Supplementary Information for

Insights into histidine kinase activation mechanisms from the monomeric blue light sensor EL346

Igor Dikiy, Uthama R. Edupuganti, Rinat R. Abzalimov, Peter P. Borbat, Madhur Srivastava, Jack H. Freed, and Kevin H. Gardner

Email: kevin.gardner@asrc.cuny.edu

This PDF file includes:

SI Methods
Figs. S1 to S7
SI References

SI Methods

Protein expression and purification

      Escherichia coli BL21 (DE3) cells (Novagen/EMD Millipore) harboring pHis-GB1-Parallel plasmids containing WT, V115A, and V119A EL346 (UniProt ID: Q2NB77) were grown in LB media at 37 °C until reaching A_{600} of 0.8-1.2, then induced with 0.5 mM isopropyl-β-D-1-thiogalactoside and grown at 18 °C for another 16-18 hours. Cells were collected by centrifugation and resuspended in Buffer A (50 mM Tris pH 8, 100 mM NaCl) before either purifying directly or flash-freezing in liquid N_2 and storing at -80 °C for later purification.

      To purify expressed protein, resuspended cell pellets were lysed by sonication, then clarified by centrifugation at 48,000 x *g* and 4 °C for 1 hour. The supernatant was filtered with a 0.22-μm filter and loaded onto a Ni^{2+} Sepharose affinity column (GE Healthcare) equilibrated with Buffer A supplemented with 15 mM imidazole. The His-tagged protein was eluted with a 12-column volume gradient to 500 mM imidazole (eluted around 150 mM imidazole). The fusion protein was then simultaneously exchanged into imidazole-free Buffer A by overnight dialysis at 4 °C and incubated with His_6-tagged TEV protease to cleave off the His_6-Gβ1 tag. Cleaved EL346 protein was separated by flowing the sample through a Ni^{2+} Sepharose affinity column equilibrated with Buffer A supplemented with 15 mM imidazole. The flowthrough was concentrated to ~5 mL and injected onto a HiLoad 16/600 Superdex 75 column (GE Healthcare) equilibrated with HSEC Buffer (50 mM HEPES pH 7.5, 100 mM NaCl, 10 mM $MgCl_2$, 0.5 mM DTT). The peak corresponding to pure monomeric EL346 was collected and either used for further assays or flash frozen in liquid N_2 and stored at -80 °C.

      Any frozen aliquots used in other assays, except for SEC-native MS (see below), were thawed at 4 °C, then injected on a Superdex 75 10/300 GL size-exclusion column (GE Healthcare) equilibrated with HSEC Buffer. The fraction corresponding to the maximum of the peak of pure monomeric EL346 was then used. All purification steps were performed in the dark under dim red light. Protein concentration was determined from an approximate absorption coefficient $\varepsilon_{446} = 11,800$ $M^{-1} \cdot cm^{-1}$ for flavin-containing proteins, assuming all flavin was protein-bound.

SEC coupled to native-MS experiments

        A MAbPac SEC-1 2.1x150 mm column (Thermo Scientific) was placed in-line with a Bruker UHR maXis-ETD ESI-qTOF mass spectrometer and equilibrated with 50 mM ammonium acetate (pH ~ 6.8) solution (1). EL346 (40 μM) was thawed, incubated in HSEC Buffer supplemented with 5 mM ADP, and either kept in the dark or illuminated with 4 flashes of a camera flash prior to injection in 5 μL aliquots and isocratic elution with the same ammonium acetate solution at a flow rate of 50 μL/min. In separate experiments, BSA and ovalbumin (both 30 μM) were each injected as above to serve as known MW standards. Elution was monitored by UV/vis absorbance at 280, 405, and 450 nm. Average mass spectra for the major elution peak were generated and subsequently deconvoluted using maximum entropy reconstruction in Bruker DataAnalysis software.

Peptide-level HDX-MS measurements

        WT, V115A, or V119A EL346 protein in HSEC Buffer was concentrated to 50 μM then microcentrifuged. This stock was kept in the dark on ice, and aliquots of 1 μL (or 2 μL for 6 hour time points) were taken from the top in the dark and allowed to equilibrate at 20 °C under either dark or light [blue LED strip (Adafruit Industries) pulsing at 20s on, 60s off] for 5-10 minutes. Deuterium exchange was initiated by diluting with 99 μL (or 198 μL for 6 hour time points) of D_2O HDX buffer (50 mM Tris, 100 mM NaCl, 10 mM $MgCl_2$, 0.5 mM DTT, 5 mM ATP or ADP, 100% D_2O, titrated to an uncorrected pH reading of 7.1). Exchange continued for 30, 60, 180, 900 (15 minutes), 3600 (1 hour), or 21600 (6 hours) seconds at the same temperature and light conditions as equilibration, then cooled on ice for 30 seconds. The 6 hour time point samples were microcentrifuged to pellet any aggregates, and the top 100 μL removed and cooled as above. Exchange was quenched by addition of 100 μL of quench buffer [50 mM ammonium acetate, 100 mM guanidinium hydrochloride, formic acid added until a 1:1 mix with HDX buffer is pH 2.5 (~1.5%)]. All HDX buffers and quench buffer were filtered with 0.2-μm filter before use.

All subsequent steps were performed in an ice bath. The quenched protein sample was immediately injected over an Enzymate BEH pepsin column (Waters). The digestion with subsequent desalting of resulting peptides was performed at a flow rate of 0.12 mL/min using 0.25% formic acid in water as a mobile phase. About 3 minutes after injection, allowing for pepsin digestion and desalting, the mixture of peptides was resolved with an analytical C18 column (Hypersil GOLD 1x50 mm, Thermo Scientific) and quickly eluted into the Bruker maXis-ETD UHR ESI-qTOF at a flow rate of 40 µL/min. Overall time between quench and elution was 5-10 minutes depending on the peptide, all within 0-4 °C and pH 2.5, to minimize back-exchange.

All HDX experiments were accompanied by one "unlabeled" run in which D_2O HDX buffer was replaced by identical buffer in 100% H_2O, to provide masses and retention times of undeuterated peptides. One unlabeled run was performed with MS/MS fragmentation to identify peptides using Bruker DataAnalysis and BioTools software. Every matched m/z and retention time pair was assigned to only one peptide sequence based on accurate mass and MS/MS measurements. This resulted in coverage of about 67% of the protein sequence, with a majority of the LOV domain missing peptides. Resulting peptide (sequence, m/z, charge state, and retention time) lists and mass spectra were analyzed in HDExaminer (Sierra Analytics) software. All peptide matches were manually confirmed after automatic assignment by HDExaminer.

HDExaminer heat maps (and difference heat maps) were used to generate structure maps (**Figs. 2-5**), direct peptide % deuteration values were used to generate peptide difference plots (**Figs. 2-4**), and direct peptide number of deuterons were used for uptake plots (**Figs. 3,5**). Structure maps were generated in PyMol (The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC.) and plots were drawn in Matplotlib/PyPlot (2) using data exported from HDExaminer.

NMR sample generation and experiments

Our previous NMR studies on EL346 (3) showed several intense peaks, likely from C-terminal flexible residues. To remove these peaks which dominate the spectrum, we generated an EL346 1-338 construct in which a stop codon was inserted in place of residue Ile339 using the partially-overlapping primer site-directed mutagenesis scheme

outlined by Liu and Naismith (4). This truncation position was selected because there was no density modeled for residues C-terminal to Leu338 in the dark state crystal structure (3).

$^{15}N/^{1}H$ TROSY HSQC spectra were collected at 25 °C on ~200 μM EL346 1-338 in NMR buffer (50 mM Tris pH 7.5, 100 mM NaCl, 10 mM $MgCl_2$, 5% D_2O) with either ADP or AMP-PNP at 5 mM using a Bruker AVANCE III HD 600 MHz spectrometer with a cryogenically-cooled TCI probe. The spectra were collected using a standard Bruker "sensitivity-enhanced" pulse sequence (5) after allowing the sample to equilibrate to the dark state for 4 hours in the bore of the magnet, using a ^{15}N spectral width of 35 ppm centered at 117 ppm, with 64 complex points and 192 scans averaged at each point, for a total acquisition time of ~7.5 hours. NMR data were processed with NMRpipe (6) without linear prediction and analyzed with NMRViewJ (7).

Site-directed spin labeling and DEER sample preparation

To create the constructs for labeling, the C78S, C111A, and C291S mutations were first introduced into WT, V119A, and V115A EL346 to remove all cysteine residues aside from the flavin-binding cysteine (Cys55) using partially-overlapping primer site-directed mutagenesis. The I188C and V270C mutations were introduced simultaneously using the same protocol. The mutant proteins were expressed and purified as described above, except that V115A EL346 I188C V270C was not cleaved with His_6-TEV protease to reduce material loss due to low expression of this mutant.

The cysteine mutant protein was concentrated to ~50 μM and exchanged into HSEC Buffer without DTT using a PD-10 gravity column (GE Healthcare). About 4 equivalents of MTSL (70 mM stock in DMSO) were added and the sample was incubated with rocking for 45 minutes at room temperature. The solution was again exchanged into HSEC Buffer without DTT using a PD-10 column to remove excess MTSL. Samples were removed before and after labeling for MALDI confirmation of covalent spin-label addition (+184 Da for each spin-label); labeling efficiency was approximately 75% double-labeled protein.

Labeled proteins were concentrated to 50-100 μM and exchanged into DEER buffer [50 mM Tris pH 7.5, 100 mM NaCl, 10 mM $MgCl_2$, 5 mM DTT, 5 mM ADP or

5

AMP-PNP, 80% D_2O, 20% d-8 glycerol (both Cambridge Isotope Labs)] then kept at 4 °C until freezing for DEER data acquisition (usually less than 24 hours). Before freezing, protein samples were incubated at room temperature for 15 minutes in the dark, then split into dark and light samples. Dark samples were loaded into ESR sample tubes (30 μL) and flash-frozen in liquid N_2 in the dark under dim red light. Light samples were kept under blue LED light for 5 minutes, then loaded and frozen the same as dark samples. All samples were kept in the dark in liquid N_2 until data collection.

Autokinase assays

Autophosphorylation was measured by incubating EL346 ("WT," wild-type; and "MTSL," C78S C111A C291S I188C V270C labeled with MTSL) at 27 μM in 50 mM Tris pH 8.2, 100 mM NaCl, 5 mM $MnCl_2$, 10% glycerol buffer, with 3.6 μM ATP (including 10 μCi γ-[^{32}P]ATP [PerkinElmer, 6000 Ci/mmol] in a 100 μL reaction), removing 10 μL samples at 0.5, 1, 2, 4, 8, 16, and 32 minute time points, quenching with 4x Laemmli buffer (Bio-Rad) with added 2-mercaptoethanol; separating using SDS-PAGE; and imaging the gels using the phosphorimaging mode of a Typhoon FLA 9500 imager. Two reactions for each protein and each condition (dark/light) were carried out at room temperature. Standards of known concentration were blotted on filter paper and imaged together with the gels, then used to calculate the amount of autophosphorylation from the band intensity quantified in ImageJ (8). Data from both reactions for each condition were pooled and plotted, with linear fits performed to data between 0.5 and 2 minutes (inclusive) to extract the slope as the initial rate. Errors in the initial rate were calculated from the covariance matrix of the fit.

Pulsed ESR measurements (DEER) and reconstruction of distance distributions

DEER measurements were performed at cryogenic temperature (60 K) on a home-built 17.3 GHz Ku-band pulse ESR spectrometer as previously described (9, 10) using a standard four-pulse DEER sequence (11). The π/2 (π) pulse width was 16 (32) ns, with a 32 ns pump π-pulse. The frequency separation between detection and pump pulses was 70 MHz, positioning the detection at the low-field edge of the nitroxide spectrum.

Typical dipolar evolution times were 1.4–2 μs, making experiments last from 2–4 hours up to 24 hours depending on signal averaging as required.

Background was subtracted from the raw time-domain DEER data using DeerAnalysis software (12) and distance distributions reconstructed by Tikhonov regularization (13) followed by the maximum entropy regularization method (MEM) (14) (**Fig. S6, top**). To remove noise from DEER signals, a recently-developed wavelet denoising method (15, 16) was applied on the original data, baseline was subtracted in the log domain, and distance distributions reconstructed from the resulting denoised time-domain data by singular value decomposition (SVD) (17), as Tikhonov regularization applied to denoised data may result in spurious peaks. In the SVD reconstruction method, a potential source of error is which singular values are selected for the reconstruction. To quantify the possible error, we measured the peak heights at the two distances of interest using 25 different singular value cutoffs and calculated the mean and standard deviation (18). In all datasets reported here, the standard deviation was between 0.2% and 1% of the mean peak height, indicating that choice of singular values does not introduce significant errors.
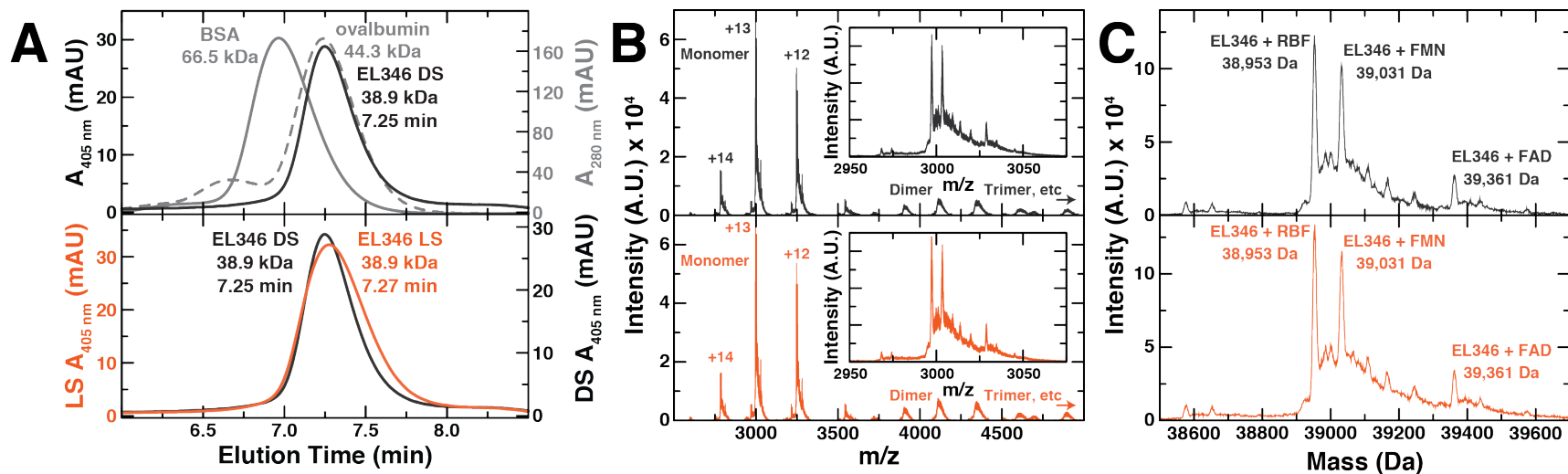
Figure S1: SEC coupled to native MS shows only flavin-bound monomeric protein. A) Visible absorbance chromatogram (405 nm) of WT EL346 injected on a MAbPac SEC-1 column without (black) and with (vermillion) illumination before injecting. Illumination resulted in a small shift in elution time from 7.25 to 7.27 min. The elution time roughly corresponds to a monomeric globular protein of 39 kDa. Elution of BSA and ovalbumin monitored by UV absorbance at 280 nm is shown in gray solid and dashed lines, respectively, as a comparison. B) Average mass spectra calculated from the monomer elution peak of the runs in (A). A charge state distribution for monomeric EL346 is seen for both dark and light samples, with very low signal corresponding to dimers and trimers. The inset shows a zoomed view of the +13 monomer peak. C) Deconvoluted (with maximum entropy) mass spectra calculated from those shown in (B). The leftmost peak corresponds to EL346 + riboflavin (RBF), while the second highest intensity peak corresponds to EL346 + FMN and the third to EL346 + FAD.
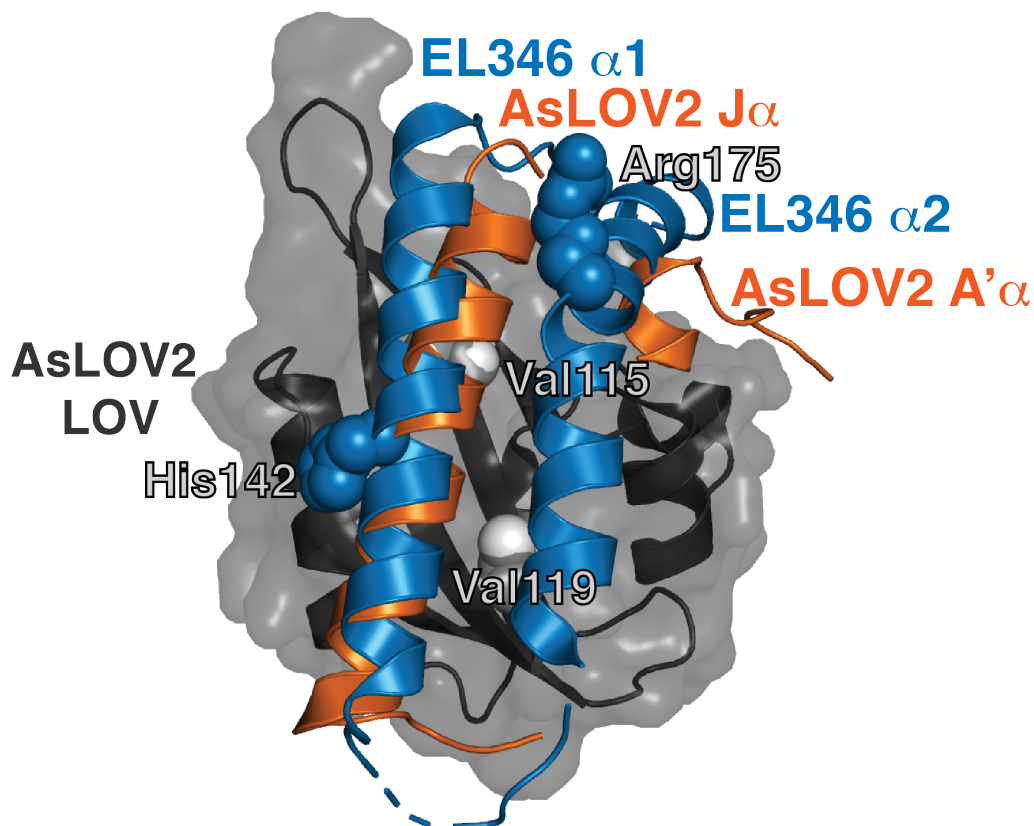
Figure S2: EL346 LOV:α1 interactions are analogous to AsLOV2:Jα helix interactions. Crystal structure of AsLOV2 (PDB: 2V0U) (19) showing the LOV domain in dark gray cartoon and surface representation with A'α and Jα helices as vermillion cartoons. EL346 coordinates were superimposed on this, based off an alignment of the EL346 dark state structure (PDB: 4R3A) (3) LOV domain with the AsLOV2 LOV domain, showing the α1 and α2 helices of the DHpL domain as blue cartoons. The DHpL helices partially occupy the same binding sites as the N- and C-terminal helices in AsLOV2. EL346 Val115 and Val119 residues are shown as light gray spheres and mostly contact α1 and α2 helices, respectively; EL346 His142 and Arg175 residues are shown as spheres in the DHpL helices.
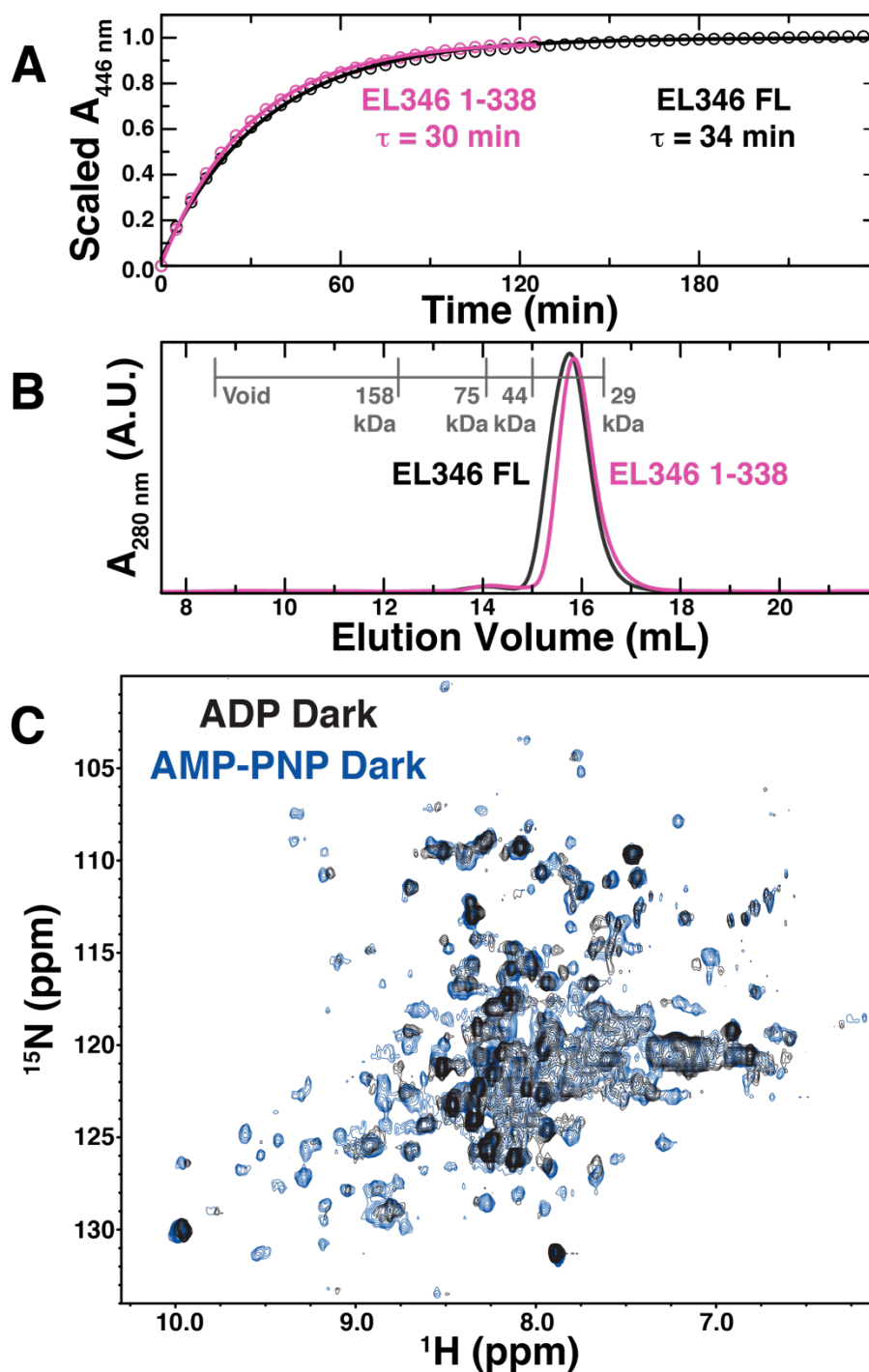
Figure S3: EL346 is less stable in the presence of ADP by NMR. A) Dark state reversion kinetics measured by gain of absorbance at the dark-state flavin maximum at 446 nm for EL346 FL (black) and 1-338 (purple). Experimental data (circles) are fit to exponential decay functions (lines) with time constants of 34 and 30 minutes for EL346 FL and 1-338, respectively. B) SEC elution chromatogram of EL346 FL (black) and 1-338 (purple) on a Superdex 200 10/300 GL column. Both proteins elute at a very similar volume. Globular protein standards are shown in gray as a comparison. C) $^{15}N/^{1}H$ TROSY HSQC spectra of EL346 1-338 in the dark in the presence of ADP (black) and AMP-PNP (blue). The structured β-strand region (left) has missing peaks in the ADP spectrum.
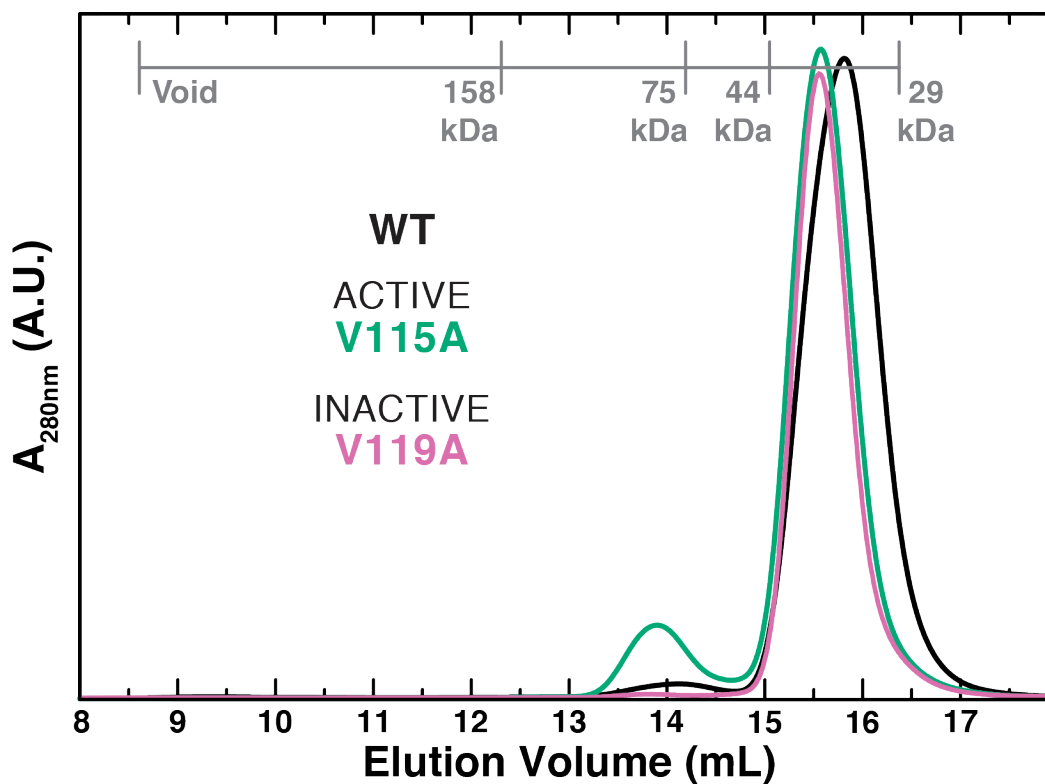
Figure S4: V115A and V119A mutations have minor effects on overall EL346 shape. SEC elution chromatograms of EL346 WT (black), active V115A (green), and inactive V119A (purple) on a Superdex 75 10/300 GL column, in the absence of nucleotide and in the dark. All proteins elute at a very similar volume, although mutants are shifted to a slightly earlier elution, indicating a similar, if slightly expanded, molecular shape of the mutant proteins. Globular protein standards are shown in gray as a comparison.
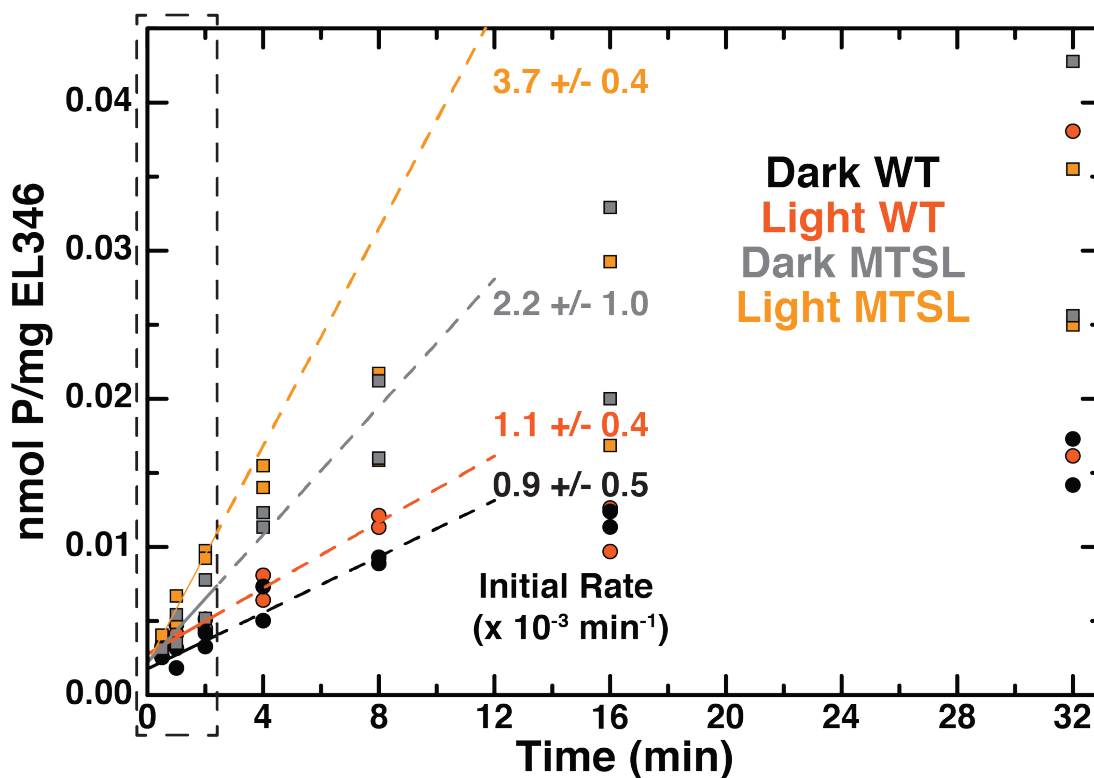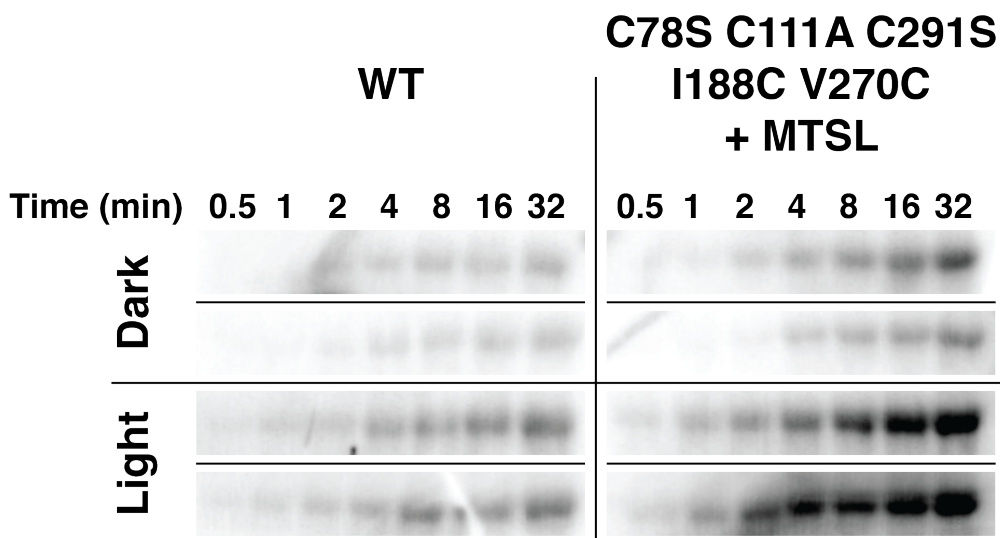
Figure S5: MTSL-labeled DEER sample shows slightly increased autophosphorylation activity. *Top* –Autoradiograms from autophosphorylation assays on WT EL346 and C78S C111A C291S I188C V270C MTSL-labeled ("MTSL") EL346 in the dark and light. *Bottom* – Plots of autophosphorylation of WT and MTSL-labeled mutant EL346 in the dark and light. Points are combined data from two reactions. Colors: WT, dark (black); WT, light (vermillion); MTSL-labeled, dark (gray); and MTSL-labeled, light (orange). Lines show linear fits to data between 0.5 and 2 minutes (inclusive, box), with the slope indicated as the initial rate next to the lines. Error is calculated from the covariance matrix of the fit.
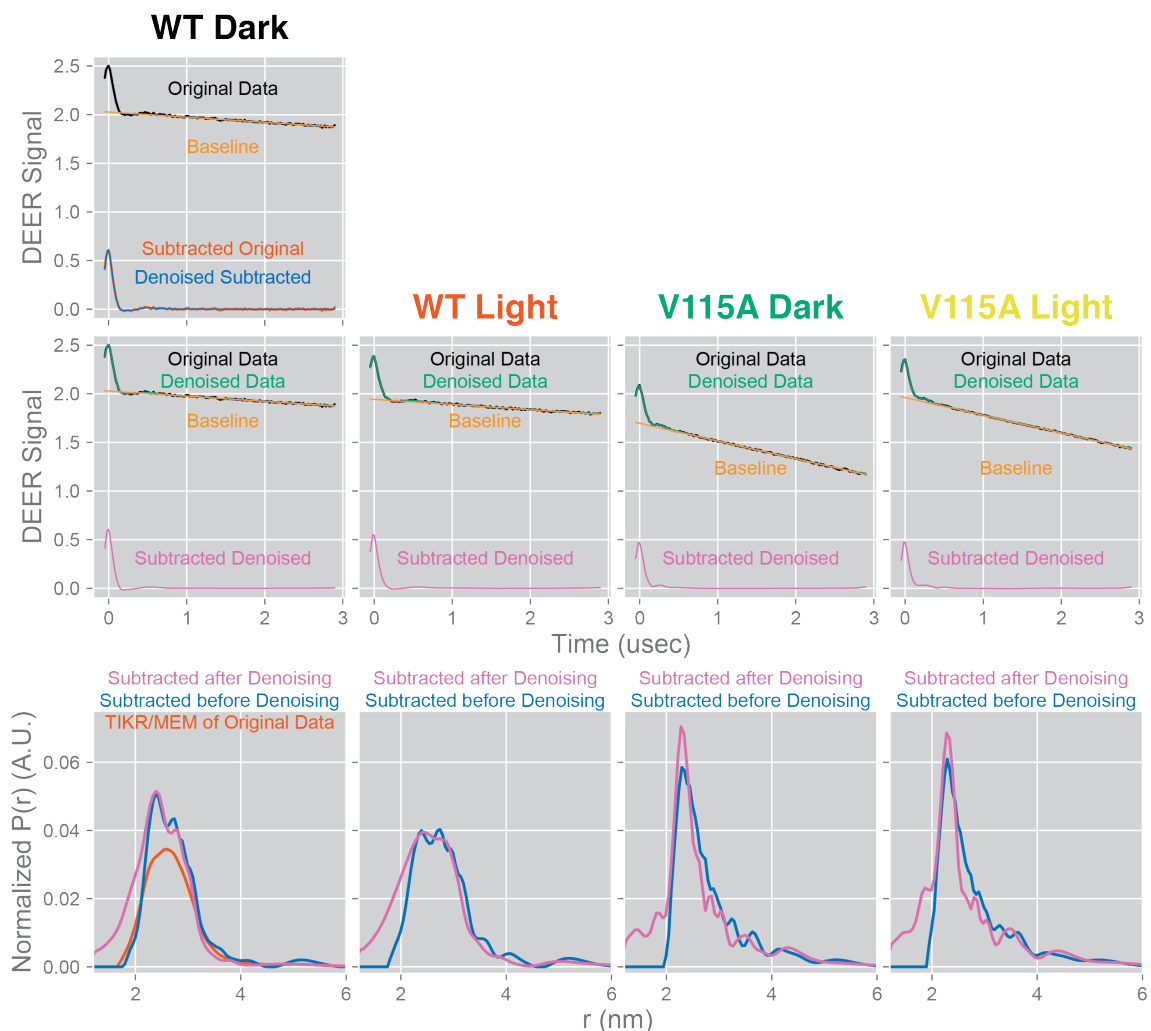
Figure S6: Examples of DEER data processing. Top – Baseline subtraction before denoising: Raw DEER signal (black), calculated baseline (orange), subtracted signal (vermillion), and denoised subtracted signal (blue). Middle – Baseline subtraction after denoising: Raw DEER signal (black), denoised DEER signal (green), calculated baseline (orange), and subtracted denoised signal (purple) (15, 16). Bottom – Distance distributions reconstructed by SVD (17, 18) on data subtracted after denoising (purple) or before denoising (blue), and by Tikhonov regularization with MEM (13, 14) on subtracted original data (vermillion). Middle panels show baseline subtraction for four of the distance distributions shown in Figure 6. Top panel illustrates denoising after subtracting for one dataset. All data are from DEER experiments of WT and V115A EL346 in the dark and light labeled at position 188 (DHpL) and 270 (CA) in the presence of ADP.
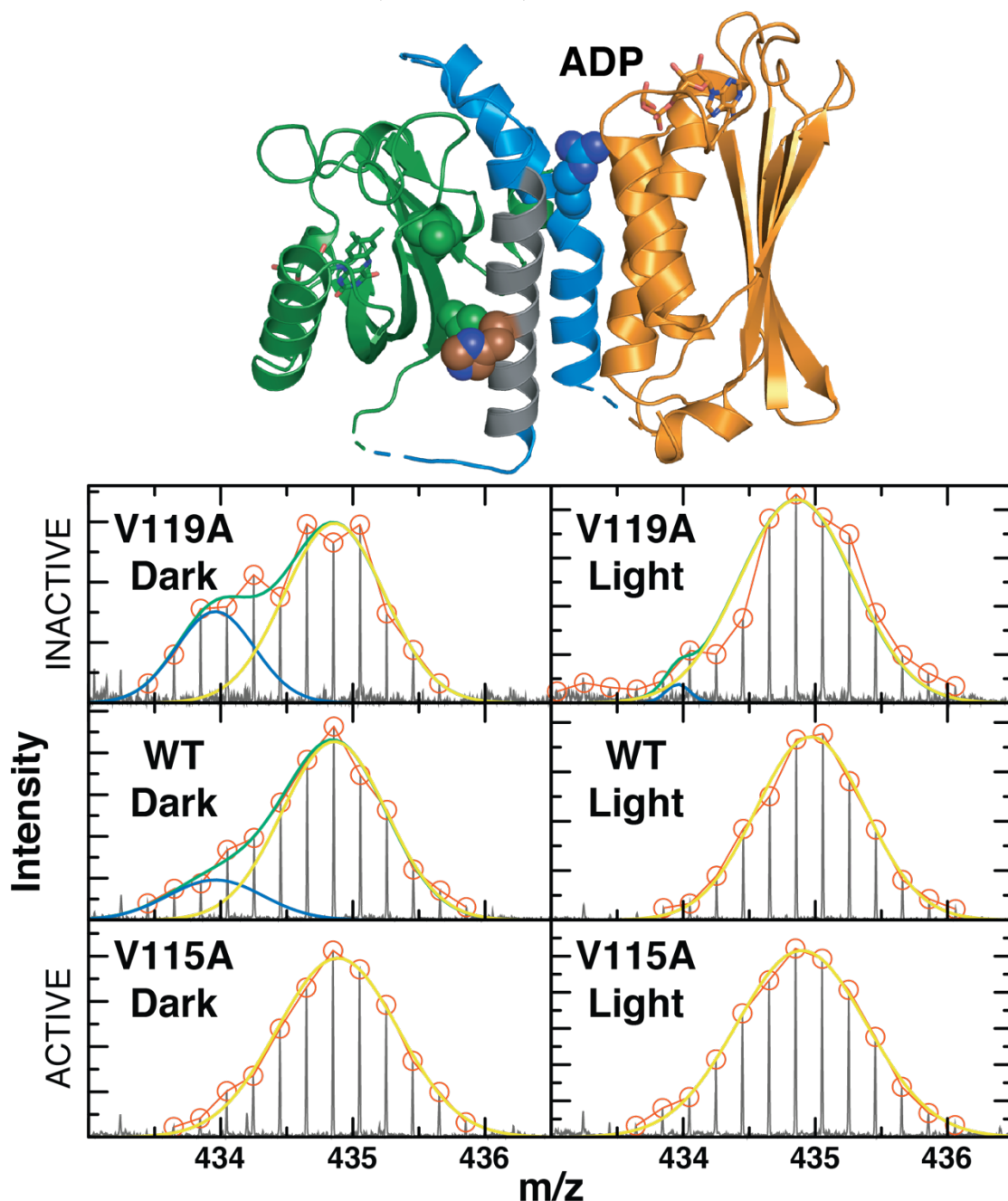
Figure S7: Bimodal peptide exchange profiles indicate slow conformational exchange.
Sample extracted MS spectra of a peptide corresponding to residues 134-151, including the
phosphoacceptor His142, highlighted in vermillion, in the dark (left) and light (right) for EL346
V119A (top), WT (middle), and V115A (bottom) in the presence of ADP. Peptide is highlighted
on the dark state crystal structure in dark gray. Spectra are shown in gray and peaks highlighted
in vermillion circles. The peaks are fit to either a sum (green) of two gaussian peaks centered at
m/z of 433.9 (blue) and 434.9 (yellow) or a single gaussian centered at m/z of 434.9 (yellow).
Dark state WT peptide shows bimodal exchange in EX1, as does inactive V119A, while light WT
and active V115A show a single distribution at the higher exchange value.

References

1. Konermann L (2017) Addressing a common misconception: Ammonium acetate as neutral pH "buffer" for native electrospray mass spectrometry. *J Am Soc Mass Spectrom* 28(9):1827-1835.
2. Hunter JD (2007) Matplotlib: A 2D graphics environment. *Comp. Sci. Eng.* 9(3):90-95.
3. Rivera-Cancel G, Ko WH, Tomchick DR, Correa F, & Gardner KH (2014) Full-length structure of a monomeric histidine kinase reveals basis for sensory regulation. *Proc Natl Acad Sci U S A* 111(50):17839-17844.
4. Liu H & Naismith JH (2008) An efficient one-step site-directed deletion, insertion, single and multiple-site plasmid mutagenesis protocol. *BMC Biotechnol* 8:91.
5. Czisch M & Boelens R (1998) Sensitivity enhancement in the TROSY experiment. *J Magn Reson* 134(1):158-160.
6. Delaglio F*, et al.* (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR* 6(3):277-293.
7. Johnson BA (2004) Using NMRView to visualize and analyze the NMR spectra of macromolecules. *Methods Mol Biol* 278:313-352.
8. Schneider CA, Rasband WS, & Eliceiri KW (2012) NIH Image to ImageJ: 25 years of image analysis. *Nat Methods* 9(7):671-5.
9. Borbat PP, Crepeau RH, & Freed JH (1997) Multifrequency two-dimensional Fourier transform ESR: an X/Ku-band spectrometer. *J Magn Reson* 127(2):155-167.
10. Georgieva ER, Borbat PP, Norman HD, & Freed JH (2015) Mechanism of influenza A M2 transmembrane domain assembly in lipid membranes. *Sci Rep* 5:11757.
11. Pannier M, Veit S, Godt A, Jeschke G, & Spiess HW (2000) Dead-time free measurement of dipole-dipole interactions between electron spins. *J Magn Reson* 142(2):331-340.
12. Jeschke G*, et al.* (2006) DeerAnalysis2006 - a comprehensive software package for analyzing pulsed ELDOR data. *Appl Magn Reson* 30(3-4):473-498.
13. Chiang YW, Borbat PP, & Freed JH (2005) The determination of pair distance distributions by pulsed ESR using Tikhonov regularization. *J Magn Reson* 172(2):279-295.
14. Chiang YW, Borbat PP, & Freed JH (2005) Maximum entropy: a complement to Tikhonov regularization for determination of pair distance distributions by pulsed ESR. *J Magn Reson* 177(2):184-196.
15. Srivastava M, Anderson CL, & Freed JH (2016) A new wavelet denoising method for selecting decomposition levels and noise thresholds. *IEEE Access* 4:3862-3877.
16. Srivastava M, Georgieva ER, & Freed JH (2017) A new wavelet denoising method for experimental time-domain signals: Pulsed dipolar electron spin resonance. *J Phys Chem A* 121(12):2452-2465.

17.   Srivastava M & Freed JH (2017) Singular value decomposition method to determine distance distributions in pulsed dipolar electron spin resonance. *J Phys Chem Lett* 8(22):5648-5655.

18.   Srivastava M & Freed JH (2019) Singular value decomposition method to determine distance distributions in pulse dipolar electron spin resonance: II. Estimating uncertainty. *J Phys Chem A* 123(1):359-370.

19.   Halavaty AS & Moffat K (2007) N- and C-terminal flanking regions modulate light-induced signal transduction in the LOV2 domain of the blue light sensor phototropin 1 from Avena sativa. *Biochemistry* 46(49):14001-14009.