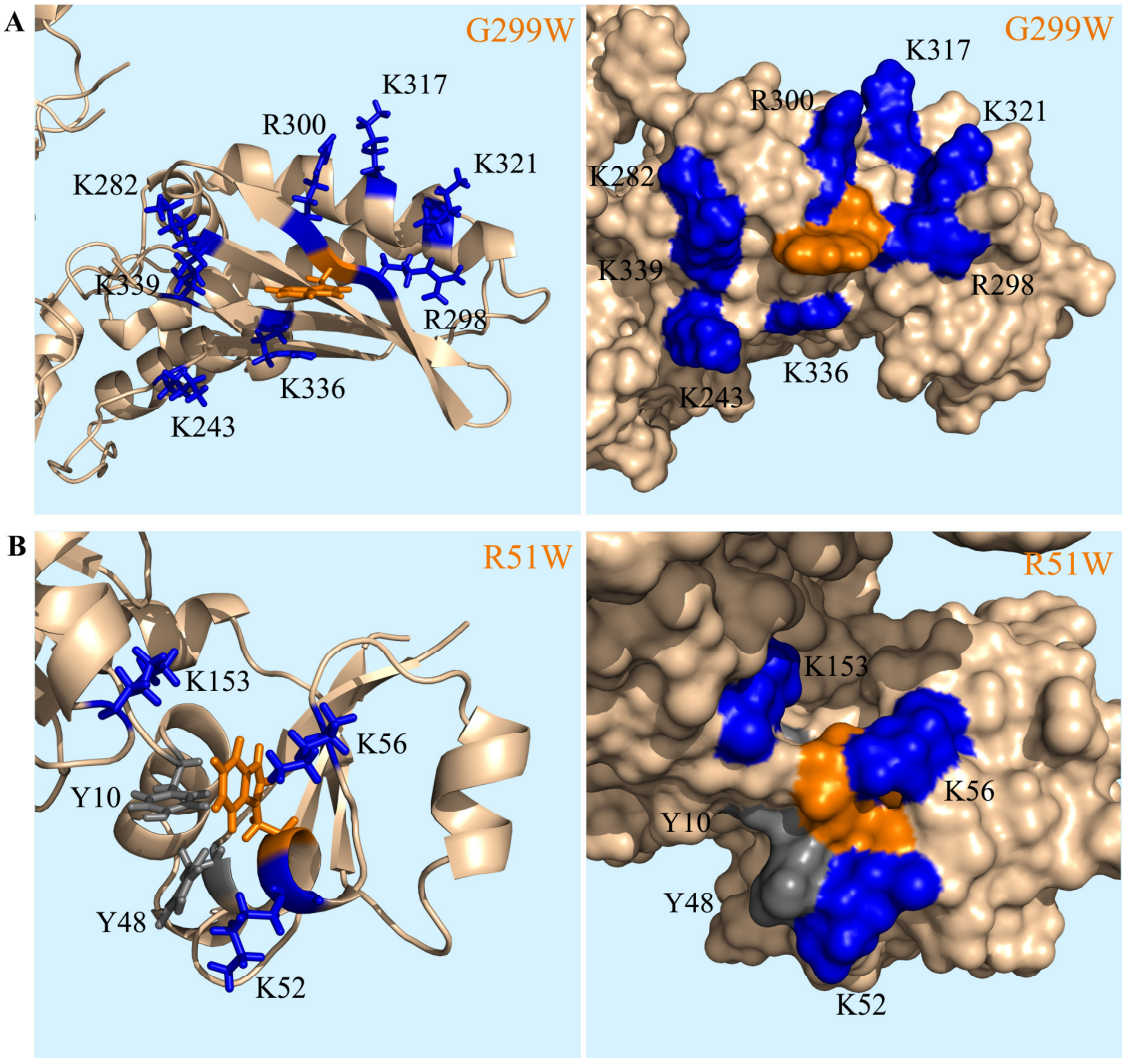

Fig. S6. Local environments of two exposed Dpo4 mutants. (**A**) Two panels show the special local environment of mutant G299W from a snapshot of MD simulations. G299W (orange) is surrounded by eight positive charged residues (blue), resulting in the increase of the hydration network rigidity. (**B**) Unlike G299W, R51W has two tyrosines (grey) and three lysines (blue) surrounding the probe tryptophan (orange), providing a relatively rigid protein environment for tryptophan. Charged residues, especially K56, can form a strong cation-π interaction with tryptophan to anchor the tryptophan position.

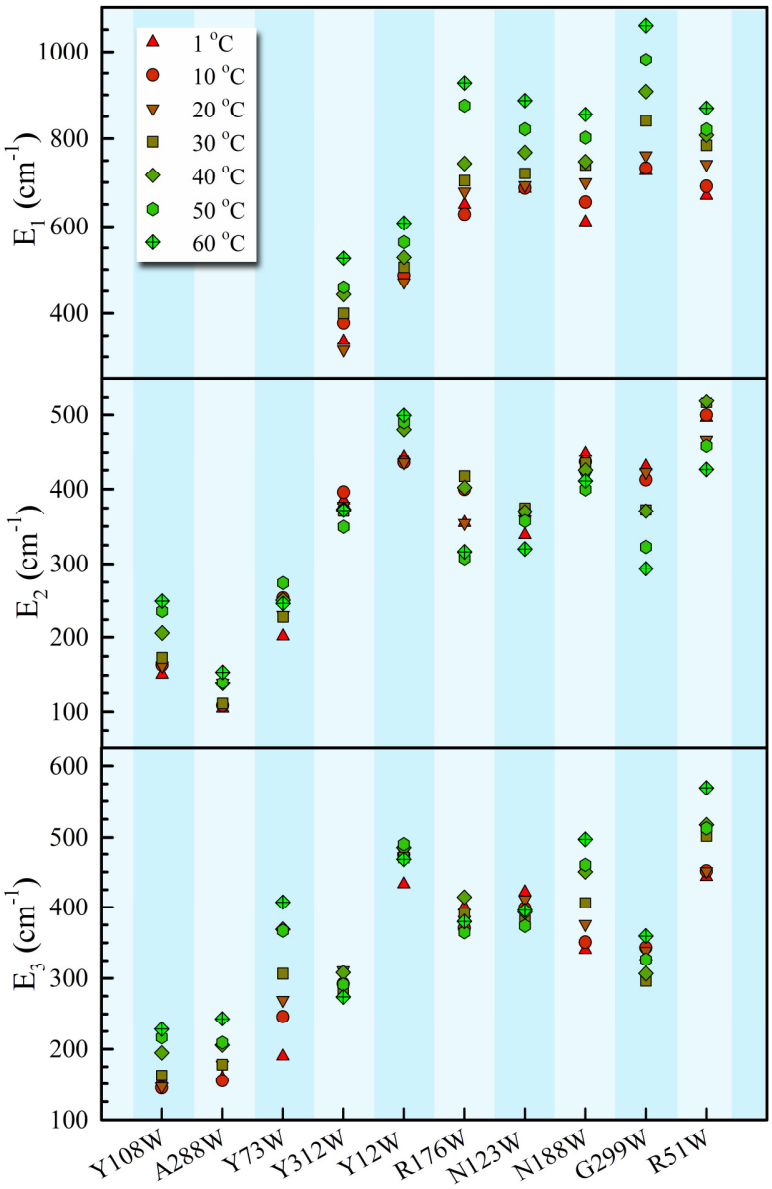

Fig. S7. Solvation energy of various components for all mutants at different temperatures. The solvation energy of the sub-picosecond component increases when temperature rises. The other two solvation energies slightly change (note the scale).

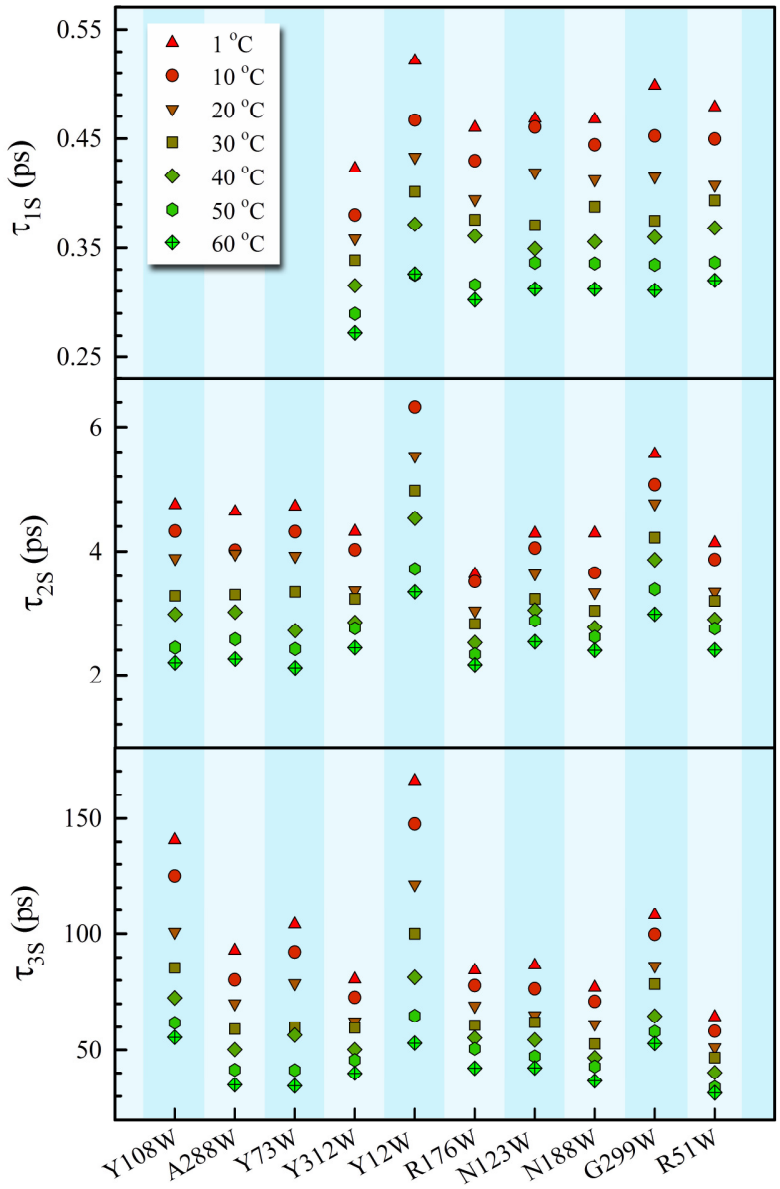

Fig. S8. Solvation times of all mutants at various temperatures. Note that the solvation times of all mutants decrease when temperature rises.

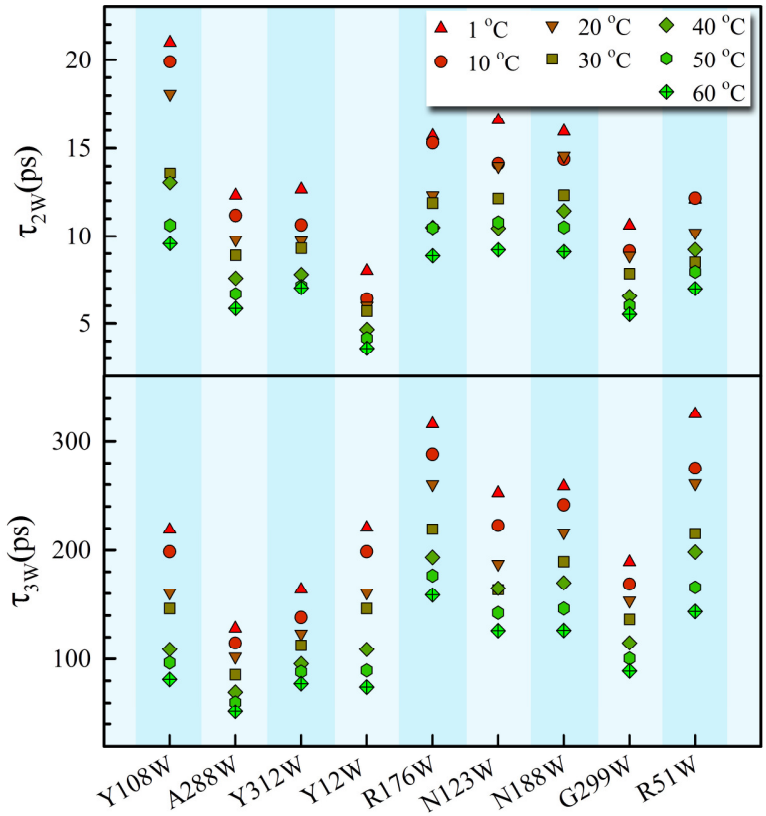

Fig. S9. Protein sidechain relaxation times for 9 mutants at various temperatures. The two relaxation times of tryptophan sidechains become shorter when temperature rises.

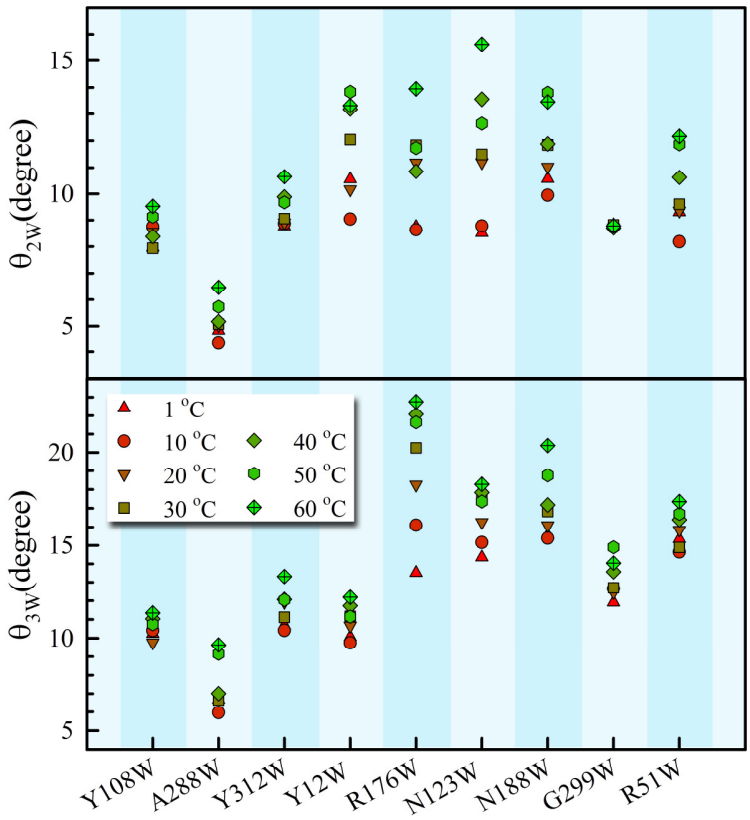

Fig. S10. Protein sidechain wobbling cone semiangles for 9 mutants at various temperatures. The two wobbling cone semiangles become larger when temperature rises.

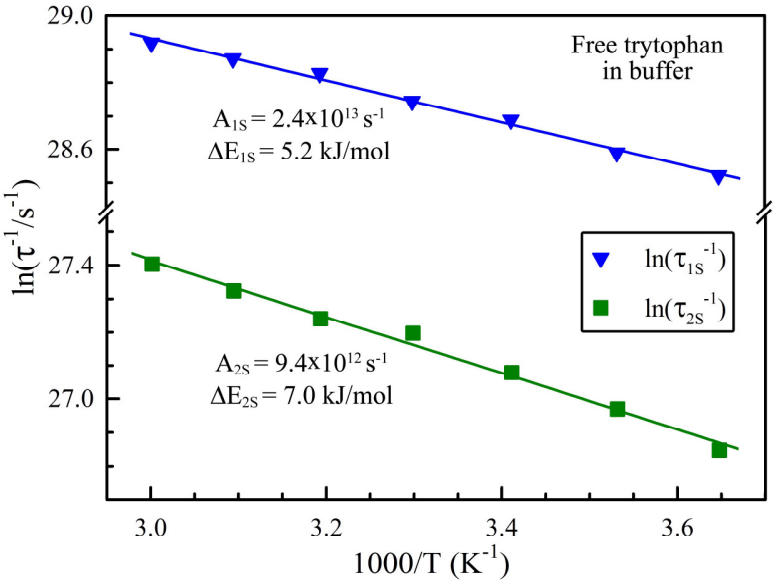

Fig. S11. Plot of tryptophan solvation rates with temperature using an Arrhenius relation. Free tryptophan in buffer probes the solvation rates of bulk water. The two solvation rates show excellent Arrhenius relations in the temperature range from 1 °C to 60 °C. The symbols are the experimental data and the solid lines show an Arrhenius fit. The activation energy (ΔE) and the prefactor (A) are also given in the figure.

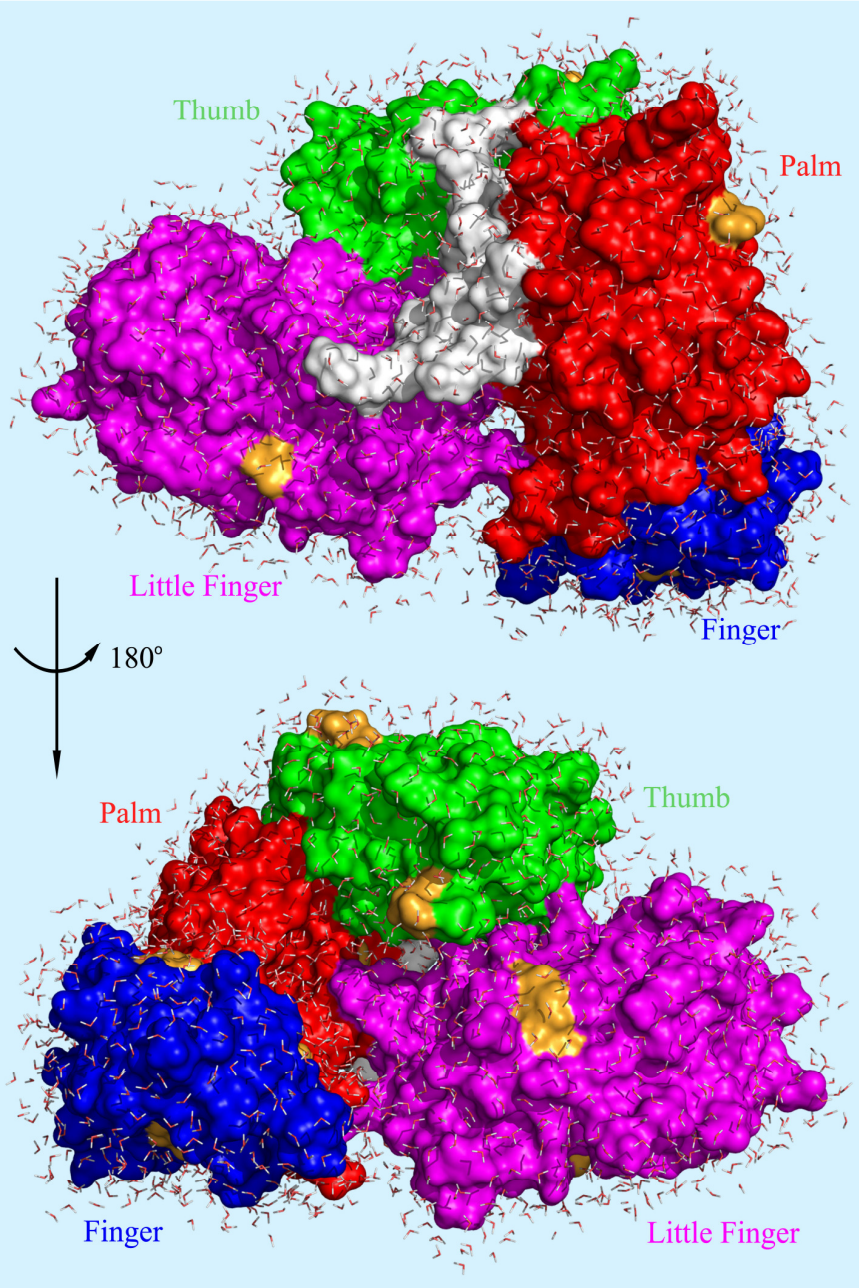

Fig. S12. Protein surface hydration for each domain of Dpo4 within 5 Å to the protein. Dpo4 consists of four domains. The little finger domain undergoes a large domain motion for binding DNA, but in the apo-form it is not separated from the other three domains. (**Top**) Shown are the four domains of Dpo4 with different colors and the white part indicates the random coil linking the thumb and little finger domains. Besides the linker, the little finger domain is also in contact with the palm and thumb domains. (**Bottom**) Shown is the protein structure by rotating 180° to show the other side and also confirms their contacts. The little finger domain closely interacts with the other three ones and the whole protein may have a collective surface hydration shell.

Table S1. Fitting parameters for hydration water and protein sidechain motions.

Mutant	ΔE_{1S} kJ/mol	ΔE_{2S} kJ/mol	ΔE_{3S} kJ/mol	A_{1S} 10^{13} s^{-1}	A_{2S} 10^{12} s^{-1}	A_{3S} 10^{11} s^{-1}	ΔE_{2W} kJ/mol	ΔE_{3W} kJ/mol	A_{2W} 10^{12} s^{-1}	A_{3W} 10^{11} s^{-1}
R51W	5.2	6.7	9.5	2.0	4.6	9.8	7.1	10.2	1.8	2.7
G299W	5.9	7.9	9.6	2.8	5.5	6.0	8.5	9.8	3.9	3.8
N188W	5.3	7.2	9.6	2.1	5.6	8.6	6.9	9.2	1.3	2.1
N123W	5.6	6.7	9.0	2.4	4.3	6.0	7.1	8.5	1.4	1.7
R176W	5.4	7.0	8.7	2.3	5.8	5.1	7.2	9.1	1.5	1.7
Y12W	6.3	9.6	14.9	3.0	9.4	39.1	9.6	14.3	8.6	22.5
Y312W	5.5	7.3	8.9	2.6	5.5	6.1	7.7	9.3	2.3	3.7
Y73W		10.8	14.4		22.5	48.9				
A288W		9.1	12.6		11.3	25.7	9.6	12.0	5.3	14.2
Y108W		10.2	12.4		17.5	16.2	10.8	13.1	5.0	13.7

ΔE_{is}, i=1-3: the solvation activation energy of each component.
A_{is}, i=1-3: the prefactor of each solvation component.
ΔE_{iw}, i=2-3: the activation energy of each tryptophan relaxation component.
A_{iw}, i=2-3: the prefactor of each tryptophan relaxation component.