Supporting Information for

The Magic of Linking Rings: Discovery of a Unique Photoinduced Fluorescent Protein Crosslink

Manman Lu[1,2], Dmitri Toptygin[3], Yufei Xiang[4,6], Yi Shi[4,6], Charles D. Schwieters[5], Emma C. Lipinski[1], Jinwoo Ahn[1,2], In-Ja L. Byeon[1,2], and Angela M. Gronenborn[1,2*]

[1]Department of Structural Biology, University of Pittsburgh School of Medicine; Pittsburgh, PA 15261, United States.

[2]Pittsburgh Center for HIV Protein Interactions, University of Pittsburgh School of Medicine; Pittsburgh, PA 15261, United States.

[3]Department of Biology, Johns Hopkins University; Baltimore, MD 21218, United States.

[4]Department of Cell Biology, University of Pittsburgh; Pittsburgh, PA 15261, United States.

[5]Computational Biomolecular Magnetic Resonance Core, Laboratory of Chemical Physics, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health; Bethesda, MD 20892-0520, United States.

[6]Present Address: Center of Protein Engineering and Therapeutics, Department of Pharmacological Sciences, Icahn School of Medicine at Mount Sinai; New York, NY 10029, United States.

*Corresponding author. Email: amg100@pitt.edu

Materials and Methods

Protein expression and purification

The cDNA encoding CypA V2Q (Met1-Glu165), purchased from GenScript, Inc. Piscataway, NJ, USA, was amplified and subcloned into pET21 vectors (EMD chemicals, Inc. San Diego, CA, USA) using NdeI and XhoI sites. The V2Q mutation was introduced to ensure that all the N-terminus is homogeneous and starts with M1 (this is not the case for the M_1-V_2 protein). For the CypA (V2Q) F60Y mutant, the single amino acid change (F60Y) was generated using partial overlapping primers (Integrated DNA Technologies, Inc. Coralville, IA, USA). All CypA proteins were expressed and purified as previously described[1-2]. In brief, CypA proteins were expressed in *E. coli* Rosetta 2 (DE3), cultured in modified M9 medium, containing 4 g/L U-$^{12}C_6$-glucose or 2 g/L U-$^{13}C_6$-glucose as the carbon source, 1 g/L $^{14}NH_4Cl$ or 1 g/L $^{15}NH_4Cl$ as the nitrogen source, and 20 mg/L 7-fluoroindole as the fluorine source, induced with 0.5 mM isopropyl b-D-thiogalactopyranoside (IPTG) for expression at 18°C for 16 hr. All CypA proteins were purified by cation exchange (HiTrap SP HP, 5 mL column) in 25 mM sodium phosphate buffer (pH 5.8), 1 mM dithiothreitol (DTT), 0.02% sodium azide (NaN_3) and eluted with a 0-1 M NaCl gradient. The final purification was performed by size-exclusion (HiLoad 26/600 Superdex 75 column) in 25 mM sodium phosphate (pH 6.5), 1 mM TCEP, 0.02% NaN_3.

For the purification of CL-CypA, full-length HIV-1 CA protein, expressed and purified as described previously[3], was added to the UV radiated mixture at a CypA:CA molar ratio of 1:4. This mixture was incubated at room temperature for at least 2 hrs to ensure complete binding, loaded onto a size-exclusion column (HiLoad 16/600 Superdex 75), equilibrated in 25 mM sodium phosphate (pH 6.5), 1 mM TCEP, 0.02% NaN_3, for separation of the CA-CypA complex (43.9 kDa), excess CA protein (25.6 kDa), and CL-CypA (18.3 kDa). Fractions containing CL-CypA were analyzed by ESI-MS and only those with a purity higher than ~77% were selected and combined.

The cDNA encoding PS1 was purchased from GenScript, Inc. Piscataway, NJ, USA, amplified and subcloned into the pET28a(+)-TEV vector using Ndel and HindIII sites. The construct contained an N-terminal 6x His-tag followed by a TEV cleavage site. U-^{15}N, 7F-Trp-PS1 was expressed in *E. coli* Rosetta 2 (DE3), cultured in modified M9 medium, containing 4 g/L $^{12}C_6$-glucose as the carbon source, 1 g/L $^{15}NH_4Cl$ as the nitrogen source, and 20 mg/L 7-fluoroindole as the fluorine source, induced with 0.5 mM IPTG for expression at 18°C for 16 hr.

The PS1 protein was purified using a 5 mL HISTrap HP column in 20 mM HEPES buffer (pH 7.5), 300 mM NaCl, 0.5 mM TCEP, and eluted with a 50-500 mM imidazole gradient. The final purification was performed by size-exclusion (HiLoad 26/600 Superdex 75 column) in 20 mM HEPES buffer (pH 7.5), 150 mM NaCl, 0.5 mM TCEP.

Trypsin digestion

Approximately 1 mg of 7F-Trp-CypA protein, either before or after UV radiation, was buffer exchanged into 100 mM Tris buffer containing 6 M urea. The CypA protein was treated with TCEP (5 mM) for 1 hr and then alkylated with 20 mM iodoacetamide for 1 hr at room temperature. The reaction mixture was diluted 10 times with ammonia bicarbonate (50 mM), which reduced the urea concentration to 0.6 M. Trypsin (Pierce™, MS grade) was added at a protease:substrate ratio of 1:50, followed by overnight incubation at 37°C to allow for complete digestion. The reaction was stopped by reducing the pH to <3 adding formic acid. The digested peptide mixture was desalted using C18 stagetips before LC/MS analysis.

Mass spectrometry

ESI-MS. The molecular masses of proteins (U-^{15}N, 7F-Trp-CypA, the F60Y variant of U-^{15}N, 7F-Trp-CypA, U-^{15}N, 7F-Trp-apo-PS1, and enriched/purified U-^{15}N-CL-CypA), before and/or after UV radiation, were determined on a LC-ESI-TOF mass spectrometry (Bruker Daltonics, Billerica, MA, USA).

nLC-MS/MS. Desalted peptides were separated online by nanoflow HPLC (Easy nLC1200, Thermo Fisher) and analyzed with a Q Exactive HF-X Hybrid Quadrupole-Orbitrap mass spectrometer (Thermo Fisher)[4-5]. Briefly, ~100 ng of digested mixture were loaded on an analytical column (C18, 3 mm particle size, 300 Å pore size, 10.5 cm; New Objective) and eluted using a 60 min LC gradient : 5% B–8% B, 0 – 5 min; 8% B – 32% B, 5 – 45 min; 32% B–100% B, 45 – 49 min; 100% B, 49 -54 min; 100% B - 5 % B, 54 min - 54 min 10 sec; 5% B, 54 min 10 sec - 60 min 10 sec; mobile phase A consisted of 0.1% formic acid (FA), and mobile phase B consisted of 0.1% FA in 80% acetonitrile. The QE HF-X instrument was operated in the data-dependent mode, where the top 6 most abundant ions (mass range 380 - 2,000, charge state 3 - 7) were fragmented by high-energy collisional dissociation (normalized collision energy 27). The target resolution was 120,000 for MS and 15,000 for MS/MS analyses. The quadrupole isolation window was 1.8 Th and the maximum injection time for MS/MS was set at 150 ms.

Fluorescence spectroscopy

Kinetic photoactivation. Slow kinetic photoactivation experiments of 7F-Trp-CypA were performed at 20°C using a SLM-48000 spectrofluorometer equipped with a 450 W Xenon arc lamp, MC-200 monochromators and calcite polarizers. To increase the photoactivation rate, the polarizer from the excitation path was removed; the excitation monochromator slits were set at the maximum width of 16 nm, and the wavelength was set at 282 nm, which corresponds to the peak of the product of two spectra, the exciting photon flux spectrum and the extinction coefficient spectrum of 7F-Trp-CypA. At these settings the UV power entering the fluorescence cell was 0.55 mW. The emission monochromator wavelength was set at 365 nm, the slits were set at 4 nm, and the emission polarizer orientation was vertical. The emission intensity was used to monitor the degree of photoactivation as a function of time. The sample was constantly mixed using a magnetic stirrer to ensure that photoactivation occurred over the entire volume, rather than just the part within the excitation beam. The total time required to fully photoactivate a 2.5 mL sample of 50 μM 7F-Trp-CypA under these conditions is about 64 min (Figure 1C). Photoactivation of 7F-Trp-CypA was consistently reproduced with different protein concentrations (from 50 μM to 300 μM) on different instruments (PTI fluorometer and FluoroMax-4, Horiba Scientific, Edison, NJ). Photoactivation of 4F- and 7F-Trp amino acids, as well as 4F- and/or 7F-Trp labeled proteins (CypA, CypA F60Y, OAA, CA CTD, CA CTD L151F, and PS1) were performed under similar conditions either on a PTI fluorometer equipped with a 75 W lamp or on a FluoroMax-4 fluorometer equipped with a 150 W lamp, where the samples were photoactivated at the excitation wavelength of 282 nm with a 16 nm slit width, with emission fluorescence intensities recorded at 365 nm (or 375 nm for 7F-Trp-CypA F60Y) with a 1 nm slit width. All samples were constantly mixed using a magnetic stirrer.

Steady state excitation-emission matrix experiments. Excitation-emission matrix data sets were obtained on two samples, the enriched/purified CL-CypA (20.4 μM) and the non-purified photoactivated 7F-Trp-CypA (50 μM). Steady-state fluorescence was measured at 20°C using a SLM-48000 spectrofluorometer with excitation monochromator slits set at 2 nm, emission monochromator slits set at 4 nm, polarizers set at a magic angle configuration (excitation and emission polarizer at 55° and 0° from vertical, respectively). 69 emission spectra were taken at exciting wavelengths from 260 nm to 350 nm with 2 nm step and from 350 nm to 442 nm with 4 nm step. Each emission spectrum consisted of 321 wavelengths from 280 nm to 600 nm, with 1

nm steps. To improve the signal/noise ratio, the intensity at each wavelength was measured four times and the four readings were averaged. The emission spectra served as the columns of a 321×69 excitation-emission matrix **M**. Several measures were taken to ensure the stability of the fluorescence sample and the instrument sensitivity during the nine-hour long data acquisition. To cancel slow excitation intensity drifts due to the aging of the Xenon arc lamp, the fluorescence emission intensity was divided by the excitation intensity measured simultaneously. Importantly, the temperature of the fluorescent sample was maintained to within ±0.05°C, the temperature of the room housing the instrument was maintained to within ±0.2°C, ensuring stable photomultiplier sensitivity. The sample was constantly mixed using a magnetic stirrer at a slow speed to prevent protein denaturation. Narrow excitation monochromator slits (2 nm) were used to minimized photodamage to the sample. Emission spectra were taken in order of descending exciting wavelengths, to avoid photodamage since short exciting wavelengths cause significantly more photodamage than the long ones.

Singular Value Decomposition (SVD). By definition, SVD of a real $m \times n$ matrix **M** consists of finding an $m \times r$ semi-orthonormal real matrix **U**, an $n \times r$ semi-orthonormal real matrix **V**, and an $r \times r$ diagonal real matrix **S** with non-negative diagonal elements, such that $\mathbf{M}=\mathbf{U}\mathbf{S}\mathbf{V}^{T}$, where r equals the smaller of m and n, and the superscript ^{T} denotes a transposed matrix. Finding the matrices **U**, **S**, and **V** for a given matrix **M** is a routine task in numerical linear algebra[6-7]. In the case where the elements of the data matrix **M** contain random noise, r represents the rank of the matrix **S**; all diagonal elements s_{ii} of the matrix **S** are positive, these elements are called the singular values and are sorted in the descending order $s_{11}>s_{22}>s_{33}>...$ It is typically observed that the first k singular values (where $k<r$) are significantly greater than the remaining r-k singular values that are rather close to one another. It is postulated that the last r-k singular values represent pure noise, whereas the first k singular values and the corresponding basis vectors (i. e., the first k columns of the matrix **U** and the first k columns of the matrix **V**) represent the signal (albeit with some noise). Next, a truncated diagonal matrix $\mathbf{S}_k$ is defined, with the first k diagonal elements being equal to those of the matrix **S** and the last r-k diagonal elements being equal to zero. The matrix product $\mathbf{U}\mathbf{S}_k\mathbf{V}^{T}$ reconstructs a reduced noise version of the original data matrix **M**; the reconstructed matrix has a reduced rank k, compared to the full rank r of the original data matrix **M**; k is usually referred to as the effective rank of the data matrix **M**.

Iterative SVD algorithm. SVD of an excitation-emission matrix requires an iterative algorithm because of the presence of scattered light in the experimental spectra. Scattered excitation light passes through emission monochromator if $|\lambda_{ex}-\lambda_{em}|\leq w_{ex}+w_{em}$, where λ_{ex} and λ_{em} are the excitation and emission wavelengths, respectively; w_{ex} and w_{em} are the full widths at half maximum of the excitation and emission monochromators, respectively. The matrix elements for which the condition $|\lambda_{ex}-\lambda_{em}|<w_{ex}+w_{em}$ is satisfied are called the affected elements; those elements cannot be used in SVD, since their use would result in a very large effective rank k, defeating the purpose of SVD. In general, for a $m\times n$ data matrix $\mathbf{M}$, a corresponding $m\times n$ Boolean array is defined, the elements of which are true for the matrix elements not affected by the scattered light and false for the affected elements that should not be used in SVD. The Boolean array can be also used to exclude inelastic Raman scattered light and elastic scattered light that passes through both monochromators due to the second-order diffraction in one of them. The iterative SVD algorithm uses the Boolean array in the following way: For the zeroth approximation the algorithm uses the matrix $\mathbf{M}^{(0)}$ instead of the experimental data matrix $\mathbf{M}$. The non-affected elements in $\mathbf{M}^{(0)}$ are the same as the elements of the matrix $\mathbf{M}$. The affected elements are replaced using a linear interpolation between nearby non-affected elements of the matrix $\mathbf{M}$. SVD of the zeroth approximation (i=0) or a subsequent i-th approximation matrix, $\mathbf{M}^{(i)}=\mathbf{U}^{(i)}\mathbf{S}^{(i)}\mathbf{V}^{(i)\mathrm{T}}$, generates three matrices $\mathbf{U}^{(i)}$, $\mathbf{S}^{(i)}$, $\mathbf{V}^{(i)}$. It is assumed that the effective rank k is known in advance and equals l (this assumption is discussed below); the truncated matrix $\mathbf{S}^{(i)}{}_l$ is used to generate the reconstructed matrix $\mathbf{U}^{(i)}\mathbf{S}^{(i)}{}_l\mathbf{V}^{(i)\mathrm{T}}$. For the i+1-th iteration the algorithm uses the matrix $\mathbf{M}^{(i+1)}$, in which the non-affected elements are the same as the elements of the matrix $\mathbf{M}$, whereas the affected elements are replaced with the corresponding elements of the rank l reconstructed matrix $\mathbf{U}^{(i)}\mathbf{S}^{(i)}{}_l\mathbf{V}^{(i)\mathrm{T}}$. If the number of affected elements is less than 5% of the total number of matrix elements, then the iterative algorithm converges after a few dozens of iterations; once the algorithm has converged, the matrices $\mathbf{U}^{(i)}$, $\mathbf{S}^{(i)}$, $\mathbf{V}^{(i)}$ do not change from one iteration to the next. To apply the iterative SVD algorithm, the effective rank k must be guessed in advance; the advance guess for k is denoted l (see above). If $l>k$, it is likely that the algorithm will not converge; if $l<k$, the final results of iterative SVD will not be accurate. In both cases the value of l needs to be changed and the iterative SVD repeated. However, if $l>k$ and the algorithm converges, the final matrices $\mathbf{U}^{(i)}$, $\mathbf{S}^{(i)}$, $\mathbf{V}^{(i)}$ are accurate, and there is no need to repeat iterative SVD. Using the final matrices $\mathbf{U}^{(i)}$, $\mathbf{S}^{(i)}$, $\mathbf{V}^{(i)}$, a "clean" version rank k excitation-emission matrix $\mathbf{M}_k=\mathbf{U}^{(i)}\mathbf{S}^{(i)}{}_k\mathbf{V}^{(i)\mathrm{T}}$ can be reconstructed; the matrix

$\mathbf{M}_k$ does not contain scattered light and has a reduced noise level compared to the original matrix **M**. The FORTRAN program svd_exem for iterative SVD of excitation-emission matrices had been used to separate contributions of unknown fluorophores in a mixture[8] and it was also used here. In the analysis of experimental excitation-emission matrices obtained with photoactivated 7F-Trp-CypA, the matrix elements that contained elastic scattered light passing through the emission monochromator due to the first- and second-order diffraction were treated as the affected ones; Raman scattering peaks were extremely small compared to fluorescence intensity and therefore did not need to be taken into consideration. Iterative SVD was carried out using the assumed effective ranks l=3, 4, and 5; the final results obtained with l=k were retained for further analysis, were k=4 for the enriched/purified CL-CypA and k=5 for the non-purified photoactivated 7F-Trp-CypA. The first eight singular values for each sample are shown in Table S4. Singular values below 0.4 were considered insignificant.

Transition from SVD basis vectors to excitation and emission spectra. Vibrational relaxation of fluorophores at room temperature takes picoseconds. The lifetimes of the singlet electronic excited states responsible for fluorescence are in the nanosecond range. This demonstrates that practically all fluorescence emission originates from the relaxed excited state. Therefore, for a single fluorophore, the fluorescence emission spectrum is independent of the excitation wavelength, and the fluorescence excitation spectrum is independent of the emission wavelength. In accordance with this physical model, a reconstructed excitation-emission matrix $\mathbf{M}_k$ of reduced rank k can be represented as a product **FG**, where the columns of an $m\times q$ matrix **F** represent fluorescence emission spectra of q fluorophores, and the rows of an $q\times n$ matrix **G** represent fluorescence excitation spectra of the same q fluorophores. Note, that if $q<k$, then $\mathbf{M}_k$ cannot be equal to **FG**, because matrix $\mathbf{M}_k$ is of rank k, while the rank of the product **FG** cannot exceed q. On the other hand, if $q>k$, then the matrix equation $\mathbf{M}_k$=**FG** will always have an infinite number of solutions, rendering it of little practical use. Therefore, one has to postulate that the number of detectable fluorophores is equal to the effective rank of the excitation-emission matrix, i. e. q=k. Thus, the number of detectable fluorophores equals k, the columns of the $m\times k$ matrix **F** represent their emission spectra, and the rows of the $k\times n$ matrix **G** represent their excitation spectra. Next, one has to calculate the unknown matrices **F** and **G,** given that the matrix $\mathbf{M}_k$=$\mathbf{US}_k\mathbf{V}^T$, as well as the matrices **U**, $\mathbf{S}_k$, **V**, are known. In the most general case the relationship between the unknown matrices **F**, **G** and the known matrices **U**, $\mathbf{S}_k$, **V** can be expressed in the form

$\mathbf{F}=\mathbf{U}'\mathbf{Q}$, $\mathbf{G}=\mathbf{R}\mathbf{V}'^{\mathrm{T}}$, where $\mathbf{Q}$ and $\mathbf{R}$ are two regular $k\times k$ matrices that satisfy the equation $\mathbf{QR}=\mathbf{S}'$, and $\mathbf{U}'$, $\mathbf{S}'$, $\mathbf{V}'$ are truncated versions of the matrices $\mathbf{U}$, $\mathbf{S}_k$, $\mathbf{V}$, in which the number of columns (and in the case of $\mathbf{S}'$ the number of rows also) is reduced from r to k. In general, the matrices $\mathbf{Q}$ and $\mathbf{R}$ can be completely arbitrary as long as they satisfy the equation $\mathbf{QR}=\mathbf{S}'$. Constraints are required to limit the choice of possible matrices $\mathbf{Q}$ and $\mathbf{R}$. The requirement that the excitation and emission spectra of the k fluorophores cannot have negative values represents a very strong constraint, according to which all elements of the matrices $\mathbf{F}$ and $\mathbf{G}$ must be non-negative. Unfortunately, this very powerful constraint is incompatible with noisy data. It therefore is necessary to introduce a penalty function for negative matrix elements and select the matrices $\mathbf{Q}$ and $\mathbf{R}$ that yield the lowest possible penalty. The following penalty function produces the most physically meaningful results:

$$\text{Penalty} = \frac{1}{2}\sum_{l=1}^{k}\sum_{i=1}^{m}\sum_{j=1}^{n}(|f_{il}| - f_{il})\,|g_{lj}| + |f_{il}|(|g_{lj}| - g_{lj})$$

Note, that the penalty defined in this way cannot be negative and it equals zero when and only when all elements of the matrices $\mathbf{F}$ and $\mathbf{G}$ are non-negative. Furthermore, the penalty does not change when the emission spectrum of any fluorophore is multiplied by any positive factor and the corresponding excitation spectrum is divided by the same factor; fluorescence intensity at any combination of excitation and emission wavelengths remains unchanged after such scaling, therefore the penalty should not change either. An iterative process was employed to find the matrices $\mathbf{Q}$ and $\mathbf{R}$ that resulted in the lowest penalty. For the zeroth approximation we took two identical diagonal matrices $\mathbf{Q}^{(0)}=\mathbf{R}^{(0)}=(\mathbf{S}')^{1/2}$, the diagonal elements of which equal square roots of the corresponding diagonal elements of the matrix $\mathbf{S}'$. The i-th iteration consists of defining a $k\times k$ matrix $\mathbf{P}^{(i)}$, all diagonal elements of which equal 1, while the off-diagonal elements are random numbers evenly distributed between -0.1 and +0.1. After calculating new matrices $\mathbf{Q}^{(i)}=\mathbf{Q}^{(i-1)}\mathbf{P}^{(i)}$, $\mathbf{R}^{(i)}=(\mathbf{Q}^{(i)})^{-1}\mathbf{S}'$, $\mathbf{F}^{(i)}=\mathbf{U}'\mathbf{Q}^{(i)}$, $\mathbf{G}^{(i)}=\mathbf{R}^{(i)}\mathbf{V}'^{\mathrm{T}}$, and calculating the new penalty, the new penalty is compared to the lowest of previously observed penalties. If the new penalty is less than the lowest previously observed, then the new matrices $\mathbf{Q}^{(i)}$, $\mathbf{R}^{(i)}$, $\mathbf{F}^{(i)}$, $\mathbf{G}^{(i)}$ are retained and become the input for the next iteration. If the new penalty is greater or equal to the lowest, then the matrices are reverted to the state they were in when the lowest penalty was observed. This is a brute force approach since the algorithm relies on random matrices $\mathbf{P}^{(i)}$ in its search. If a zero penalty is reached after a certain number of iterations, then the solution is not unique, i. e. different combinations of shapes of fluorophore spectra are possible that result in equally good fits to the data matrix. If the

penalty remains positive after tens of millions of iterations, the calculation is stopped since the solution is unique with the shapes of the excitation and emission spectra of all fluorophores uniquely defined. We observed the latter with both enriched/purified (k=4) and non-purified (k=5) photoactivated 7F-Trp-CypA samples.

Spectral correction. The excitation and emission spectra of all fluorophores were corrected for instrument-dependent wavelength sensitivity variation and for the primary and secondary inner filter effect[9]. These corrections were only carried out after iterative SVD and the transition from basis vectors to non-negative excitation and emission spectra of the actual fluorophores. If corrections are performed prior to SVD, this significantly changes the random error statistics, yielding larger standard deviations for some elements of the excitation-emission matrix **M** than for the other elements. Therefore we performed (1) iterative SVD, (2) transition from basis vectors to excitation and emission spectra, (3) spectral correction, and (4) re-balancing.

Re-balancing excitation and emission spectra. Fluorescence intensity at any combination of excitation and emission wavelengths remains unchanged when the emission spectrum of a fluorophore is multiplied by any positive factor and the corresponding excitation spectrum is divided by the same factor, which implies that even if the shapes of the excitation and emission spectra are uniquely defined, the absolute numbers still remain ambiguous. Although this ambiguity has no physical significance, its removal is desirable in order to derive completely reproducible numerical results. To this end the integral under the excitation spectrum and under the emission spectrum of each fluorophore (trapezoidal numerical integration) is calculated, as well as the square root of the ratio of the two integrals. The emission spectrum is multiplied by the square root and the excitation spectrum is divided by the square root. This re-balancing operation is carried out separately for each fluorophore. After this re-balancing the integrals under the excitation and the emission spectrum of every fluorophore are equal to each other, and the ambiguity is removed. Re-balancing is carried out after the spectral correction since spectral correction perturbs the balance between the two integrals.

TCSPC experiments. Time-Correlated Single Photon Counting (TCSPC) data were obtained at 20°C using a custom-built instrument with tunable picosecond laser excitation described previously[9]. The overall time resolution of the instrument is ~65 ps (FWHM), the excitation pulse duration is ≤11 ps (FWHM), the excitation pulse rate is 4,100,000±50 pulses per second. The mean excitation power was 9.2±0.4 μW measured at the fluorescence cell, the

excitation wavelength was 305 nm, the excitation laser bandwidth was <0.01 nm, with vertical excitation polarization and emission polarizers at 55° from vertical (magic angle). The emission monochromator bandwidths were 8 nm (FWHM) and TCSPC data were acquired at 41 emission wavelengths from 325 nm to 525 nm at 5 nm wavelength steps. To avoid pile-up problems, emission photon counting rate was less than 0.5% of the excitation pulse rate, i. e., 20,500 photons per second per one detection wing. The instrument has two identical detection wings working in parallel. Multichannel analyzers comprised 2048 total channels, 1926 of which were completely free of systematic errors and therefore used, the timing calibration was 13.298±0.001 ps/channel. Emission photons were counted by each wing for 200 s per emission wavelength. TCSPC decay curves collected near the peak emission wavelength typically had ~5 million integral photon counts and ~23 thousand peak photon counts. A separate "lamp curve", representing the impulse response function of the whole system was recorded quasi-simultaneously with the fluorescence decay curve at each of the 41 emission wavelengths, using 40 s of photon counting per wavelength and included in the data analysis. Two data sets were obtained: one with the enriched/purified CL-CypA sample (20.4 μM), one with the non-purified photoactivated 7F-Trp-CypA mixture (50 μM).

TCSPC data analysis. The TCSPC data obtained at all wavelengths were globally analyzed using the program l_global described previously[9]. Since it was known a priori that both samples (enriched/purified and non-purified) contained the same fluorophores, but in different proportions, both data sets were analyzed together. In the global fitting, the lifetimes τ_n were treated as global parameters (they were linked across all emission wavelengths and both data sets), whereas the pre-exponential factors α_n were treated as local parameters (they had individual values for every emission wavelength and every data set). Global fits were attempted with different numbers of exponentials N_{exp} between 1 and 7. For N_{exp}=7 the global χ^2 minimization algorithm failed to converge (singular Hessian matrix), irrespective of initial guesses. The reduced χ^2 values for N_{exp} between 1 and 6 are provided in Table S5. For 2×41×1926=157932 degrees of freedom, the desired reduced χ^2 value should lie between 0.9908 and 1.0092 with 99% probability; none of the values in Table S5 fell in this range. Therefore, manual inspection of the weighted residuals and autocorrelations for the 2×41 decay curves was performed and revealed a systematic error resulting from the presence of Raman scattered light at wavelengths from 330 nm to 350 nm (peak at 340 nm). The extremely narrow peak of the autocorrelation function suggested that exponentials with τ>65 ps were not affected by this systematic error. Given that all lifetimes determined for the

fluorophores were significantly greater than 65 ps, we are confident that this systematic error did not affect our results. Increasing N_{exp} by 1 adds 83 new fitting parameters (1 global + 2×41 local). Therefore, a decrease in the reduced χ^2 upon the addition of one exponential must exceed 0.0015 in order to be statistically significant (95% confidence). This implies that increasing N_{exp} from 1 to 5 is statistically justified, while a further increase of N_{exp} from 5 to 6 is not. The values of the five lifetimes are: τ_1=0.228±0.011 ns, τ_2=1.223±0.036 ns, τ_3=2.384±0.008 ns, τ_4=3.801±0.019 ns, τ_5=18.15±0.39 ns.

Intensity spectra associated with the five exponentials. Steady-state emission spectra of both enriched/purified and non-purified samples excited at 305 nm were measured using a SLM-48000 spectrofluorometer. The emission spectra were corrected for the wavelength variation in instrument sensitivity and for the inner filter effect as previously described[9]. The program trspectr[9] was used to combine the corrected steady-state emission spectra with the results of the global fitting of the TCSPC data. Outputs of the program trspectr are shown in Figure 2F and S5B. Intensities obtained directly from the global analysis of TCSPC data by re-normalizing $\tau_n\alpha_n$ values using steady-state emission intensities are shown in filled circles. The smooth lines were obtained by polynomial fitting as described previously[9]. Note, all spectra associated with lifetimes τ_1=0.228 ns and τ_2=1.22 ns go negative at wavelengths above 404 nm and 412 nm, respectively, but only for the enriched/purified sample (bottom panel of Figure S5E). This may be the result of a time-dependent red shift (TDRS) in the emission of one or more of the major fluorophores with the main lifetimes τ_3=2.38 ns and/or τ_4=3.80 ns. It is possible, therefore, that the two shortest lifetimes do not represent two separate fluorophores, but rather represent the relaxation of the electrostatic interactions around the fluorophores with the lifetimes τ_3 and/or τ_4. It is also possible that τ_1 and τ_2 represent unresolved contributions from minor impurities plus the relaxation of the major fluorophores. However, the main lifetimes τ_3=2.38 ns, τ_4=3.80 ns, and τ_5=18.1 ns are unequivocally associated with three different fluorophores.

QY measurement. The quantum yield of enriched/purified crosslinked CypA was measured at 20°C using a 305 nm exciting wavelength. A solution of 2,5-diphenyl-1,3-oxazole (PPO) in cyclohexane was employed as a quantum yield reference. The latter possesses a very high quantum yield (near 100%)[10] that varies little with temperature and other experimental conditions, absorbs light at the chosen exciting wavelength, and shows two main emission peaks

with wavelengths of 357 nm and 374 nm, close to the peak emission wavelength of enriched/purified crosslinked CypA (370 nm). Steady-state emission spectra of the sample and the reference were measured using a SLM-48000 spectrofluorometer with polarizers at the magic angle configuration. From each emission spectrum the corresponding solvent spectrum was subtracted. Spectra were corrected for instrument-dependent wavelength sensitivity variation and for the primary and secondary inner filter effects. The corrected spectra were numerically integrated (trapezoidal integration) over the wavelength range from 313 nm to 590 nm. Quantum yields were calculated using the equation[11]

$$\eta_S = \eta_R \frac{I_S}{I_R} \frac{A_R}{A_S} \frac{n_S^2}{n_R^2}$$

where η_S and η_R denote the quantum yields, I_S and I_R the integrals under the corrected emission spectra, A_S and A_R the absorbances at the exciting wavelength, n_S and n_R the refractive indices of the solvents, and the subscripts S and R refer to the sample and the reference. The integrals I_S and I_R were obtained as described above. Absorbances A_S and A_R were measured against the corresponding solvent in the reference path using a double-beam Shimadzu spectrophotometer. Refractive indices at 20°C and 370 nm of water n_S=1.3468 and cyclohexane n_R=1.4450 were taken from the literature[12-13]. The quantum yield of the enriched/purified CL-CypA sample was determined as 0.395±0.003. The quantum yield of the non-purified photoactivated 7F-Trp-CypA sample measured using the same reference measured as 0.278±0.002. The quantum yield of the non-purified sample was also measured with a solution of p-terphenyl in cyclohexane (not a perfect wavelength match) as a reference and yielded 0.276±0.002.

NMR spectroscopy

Samples of U-^{15}N, 7F-Trp-labeled CypA (before and after UV radiation) contained ~100 µM protein in NMR buffer (25 mM sodium phosphate, pH 6.5, 1 mM TCEP, 0.02% NaN_3, and 7% D_2O). The enriched/purified sample of U-^{13}C,^{15}N-labeled CL-CypA contained ~390 µM protein in the same NMR buffer. All NMR experiments were carried out at 25°C. Backbone and sidechain resonance assignments were performed using a set of 2D and 3D experiments, including 2D ^{1}H-^{15}N HSQC, 2D ^{1}H-^{13}C HSQC, 2D ^{1}H-^{13}C aromatic TROSY, 3D HNCA, 3D HNCO, 3D HN(CO)CA, 3D H(CCO)NH, and 3D C(CO)NH, which were collected on a 14.1 T Bruker AVANCE III HD spectrometer, equipped with a 5 mm CP TCI H-C/N-D triple-resonance, z-axis

gradient cryoprobe. ^{1}H-^{1}H NOE restraints were extracted from 3D simultaneous ^{13}C- and ^{15}N-edited NOESY experiments collected with mixing time of 100 ms on a 21.1 T Bruker AVANCE II spectrometer, equipped with a 5 mm CP TCI H-C/N-D triple-resonance, z-axis gradient cryoprobe. ^{19}F NMR experiments were acquired on a 14.1 T Bruker AVANCE II spectrometer, equipped with a 5 mm CP TXO F/C–H–D triple-resonance, z-axis gradient cryoprobe. ^{1}H and ^{19}F chemical shifts were referenced with respect to DSS (4,4-dimethyl-4-silapentane-1-sulfonic acid) and TFA (trifluoroacetic acid). All data were processed with either TopSpin (Bruker) or NMRPipe[14] and analyzed in CcpNmr Analysis V2[15].

NMR structure calculation

A total of 2611 intramolecular ^{1}H-^{1}H NOEs were manually assigned unambiguously from the 3D ^{1}H-^{15}N NOESY-HSQC (2330 NOEs) and 3D ^{1}H-^{13}C aromatic NOESY-HSQC (281 NOEs) spectra. All NOE peaks were initially binned into three groups (i.e., strong, medium, and weak) based on peak intensities. The bounds of the distance restraints were set as following: (1) in 3D ^{1}H-^{15}N NOESY-HSQC, 1.7-3.1 Å for strong, 1.7-4.7 Å for medium, and 1.7-6.0 Å for weak; (2) in 3D ^{1}H-^{13}C aromatic NOESY-HSQC, 1.7-3.1 Å for strong, 1.7-3.7 Å for medium, and 1.7-6.0 Å for weak. Violated NOEs in the initial calculation were manually inspected, error-prone NOEs were discarded, and incorrect distance bounds were adjusted. In addition, a total of 262 backbone dihedral restraints in the form of ϕ/ψ angles were generated by TALOS-N[16] predictions based on experimental chemical shifts.

Structure calculations were performed in Xplor-NIH version 2.52[17-18]. A topology file for para-Phe-Trp crosslink was generated by jointing Cζ of Phe and Cζ2 of Trp with a bond length of 1.54 Å, while maintaining the topology and planarity of the aromatic rings. The crystal structure of 7F-Trp-CypA was used as the initial model and edited with the para-Phe-Trp topology. An equivalent model was generated using the WT CypA crystal structure. 200 structures were calculated by simulated annealing in torsion angle space, from 3000 K to 25 K in steps of 12.5 K, followed by a final gradient minimization in Cartesian space, using 2611 NOE restraints and 262 torsion angle restraints. The final ensemble contains the 10% (20 out of 200) lowest-energy structures. RMSD values were calculated using routines in Xplor-NIH version 2.52[17-18]. Visualization of structures and ensembles were generated in PyMOL[19].

PDB search

A search of all protein structures in the PDB for proteins with Trp-Phe pairs that resemble the local configuration in CypA was performed by the Hoch laboratory using python scripts adapted from previously published codes[20]. Two criteria were applied in the search: (1) the distance between the center of the Trp six-membered ring to the center of the Phe ring was restricted to lie within 4-6 Å; (2) the angle defined by three points, from the center of the Phe ring to the center of the six-membered Trp ring to the C7 atom of the Trp, was restricted within 0-10 degrees to ensure that the Phe ring approaches the Trp from the direction of the C7 position as seen in CypA. For reference, the W121-F60 pair in CypA has a defined distance of 5.8 Å and a defined angle of 9.2° (Fig. S8A and S8B). A total of 589 entries with Trp-Phe pair(s) satisfying the above criteria were obtained in the search.

References

(1) Liu, C.; Perilla, J. R.; Ning, J. Y.; Lu, M.; Hou, G.; Ramalho, R.; Himes, B. A.; Zhao, G.; Bedwell, G. J.; Byeon, I. J.; Ahn, J.; Gronenborn, A. M.; Prevelige, P. E.; Rousso, I.; Aiken, C.; Polenova, T.; Schulten, K.; Zhang, P., Cyclophilin A Stabilizes the Hiv-1 Capsid through a Novel Non-Canonical Binding Site. *Nat. Commun.* **2016,** *7, 10714*.

(2) Lu, M.; Ishima, R.; Polenova, T.; Gronenborn, A. M., ^{19}F NMR Relaxation Studies of Fluorosubstituted Tryptophans. *J. Biomol. NMR* **2019,** *73* (8-9), 401-409.

(3) Lu, M.; Hou, G.; Zhang, H.; Suiter, C. L.; Ahn, J.; Byeon, I. J. L.; Perilla, J. R.; Langmead, C. J.; Hung, I.; Gor'kov, P. L.; Gan, Z. H.; Brey, W.; Aiken, C.; Zhang, P. J.; Schulten, K.; Gronenborn, A. M.; Polenova, T., Dynamic Allostery Governs Cyclophilin A-HIV Capsid Interplay. *Proc. Natl. Acad. Sci. U. S. A.* **2015,** *112* (47), 14617-14622.

(4) Shi, Y.; Pellarin, R.; Fridy, P. C.; Fernandez-Martinez, J.; Thompson, M. K.; Li, Y. Y.; Wang, Q. J.; Sali, A.; Rout, M. P.; Chait, B. T., A Strategy for Dissecting the Architectures of Native Macromolecular Assemblies. *Nat. Methods* **2015,** *12* (12), 1135-1138.

(5) Xiang, Y. F.; Sang, Z.; Bitton, L.; Xu, J. Q.; Liu, Y.; Schneidman-Duhovny, D.; Shi, Y., Integrative Proteomics Identifies Thousands of Distinct, Multi-Epitope, and High-Affinity Nanobodies. *Cell Syst.* **2021,** *12* (3), 220-234.

(6) Golub, G. H.; Reinsch, C., Singular Value Decomposition and Least Squares Solutions. *Numer. Math.* **1970,** *14* (5), 403-420.

(7) Chan, T. F., An Improved Algorithm for Computing the Singular Value Decomposition. *ACM Trans. Math. Softw.* **1982,** *8* (1), 72-83.

(8) Brown, B. C.; Brice, S.; Hawker, E.; Maza, S.; Meyer, H.; Pla-Dalmau, A.; Tayloe, R.; Tanaka, H. A.; Toptygin, D. *Study of Scintillation, Fluorescence and Scattering in Mineral Oil for the Miniboone Neutrino Detector*, IEEE Symposium Conference Record Nuclear Science 2004., 16-22 Oct. 2004; 2004; pp 652-656 Vol. 1.

(9) Toptygin, D., Analysis of Time-Dependent Red Shifts in Fluorescence Emission from Tryptophan Residues in Proteins. *Methods Mol. Biol.* **2014,** *1076*, 215-56.

(10) Berlman, I. B., *Handbook of Fluorescence Spectra of Aromatic Molecules*. Academic Press: New York, 1966.

(11) Lakowicz, J. R., *Principles of Fluorescence Spectroscopy*. Kluwer Acad./Plenum Publ.: New York, NY, 1999.

(12) Kerl, K.; Varchmin, H., Refractive-Index Dispersion (RID) of Some Liquids in the UV/VIS Between 20°C and 60°C. *J. Mol. Struct.* **1995,** *349*, 257-260.

(13) Daimon, M.; Masumura, A., Measurement of the Refractive Index of Distilled Water from the near-Infrared Region to the Ultraviolet Region. *Appl. Opt.* **2007,** *46* (18), 3811-3820.

(14) Delaglio, F.; Grzesiek, S.; Vuister, G. W.; Zhu, G.; Pfeifer, J.; Bax, A., NMRPipe - a Multidimensional Spectral Processing System Based on Unix Pipes. *J. Biomol. NMR* **1995,** *6* (3), 277-293.

(15) Vranken, W. F.; Boucher, W.; Stevens, T. J.; Fogh, R. H.; Pajon, A.; Llinas, P.; Ulrich, E. L.; Markley, J. L.; Ionides, J.; Laue, E. D., The CCPN Data Model for NMR Spectroscopy: Development of a Software Pipeline. *Proteins* **2005,** *59* (4), 687-696.

(16) Shen, Y.; Bax, A., Protein Structural Information Derived from NMR Chemical Shift with the Neural Network Program TALOS-N. *Methods Mol. Biol.* **2015,** *1260*, 17-32.

(17) Schwieters, C. D.; Kuszewski, J. J.; Tjandra, N.; Clore, G. M., The Xplor-NIH NMR Molecular Structure Determination Package. *J. Magn. Reson.* **2003,** *160* (1), 65-73.

(18) Schwieters, C. D.; Kuszewski, J. J.; Clore, G. M., Using Xplor-NIH for NMR Molecular Structure Determination. *Prog. Nucl. Magn. Reson. Spectrosc.* **2006,** *48* (1), 47-62.
(19) *The Pymol Molecular Graphics System.* Schrödinger, **2000**, Version 2.0.
(20) Baskaran, K.; Wilburn, C. W.; Wedell, J. R.; Koharudin, L. M. I.; Ulrich, E. L.; Schuyler, A. D.; Eghbalnia, H. R.; Gronenborn, A. M.; Hoch, J. C., Anomalous Amide Proton Chemical Shifts as Signatures of Hydrogen Bonding to Aromatic Sidechains. *Magn. Reson.* **2021,** 2 (2), 765-775.

Supplementary Figures

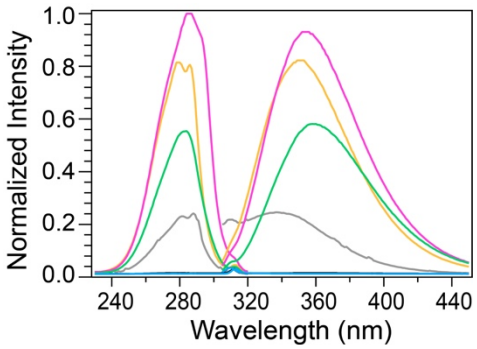

Figure S1. Fluorescence spectra of Trp, F-Trps, and wild-type CypA. Excitation and emission spectra of Trp (yellow), 4F-Trp (black), 5F-Trp (magenta), 6F-Trp (green), 7F-Trp (blue), and wild-type CypA (gray) are superimposed. Excitation spectra were measured with an emission wavelength of 350 nm. Emission spectra were measured with an excitation wavelength of 280 nm. Measurements were performed on 10 μM samples in aqueous 25 mM NaPhosphate buffer, pH 6.5 at room temperature on a PTI fluorometer.

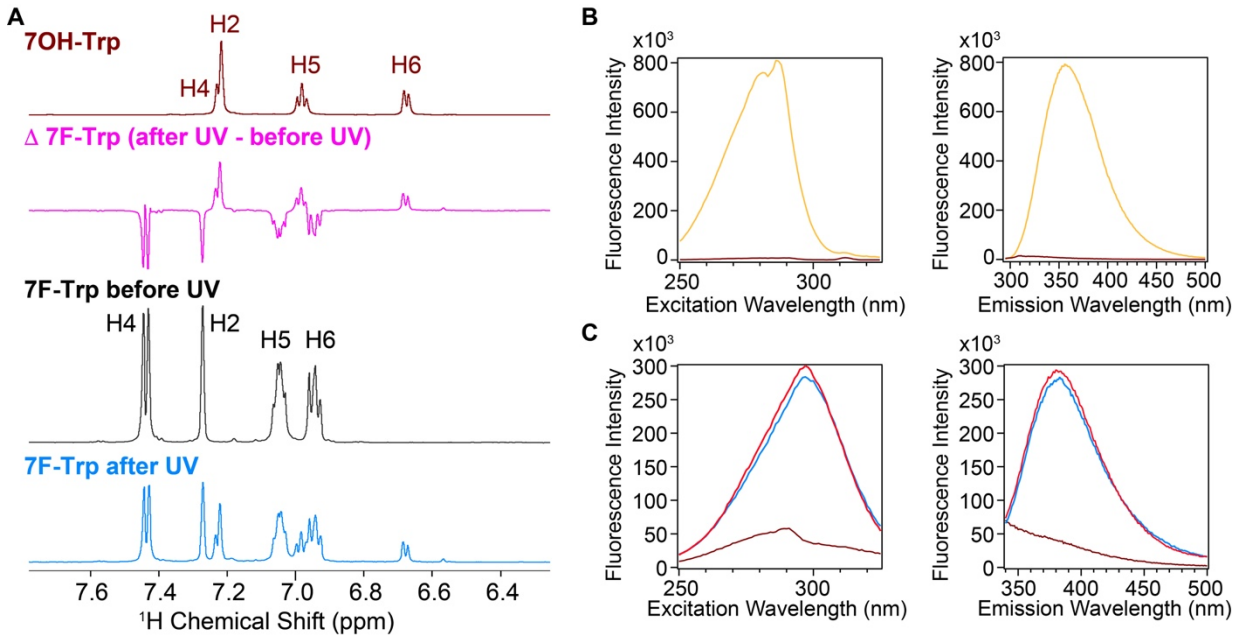

Figure S2. Generation of 7OH-Trp by UV radiation. (A) ^{1}H NMR spectra of 7F-Trp before (black) and after (blue) UV radiation, with the difference spectrum shown in magenta. The ^{1}H NMR spectrum of 7OH-Trp is shown in maroon. The difference spectrum establishes that 7F-Trp disappears and 7OH-Trp is generated during UV radiation. (B) Excitation and emission spectra of Trp (yellow) and 7OH-Trp (maroon). Excitation spectra were measured with an emission wavelength of 350 nm. Emission spectra were measured with an excitation wavelength of 280 nm. These fluorescence spectra were collected on a FluoroMax-4 fluorometer with 2 nm slit widths. (C) 7OH-Trp before (maroon) and after (red) UV radiation, and 7F-Trp after UV radiation (blue). These fluorescence spectra were collected on a FluoroMax-4 fluorometer with 4 nm slit widths. Note that 7OH-Trp is virtually nonfluorescent but can generate a low fluorescent species upon UV radiation. The fluorescence spectra of 7F-Trp and 7OH-Trp after UV radiation are virtually the same, indicating that identical products were generated.

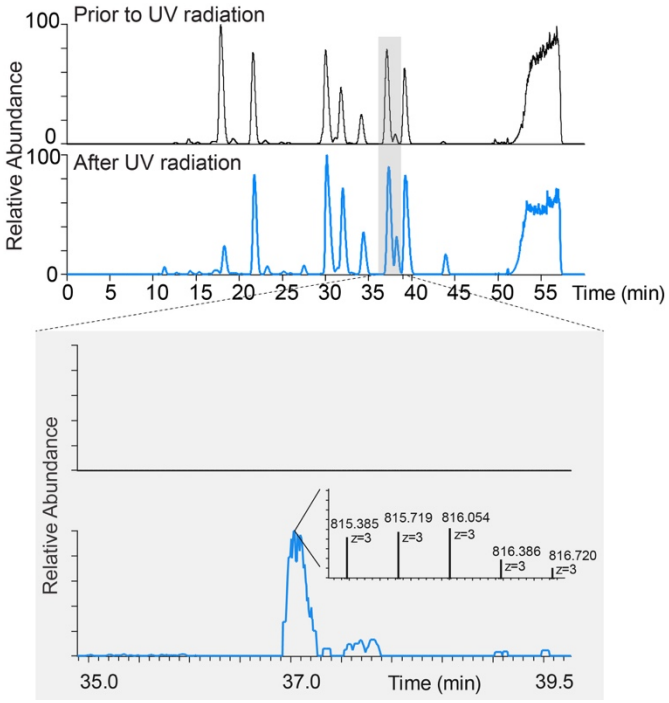

Figure S3. Identification of the crosslinked tryptic peptide. nLC-MS chromatography and spectra of trypsin digested 7F-Trp-CypA prior to (black) and after (blue) UV radiation. Note that the tryptic peptide with the crosslink (theoretical mass, 2444.7781 Da) is only observed in the sample after UV radiation.

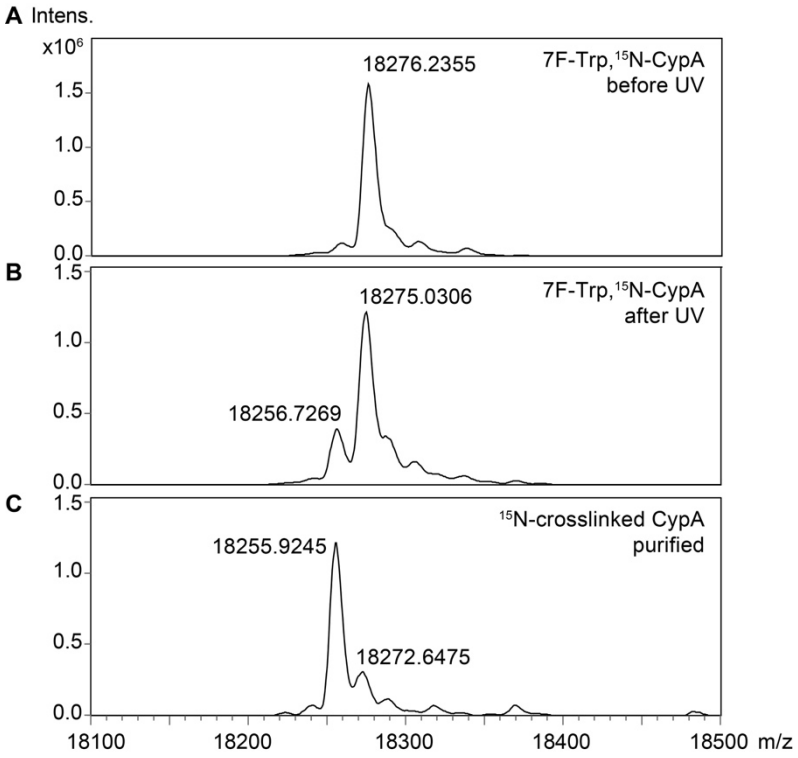

Figure S4. Determination of molecular masses by ESI-MS. (A) Deconvoluted ESI-MS spectrum of U-^{15}N, 7F-Trp-CypA before UV radiation. The theoretical mass is 18275.83 Da, indicating a mass error of 22 ppm. (B) Deconvoluted ESI-MS spectrum of U-^{15}N, 7F-Trp-CypA after UV radiation. (C) Deconvoluted ESI-MS spectrum of enriched/purified U-^{15}N-crosslinked CypA. The theoretical mass is 18255.82 Da, indicating a mass error of 6 ppm.

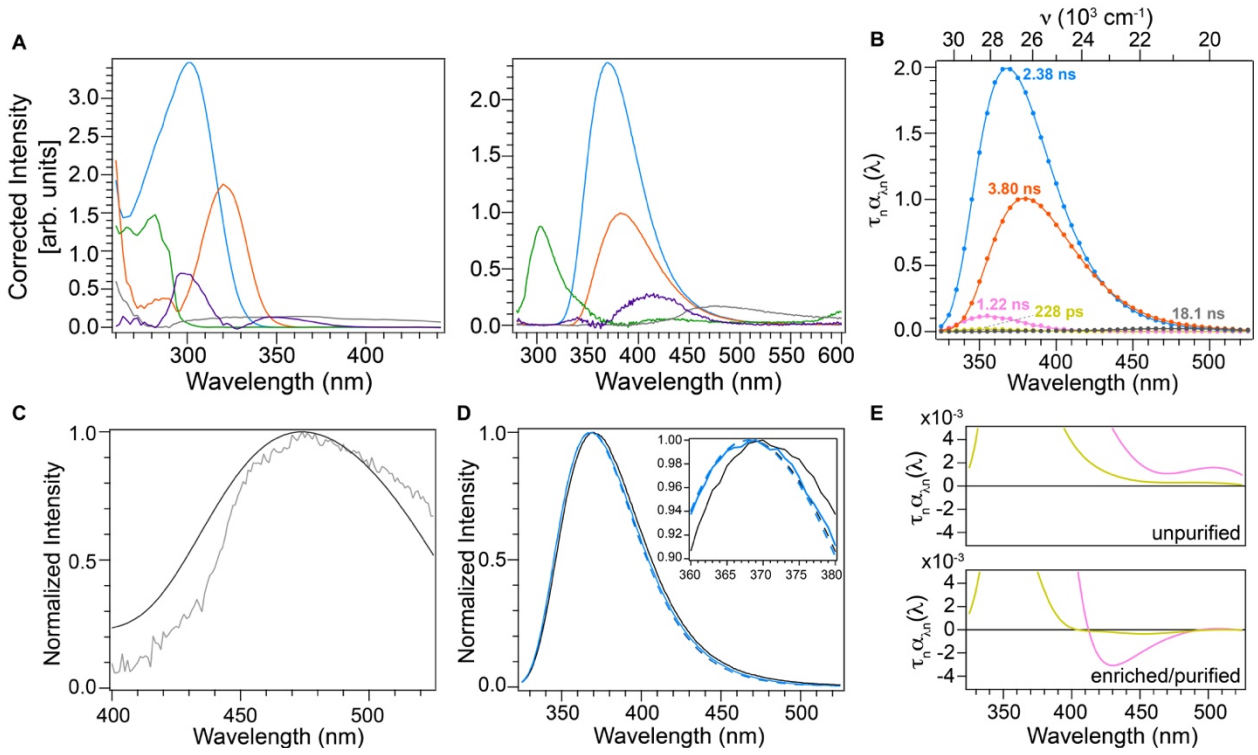

Figure S5. Fluorescence characterization of 7F-Trp-CypA after photoactivation. (A) Fluorescence excitation (left) and emission (right) spectra derived by SVD of the excitation-emission matrix measured for the non-purified UV-radiated mixture. Five distinct fluorophores (I-V) were identified and are colored in blue, orange, green, purple, and gray, respectively. (B) Corrected emission intensity spectra associated with the five exponential terms determined in TCSPC measured for the non-purified UV-radiated mixture. Spectra corresponding to lifetimes of 2.38 ns, 3.80 ns, 1.22 ns, 288 ps, and 18.1 ns are colored in blue, orange, pink, sand, and gray. Intensities obtained by re-normalizing $\tau_n\alpha_n$ values using steady-state emission intensities are shown by filled circles. Smooth lines were generated using degree 8 polynomials. (C) Normalized fluorescence emission spectra of the red-emitting minor species (fluorophore V) in the UV-radiated mixture, extracted by SVD (gray line, peak wavelength 474 nm) and TCSPC (black line, τ =18.1±0.39 ns). (D) Normalized fluorescence emission spectra of crosslinked CypA, extracted by SVD of non-purified (solid black line) and enriched/purified (solid blue line) samples, and TCSPC experiments with non-purified (dashed black line) and enriched/purified (dashed blue line) samples. An expanded view across the top of the spectra is shown in the inset. (E) Expanded view of TCSPC spectra of unpurified photoactivated mixture (top) and enriched/purified CL-CypA (bottom). Spectra corresponding to lifetimes of 1.22 ns and 288 ps are colored in pink and sand, respectively.

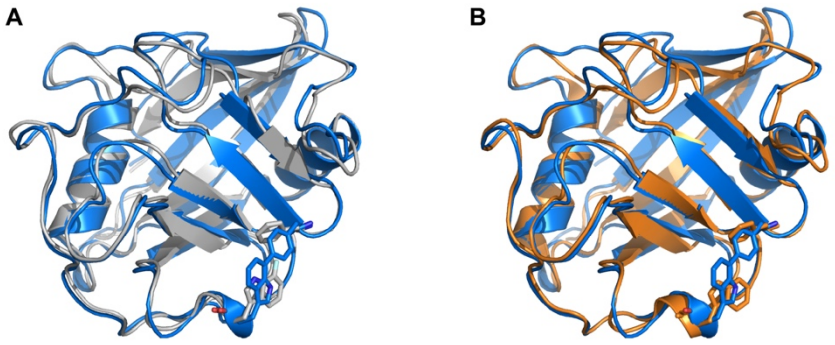

Figure S6. Superpositions of the structures of crosslinked CypA, 7F-Trp-CypA, and WT-CypA. Structures are shown in ribbon representation. Sidechains of residue 60F and 121W are shown in stick representations. (A) Coordinates of the 7F-Trp-CypA crystal structure (gray) and those of the lowest-energy NMR structure of crosslinked CypA (blue) were best-fit. (B) Coordinates of the WT-CypA crystal structure (orange) and those of the lowest-energy NMR structure of crosslinked CypA (blue) were best-fit.

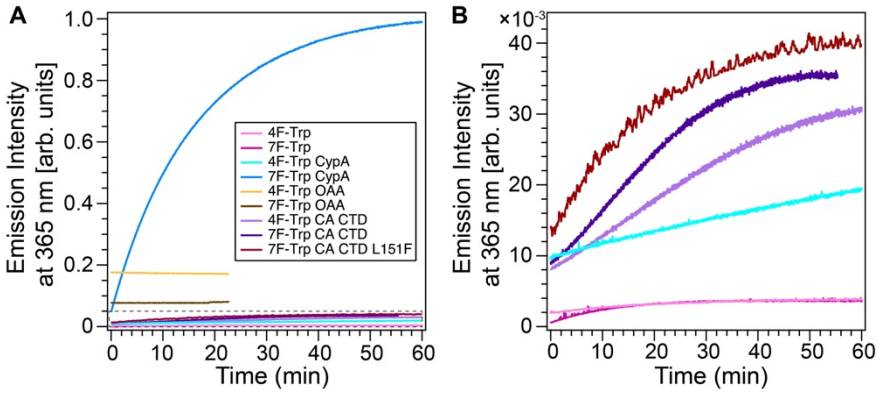

Figure S7. Photoactivation of free amino acids 4F- and 7F-Trp and 4F- and 7F-Trp containing proteins. (A) Kinetics of photoactivation of 4F-Trp (light pink), 7F-Trp (dark pink), 4F-Trp- CypA (cyan), 7F-Trp- CypA (blue), 4F-Trp-OAA (yellow), 7F-Trp-OAA (brown), 4F-Trp-CA CTD (light purple), 7F-Trp-CA CTD (dark purple), and 7F-Trp-CA CTD L151F (maroon). The low-intensity area enclosed by dashed lines is enlarged in (B). Note, only 7F-Trp-CypA yielded a highly fluorescent species upon UV radiation.

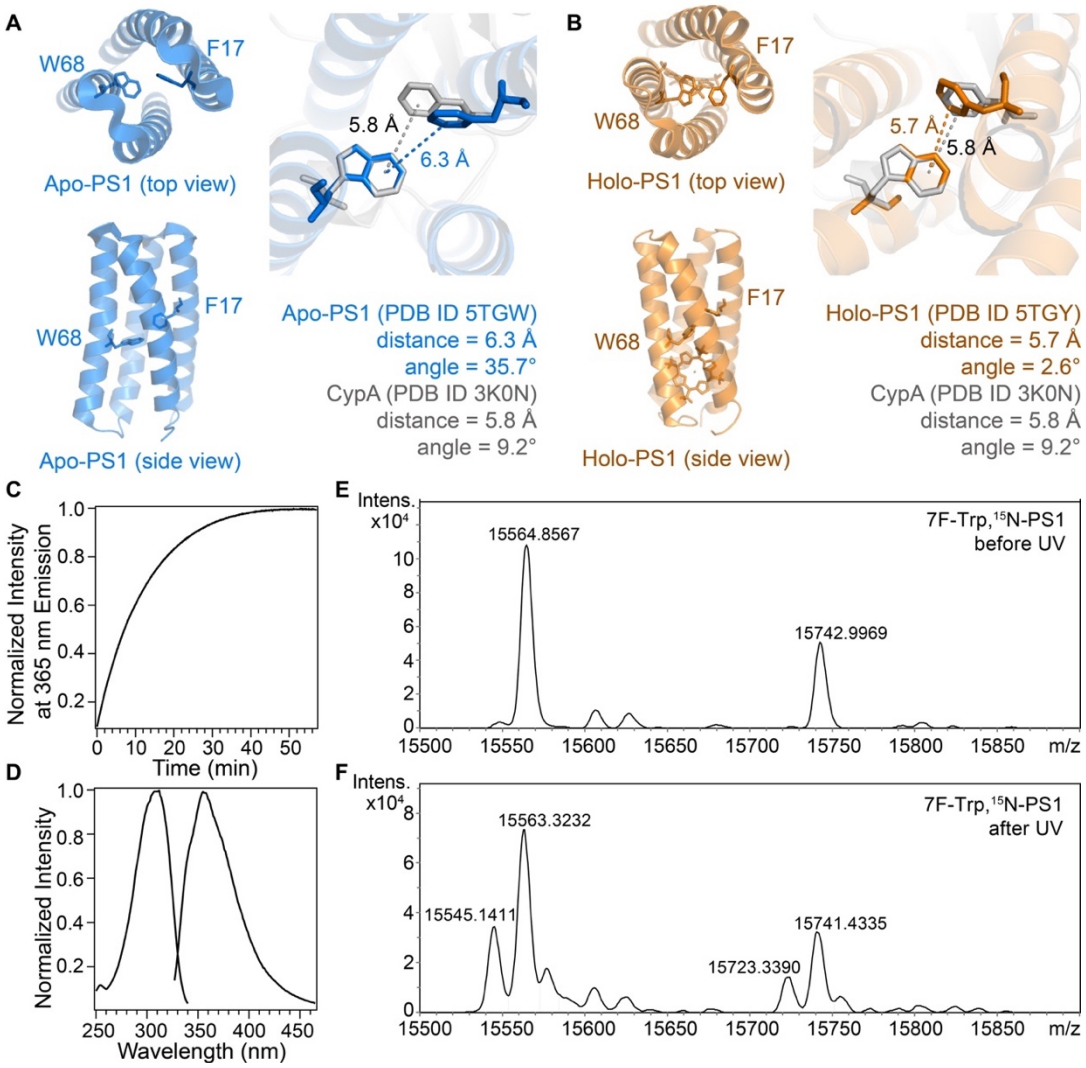

Figure S8. Photoactivation of U-^{15}N, 7F-Trp-apo-PS1. (A) Two views of the structure of apo PS1 in ribbon representation (blue), with residues W68 and F17 in stick representation. A superposition of W68-F17 pair in apo-PS1 (blue) and the W121-F60 pair in CypA (gray) is shown at the right with the distances between the Trp and Phe rings indicated by the dotted lines. (B) Two views of the structure of holo PS1 in ribbon representation (orange), with residues W68 and F17 in stick representation. A superposition of the W68-F17 pair in holo-PS1 (orange) and the W121-F60 pair in CypA (gray) is shown at the right. Distance and angle were defined as described in Materials and Methods. (C) Time-dependent fluorescence of U-^{15}N, 7F-Trp-apo-PS1 upon photoactivation at 282 nm. (D) Excitation and emission spectra of U-^{15}N, 7F-Trp-apo-PS1 after photoactivation. (E) Deconvoluted ESI-MS spectra of U-^{15}N, 7F-Trp-apo-PS1 before UV illumination. The PS1 protein contains a N-terminal His tag followed by a TEV cleavage site. The peak at 15743.0 Da indicates a minor species with gluconoylation (+178 Da), which is commonly seen in His tagged proteins. (F) Deconvoluted ESI-MS spectra of U-^{15}N, 7F-Trp-apo-PS1 after UV illumination. Note the new peak at 15545.1 Da (~ 20 Da less than the main mass peak before UV), indicating that a crosslinked species was generated after photoactivation.

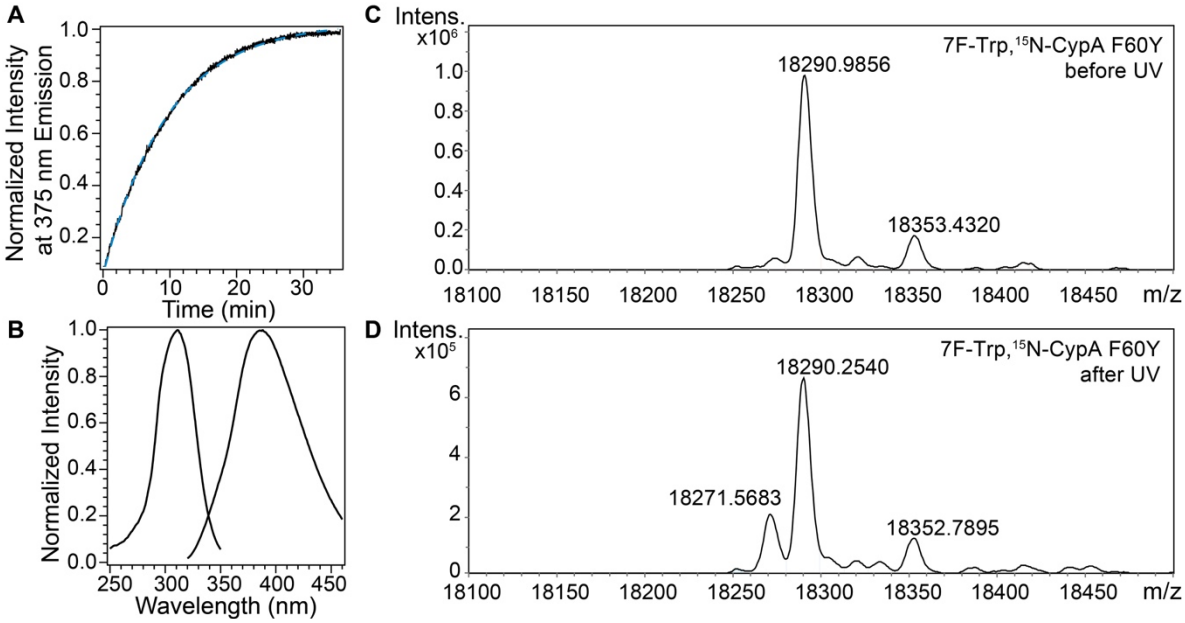

Figure S9. Photoactivation of the F60Y variant of U-^{15}N, 7F-Trp-CypA. (A) Time-dependent fluorescence upon photoactivation. The experimental emission intensity (solid black curve) was fit to a single-exponential decay function (blue dashed line). (B) Excitation and emission spectra of the F60Y variant of U-^{15}N, 7F-Trp-CypA after photoactivation. (C) Deconvoluted ESI-MS spectra of the F60Y variant of U-^{15}N, 7F-Trp-CypA before UV radiation. (D) Deconvoluted ESI-MS spectra of the F60Y variant of U-^{15}N, 7F-Trp-CypA after UV radiation. Note that a new peak appears at 18271.6 Da (~ 20 Da less than the main mass peak before radiation), indicating a crosslinked species was generated.

Supplementary Tables

Table S1. Amounts of different species in the UV radiated mixture [a]

Res#	7F-Trp-CypA	7OH-Trp-CypA	CL-CypA	X-Trp-CypA
60	28%	34%	24%*	14%
62	26%	29%	28%	17%
120	25%	30%	31%	14%
121	22%	33%	35%	10%
122	27%	32%	28%	14%
Average**	25%	31%	30%	14%

[a] Amounts were estimated based on peak intensities in the 2D ^{1}H-^{15}N HSQC spectrum
** Inaccurate intensity due to overlapping peak; therefore, the F60 resonance was excluded from the average*
*** The average was calculated taking residues 62, 120, 121, and 122 into account.*

Table S2. Fluorescence intensities derived by SVD

Sample	Fluorophore	Excitation		Emission		Total intensity[a]	Sample conc. (μM)	Adjusted intensity[b]
		λ_{ex} (nm)	$G_n(\lambda_{ex})$	λ_{em} (nm)	$F_n(\lambda_{em})$			
Enriched/ Purified	I (CL-CypA)	302	3.34	369	2.30	7.67	20.4	0.3759
	II	322	1.22	381	0.52	0.64		0.0312
	III (Tyr)	282	0.88	303	0.57	0.50		0.0246
	IV	294	0.36	410	0.12	0.05		0.0022
Un-purified	I (CL-CypA)	302	3.47	370	2.33	8.09	50.0	0.1618
	II	320	1.88	382	1.00	1.87		0.0374
	III (Tyr)	282	1.48	303	0.88	1.30		0.0259
	IV	296	0.71	415	0.28	0.20		0.0040
	V	366	0.14	474	0.17	0.02		0.0005

[a] Total intensity equals $G_n(\lambda_{ex}) \times F_n(\lambda_{em})$

[b] Fluorescence intensity was adjusted according to sample concentration (total intensity divided by sample concentration)

Table S3. Summary of NMR restraints and structural statistics

Number of NOE restraints	
Intraresidue (i-j=0)	591
Sequential (|i-j|=1)	785
Medium range (2≤|i-j|≤4)	364
Long range (|i-j|≥5)	871
Total	2611
Number of dihedral angle restraints	
φ	131
ψ	131
Structural statistics (20 lowest energy structures)	
Violations	
NOE distance restraints (Å)	0.043±0.003
Dihedral angles restraints (°)	1.362±0.291
Deviations from idealized geometry	
Bond lengths (Å)	0.003±0.000
Bond angles (°)	0.587±0.013
Improper torsions (°)	0.485±0.017
Average pairwise r.m.s. deviation (Å)	
All heavy atoms	1.5±0.2
Backbone heavy atoms	0.9±0.2

Table S4. The largest 8 singular values determined by SVD of the excitation-emission matrices for the enriched/purified sample and non-purified sample

	s_{11}	s_{22}	s_{33}	s_{44}	s_{55}	s_{66}	s_{77}	s_{88}
Enriched/purified	225.588	7.656	3.670	0.457	0.351	0.340	0.338	0.312
Non-purified	231.370	15.609	6.667	2.495	0.742	0.378	0.359	0.340

Table S5. Reduced global χ2 values obtained by simultaneous fitting the TCSPC data obtained on both enriched/purified and non-purified sample at 41 emission wavelengths as a function of the number of exponential terms

N_{exp}	Reduced χ^2
1	10.963529
2	2.009542
3	1.128503
4	1.023267
5	1.017909
6	1.017763