Dissecting The Interaction between Cryptochrome and Timeless Reveals Underpinnings of Light-Dependent Recognition

Connor M. Schneps‡, Robert Dunleavy & Brian R. Crane*

Cornell University, Department of Chemistry & Chemical Biology, Ithaca, NY 14853

‡Current Location: Dana-Farber Cancer Institute, Department of Medical Oncology, Boston, MA 02115

*For correspondence: bc69@cornell.edu

ABSTRACT

Circadian rhythms are determined by cell-autonomous transcription-translation feedback loops that entrain to environmental stimuli. In the model circadian clock of *Drosophila melanogaster*, the clock is set by the light-induced degradation of the core oscillator protein Timeless (TIM) by the principal light-sensor Cryptochrome (CRY). The cryo-EM structure of CRY bound to TIM revealed that within the extensive CRY:TIM interface the TIM N-terminus binds into the CRY FAD pocket, in which FAD and the associated phosphate binding loop (PBL) undergo substantial rearrangement. The TIM N-terminus involved in CRY binding varies in isoforms that facilitate adaptation of flies to different light environments. Herein we demonstrate through peptide binding assays and pulsed-dipolar ESR spectroscopy that the TIM N-terminal peptide alone exhibits light-dependent binding to CRY and that the affinity of the interaction depends on the initiating methionine residue. Extensions to the TIM N-terminus that mimic less light-sensitive variants have substantially reduced interaction with CRY. Substitutions of CRY residues that couple to the flavin rearrangement in the CRY:TIM complex have dramatic effects on CRY light activation. CRY residues Arg237 on α8, and Asn253, and Gln254 on the PBL are critical for the release of the CRY autoinhibitory C-terminal tail (CTT) and subsequent TIM binding. These key light-responsive elements of CRY are well conserved throughout Type I cryptochromes of invertebrates, but not by cryptochromes of chordates and plants, which likely utilize a distinct light-activation mechanism.

INTRODUCTION

In eukaryotes, circadian clocks are molecular oscillators comprised of transcription-translation feedback loops (TTFLs) that are sensitive to the diurnal cycle through the action of specific photoreceptors[1-3]. In the so-called “positive arm” of the clock, transcription factor complexes bind to the regulatory elements of clock-controlled genes, promoting the expression of repressor proteins[2,4]. In the negative arm of the clock, repressor proteins translocate to the nucleus after a delay and inhibit the positive elements from further promoting transcription of their respective genes[2,4]. As the repressors accumulate in the cytoplasm over the course of a circadian day, they undergo rhythmic post-translational modifications that contribute to their degradation and the re-initiation of the cycle[4-7]. The molecular architecture of the TTFL is conserved across many species, including fungi, insects and mammals[1-3,8].

In the model circadian clock of the fruit fly *Drosophila melanogaster*, the transcription factors Clock and Cycle (CLK and CYC) form a heterodimeric complex to promote the transcription of *period* (*per*) and *timeless* (*tim*) genes. PER and TIM heterodimerize, translocate to the nucleus, and inhibit CLK/CYC[3]. The PER and TIM proteins then undergo post-translational modifications that induce their respective degradation through the ubiquitin-proteasome pathway. TIM degradation is induced by light through action of the flavoprotein Cryptochrome (CRY).

CRYs are UV-A/blue light photoreceptors that are found across the kingdoms of life[3]. Structurally, CRYs are closely related to the DNA-repair photolyase (PL) enzymes; they contain an α/β Rossman fold and an α-helical domain that binds flavin adenine dinucleotide (FAD). A variable CRY C-terminal extension (CCE) distinguishes CRYs from photolyases and each other. In invertebrates, Type I CRYs act as light sensors[9]. For example, *Drosophila* CRY, serves as the principal photosensor that entrains the fly rhythm to light[10,11-14]. The CRY CCE contains a C-terminal tail (CTT) helix that acts as an autoinhibitory element in the dark[10,12-15]. Upon photoreduction of FAD to the anionic semiquinone (ASQ) the CTT undocks from the FAD binding pocket, allowing CRY to bind TIM[12,17]. This complex recruits the E3 ubiquitin ligase Jetlag (JET) that then ubiquitinates CRY:TIM and directs the proteins down the proteasomal pathway[18,19]. Removal of the CTT from CRY (CRYΔ) constitutively activates the protein, and prevents accumulation of TIM in dark or light[19-21]. CRYΔ has been shown to be stable and bind flavin[17,22], however; the biophysical properties of the variant have not been extensively studied.

The cryo-EM structure of CRYΔ in complex with TIM revealed a central core of TIM armadillo (ARM) repeats, of which the N-terminal three repeats form an extensive interface with CRY (Figure 1)[23]. The N-terminal helix of the first ARM repeat replaces the CTT within the FAD binding pocket and binds in a mode that resembles how photolyases recognize a DNA lesion (Figure 1A). Notably, the initiating Met residue of TIM engages in several hydrophobic and hydrogen bonding interactions within the flavin pocket, adjacent to the cofactor (Figure 1B). Initiator Met residues are often modified in flies[24], and although the N-terminal Met is clearly present in the structure, chemical modifications are difficult to discern at the local resolution of ~2.8 Å. Moreover, flies produce a long (L) isoform of TIM (L-TIM) that appends 23 N-terminal residues in response to varied light conditions at different latitudes[25,26]. L-TIM interacts more weakly than short TIM (S-TIM) with CRY and this attenuated interaction results in reduced photosensitivity of the clock[27]. The CRY:TIM structure suggests that the addition of residues N-terminal to the initiating Met may indeed alter the ability of CRY to bind to TIM[23].

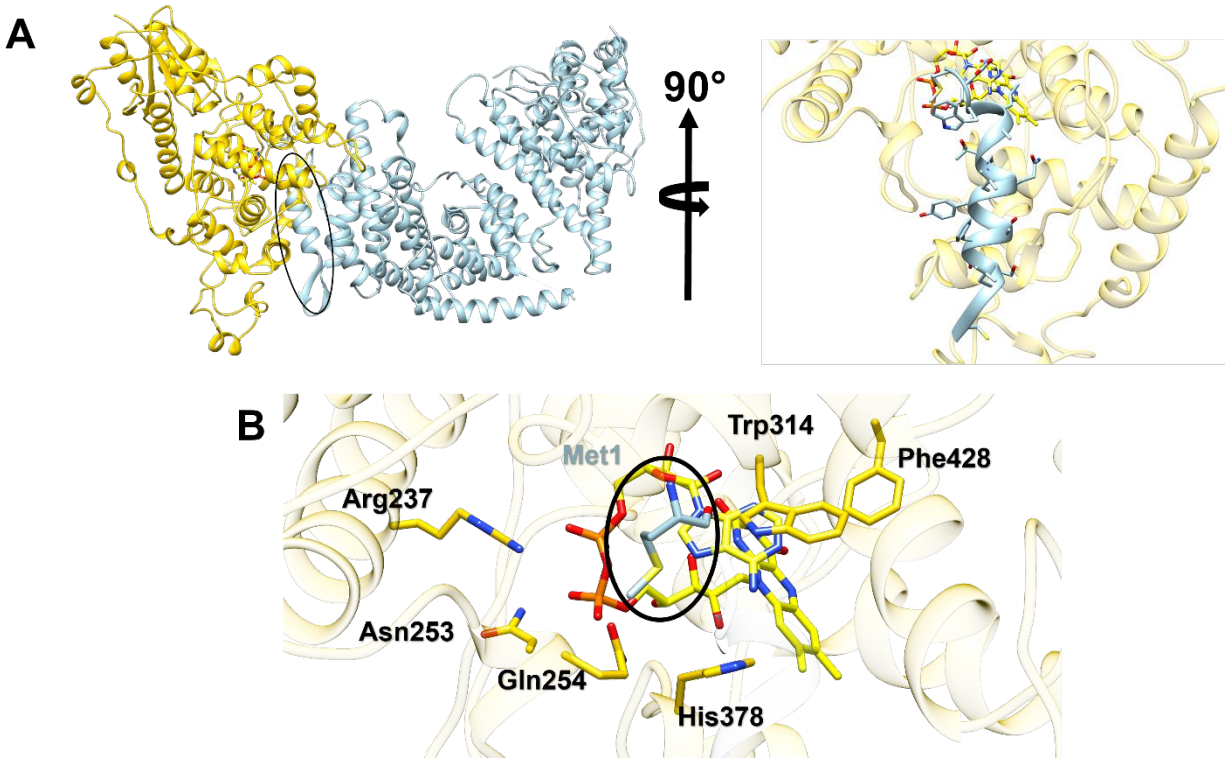


Figure 1. Binding environment of the TIM N-terminus within CRY. (A) The structure of the CRY:TIM complex (CRYΔ in gold; FAD in gold; TIM in light blue; PDB 8DD7) reveals an extensive interface between the CRY primary flavin pocket and the TIM N-terminal ARM repeats. The first 19 residues of the N-terminus of TIM (circled) replace the CRY CTT by binding into the flavin pocket (closeup is rotated by 90° compared to left view). (B) The initiating Met residue of TIM (light blue; circled) makes extensive contacts at the center of the interface with several highlighted CRY residues.

The CRY flavin pocket undergoes considerable structural rearrangement in binding to the TIM N-terminus. In particular, the diphosphate moiety of FAD changes conformation and a coordinating Mg^{2+} ion is replaced by Arg237, which shifts to hydrogen bond with the flavin phosphates in concert with a tilt of the α8-helix on which it resides. The so-called “phosphate-binding loop” (PBL, residues 249-263, following CRY α8) collapses into the FAD pocket and restructures the surrounding surface. Asn253 and Gln254 on the PBL hydrogen bond directly with the FAD diphosphate moiety in the complex, whereas they reside at the protein surface in unbound CRY. Although the conformational changes involving Arg237, Asn253 and Gln254 clearly couple to TIM binding, it is unclear if they are promoted by interaction with TIM, result from release of the CTT, or themselves respond to flavin photoreduction.

Herein, we apply a fluorescent peptide binding assay and double electron electron resonance ESR spectroscopy (DEER) to establish that the N-terminal ARM helix of TIM constitutes the essential light-sensitive binding interaction with CRY. Moreover, TIM recognition is indeed sensitive to the nature of the N-terminus, thereby rationalizing the biological impact of the *ls-tim* polymorphism. In addition, light-dependent TIM recognition involves more than autoinhibition by the CTT because affinity for the TIM N-terminus increases with light in a CRY variant that does not contain the autoinhibitory region (CRYΔ). Surprisingly, Arg237, Asn253 and Gln254 are shown to be critical factors in CTT release, a step that must precede TIM binding. Residue conservation and structural analyses suggest that this light activation mechanisms is specific for Type I invertebrate cryptochromes.

METHODS

Expression and Purification of Full-Length CRY, CRYΔ, Sort-Tagged CRY Variants

Protein expression and purification of full-length *Drosophila melanogaster* CRY (Uniprot ID: O77059, residues 1-539; unconserved residues 540-542 were deleted to improve expression[10,11]), CRYΔ (residues 1-520), and sort-tagged constructs were produced in a similar manner as previously described[12,17]. All constructs were cloned in frame with an N-terminal twin-strep tag and TEV proteolytic cleavage site between NdeI and XhoI in pet28a(+). The deletion of the CTT for CRYΔ was introduced via Kinase-Ligase-DpnI (KLD) reaction (New England BioSciences), the sort-tag (LPGTG) on the C-terminus was introduced via Restriction-Free (RF) cloning, and point mutations were made using QuikChange (Novagen; Arg237Ala) and KLD (New England Biosciences) for Asn253Ala/Gln254Ala. All constructs were expressed in CmpX13 cells, an engineered strain of *E. coli* with an upregulated riboflavin transporter[27], induced with 0.4 mM isopropyl-β-D-1-thiogalactopyranoside (IPTG) at 17 °C when the OD_{600} reached 0.6-0.8, followed by growth overnight. Proteins were purified in a similar manner as previously reported, except the elution buffer for CRYΔ contained 50 mM HEPES (pH 8), 150 mM NaCl, 10% (vol/vol) glycerol, 1 mM ethylenediaminetetraacetic acid (EDTA), 50 mM biotin, and 4 mM Tris(2-carboxyethyl)phosphine hydrochloride (TCEP-HCl) to mitigate oxidative aggregate formation. The resin used for all CRY-related purifications was the Strep-Tactin XT 4-Flow resin (IBA Life Sciences).

Expression, Purification, and Modification of N-terminal TIM Peptides

The first 19 amino acids of *Drosophila melanogaster* TIM (MDWLLATPQLYSAFSSLGC; S-TIM_N) was fused to the C-terminus of the small ubiquitin like modifier protein (SUMO) in

pet28a(+) in frame with a N-terminal $(His)_6$ tag. Deletion of the initiating Met residue for the S-TIM_N-NoM construct (DWLLATPQLYSAFSSLGC) was introduced via KLD. Cells were grown in Miller's Luria-Bertani Broth (Difco) at 37 °C and induced with 0.4 mM IPTG at 25 °C when the OD_{600} reached 0.6-0.8, after which they continued to grow overnight. The cells were harvested and sonicated in lysis buffer containing 50 mM Tris (pH 8), 150 mM NaCl, 1 mM $MgCl_2$, 10 mM imidazole, and 1 µL of 250 U/µL benzonase. The lysis mixture was centrifuged at 30,000 × *g* for 45 minutes to remove cell debris. The lysate was added directly to Nickel-NTA agarose beads (Gold Biotechnology) that was pre-equilibrated with the lysis buffer. The resin was washed with buffer containing 50 mM Tris (pH 8), 150 mM NaCl, and 20 mM imidazole. Prior to elution, the resin with the SUMO fusion was incubated with ULP1 (to cleave SUMO) and molar excess of fluorescein-5-maleimide (FAM) overnight at 4 °C. The eluent was taken and ran through a HiPrep desalting column in nanopure H_2O, then dried overnight and resuspended in 50 mM Tris (pH 8) and 150 mM NaCl to concentrate. The S-TIM_N-SL (SL = spin-label) was prepared in the same fashion, except the resin was mixed with (1-oxyl-2,2,5,5-tetramethylpyrroline-3-methyl) methanethiosulfonate) (MTSL) in molar excess, desalted into nanopure H_2O, evaporated solvent overnight, and resuspended in D_2O. To test for the presence of the SUMO fusion, the resin was washed with buffer containing 50 mM Tris (pH 8), 150 mM NaCl, and 300 mM imidazole for SDS-PAGE analysis.

The S-TIM_N+3 peptide (**KSV**MDWLLATPQLYSAFSSLGC) that contains C-terminal residues from extension of the long-TIM isoform (in bold) was purchased at a purity of greater than 90% from the Biomatik Corporation. It was labeled with FM in 50 mM Tris (pH 8) and 150 mM NaCl in a 1:1.5 (peptide:FM) molar ratio overnight at 4 °C. The mixture was further purified on a Superdex 30 Size Exclusion Column (10/300 GL; GE Healthcare) in buffer containing 50 mM Tris (pH 8) and 150 mM NaCl. The buffer was evaporated overnight to concentrate the peptide, then resuspended in the same buffer.

Fluorescence Polarization Assays

Fluorescence polarization assays were performed using the SpectraMax M5 Microplate Reader with Costar 3694 96-Well plates. All protein samples and peptides were diluted into buffer containing 50 mM Tris (pH 8) and 150 mM NaCl at a final volume of 100 µL. The fluorescently labeled peptides were added prior to protein samples for a final concentration of 0.1 µM. For WT CRY samples in the light, a 100 µM stock was illuminated under low intensity blue light (GE LED+ Color; 8.5 W; 650 lumens) for 70 minutes before transfer to the 96-well plate. For CRYΔ and Asn253Ala/Gln254Ala in the light, the protein was pre-incubated with the fluorescent S-

TIM_N peptide and illuminated for 10 minutes at high intensity (2.79 mW/cm^2) in the OptoWell (OptoBioLabs) system prior to transfer to the 96-well plate. The polarization assays on the SpectraMax utilized the fluorescence polarization protocol, with a g-factor of 1.000, an excitation wavelength of 485 nm, and an emission wavelength of 525 nm with an emission wavelength cutoff at 515 nm. Three replicates per protein concentration were taken, averaged, and fitted in Matlab using the following equation, which assumes that the peptide concentration is much smaller than the K_D values:

$$FP = \frac{A[P]}{K_D + [P]} + \Delta FP \quad (1)$$

where FP is the experimentally determined average fluorescence polarization, K_D is the affinity constant of the TIM peptides for CRY, $[P]$ is the protein concentration, ΔFP accounts for the change in fluorescence polarization relative to 0, and A is a constant. The error bars reflect the standard error in n=3 for each protein concentration. The above binding model[28,29] assumes that the probe concentration (0.1 μM) is much less than the measured dissociation constants (15-30 μM) and that there is no appreciable nonspecific binding. Fitting to either a more generally representative quadratic equation or an exact solution incorporating non-specific binding produced the same dissociation constants within error[30].

Sortylation of Arg237Ala and Asn253Ala/Gln254Ala CRY Variants

The sortylation procedure was carried out as previously reported[12,17]. Briefly, a sortase recognition motif (Leu-Pro-Gly-Thr-Gly) was cloned onto the C-terminus of CRY at residue 539. The glycine-linked donor peptide was spin-labeled at the C-terminal Cys residue by MTSL (1-oxyl-2,2,5,5-tetramethylpyrroline-3-methyl) methanethiosulfonate) producing the labeled peptide Gly-Gly-Gly-Gly-Cys-SL, which was enzymatically added to the CRY C-terminal sortase motif with excision of one Gly residue. Hence, the spin label is disulfide linked to the C-terminal Cys residue at what would be position 548. All reaction mixtures were further purified on an analytical Superdex 200 Size Exclusion Column (10/300 GL, GE Healthcare) in buffer containing 50 mM HEPES (pH 8), 150 mM NaCl, and 10% glycerol (vol/vol). No reductants were added to the gel filtration buffer to avoid reduction of the SORTC-SL nitroxide radical or reduction of the disulfide bond between the spin probe and the peptide. Spin-labeled samples were buffer exchanged into deuterated buffer (50 mM HEPES pH 8, 150 mM NaCl) with 25% d8-glycerol utilizing a 50 kDa spin concentrator tube under ambient conditions.

Preparation of CRYΔ and S-TIM$_N$-SL for Double Electron-Electron Resonance (DEER) Experiments

For DEER experiments, CRYΔ samples were exchanged into deuterated buffer (50 mM HEPES pH 8, 150 mM NaCl) with 25% d8-glycerol in a 50 kDa spin concentrator tube. S-TIM$_N$-SL (dissolved in D_2O; 160 μM stock) was added to the deuterium exchanged CRYΔ in a 1:2 vol/vol ratio and concentrated slowly after the addition of the d8-glycerol.

ESR Spectroscopy Experiments (cwESR and 4-pulse DEER)

All ESR measurements were performed in a similar fashion to what has been previously reported[12,17]. Continuous-wave ESR (cwESR) experiments were performed at X-band (~9.4 GHz) at room temperature with a modulation amplitude of 2 G on a Bruker E500 spectrometer equipped with a super Hi-Q resonator. Samples were illuminated with a blue laser (TECBL-440, 30 mW, 440 nm) for 5-10 s in the resonator cavity. cwESR spectra were taken pre and post irradiation at X-band (~9.4 GHz) prior to DEER measurements. All DEER measurements were carried out at Q-band (~34 GHz) on a Bruker E580 spectrometer equipped with a 10 W solid-state amplifier (150 W equivalent TWTA), 150W RF amplifier, and an arbitrary waveform generator. DEER spectra were measured at 60 K in an EN 5107D2 Cavity with a cryogen-free insert/temperature controller. The measurements were performed using four pulses (π-τ_1- π_{pump}-τ_2-π- τ_2-echo) with 16-step phase cycling and a $\pi/2$ pulse length of 18 ns. The pump and probe pulses (flavin and nitroxide respectively) were separated by 84 MHz (~30 G). Raw data was background corrected with an exponential decay function.

DEER Spectroscopy Data Analysis

Data analysis for DEER was performed utilizing DD version 7B developed at Vanderbilt University (https://lab.vanderbilt.edu/hustedt-lab/dd/)[31] and the Singular Value Decomposition (SVD) method developed at Cornell University by ACERT (https://denoising.cornell.edu/)[32]. In DD, depending on the radical formed, the average distance ($\langle R \rangle$)and width (σ) for either one or two components were fixed while varying the population (in the two-component case) in order to find the best fit in accordance with the docked and undocked states established previously[12,17]. In addition to the percentage of docked and undocked components, noise corrected error values (χ_v^2) were calculated using the equation:

$$\chi_v^2 = \frac{1}{N-q}\sum_{i=1}^{N}\frac{[V(t_i)-F(t_i)]^2}{s_i^2} \quad (2)$$

where $V(t)$ and $F(t)$ are the experimental and the fit data respectively, N is a number of points, q is a number of variables changed and s_i is the estimated noise level for the *i*th point. Such values serve as an indication of the fit between the experimental data and the distance reconstruction; values below 2 indicate good fits. The SVD method was used to determine the distance distributions (P(r)) by obtaining an approximate solution to the Fredholm Equation (K P = S), where P is distance distribution, S is the dipolar signal in the time domain and K is the kernel representing the dipolar interaction between two spins. The SVD method obtains a solution, P(r), by choosing only the large singular values that correspond to the signal. The flavin-nitroxide distance was estimated by calculating the distance of the flavin to the sulfhydryl group of Cys19 in the CRY:TIM structure and assuming that Cys19 bonded to MTSL was a worm-like chain with a persistence length of 1.4 Å (https://chemeslab.myqnapcloud.com:3939/CeffApp/)[33].

UV-Vis Spectroscopy

UV-visible spectra of CRYΔ [in 50 mM HEPES (pH 8), 150 mM NaCl, 10% glycerol (vol/vol)] were taken in a quartz cuvette with a pathlength of 0.2 cm. The spectra were measured using an Agilent 8534 diode-array spectrophotometer with a single reference wavelength set to 800 nm for background correction. All samples were irradiated using a blue laser (TECBL-440, 30 mW, 440 nm, World Star Tech) for 10 seconds.

Small Angle X-Ray Scattering of CRYΔ

Samples of CRYΔ were prepared by cleaving the twin-strep-tag off with TEV protease in a 1:20 mg (TEV:CRYΔ) ratio prior to further purification on the Superdex 200 26/60 (GE Healthcare) size exclusion column, with the final concentration of the protein at 4 mg/mL [50 mM HEPES (pH 8), 150 mM NaCl, 10% glycerol (vol/vol), and 4 mM TCEP-HCl]. Small Angle X-Ray Scattering (SAXS) data were collected at the Cornell High Energy Synchrotron Source beamline ID7A1. The accessible q range was from 0.005 - 0.7 $Å^{-1,}$ data were collected using an Eiger 4M Detector. Dark samples were covered with aluminum foil until time of collection, and light samples were illuminated at 440 nm (TECBL-440, 30 mW, World Star Tech) for 30 seconds prior to collection. The data were processed using RAW[34].

MALDI-TOF of S-TIM$_N$-NoM Peptide

The S-TIM$_N$-NoM sample was first diluted to a concentration of 1 μM in its storage buffer [50 mM Tris (pH 8), 150 mM NaCl]. A matrix solution containing saturated α-cyano-4-hydroxycinnaminic acid (HCCA; Sigma Aldrich) in TA30 solvent (30:70 [vol/vol] acetonitrile:

0.1% trifluoroacetic acid in water) was mixed 1:1 with the peptide solution. 0.5 µL of this mixture was then deposited onto a ground steel target and allowed to dry. The final concentration of the peptide on the target was 500 nM. The mass spectrum of S-TIM$_N$-NoM was collected on a Bruker Autoflex Max MALDI-TOF Mass Spectrometer.

RESULTS

TIM is a large protein with extensive stretches of disorder and many sites of posttranslational modification[3,17-19]. These features make it difficult to isolate TIM in quantity for *in vitro* interaction experiments. Thus, we sought to follow light-dependent binding of CRY to key TIM recognition elements that did not depend on purifying the entire TIM protein. We investigated the binding of CRY to TIM-derived peptides labeled with fluorescent probes and spin-reporters by fluorescence anisotropy and DEER measurements, respectively. To delineate steps in the light-sensing mechanism these assays were carried out with full-length CRY in dark and light, as well as with CRY variants, such as CRYΔ, which lacks the autoinhibitory CTT[12,17].

Characterization of CRYΔ

CRYΔ binds TIM in both dark and light[20-22]. To quantify TIM binding to CRYΔ, CRYΔ was produced following a similar strategy as that used for full-length CRY[12,17,21,26]. The reduced ASQ state of CRYΔ was less stable than that of CRY, likely owing to the absence of the CTT. Higher flavin exposure increased ROS generation and promoted the formation of oxidative protein aggregates. However, twin-strep-tag addition and a higher concentration of reductant during purification mitigated aggregation and produced CRYΔ as a reasonably stable monomer, even after TEV protease cleavage of the twin-strep tag (Figure 2A).

As has been shown previously[35], CRYΔ has a similar flavin UV/Vis absorption spectrum as the full length protein. Also, like its full-length counterpart, CRYΔ binds oxidized FAD in the dark and is readily photoreduced in the light to the ASQ state, showing a characteristic UV/Vis peak at 403 nm (Figure 2B). To further validate the formation of the ASQ by CRYΔ, continuous wave electron spin resonance (cwESR) spectroscopy gave a flavin radical signal with a linewidth between 16 and 18 G that is characteristic of the ASQ[12](Figure 2C). Small angle X-ray scattering (SAXS) reveals similar scattering profiles for CRYΔ in dark and light that are characteristic of monomeric, globular proteins (Figures 2D/E). Hence CRYΔ does not undergo

large-scale conformational changes or oligomerization upon photoreduction.

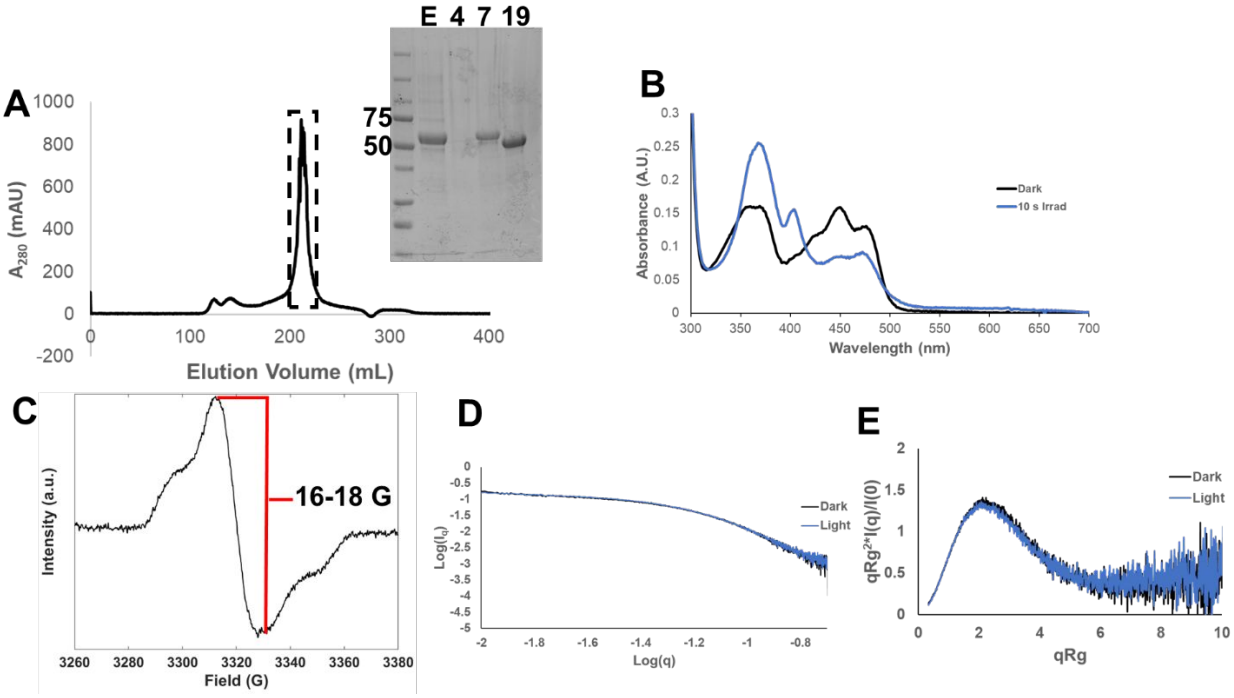


Figure 2. Characterization of CRYΔ. (A) Size exclusion chromatogram of recombinantly expressed CRYΔ after TEV cleavage. The main peak is denoted by a dashed black box and corresponds to fraction 19 of the SEC column as run on the SDS-PAGE gel (inset). The other gel lanes represent the fractions of the strep-affinity column (E), void (fraction 4) and CRYΔ prior to TEV cleavage of the twin-strep-tag (fraction 7). (B) UV/Vis spectra of CRYΔ in the dark (black) and after 10 seconds of irradiation with a 440 nm laser (blue). These spectra are virtually identical to that of WT CRY[12]. (C) X-band continuous wave electron spin resonance (cwESR) spectrum of CRYΔ post irradiation with blue light for 10 seconds. The peak-to-peak width is indicative of the ASQ state[12]. (D)/(E) Scattering profiles (log-log plot) (D) and dimensionless Kratky plots (E) of CRYΔ in dark (black) and post-illumination with blue light after 30 seconds (blue). Both scattering profiles are similar, and the dimensionless Kratky plots indicate globular, well-folded proteins in both states.

CRY Binds to the N-terminal ARM Helix of TIM

The CRYΔ/TIM complex structure revealed the N-terminal 19-residue ARM helix of TIM to be at the center of the binding interface, effectively replacing the CTT in the dark-state full-length structure (Figure 1A). Given that this TIM helix replaced the CTT in the flavin binding pocket, we tested if a free peptide composed of this sequence would alone bind to CRY. A peptide

corresponding to this helix, S-TIM$_N$ (MDWLLATPQLYSAFSSLGC), was recombinantly produced as a larger fusion protein, isolated and labeled with fluorescein-5-maleimide (FM) at the native cysteine residue for fluorescence polarization binding assays (Figures S1-S2). CRYΔ bound S-TIM$_N$ in both dark and light, but with somewhat higher affinity in the light (Figure 3A, Table 1). In contrast S-TIM$_N$ bound to full-length CRY (residues 1-539) only in light with an affinity similar to that seen for CRYΔ in light (Figure 3B). To further validate whether the binding modality of S-TIM$_N$ recapitulates that of TIM in the complex structure, S-TIM$_N$ was spin-labeled with MTSL ((1-oxyl-2,2,5,5-tetramethylpyrroline-3-methyl) methanethiosulfonate); S-TIM$_N$-SL; Figure S1B) and 4-pulse double electron-electron resonance (4P-DEER) experiments were performed on the CRY-peptide complex, similar to those previously described[12,17]. CwESR gave a strong nitroxide-only signal in the dark whereas the ASQ and nitroxide signals overlapped in the light (Figure S3A-B). The time-domain dipolar data from 4P-DEER displayed a rapid decay with strong oscillations (Figure S3C), indicating that the ASQ and nitroxide are relatively close together in space. SVD-based analysis of the data (Figure S3D)[31] yielded a distance maximum of 34.3 Å (Figure 3C). This distance is close to what is predicted from the CRY:TIM complex structure, provided that the peptide extends by the several flexible residues within the label (Figure 3C inset). These data affirm that binding of S-TIM$_N$ to CRY is indeed light-dependent, and that the CTT is important for regulating the interactions between the peptide and the binding pocket of CRY.

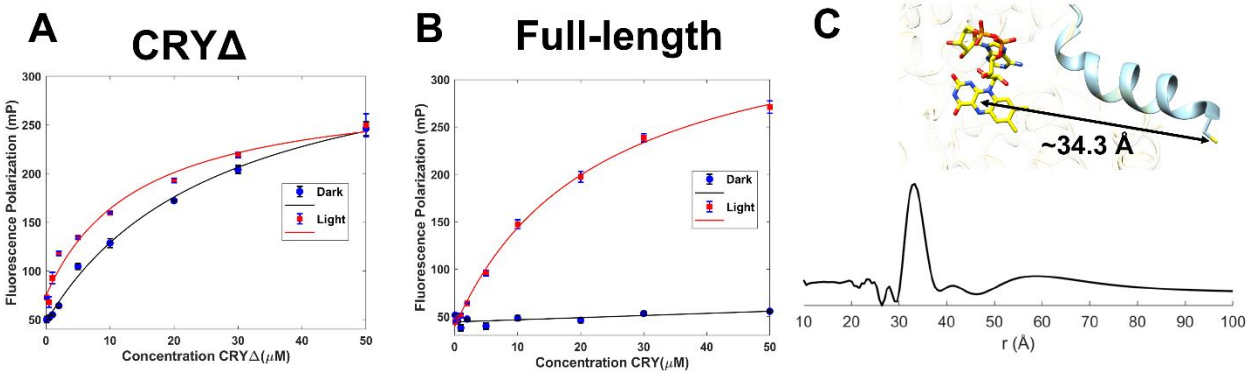


Figure 3. S-TIM$_N$ alone binds to CRY. (A) Fluorescence polarization assays for S-TIM$_N$ binding to CRYΔ in dark (blue) and light (red). (B) Binding curves for S-TIM$_N$ to full-length CRY in dark (blue) and light (red). Each binding experiment was performed in triplicate; error bars are commensurate with the point sizes. (C) DEER-derived distance distribution of CRYΔ-S-TIM$_N$-SL (below) compared to the structural model of the spin label relative to the flavin based on the cryo-EM structure of CRY:TIM. Binding constants are given in Table 1.

Initiating Met of TIM is Important for Light-Dependent Recognition by CRY

Met1 of TIM engages in a network of hydrophobic and hydrogen bonding interactions within the CRY binding pocket (Figure 1B). We sought to test the impact of Met1 on CRY recognition of TIM in two ways: 1) by deleting Met1 (S-TIM$_N$-NoM; DWLLATPQLYSAFSSLGC-) and 2) by adding residues to the N-terminus of S-TIM$_N$ from the *D. melanogater ls-tim* allele, which is less light-sensitive and advantageous for flies living at high latitudes[18,26,36] (S-TIM$_N$+3; **KSV**MDWLLATPQLYSAFSSLGC-). These three residues (Lys-Ser-Val) from the 23-residue L-TIM extension are predicted to interact within the CRY:TIM interface, with the additional residues extending beyond the immediate pocket (Figure S4). S-TIM$_N$ peptides containing these modifications were either recombinantly expressed and purified or purchased, then fluorescently labeled with FM (Figures S5-S6). CRYΔ bound to neither of these peptides (Figure 4A). Similarly, neither the S-TIM$_N$-NoM-FM nor the S-TIM$_N$+3-FM peptides bound appreciably to full-length CRY in either dark or light (Figure 4B-C, Table 1).

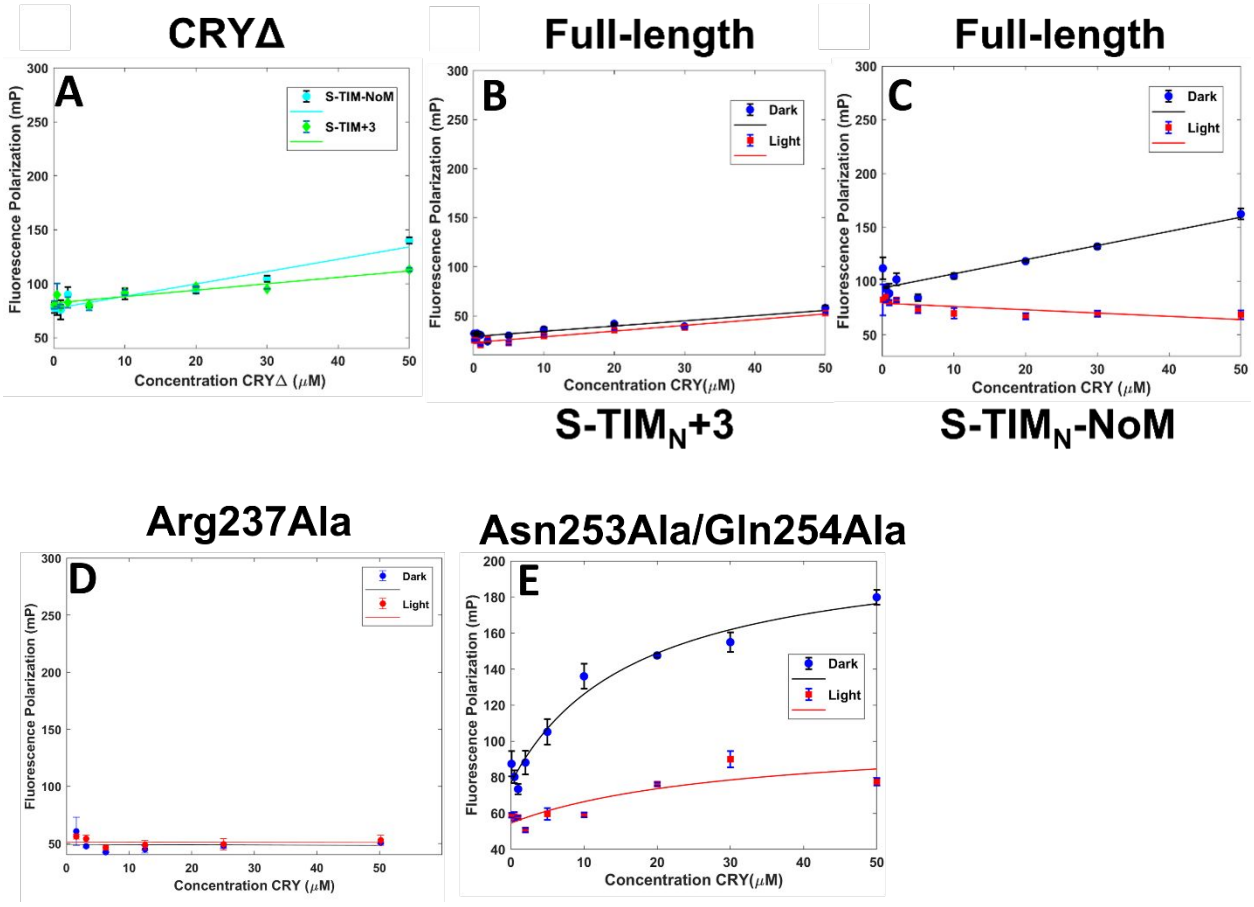


Figure 4. Interactions between modifed S-TIM$_N$ peptides and CRY variants. (A) Fluorescence polarization binding curves for the S-TIM$_N$-NoM (cyan) and S-TIM$_N$+3 (green) to CRYΔ. (B)/(C)

Binding curves of the S-TIM$_N$-NoM and S-TIM$_N$+3 to full-length CRY in dark (blue) and light (red). (D)/(E) Binding curves of S-TIM$_N$ to Arg237Ala (D) and Asn253Ala/Gln254Ala (E) in dark (blue) and light (red). All experiments were performed in triplicate.

Variant		S-TIM$_N$ (K_D in μM)	S-TIM$_N$-NoM (K_D in μM)	S-TIM$_N$+3 (K_D in μM)
CRYΔ	**Dark**	27 ± 2*	**Dark** >10^3	**Dark** >10^3
	Light	14 ± 2*	**Light** >10^3	**Light** >10^3
CRY Full-Length	**Dark**	>10^3	**Dark** >10^3	**Dark** >10^3
	Light	23 ± 6*	**Light** >10^3	**Light** >10^3
Arg237Ala	**Dark**	>10^3	ND	ND
	Light	>10^3		
Asn253Ala/Gln254Ala	**Dark**	20 ± 3	ND	ND
	Light	30 ± 20		

Table 1. Binding affinities of peptides to CRY variants. *Uncertainty represents the 95% confidence interval across n=3 measurements all protein concentrations. ND – not determined.

Arg237, Asn253, and Gln254 are Critical Residues for CTT Release

Given that Arg237, Asn253 and Gln254 form new interactions with FAD in TIM-bound CRY (Figure 5A), we tested if Ala substitutions in these residues affected binding of S-TIM$_N$ and displacement of the CTT. Fluorescent anisotropy measurements of S-TIM$_N$ showed surprising behavior for both Arg237Ala and Asn253Ala/Gln254Ala (Figure 4D/E). Arg237Ala displayed no detectable binding of S-TIM$_N$ in either dark or light (Table 1). In contrast, the Asn253Ala/Gln254Ala variant bound both in dark and light, with stronger binding and a larger fraction of binding in the dark. These results suggested that Arg237Ala and Asn253Ala/Gln254Ala were modulating the conformation of the CTT and accessibility of the flavin pocket.

The sortase-mediated spin-labeling method that places a nitroxide spin label at an engineered C-terminal Cys residue (Cys 548)[12,17] was utilized to quantify the level of CTT undocking in the Arg237Ala and Asn253Ala/Gln254Ala variants. DEER spectroscopy and subsequent fitting of the proportions of docked (29.3 ± 2.0 Å) and undocked (39.5 ± 6.7 Å) CTT conformations[12] with

the DEER analysis program DD[32] yielded the degree of CTT release in each case (Figure 5B leftmost). As an alternative analysis approach, singular-value decomposition (SVD)-based fitting improved the fits to the time-domain data sets and also gave similar distance maxima (Figures S7-S8). Qualitatively, the long-distance components obtained by SVD resemble the distributions of undocked states obtained by DD[32] (Figure S8). DEER data was also measured for CRY with the Asn253Ala/Gln254Ala substitutions in the context of the Leu405Glu/Cys416Asn (EN) variant, which forms a neutral semiquinone, does not release the CTT and hence serves as a dark-state proxy that contains a reference spin on the flavin[12]. These analyses indicated that, unlike with WT, the Arg237Ala CTT remained almost entirely docked in the photoreduced state (Figure 5B). Similarly, the CTT of the Asn253Ala/Gln254Ala variant remained mostly docked in the photoreduced state (~70%, Figure 5B), a proportion that did not change greatly (~77%) in the dark state EN proxy of this variant (Asn253Ala/Gln254Ala/Leu405Glu/Cys416Asn). Thus, Arg237Ala prevents CTT release in the light and as a result S-TIM_N does not bind to CRY in either dark or light. In contrast, S-TIM_N binds Asn253Ala/Gln254Ala in both dark and light, despite the CTT having a primarily docked conformation in each that also does not respond to light. To investigate Asn253Ala/Gln254Ala further, the interaction between the spin-labeled S-TIM_N peptide and the Asn253Ala/Gln254Ala light state and dark-state proxy was measured by DEER and found to be considerably different than that observed for WT CRYΔ (Figure S9). Thus, the Asn253Ala/Gln254Ala substitutions have altered CRY in a manner that allows for some degree of non-native binding of the TIM peptide and loss of proper gating by flavin photoreduction. The weaker interaction implied by the partially undocked CTT of this variant in the dark state (Asn253Ala/Gln254Ala/Leu405Glu/Cys416Asn, Figure 5B) may also allow S-TIM_N to partially compete with the CTT of Asn253Ala/Gln254Ala for the flavin pocket in the unactivated protein.

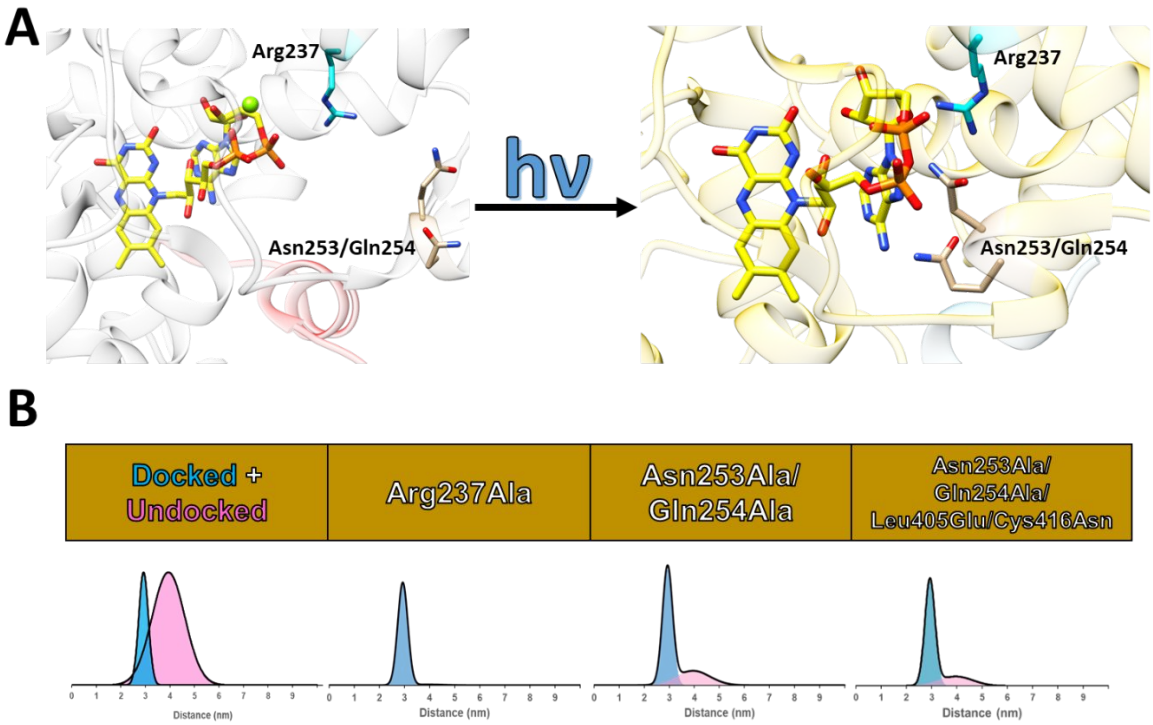


Figure 5. Arg237, Asn253, and Gln254 influence CTT behavior in light. (A) Comparision of dark-state CRY structure (4GU5; grey and CRY CTT in red with Mg ion shown as a green sphere) and TIM-bound CRY structure (8DD7; "light"; gold and light blue for N-terminal helix of TIM). Note the structural rearrangement of R237 (cyan), N253 and Q254 on the phosphate-binding loop (tan) in the TIM-bound light state. B) Distance distributions for CTT spin-labeled samples after illumination (Leu405Glu/Cys416Asn [dark state mimic] and WT[12,17]; leftmost), Arg237Ala (middle-left; % undocked = 2.2 ± 0.6; χ_v^2 = 5.3), and Asn253Ala/Gln254Ala (middle-right; % undocked = 29.6 ± 0.8; χ_v^2 = 2.2), and Asn253Ala/Gln254Ala/Leu405Glu/Cys416Asn (rightmost; % undocked = 23.2 ± 0.005; χ_v^2 = 3.6) with docked populations colored in light blue and undocked colored in light pink.

DISCUSSION

Light-dependent formation of the CRY:TIM complex entrains the Drosophila TTFL to light. Although the interaction between CRY and TIM is extensive (over 2300 Å^2 of buried surface area on TIM) and involves many structural elements of both proteins, the principal contact involves the 16-residue N-terminal helix of TIM, which accounts for 970 Å^2 of buried surface area in the interface[23]. The TIM N-terminus inserts directly into the FAD binding pocket, replacing the CRY CTT (Figure 5, Figure 5A)[23]. TIM has only been produced within insect cells,

usually in co-expression with CRYΔ and at low levels. Thus, the finding that the S-TIM$_N$ peptide alone binds CRY in a light-dependent manner is both useful and convenient for studying the light-regulation of this critical binding interaction. S-TIM$_N$ bound to CRYΔ with reasonably high affinity in the dark and even tighter affinity in the light (Figure 3A-B, Table 1). These data are consistent with CRYΔ constitutively interacting with TIM in cells[17,21]. 4P-DEER spectroscopy with CRYΔ and spin-labeled S-TIM$_N$ produced a distance of ~34 Å between the spin-label and the flavin, which is close to that predicted by the structure of the CRY:TIM complex (Figure 3B). It follows that the position of the peptide in the binding pocket is likely similar to the N-terminus of full-length TIM. Although it was expected for CRYΔ to bind S-TIM$_N$ in both dark and light because there is no CTT to block the flavin pocket in CRYΔ, the higher affinity in light indicates that conformational changes within CRY, independent of CTT displacement, also contribute to light-state affinity. The CRY:TIM structure reveals several regions of CRY that change conformation relative to the dark state that include the PBL, but also the C-terminal lid, protrusion motif and flavin interacting residues in addition to Arg237[23]; FAD photoreduction may indeed promote such restructuring of the flavin pocket to increase affinity for S-TIM$_N$, even in the absence of the CTT. Nonetheless, S-TIM$_N$ does not greatly differentiate the light vs. dark state of CRYΔ, and hence the change in affinity of the CTT for the flavin pocket in light is likely the primary driver of target discrimination.

Removal of, or additions to, the N-terminal residue of the S-TIM$_N$ peptide dramatically affect affinity for CRY. Within the center of the CRY:TIM contact, the Met1 residue of TIM participates in extensive interactions that include backbone hydrogen bonding and hydrophobic sidechain packing against the CRY PBL (Figure 1B)[23]. In *Drosophila*, initiating Met residues are often post-translationally modified by removal, N-terminal acetylation[24,37,38] or oxidation[39]. Furthermore, the *ls-tim* allele, which correlates with adaptation to high-latitudes, encodes for a L-TIM isoform that adds 23 residues to the N-terminus and is less light sensitive than S-TIM due to weaker interactions with CRY[17,23,24]. Given these considerations, we applied the S-TIM$_N$ binding assay to test the importance of the Met1 to CRY recognition. Met removal to generate the S-TIM$_N$-NoM peptide abrogates binding to either CRY or CRYΔ . Thus, it seems unlikely that the TIM N-terminal Met could be proteolytically processed without affecting the interaction with CRY. Furthermore, additions to the TIM N-terminus also reduce affinity, S-TIM$_N$+3 also shows no appreciable binding to either CRY or CRYΔ. Although the full L-TIM addition is longer than that studied here, even adding three N-terminal residues from this isoform greatly curtails peptide binding to CRY or CRYΔ. These three residues are constrained to lie within the

interface, but additional residues from L-TIM would extend the N-terminus enough for it to project free of the pocket formed with CRY (Figure S4). Although the complete L-TIM modification may interact elsewhere on the surface of the complex, the ability of full-length L-TIM to still bind CRY, albeit with weakened affinity compared to S-TIM[18,26,36], more likely reflects contributions from the binding elements provided by the entire TIM protein that do not involve the N-terminus.

The CRY:TIM structure indicates that binding of TIM involves a conformational change within the FAD itself, hydrogen bonding alterations between the cofactor and protein and a large structural rearrangement of the PBL (Figure 1, Figure 5A)[23]. Particularly Arg237, Asn253, and Gln254 undergo the most drastic rearrangements relative to the dark-state structure of CRY alone (Figure 5A). Arg237 forms new hydrogen bonds with the FAD phosphate groups, replacing Mg^{2+}, whereas the helical portion of the phosphate binding loop (where Asn253 and Gln254 reside) translates and rotates inward with the sidechains pointing toward where the CTT would be in the dark state (Figure 5A).

It is reasonable to assume that Arg237, Asn253 and Gln254 change conformation in concert with TIM binding because these residues do not contact the flavin nor the CTT in dark-state CRY. However, the Arg237Ala variant cannot release the CTT in the light, which rather suggests that the replacement of the Mg^{2+} counter ion with Arg237 along with the coordinated movement of α8 propagates changes from the flavin redox state to CTT undocking. Ala substitutions at N253/Q254 also destabilize the CTT against the flavin pocket in the dark state, despite those residues not directly contacting the CTT. Like Arg237Ala, the Asn253Ala/Gln254Ala variant shows little increase in the proportion of the undocked conformation with light; and like Arg237Ala, this lack of photoresponsiveness indicates that the PBL conformational change derives from flavin photoreduction and perhaps the altered conformation of the diphosphate moiety. Moreover, PBL rearrangement likely encourages CTT displacement because the TIM-bound, light-state conformation of the PBL overlaps with Trp536 of the CTT in the dark state (Figure 6A). The altered PBL conformation is further stabilized by His377 and His378, two residues also involved in CTT displacement and TIM binding[17,40]. Thus, the PBL, α8 and Arg237 respond to structural perturbations in FAD, which likely accompany ASQ formation and the new PBL conformation then competes with the CTT for the flavin pocket. The ability of S-TIM_N to bind to both the dark and light states of Asn253Ala/Gln254Ala CRY, but with a conformation that differs from that of WT indicates that these substitutions have perturbed the protein sufficiently to decouple the light response, yet still allow a binding interaction that

likely resembles the native interaction. S-TIM$_N$ may also compete more effectively with the CTT in the variant than in the WT. Thus, the CTT in Asn253Ala/Gln254Ala may be both more weakly bound in the dark state and decoupled from the photoactivation of the FAD cofactor.

The CRY regions revealed to be critical for light activation are well conserved in Type-I light sensitive cryptochromes of Protostomia (arthropods, mollusks, annelids, brachiopods) and echinoderms, but not chordates, bacteria, archaea or plants (Figures 6B and 7). Indeed, CRY2 of *Arabidopsis thaliana*, which has been structurally characterized in a light-activated tetrameric state[41,42], has little sequence similarity with Drosophila CRY in the PBL and also does not conserve Arg237. As such, the corresponding regions of CRY2 interact differently with FAD compared to *Drosophila* CRY (Figure 7). For example, in plant CRY2 the residue analogous to Arg237, Phe224, does not interact with FAD, instead Ser245 hydrogen bonds to the diphosphate group (Figure 7A). In Drosophia CRY, the region that harbors Ser245 and couples to the PBL has a very different sequence and conformation than in plant CRY (Figure 7B). Drosophia CRY also does not form high order homooligomers upon light activation, as does CRY2, and hence the overall conformational response to the change in cofactor state differs considerably in the two families of proteins.

CONCLUSION

In summary, CRY exhibits light dependent binding to a peptide composed of the N-terminal TIM sequence. The S-TIM$_N$ peptide does bind CRYΔ more strongly in light than dark, but the change in affinity is much less than the differential binding in light and dark demonstrated by full-length CRY. Thus, CTT undocking in response to flavin photoreduction in the light is the primary mechanism of gating interactions with TIM. Removal of the N-terminal Met or the addition of N-terminal residues abolishes binding of the S-TIM$_N$ peptide, the latter providing a detailed biochemical rationale for how the LS-TIM isoform modulates light sensitivity in flies by weakening the affinity of the N-terminal binding element. CRY variants with residue substitutions in α8 and the PBL fail to fully undock the CTT in light and do not properly bind and discriminate the S-TIM$_N$ peptide. The involvement of flavin-dependent PBL rearrangements in CTT undocking is explained by steric clash between the PBL in the CRY light state and the CTT in the dark state. Although the dissociation constant between CRY and S-TIM$_N$ is only in the ~10 μM range, which is considerably weaker than that estimated for full-length TIM[17] (< 1 μM), it provides a starting point to consider using the N-terminal peptide of TIM as a tag for engineering optogenetic tools[43-45] that allow control of protein interactions with light[41,42]. Indeed, CRY2 of *Arabidopsis thaliana* has been used with its partner CIB1 to recruit heterologous proteins

together with light[46,47]. In conclusion, our results here characterize important steps in CRY photoactivation and further provide a rationale to develop CRY and S-TIM$_N$ as transferable recognition elements for the control of protein-protein interactions.

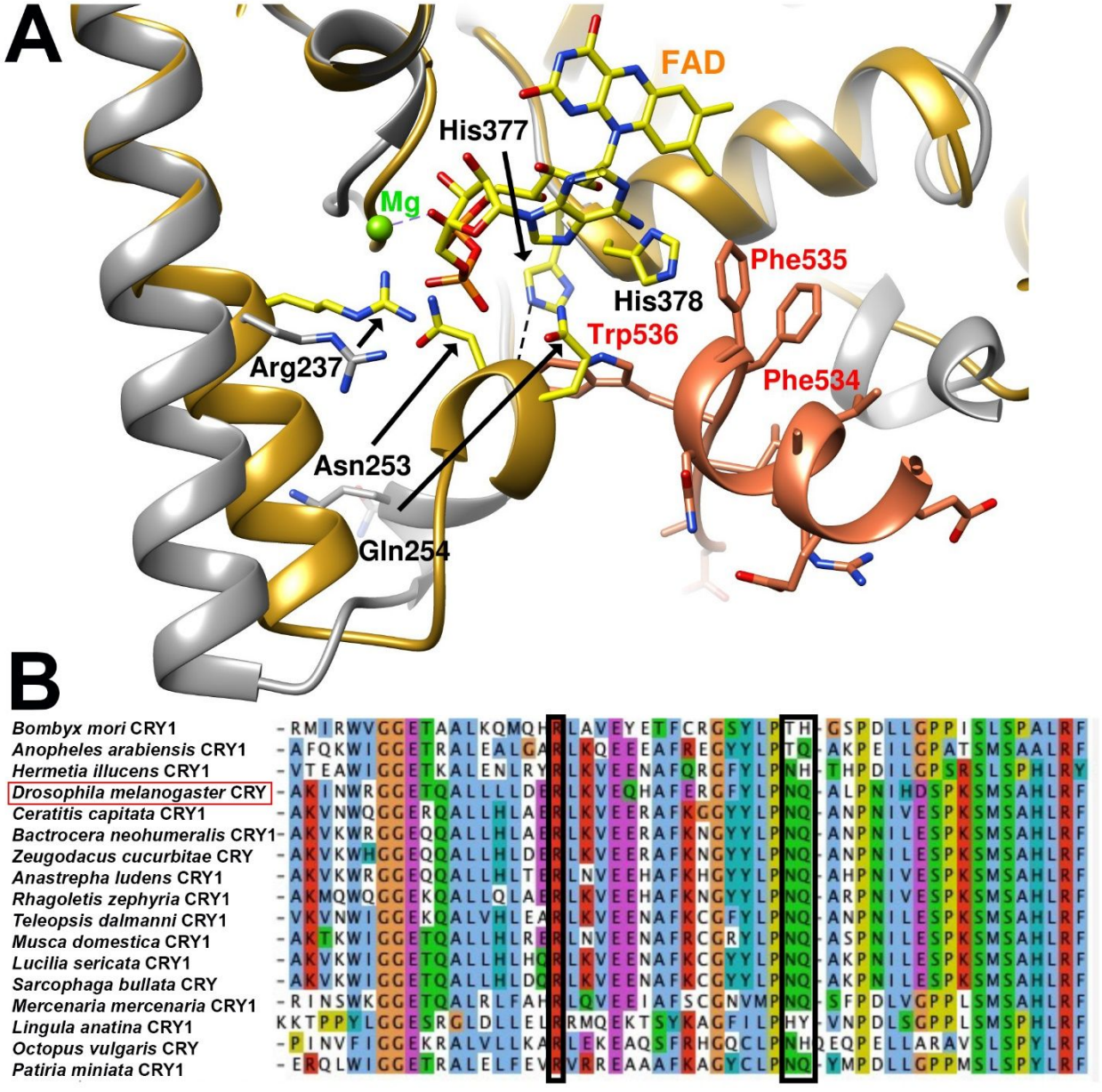


Figure 6. Structure and conservation of residues involved in CRY light activation. (A) Superposition of dark-state CRY (4GU5, grey ribbons) and light-state CRY bound to TIM (8DD7, golden ribbons) showing the movement of Arg237, α8 and PBL residues Asn253,Gln254 (yellow bonds and black arrows) . The PBL in the light-state overlaps with the CTT residue Trp536 (red ribbons and bonds). (B) Sequence alignment of Type-1 invertebrate CRYs showing the

conservation of Arg237, Asn253 and Gln254 (black boxes). Sequences represent arthropod, mollusc, brachiopod, and echinoderm CRY sequences, with Drosophila CRY boxed in red. Protein sequence accession numbers: NP_001182628- *Bombyx mori* CRY1; XP_040162237- *Anopheles arabiensis* CRY1; XP_037923127 – *Hermetia illucens* CRY1; NP_732407 – *Drosophila melanogaster* CRY; XP_004529290 – *Ceratitis capitata* CRY1; XP_050318184 – *Bactrocera neohumeralis* CRY1; BAI67363 – *Zeugodacus cucurbitae* CRY; XP_053955954 – *Anastrepha ludens* CRY1; XP_017479296 – *Rhagoletis zephyria* CRY1 ; XP_037947512 – *Teleopsis dalmanni* CRY1; XP_058985859 – *Musca domestica* CRY1; XP_037810294 – *Lucilia sericata* CRY1; TMW52653 – *Sarcophaga bullata* CRY; XP_053397012 – *Mercenaria mercenaria* CRY1; XP_013384752 – *Lingula anatina* CRY1; CAI9721929 – *Octopus vulgaris* CRY- XP_038072864 – *Patiria miniata* CRY1.

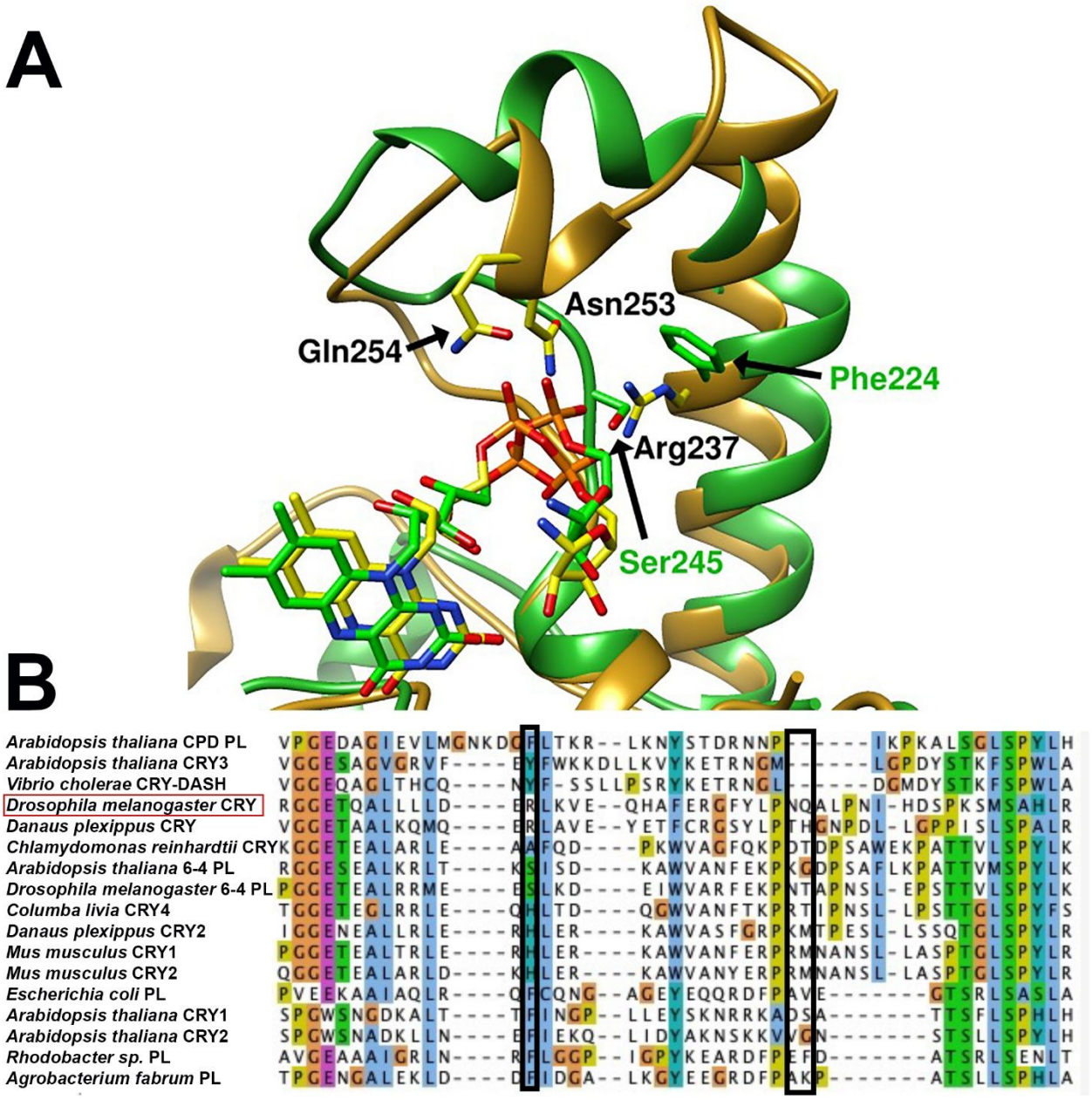


Figure 7. Role of the phosphate binding loop in light activation by invertebrate and plant CRYs. (A) Superposition of light-state CRY bound to TIM (8DD7, golden ribbons) and light-state Arabidopsis CRY2 (PDB code: 6M79)[37] showing the differing conformation and interaction of the phosphate binding loop. CRY Arg237 is replaced by a Phe residue that does not interact with the FAD diphosphate group; instead CRY2 Ser245 takes on this role. (B) Sequence alignment of a representative group of invertebrate, chordate and plant cryptochromes and photolyases showing that Arg237, Asn253 and Gln254 are not conserved and that their surrounding sequence regions have considerable variability. Drosophila CRY sequence is boxed in red,

residues equivalent to CRY Arg237, Asn253 and Gln254 are boxed in black. Protein sequence accession numbers: AAC08008 – *Arabidopsis thaliana* CPD PL; NP_568461.3 – *Arabidopsis thaliana* CRY3; WP_071197283 – *Vibrio cholerae* CRY-DASH; NP_732407 – *Drosophila melanogaster* CRY; AAX58599 – *Danaus plexippus* CRY; XP_042923874 – *Chlamydomonas reinhardtii* CRY; BAA24449 – *Arabidopsis thaliana* 6-4 PL; NP_001260633 – *Drosophila melanogaster* 6-4 PL; AYE53930 – *Columba livia* CRY4; ABA62409 – *Danaus plexippus* CRY2; EDL21416- *Mus musculus* CRY1; EDL27606 – *Mus musculus* CRY2; WP_062883603- *Escherichia coli* PL; NP_567341 – *Arabidopsis thaliana* CRY1; NP_849588 – *Arabidopsis thaliana* CRY2; KRO89520 – *Rhodobacter sp.* PL; SEQ40124 – *Agrobacterium fabrum* PL. CPD PL – cyclobutane pyrimidine dimer photolyase, 6-4 PL- 6-4 photolyase, CRY-DASH – Drosophila, Arabidopsis, Synechocystis, human (**DASH**)-type cryptochromes.

ASSOCIATED CONTENT

Supporting Information

The supporting information is available free of charge at (insert link from journal).

Purification and labeling of S-TIM$_N$; Mass spectrometry characterization of the S-TM$_N$ peptide; CRYΔ+S-TIM$_N$-SL ESR data; Purification and labeling of S-TIM$_N$-NoM; Purification and labeling of S-TIM$_N$+3; ESR data for Arg237Ala, Asn253Ala/Gln254Ala, and Asn253Ala/Gln254Ala/Leu405Glu/Cys416Asn variants; Distributions of CTT-SL to flavin radical distances derived by SVD for CRY variants. S-TIM$_N$-SL peptide interactions with Asn253Ala/Gln254Ala variants.

Notes

The authors declare no competing interests.

Author Contributions

C.M.S. and B.R.C. designed research; C.M.S. and R.D. performed research; C.M.S., R.D. and B.R.C. carried out data analysis; C.M.S. and B.R.C. wrote the paper.

ACKNOWLEDGEMENTS

This work was financially supported by NIH grant R35GM122535 (B.R.C.), NSF grant MCB-1715233 (B.R.C.). ESR measurements were carried out at ACERT which is supported by NIH/NIGMS awards P41 GM103521 and 1S1 0OD021543.

We wish to thank Siddarth Chandrasekaran for initial cloning of S-TIM$_N$ into pet28a(+), Mike Lynch for assistance with MALDI, and the Cornell Biotechnology Resource Center (Ruchika Bhawal, Elizabeth Anderson, Qin Fu, and Sheng Zhang) for assistance with the mass spectrometry data collection and processing.

References

1. Yakir, E., Hilman, D., Harir Y. and Green, R. M. (2006). Regulation of output from the plant circadian clock. *FEBS Journal* **274**, 335-345.

2. Takahashi, J. S. (2017). Transcriptional architecture of the mammalian circadian clock. *Nature Review Genetics* **18**, 164-179.

3. Crane, B. R., and Young, M. W. (2014). Interactive features of proteins composing eukaryotic circadian clocks. *Annual Reviews in Biochemistry* **83**, 191-219. 10.1146/annurev-biochem-060713-035644.

4. Harmer, S. L., Panda, S. and Kay, S. A. (2001). Molecular bases of circadian rhythms. *Annual Review of Cell and Developmental Biology* **17**, 215-253.

5. Crosthwaite, S. K., Dunlap, J. C. and Loros, J. J. (1997). *Neurospora wc-1* and *wc-2* transcription, photoresponses, and the origions of circadian rhythmicity. *Science* **276**, 763-769.

6. Best, J. D., Maywood, E. S., Smith, K. L. and Hastings, M. H. (1999). Rapid resetting of the mammalian circadian clock. *Journal of Neuroscience* **19**, 828-835.

7. Pittendrigh, C. S. (1993). Temporal organization: reflections of a Darwinian clock-watcher. *Annual Reviews in Physiology* **55**, 17-54.

8. Stanewsky, R. (2003). Genetic Analysis of the Circadian System in *Drosophila melanogaster* and Mammals. *Journal of Neurobiology* **54**, 111-147.

9. Ozturk, N. (2017). Phylogenetic and Functional Classification of the Photolyase/Cryptochrome Family. *Photochemistry and Photobiology* **93**, 104-111.

10. Zoltowski, B. D., Vaidya, A. T., Top, D., Widom J., Young, M. W. and Crane, B. R. (2011). Structure of full-length *Drosophila* cryptochrome. *Nature* **480**, 396-399.

11. Levy, C., Zoltowski, B. D., Jones, A. R., Yaidya, A. T., Top, D., Widom, J., Young, M. W., Scrutton, N. S., Crane, B. R. and Leys, D. (2013). Updated structure of *Drosophila* cryptochrome. *Nature* **495**, E3-E4.

12. Chandrasekaran, S., Schneps, C. M., Dunleavy, R., Lin, C., DeOliveira, C. C., Ganguly, A. and Crane, B. R. (2021). Tuning flavin environment to detect and control conformational switching in *Drosophila* Cryptochrome. *Communications Biology* **4**, 1-12.

13. Chaves, I., Pokorny, R., Byrdin, M., Hoang, N., Ritz, T., Brettel, K., Essen, L.-O., van der Horst, G. T. J., Batschauer, A. and Ahmad, M. (2011) The cryptochromes: blue light photoreceptors in plants and animals. *Annual Reviews in Plant Biology* **62**, 335–364.

14. Öztürk, N. Hong, S.-H., Özgür, S., Selby, C. O., Morrison, L., Partch, C., Zhong, D. and Sancar, A. (2007). Structure and function of animal cryptochromes. *Cold Spring Harb. Symposia Quant. Biol*. **72**, 119–131.

15. Emery, P., So, W. V., Kaneko, M., Hall, J. C. and Rosbash, M. (1998). Cry, a Drosophila clock and light-regulated cryptochrome, is a major contributor to circadian rhythm resetting and photosensitivity. *Cell* **95**, 669–679.

16. Emery, P., Stanewsky, R., Hall, J. C. and Rosbash, M. (2000). A unique circadian-rhythm photoreceptor. *Nature* **404**, 456–457.

17. Lin, C., Schneps, C. M., Chandrasekaran, S., Ganguly, A. and Crane, B. R. (2022). Mechanistic insight into light-dependent recognition of Timeless by *Drosophila* Cryptochrome. *Structure*. doi: 10.1016/j.str.2022.03.010.

18. Peschel, N., Chen, K. F., Szabo, G. and Stanewsky, R. (2009). Light-dependent interactions between the *Drosophila* circadian clock factors cryptochrome, jetlag, and timeless. *Current Biology* **19**, 241-247.

19. Koh, K., Zheng, X. Z. and Sehgal, A. (2006). JETLAG resets the *Drosophila* circadian clock by promoting light-induced degradation of TIMELESS. *Science* **312**, 1809-1812.

20. Dissel, S., Codd, V., Fedic, R., Garner, K. J., Costa, R., Kyriacou, C. P. and Rosato, E. (2004). A constitutively active cryptochrome in *Drosophila melanogaster*. *Nature Neuroscience* **7**, 834-840

21. Hemsley, M. J., Mazzotta, G. M., Mason, M., Dissel, S., Toppo, S., Pagano, M. A., Sandrelli, F., Meggio, F., Rosato, E., Costa, R. and Tasatto, S. C. E. (2007). Linear motifs in the C-terminus of *D. melanogaster* cryptochrome. *Biochemical and Biophysical Research Communications* **355**, 531-537.

22. Busza, A., Emery-Le, M., Rosbash, M. and Emery, P. (2004). Roles of the Two *Drosophila* CRYPTOCHROME Structural Domains in Circadian Photoreception. *Science* **304**, 1503-1506.

23. Lin, C., Feng, S., DeOliveira, C. C. and Crane, B. R. (2023) Cryptochrome-Timeless structure reveals circadian clock timing mechanisms. *Nature* **617**(7959), 194-199.

24. Goetze, S., Qeli, E., Mosimann, C., Staes, A., Gerrits, B., Roschitzki, B., Mohanty, S., Niederer, E. M., Laczko, E., Timmerman, E., Lange, V., Hafen E., Aebersold, R., Vandekerckhove, J., Basler, K., Ahrens, C. H., Gevaert, K. and Brunner, E. (2009). Identification and Functional Characterization of N-Terminally Acetylated Proteins in *Drosophila melanogaster*. *PLoS Biology* **7**(11), e1000236.

25. Tauber, E., Zordan, M., Sandrelli, F., Pegoraro, M., Breda, C., Daga, A., Selmin, A., Monger, K., Benna, C., Rosato, E., Kyriacou, C. P. and Costa, R. (2007). Natural selection favors a newly derived *timeless* allele in *Drosophila melanogaster*. *Science* **316**(5833), 1895-1898.

26. Deppisch, P., Prutscher, J. M., Pegoraro, M., Tauber, E., Wegener, C. and Helfrich-Förster, C. (2022). Adaptation of *Drosophila melanogaster* to Long Photoperiods of High-Latitude Summers is Facilitated by the ls-Timeless Allele. *Journal of Biological Rhythms* **37**(2), 185-201.

27. Mathes, T., Vogl, C., Stolz, J. & Hegemann, P. (2009). *In vivo* generation of flavoproteins with modified cofactors. *Journal of Molecular Biology* **385**, 1511–1518.

28. Rossi, A. M. & Taylor, C. W. (2011). Analysis of protein-ligand interactions by fluorescence polarization. *Nature Methods* **6**, 365-387.

29. Moerke, N. J. (2009). Fluorescence Polarization (FP) Assays for Monitoring Peptide-Protein or Nucleic Acid-Protein Binding. *Current Protocols in Chemical Biology* **1**, 1-15.

30. Roehrl, M. H. A., Wang, J. Y. and Wagner, G. (2004). A General Framework for Development and Data Analysis of Competitive High-Throughput Screens for Small-Molecule Inhibitors of Protein-Protein Interactions by Fluorescence Polarization. *Biochemistry* **43**(51), 16056-16066.

31. Srivastava, M. and Freed, J. H. (2017). Singular Value Decomposition Method to Determine Distance Distributions in Pulsed Dipolar Electron Spin Resonance. *Journal of Physical Chemistry Letters* **8**, 5648–5655.

32. Stein, R. A., Beth, A. H. & Hustedt, E. J. (2015). A straightforward approach to the analysis of double electron-electron resonance data. *Methods in Enzymology* **563**, 531–567.

33. Kjaergaard, M., Glavina, J., Chemes, L.B. (2021) Predicting the effect of disordered linkers on effective concentrations and avidity with the “Ceff calculator” app. *Methods in Enzymology* **647**, 145-171.

34. Hopkins, J. B., Gillilan, R. E. and Soren, S. (2017). BioXTAS RAW: improvements to a free open-source program for small-angle X-ray scattering data reduction and analysis. *Journal of Applied Crystallography* **50**, 1543-1553.

35. Ozturk, N., Selby, C. P., Annayev, Y., Zhong, D. and Sancar A. (2010). Reaction mechanism of *Drosophila* cryptochrome. *Proceedings of the National Academy of Sciences USA* **108**(2), 516-521.

36. Sandrelli, F., Tauber, E., Pegoraro, M., Mazzotta, G., Cisotto, P., Landskron, J., Stanewsky, R., Piccin, A., Rosato, E., Zordan, M., Costa, R. and Kyriacou, C. P. (2007). A molecular basis for natural selection at the *timeless* locus in *Drosophila melanogaster*. *Science* **316**(5833), 1898-1900.

37. Bonet-Costa, C., Vilaseca, M., Diema, C., Vujatovic, O., Vaquero, A., Omeñaca, N., Castejón, L., Bernués, Gurealt, E. and Azorín, F. (2012). Combined bottom-up and top-down mass spectrometry analyses of the pattern of post-translational modifications of *Drosophila melanogaster* linker histone H. *Journal of Proteomics* **75**, 4124-4138.

38. Polevoda, B. and Sherman M. (2003). N-terminal acetyltransferases and sequence requirements for N-terminal acetylation of eukaryotic proteins. *Jounral of Molecular Biology* **325**, 595-622.

39. Lim, D.-H., Han, J. Y., Kim, J.-R., Lee, Y. S. and Kim, H-Y. (2012). Methionine sulfoxide reductase B in the endoplasmic reticulum is critical for stress resistance and aging in *Drosophila*. *Biochemical and Biophysical Research Communications* **419**, 20-26.

40. Palayam, M., Ganapathy, J., Guercio, A. M., Tal, L., Deck, S. L. and Shabek, N. (2021). Structural insights into photoactivation of plant Cryptochrome-2. *Communications Biology* **4**(28), doi.org/10.1038/s42003-020-01531-x.

41. Ma, L., Guan, Z., Wang, Q., Yan, X., Wang, J., Wang, Z., Cao, J., Zhang, D., Gong, X. and Yin, P. (2020). Structural insights into the photoactivation of *Arabidopsis* CRY2. *Nature Plants* **6**, 1432-1438

42. Kolarski, D., Miller, S., Oshima, T., Nagai, Y., Aoki, Y., Kobauri, P., Srivastava, A., Sugiyama, A., Amaike, A., Sato, A., Tama, F., Szymanski, W., Feringa, B. L., Itami, K. and Hiorta, T. (2021) Photopharmacological Manipulation of Mammalian CRY1 for Regulation of the Circadian Clock. *Journal of the American Chemical Society* **143**, 2078-2087.

43. Kolarski, D., Miró-Vinyals, C., Sugiyama, A., Sirvastava, A., Ono, D., Nagai, Y., Iida, M., Itami, K., Tama, F., Szymanski, W., Hirota, T. and Feringa, B. L. (2021). Reversible modulation of circadian time with chronophotopharmacology. *Nature Communications* **12**. doi.org/10.1038/s41467-021-23301-x.

44. Kolarski, D., Sugiyama, A., Breton, G., Rakers, C., Ono, D., Schulte, A., Tama, F., Itami, K., Szymanski, W., Hirota, T. and Feringa, B. L. (2019). Controlling the Circadian Clock with High Temporal Resolution through Photodosing. *Journal of the American Chemical Society* **141**, 15784-15791.

45. Taslimi, A., Vrana, J. D., Chen, D., Borinskaya, S., Mayer, B. J., Kennedy, M. J. and Tucker, C. L. (2014). An optimized optogenetic clustering tool for probing protein interaction and function. *Nature Communications* **5**. DOI: 10.1038/ncomms5925.

46. Pathak, G. P., Spiltoir, J. I., Höglund, C., Polstein, L. R., Heine-Koskinen, S., Gersbach, C. A., Rossi, J. and Tucker, C. L. (2017). Bidirectional approaches for optogenetic regulation of gene expression in mammalian cells using *Arabidopsis* cryptochrome 2. *Nucleic Acids Research* **45**(20), doi.org/10.1093/nar/gkx260.

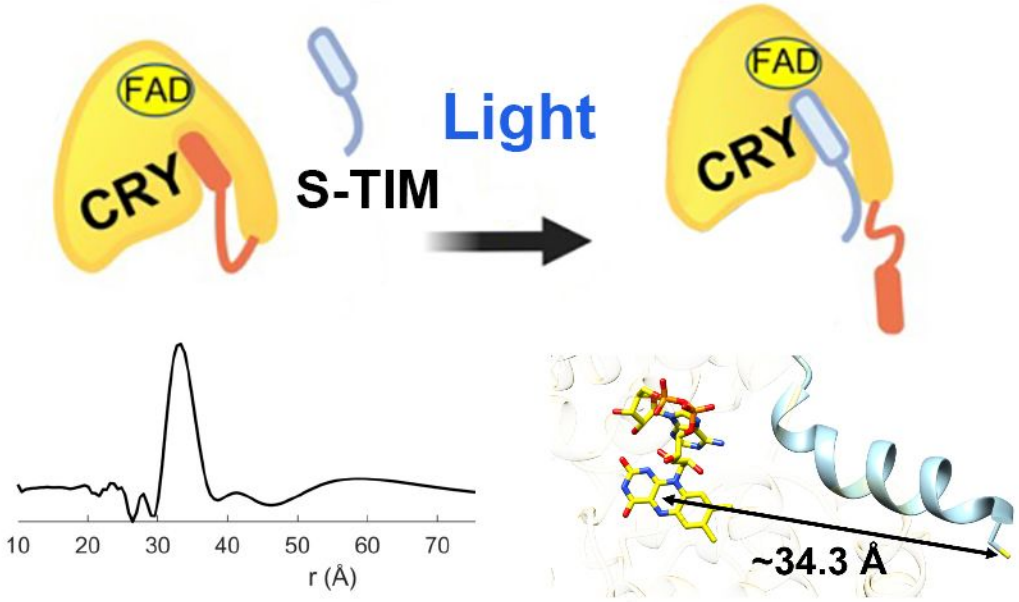
47. Kennedy, M. J., Hughes, R. M., Peteya, L. A., Schwartz, J. W., Ehlers, M. D. & Tucker, C. L. (2010). Rapid blue-light-mediated induction of protein interactions in living cells. *Nature Methods* **7**, doi.org/10.1038/nmeth.1524.

Accession IDs of Proteins from UniProt

Drosophila melanogaster Cryptochrome: O77059

Drosophila melanogaster Timeless: P49021

For Table of Contents use only



Supporting Information for:

Dissecting The Interaction between Cryptochrome and Timeless Reveals Underpinnings of Light-Dependent Recognition

Connor M. Schneps[‡], Robert Dunleavy & Brian R. Crane[*]

Department of Chemistry and Chemical Biology, Cornell University, Ithaca, NY 14850

[‡]Current Location: Dana-Farber Cancer Institute, Department of Medical Oncology, Boston, MA 02115

[*]Correspondence: bc69@cornell.edu

This PDF file includes:

Supplemental Figure S1 – Purification and labeling of S-TIM$_N$.

Supplemental Figure S2 – Mass spectrometry characterization of the S-TM$_N$ peptide.

Supplemental Figure S3 – CRYΔ+S-TIM$_N$-SL ESR spectroscopy data.

Supplemental Figure S4 – Modeling of S-TIM$_N$+3 at the CRY:TIM interface.

Supplemental Figure S5 – Purification and labeling of S-TIM$_N$-NoM.

Supplemental Figure S6 – Purification and labeling of S-TIM$_N$+3.

Supplemental Figure S7 – ESR spectroscopy data for Arg237Ala, Asn253Ala/Gln254Ala, and Asn253Ala/Gln254Ala/Leu405Glu/Cys416Asn variants spin-labeled on the CTT.

Supplemental Figure S8 – Distributions of CTT-SL to flavin radical distances derived by SVD for CRY variants.

Supplemental Figure S9 – ESR spectroscopy data for the S-TIM$_N$-SL peptide bound to the Asn253Ala/Gln254Ala and Asn253Ala/Gln254Ala/Leu405Glu/Cys416Asn variants.

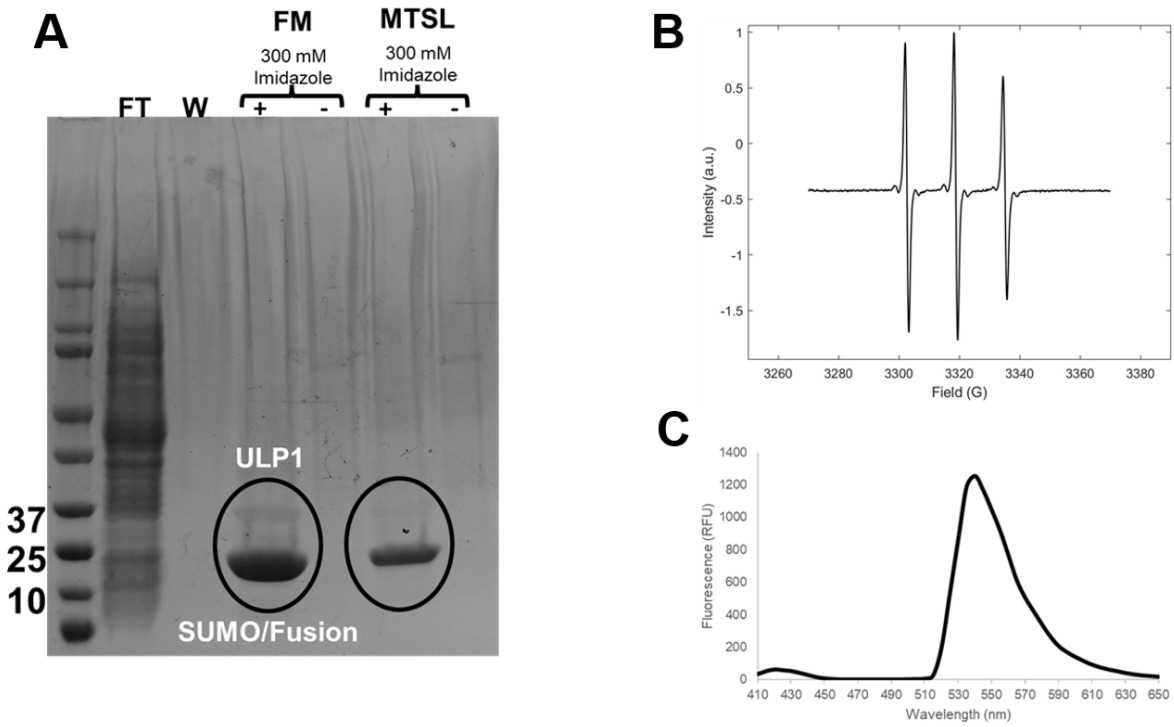
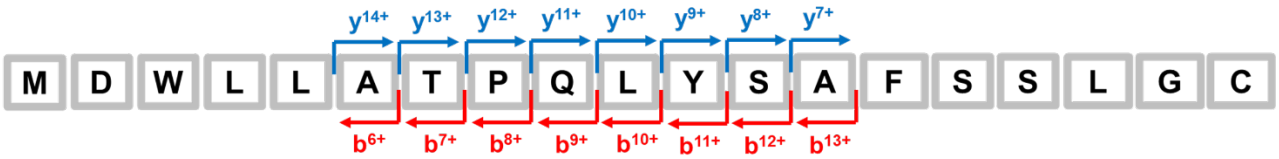


Figure S1. Purification and labeling of S-TIM$_N$. (A) SDS-PAGE gel of fractions from the Ni^{2+}-NTA affinity column with lanes for flow-through (FT), wash (W), and elutions for the fluorescein-5-maleimide labeled (FM) or spin-labeled (MTSL) peptides with or without (+/-) 300 mM imidazole. The eluent with imidazole was taken for further purification and labeling. Circles identify bands for the SUMO fusion (lower, darker) and ULP1 (higher, lighter). (B) X-band cw-ESR spectrum of S-TIM$_N$-SL. The nitroxide displays three characteristic lines with narrow linewidths that indicate high mobility in solution. (C) Fluorescence emission spectrum of S-TIM-FM.



Fragment	b-ions	m/z	Fragment	y-ions	m/z
MD	b^{2+}	132.0385	G^{F}C	y^{2+}	606.1177
MDW	b^{3+}	225.0781	LG^{F}C	y^{3+}	719.2018
MDWL	b^{4+}	281.6201	SLG^{F}C	y^{4+}	806.2338
MDWLL	b^{5+}	338.1622	SSLG^{F}C	y^{5+}	893.2658
MDWLLA	b^{6+}	373.6807	FSSLG^{F}C	y^{6+}	1040.3343
MDWLLAT	b^{7+}	424.2046	AFSSLG^{F}C	y^{7+}	1111.3713
MDWLLATP	b^{8+}	472.7309	SAFSSLG^{F}C	y^{8+}	1198.4034
MDWLLATPQ	b^{9+}	536.7602	YSAFSSLG^{F}C	y^{9+}	1361.4667
MDWLLATPQL	b^{10+}	593.3023	LYSAFSSLG^{F}C	y^{10+}	1474.5508
MDWLLATPQLY	b^{11+}	674.8339	QLYSAFSSLG^{F}C	y^{11+}	1602.6093
MDWLLATPQLYS	b^{12+}	718.3499	PQLYSAFSSLG^{F}C	y^{12+}	1699.6621
MDWLLATPQLYSA	b^{13+}	753.8685	TPQLYSAFSSLG^{F}C	y^{13+}	1800.7098
MDWLLATPQLYSAF	b^{14+}	827.4027	ATPQLYSAFSSLG^{F}C	y^{14+}	1871.7469
MDWLLATPQLYSAFS	b^{15+}	870.9127	LATPQLYSAFSSLG^{F}C	y^{15+}	1984.8310
MDWLLATPQLYSAFSS	b^{16+}	914.4347	LLATPQLYSAFSSLG^{F}C	y^{16+}	2097.9150
MDWLLATPQLYSAFSSL	b^{17+}	970.9768	WLLATPQLYSAFSSLG^{F}C	y^{17+}	2283.9943
MDWLLATPQLYSAFSSLG	b^{18+}	998.4875	DWLLATPQLYSAFSSLG^{F}C	y^{18+}	2399.0213

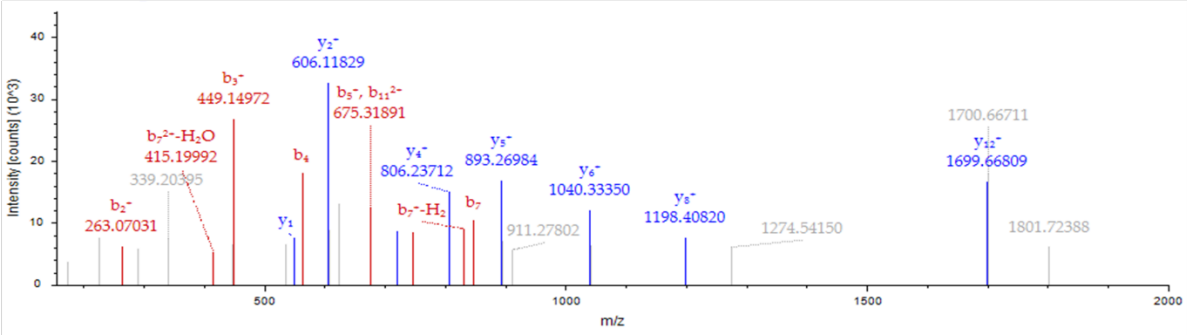


Figure S2. Mass spectrometry characterization of the S-TM$_N$ peptide. NanoLC/MS-MS was used to validate the expression of the S-TIM$_N$ peptide (MDWLLATPQLYSAFSSLGC) labeled with fluorescein-5-maleimide. Mass spectrum containing peptide with labeled cysteine residue (denoted as ^{F}C) and characteristic fragmentation pattern of y and b ions (defined above) after AspN digestion. LC/MS-MS performed at the Cornell Proteomics Facility.

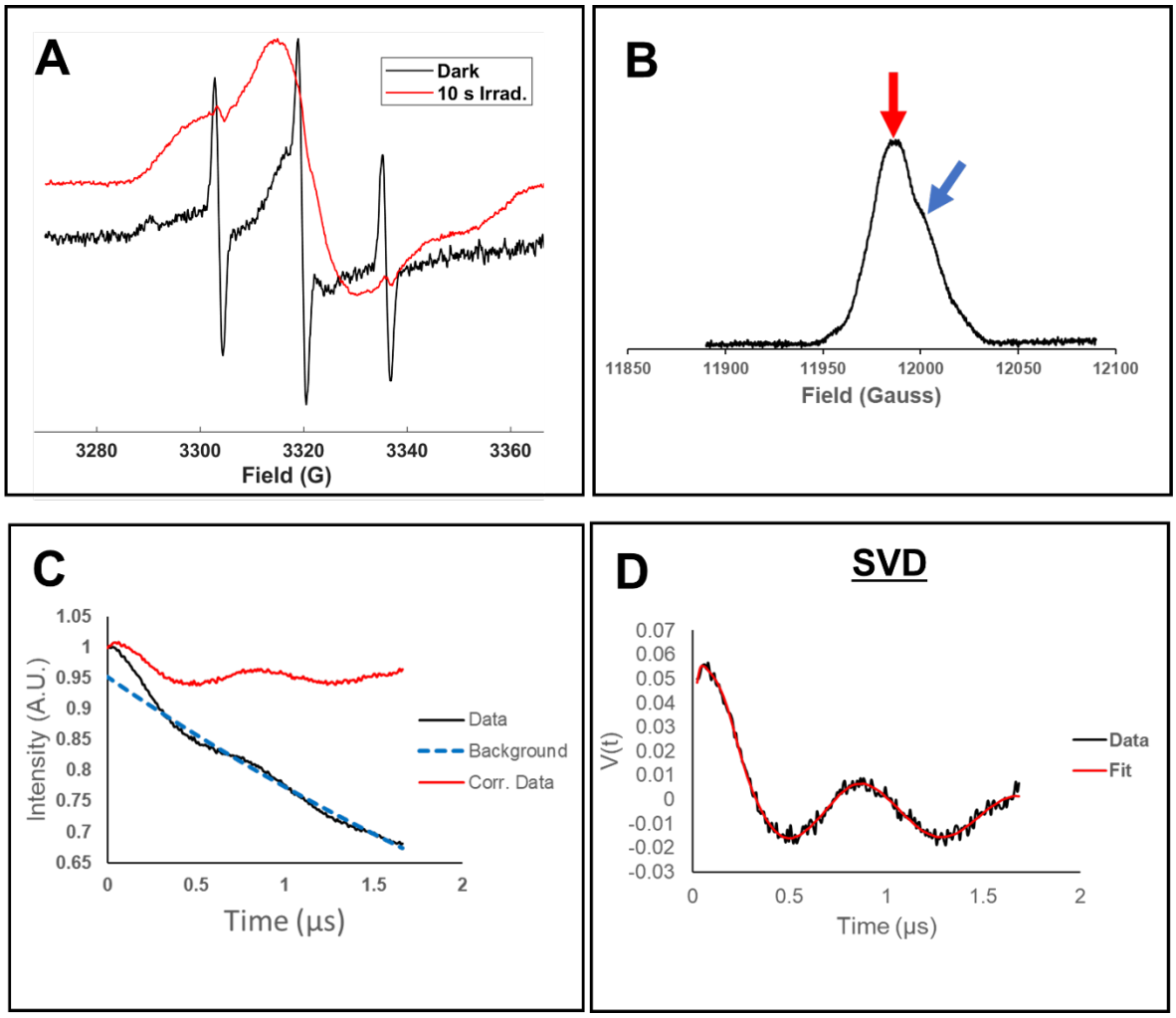


Figure S3. CRYΔ+S-TIM$_N$-SL ESR spectroscopy data. (A) X-band cw-ESR spectrum of CRYΔ+S-TIM-SL in dark (black) and after 10 seconds of irradiation with blue light (red). (B) Representative field swept echo (FSE) with pump (flavin) and probe (nitroxide) pulses marked with red and blue arrows, respectively. (C) Primary DEER trace, background subtraction and corrected (corr.) data. (D) Time domain fitting using the SVD method for distance reconstruction given in Figure 3C.

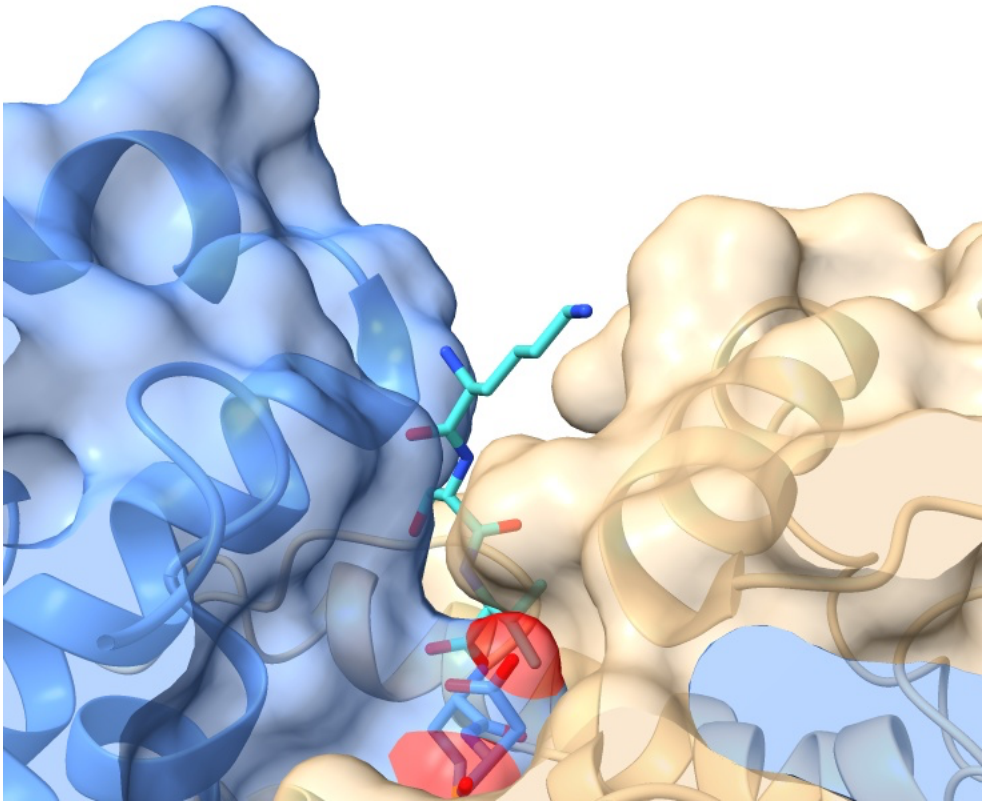


Figure S4. Modeling of Modeling of S-TIM$_N$+3 at the CRY:TIM interface. The L-TIM 3-residue N-terminal extension (Lys-Ser-Val, cyan) onto the structure of S-TIM (blue) in complex with CRY (tan; PDB:8DD7). Residues in addition to these three at the N-terminus of L-TIM provide sufficient length to limit further interactions with other regions that compose the interface.

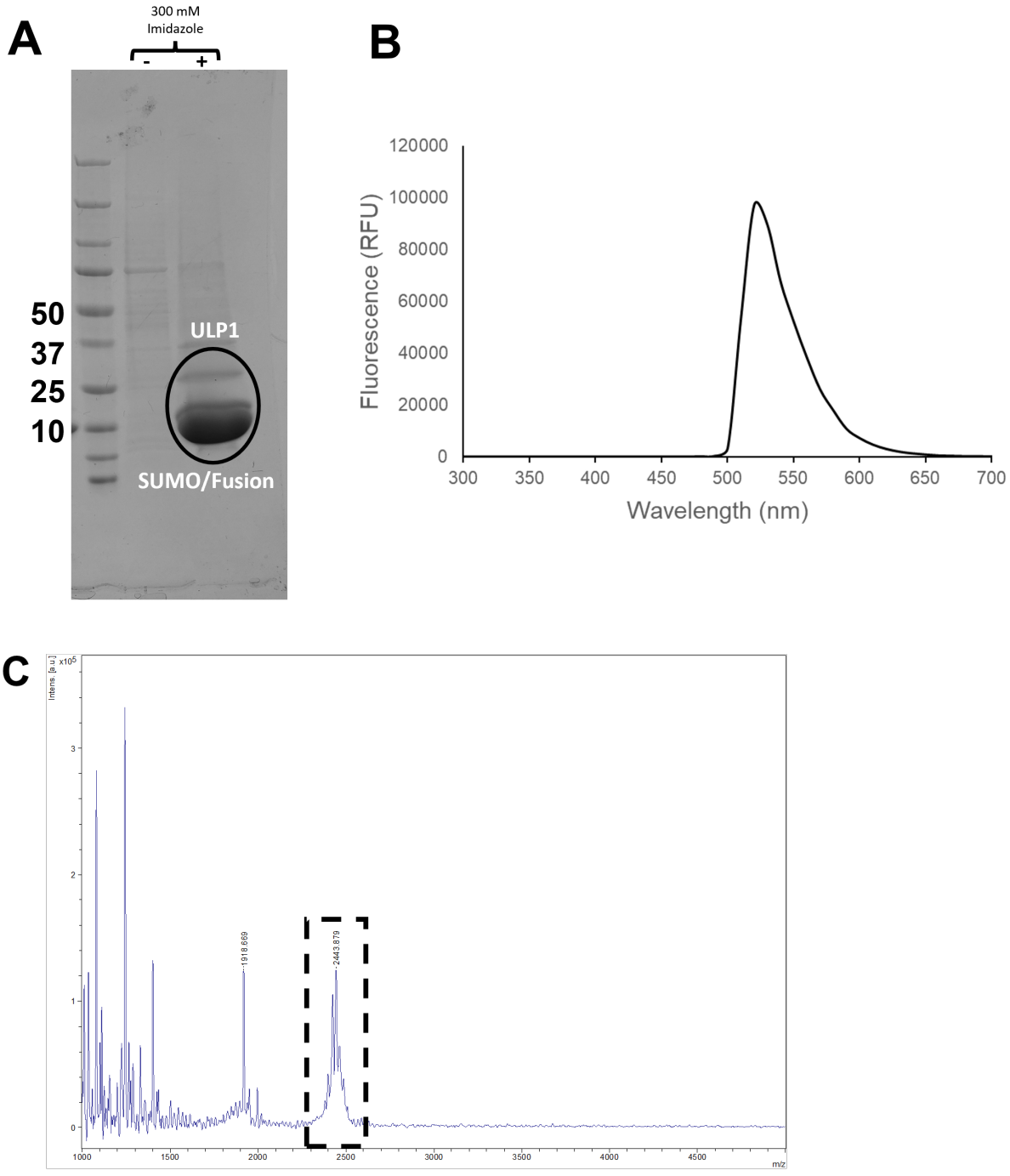


Figure S5. Purification and labeling of S-TIM$_N$-NoM. (A) SDS-PAGE gel of Ni^{2+}-NTA affinity column fractions with lanes representing elution for the fluorescently labeled peptide with or without (+/-) 300 mM imidazole. The eluent with imidazole was taken for further purification and labeling. Circled bands are for the SUMO fusion and ULP1. (B) Fluorescence emission spectrum of S-TIM$_N$-NoM labeled with fluorescein-5-maleimide. (C) MALDI of S-TIM$_N$-NoM labeled with fluorescein-5-maleimide. Intense peaks earlier in spectrum are likely degradation products.

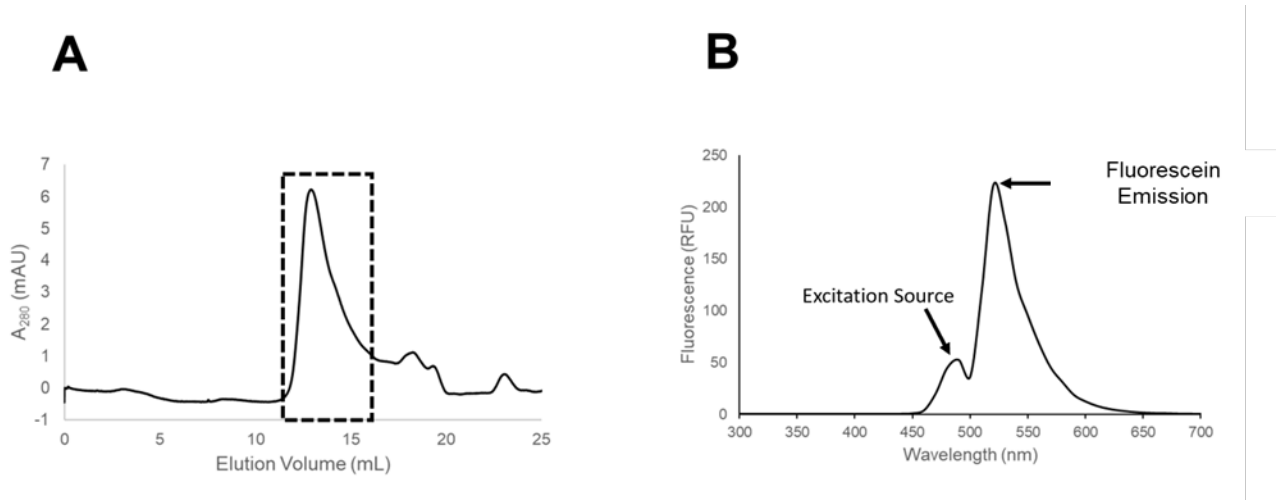


Figure S6. Purification and labeling of S-TIM_N+3. (A) Analytical S30 trace post reaction with fluorescein-5-maleimide. (B) Fluorescence emission of labeled peptide.

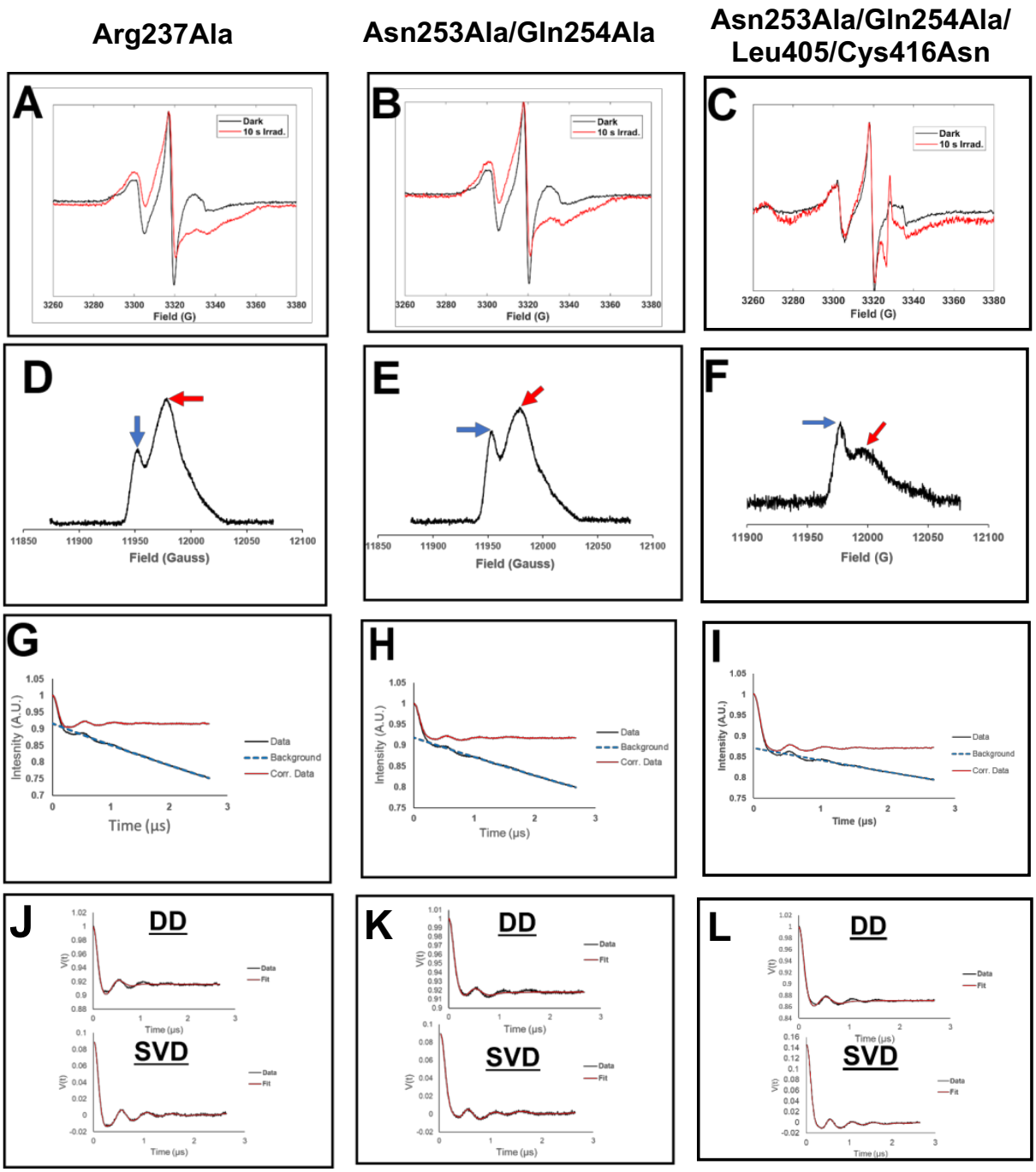


Figure S7. ESR spectroscopy data for Arg237Ala [A,C,E,G], Asn253Ala/Gln254Ala [B,D,F,H] and Asn253Ala/Gln254Ala/Leu405Glu/Cys416Asn [C,F,I,L] variants spin-labeled on the CTT. (A)/(B)/(C) X-band cwESR spectra of sort-tagged Arg237Ala and Asn253Ala/Gln254Ala in dark (black) and after 10 seconds of irradiation with blue light (red). (D)/(E)/(F) Representative field swept echo (FSE) with pump (flavin) and probe (nitroxide) pulses marked with red and blue arrows, respectively. (G)/(H)/(I) Primary DEER trace, background subtraction and corrected (corr.) data. (J)/(K)/(L) Time domain fitting using DD and SVD methods (as previously reported[12,17]) for distance reconstruction seen in Figure 5 and Figure S7.

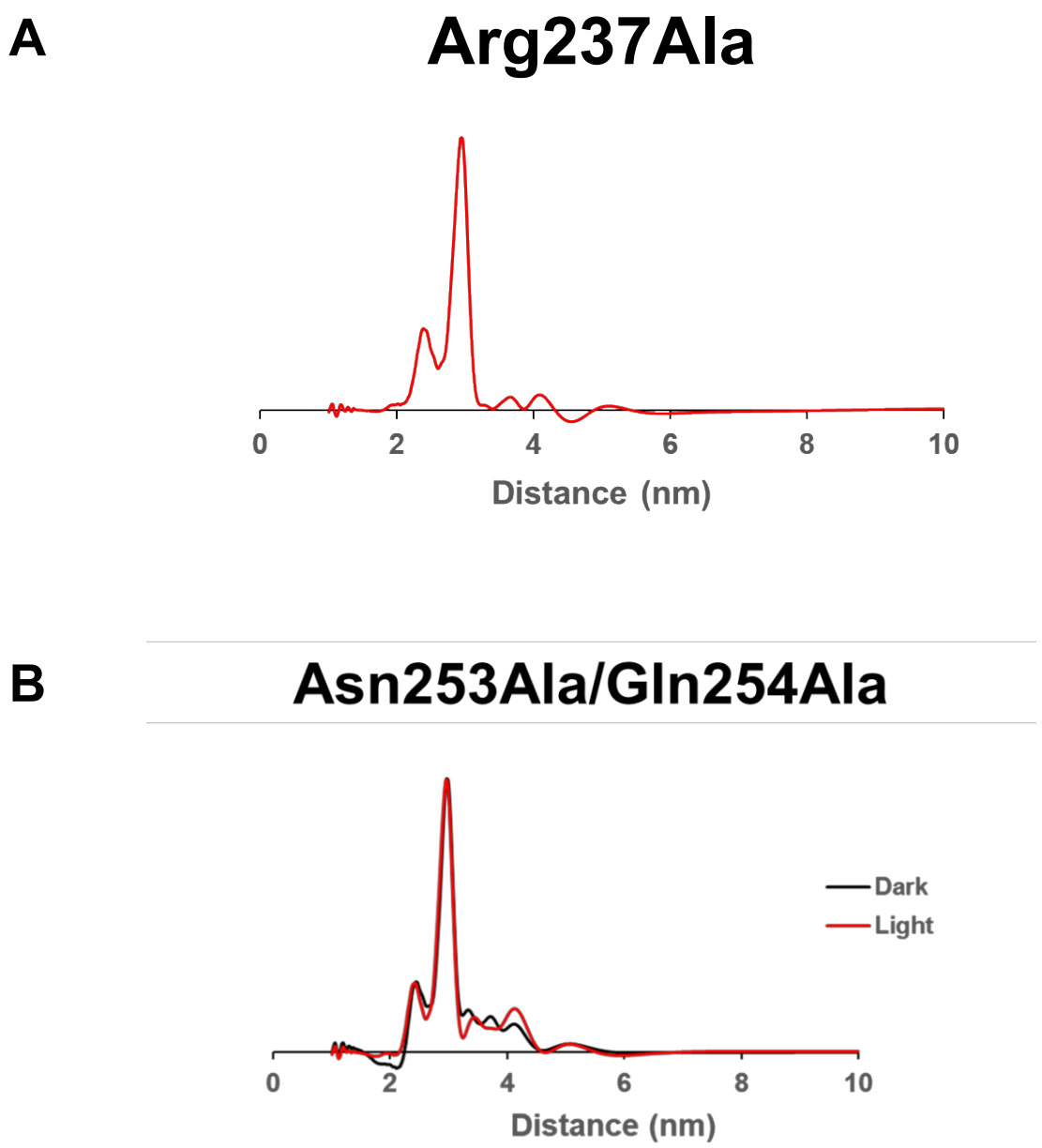


Figure S8: Distributions of CTT-SL to flavin radical distances derived by SVD for (A) Arg237Ala and (B) Asn253Ala/Gln254Ala with the flavin ASQ (Light) and flavin NSQ dark state proxy (Leu495Glu/Cys416Asn, Dark). The distance maxima are at 29.5 Å for both samples, with varying populations of longer distance components, commensurate with results calculated by the DD method in Figure 5. Note the increased populations in the longer distance components of Asn253Ala/Gln254Ala relative to Arg237Ala. Time domain data are given in Figure S6.

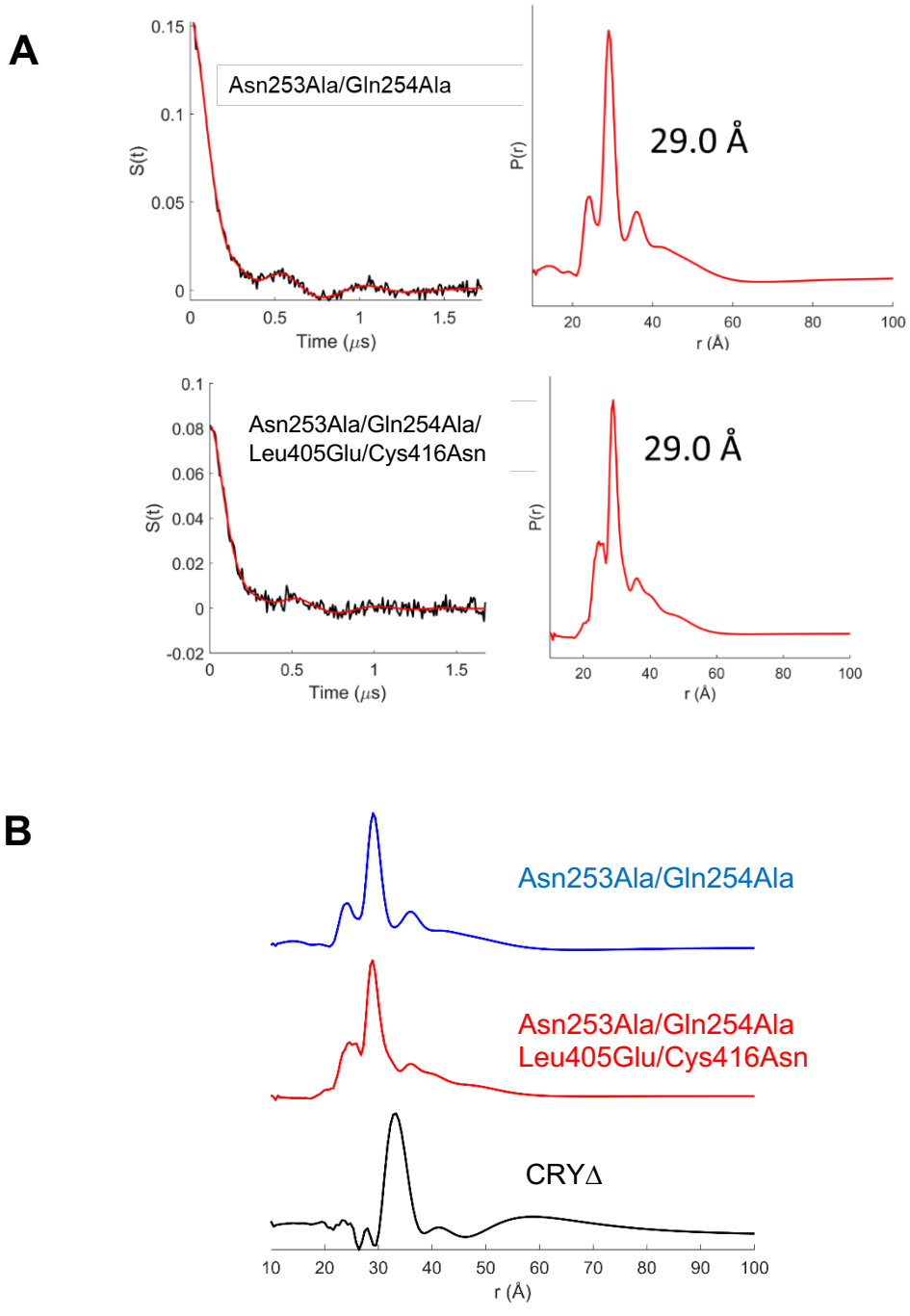


Figure S9. S-TIM$_N$ peptide binding to the CRY Asn253Ala/Gln254Ala variants (A) Time domain data (left) and corresponding SVD-derived distributions (right) for S-TIM$_N$-SL to flavin radical interactions derived by SVD. Fits to the time domain data based on the distance distributions are shown in red. Light-state ASQ data given at top and dark-state proxy "EN" data (Leu405Glu/Cys416Asn) given at the bottom. (B) Comparisons of the S-TIM$_N$ binding distributions for CRYΔ and Asn253Ala/Gln254Ala variants.