Structural basis of photosensitivity in a bacterial light-oxygen-voltage/helix-turn-helix (LOV-HTH) DNA-binding protein

Abigail I. Nash[a,1], Reginald McNulty[b,1], Mary Elizabeth Shillito[b], Trevor E. Swartz[c,2], Roberto A. Bogomolni[c], Hartmut Luecke[b,3], and Kevin H. Gardner[a,3]

[a]Departments of Biochemistry and Pharmacology, University of Texas Southwestern Medical Center, 5323 Harry Hines Boulevard, Dallas, TX 75390–8816; [b]Center for Biomembrane Systems, Department of Molecular Biology and Biochemistry, Department of Physiology and Biophysics, Department of Information and Computer Science, University of California, Irvine, CA 92697–3900; and [c]Department of Chemistry and Biochemistry, University of California, Santa Cruz, CA 95064

Edited by J. Clark Lagarias, University of California, Davis, CA, and approved April 22, 2011 (received for review January 11, 2011)

Light-oxygen-voltage (LOV) domains are blue light-activated signaling modules integral to a wide range of photosensory proteins. Upon illumination, LOV domains form internal protein-flavin adducts that generate conformational changes which control effector function. Here we advance our understanding of LOV regulation with structural, biophysical, and biochemical studies of EL222, a light-regulated DNA-binding protein. The dark-state crystal structure reveals interactions between the EL222 LOV and helix-turn-helix domains that we show inhibit DNA binding. Solution biophysical data indicate that illumination breaks these interactions, freeing the LOV and helix-turn-helix domains of each other. This conformational change has a key functional effect, allowing EL222 to bind DNA in a light-dependent manner. Our data reveal a conserved signaling mechanism among diverse LOV-containing proteins, where light-induced conformational changes trigger activation via a conserved interaction surface.

allosteric regulation | photosensing | PER-ARNT-SIM domain

Environmental sensory proteins play a crucial function for cellular adaptation in response to changing conditions. These proteins frequently contain effector domains whose activity is regulated by specialized sensory domains sensitive to various stimuli. One widely distributed class of such sensory domains is the PAS (PER-ARNT-SIM) family, whose members typically regulate protein/protein interactions in response to changing environmental cues (1). A subset of PAS domains, called light-oxygen-voltage (LOV) domains, use flavin cofactors to detect changes in blue light intensity or redox state (2). LOV domains are found in regulatory proteins for phototropism (3), seasonal gene transcription (4), bacterial stress responses (5, 6), and many other diverse biological responses. Within these pathways, LOV domains control a wide range of effector domains, including kinases, F boxes, and DNA-binding domains (7). Recently, these natural proteins have been joined by engineered LOV fusions that confer in vitro and in vivo LOV-based photoregulation to a range of protein targets (8–11).

This raises the question: How can a class of light-regulated domains with similar tertiary structures control such a wide variety of effectors? What is clear is that LOV domains all share similar architectures and photochemical responses to illumination, harnessing the energy of incoming blue light photons to form a covalent adduct between the Sγ sulfur on a conserved cysteine residue and the C4a carbon of a flavin cofactor (12, 13). Formation of this bond generates structural changes that propagate to the domain surface, altering the interactions of the core LOV domain with intra- or interprotein partners (14–18). For example, structural studies on *Avena sativa* phototropin 1 LOV2 (AsLOV2) demonstrated light-induced unfolding of the Jα-helix located C-terminal to the canonical LOV domain (15). Similarly, the *Neurospora crassa* VIVID protein reorients an N-terminal α-helix, β-strand extension of its LOV domain upon illumination (18). In both cases, the external structures interact with the β-sheet surface of the LOV domain, suggesting a site for signal propagation common between them. The functional importance of regulated interactions at this site have been validated by the ability of point mutations on the β-sheet or interacting effector surfaces to decouple changes in effector activity from adduct formation (18, 19).

Among the known LOV-containing proteins are several transcription factors, such as the zinc-finger containing *N. crassa* white collar-1 (WC-1) (20) and the algal basic leucine zipper AUREOCHROMEs (21). Although light controls the binding of these proteins to DNA, the mechanism(s) of this regulation is not understood at a molecular level. Here we address this shortcoming by examining how a LOV domain directly regulates DNA binding, establishing the generality of LOV signaling. Our studies focus on EL222, a 222 amino acid protein isolated from the marine bacterium *Erythrobacter litoralis* HTCC2594. In addition to an N-terminal LOV domain, EL222 also contains a C-terminal helix-turn-helix (HTH) DNA-binding domain representative of LuxR-type DNA-binding proteins (22). Combining regulatory models from a diverse group of LOV-based photosensors (15) and LuxR-type proteins (23), we hypothesized that the EL222 N-terminal LOV domain represses DNA-binding activity of the C-terminal domain in the dark, and that this inhibition would be released with blue light illumination.

Results

Dark-State Crystal Structure of EL222 Suggests Mode to Inhibit DNA Binding. As an initial step to examining this model, we solved the 2.1-Å resolution crystal structure of EL222 in the dark state (Table S1), observing interactions between the LOV and HTH domains consistent with our hypothesis (Fig. 1). The EL222 structure contains both of the two expected domains, an N-terminal

Author contributions: A.I.N., R.M., H.L., and K.H.G. designed research; A.I.N., R.M., M.E.S., H.L., and K.H.G. performed research; A.I.N., R.M., M.E.S., T.E.S., and R.A.B. contributed new reagents/analytic tools; A.I.N., R.M., H.L., and K.H.G. analyzed data; and A.I.N., R.M., H.L., and K.H.G. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Data deposition: The atomic coordinates and structure factor amplitudes of EL222 have been deposited in the Protein Data Bank, www.pdb.org (PDB ID code 3P7N) and NMR chemical shifts with the BioMagResBank, www.bmrb.wisc.edu (accession no. 17640).

[1]A.N. and R.M. contributed equally to this work.

[2]Present address: Early Stage Pharmaceutical Development, Genentech, Inc., South San Francisco, CA 94080.

[3]To whom correspondence may be addressed. E-mail: hudel@uci.edu or Kevin.Gardner@utsouthwestern.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1100262108/-/DCSupplemental.

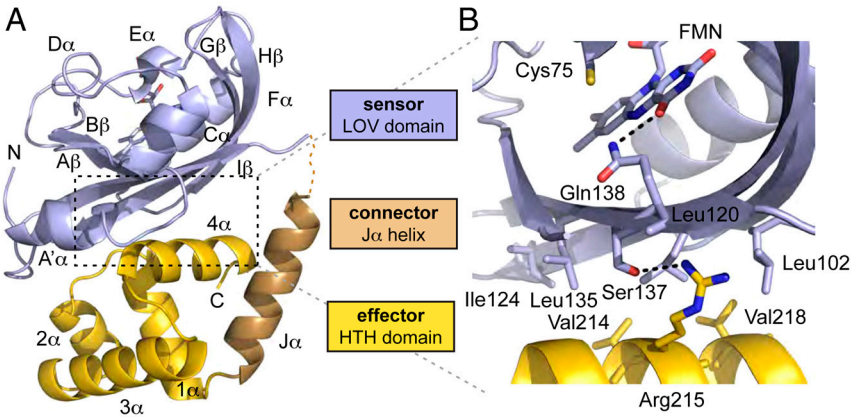

Fig. 1. The dark-state crystal structure of EL222 reveals extensive LOV-HTH interactions predicted to inhibit HTH DNA-binding activity. (*A*) Overview of EL222 structure, highlighting locations of the LOV (blue) and HTH (gold) domains and the Jα-helix (bronze) connecting the two. The LOV domain binds to the HTH domain using the LOV β-sheet surface, consistent with other LOV-effector complexes. (*B*) Expansion of the LOV/HTH interface as observed in chain A of the EL222 structure, as indicated by the boxed region in *A*. To bind the LOV domain, the HTH domain presents the 1α-2α linker and 4α-helix, the latter of which typically provides a dimerization interface for DNA-bound HTH domains. Thus sequestered, the 4α-helix is unable to participate in HTH/HTH interactions observed in many DNA-bound HTH complexes.

α/β LOV domain and a C-terminal all-helical HTH domain. A single FMN chromophore was observed within the LOV domain, orienting the critical isoalloxazine C4a atom only 3.9 Å from the cysteine (Cys) 75 Sγ atom that is expected to form the photochemical adduct. The LOV domain is followed by a C-terminal Jα-helix as observed in other LOV structures (15, 24), but here serves as an interdomain linker that associates more closely with the HTH effector domain rather than docking onto the LOV β-sheet surface as in AsLOV2 (15). This arrangement allows the EL222 LOV β-sheet surface to directly interact with the 4α-helix and 1α-2α loop of the HTH domain. This β-sheet interface is analogous to that used by other LOV and PAS domains to bind their effectors (25) (Fig. S1), burying approximately 700 $Å^2$ of surface area between the EL222 LOV and HTH domains. Notably, we observed differences in the relative arrangement of the LOV and HTH domains in the two molecules of EL222 found in the asymmetric unit, due to the translation of the HTH domain by approximately 2.5 Å parallel to the axis of helix 4α (Fig. S1*E*). Although this translation slightly alters the particular interactions between domains (Table S2), both molecules still fundamentally use a similar mix of hydrophobic and polar contacts at the LOV/HTH interface (Fig. 1*B*). The plasticity of this interface is consistent with a signaling role, poised for the facile conversion between conformations via allosteric change within the LOV domain (26). As in the structures of NarL and DosR (27, 28), the regulatory LOV domain of EL222 contacts the HTH dimerization helix (4α), but it does not also directly contact with the DNA-binding recognition helix (3α) as observed with the regulatory domains of these other structures. Such structural comparisons supported our hypothesis that EL222 fails to bind DNA in the dark by both sequestering the likely dimerization interfaces (LOV β-sheet; HTH 4α-helix) and the LOV domain interfering with HTH-DNA interactions (Fig. S2).

Photoactivation of EL222 Leads to Adduct Formation and Domain-Scale Rearrangements. Turning from structure to function, we examined light-induced changes in the visible absorbance spectrum to establish that EL222 can undergo LOV photochemistry. As expected for a flavin-containing LOV domain, we observed significant absorbance around 450 nm with vibrational fine structure (29) (Fig. 2*A*). This absorbance diminished significantly after illuminating samples, with three isosbestic points at 330, 384, and 407 nm, consistent with formation of the Cys-FMN adduct. After ceasing illumination, we observed subsequent dark-state recovery of the characteristic absorbance profile with first order exponential kinetics ($\tau = 25.5$ s for 25 °C, pH 6.0).

Having established the photosensitivity of EL222, we probed the ability of adduct formation to generate large-scale conformational changes using limited proteolysis. Both dark and lit EL222 treated with chymotrypsin demonstrated an initial cleavage removing the N-terminal His_6-tag within the first 5 min (Fig. 2*B*), but exhibited different behavior with extended incubation times. Dark-state samples underwent little additional proteolysis, consistent with a well-folded, compact protein. In contrast, lit-state samples were more quickly and extensively proteolyzed, with little full-length protein remaining intact after 60 min. Notably, chymotrypsin treatment of lit-state samples generated stable fragments, one of which was consistent with an intact LOV domain (Fig. 2*B*, species C). Mass spectrometry established that this fragment was generated by cleavage within the interdomain Jα-helical linker at Met159, which packs against the HTH 4α-helix in the dark-state structure. These data, together with our observation of protease-resistant fragments in dark conditions, suggest that light-induced conformational changes increase the accessibility of the Jα-linker via reorientation of the ordered LOV and HTH domains. This is supported by limited differences between CD spectra recorded under dark and lit conditions (Fig. S3).

NMR Studies of EL222 Photoactivation Establish Long-Range Light-Induced Conformational Changes. To probe these light-induced changes at higher resolution, we used solution NMR spectroscopy. Using a combination of triple resonance and NOESY data, we assigned ^{15}N, ^{13}C, and ^{1}H chemical shifts of EL222 in the dark state (81% of the backbone, 40% of the side chain). TALOS analyses of these chemical shifts (30), combined with through-space ^{1}H-^{1}H NOE data, let us confirm that EL222 has very similar secondary and tertiary structures in the crystal and solution states. Notably, solution measurements confirmed that EL222 is monomeric under these conditions. Upon illumination, we observed chemical shift and peak intensity changes at many backbone and side chain positions as observed in $^{15}N/^{1}H$ heteronuclear single quantum coherence (HSQC) (Fig. 3*A*) and $^{13}C/^{1}H$ HSQC spectra (Fig. S4). Such changes reflect alterations in the local electronic environments around NMR-active nuclei. Critically, all spectra maintained comparable chemical shift dispersion in the dark and lit states, consistent with LOV photochemistry inducing a domain reorientation, but not unfolding as observed in AsLOV2 (15).

To identify which sites experienced significant changes, we compared $^{15}N/^{1}H$ HSQC spectra recorded under dark and lit conditions (Fig. 3*a*), using $^{15}N/^{1}H$ Scotch exchange spectroscopy to assign lit-state chemical shifts by correlating dark-state ^{15}N shifts with lit-state ^{1}H shifts (31). From the 109 pairs of dark- and lit-state chemical shifts unambiguously assigned with this analysis, we established that chemical shift changes occur throughout the length of the protein (Fig. 3*B*). Although clusters of perturbed residues in the LOV domain likely report on adduct-induced configurational changes in the FMN chromophore and resulting conformational changes in the surrounding protein, we also clearly observed long-range (>15 Å from the flavin C4a atom) effects at sites outside the LOV domain as well. These include changes in

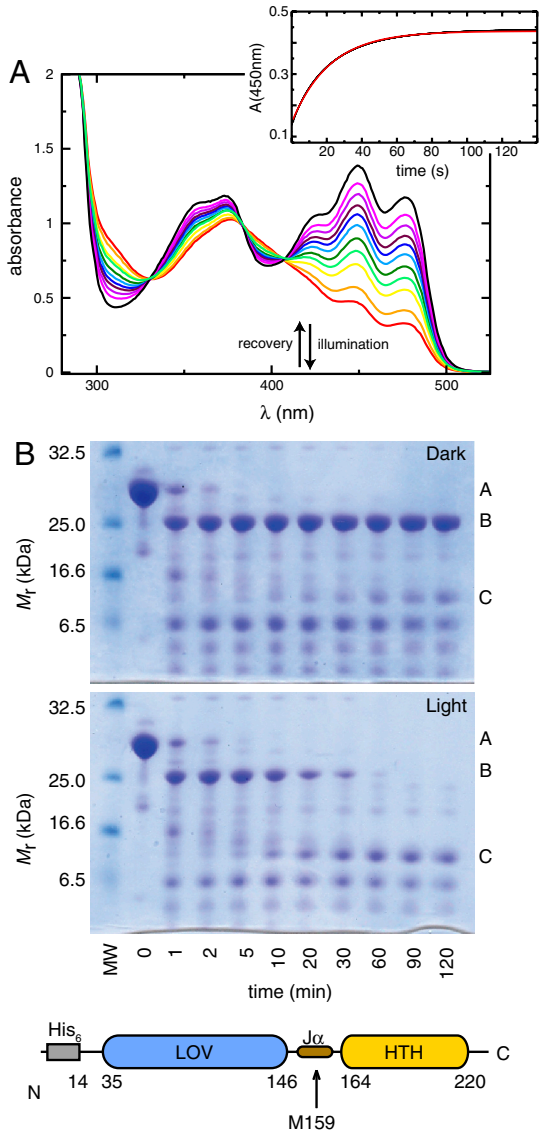

Fig. 2. Photochemical formation of Cys-FMN adduct in EL222 is correlated with domain-level reorganization. (*A*) UV-visible absorbance spectra of EL222, showing the expected absorbance near 450 nm with fine structure from protein-FMN interactions for dark-state EL222 (black). Illumination induces covalent adduct formation with a loss of absorbance above 400 nm (red), which gradually returns in the dark with spontaneous adduct decay (orange through purple, spectra recorded approximately every 5 s). The rate of dark-state recovery was determined by fitting the absorbance at 450 nm following illumination. (*Inset*) Data shown in black and fit to first-order exponential in red. (*B*) SDS-PAGE analysis of chymotrypsin limited proteolysis experiments shows kinetics of degradation are affected by illumination, with the Jα-helical linker becoming more accessible upon illumination. Significantly populated species include His_6-EL222 (14–222) (*A*), EL222 (14–222) (*B*), and EL222 (14–156, corresponding to an isolated LOV domain) (*C*).

the N-terminal A'α helix that precedes the LOV domain, plus multiple residues in the HTH domain that are significantly shifted ($\Delta\delta > 0.05$ ppm; Fig. 3*C*). These include several residues in the 4α-helix, including Leu213, Arg215, Ile216, and Glu219, plus sites in the 1α-2α (Leu178) and 3α-4α (Lys208, Thr209) loops. All of these residues are proximal to the LOV β-sheet, supporting our limited proteolysis findings that localized light-driven cysteinyl adduct formation triggers structural alterations beyond the LOV domain itself and fully throughout EL222. In addition, the significant perturbation of residues at the interface between the domains further supports an interdomain reorientation upon blue light illumination.

To complement this view from chemical shift changes, we used NMR-based measurements of backbone amide deuterium exchange rates to establish light-induced changes in domain structure and stability. We obtained these data by resuspending uniformly ^{15}N-labeled samples in D_2O-containing buffer, monitoring exchange by loss of intensity in consecutively recorded $^{15}N/^{1}H$ HSQC spectra. As we have not assigned the lit-state chemical shifts, ^{2}H exchange measurements under illumination relied on duty-cycling the sample between the dark and lit states, using assigned dark-state spectra to measure rates. Converting these exchange rate data into protection factors (32), we found that numerous sites across the protein exchanged very slowly in the dark state, consistent with stable hydrogen bonding as expected from regular secondary structure (Fig. 3*D*). Many amides within the LOV domain β-sheet surface are very well protected as expected for PAS domains (15, 33) and specific residues within the first and fourth helices of the HTH (1α and 4α) appear refractory to exchange. Upon illumination, these highly protected regions showed an overall decrease in protection factor, suggestive of distortion in the LOV structure as previously observed in AsLOV2 (15). The fact that these sites remained protected from exchange overall is consistent with both the LOV and HTH domains remaining stably folded, and with light inducing a separation or relative reorientation of the LOV and HTH domains as suggested by limited proteolysis and chemical shift analyses.

Photoactivation of EL222 Promotes DNA-Binding Activity. These light-induced structural changes imply a corresponding functional change, which we presumed to be a light-activated DNA-binding activity, given our data above and the domain architecture of EL222. Without a preestablished biological role of this protein, we started without any validated DNA-binding site. To address this issue, we used a candidate-based approach, assuming that EL222 might be autoregulatory and bind to a DNA sequence upstream of its own coding sequence. Scanning through the 350-bp region located 5′ to the start of EL222 translation with a series of 21 overlapping 45-bp candidate sequences tested, none bound EL222 as assessed by gel shift assays conducted under dark conditions. However, all of the candidate sequences bound EL222 under illumination at or above 70-μM protein (Fig. S5*A*), suggesting light-dependent activation of nonspecific DNA binding. Titrating to lower protein concentrations, we found two sequences that bound EL222 at concentrations as low as 7 μM (Fig. 4 for results of one of these sequences, oligomer 1). In both instances, DNA binding only occurred when the protein:DNA mix was incubated under white light. Binding was cooperative with respect to protein concentration, with a Hill coefficient of approximately four, suggesting that a pair of dimers bound within this 45-bp section. No binding occurred under dark-state conditions, even at protein concentrations capable of nonspecific DNA binding in the light. Protein previously exposed to bright light, then allowed to recover to dark state overnight at 4 °C also demonstrated the same minimal residual DNA-binding activity as protein that was not exposed to light, indicating the activity is reversible and light dependent (Fig. S5*B*). From these data, we can conclude that EL222 demonstrates light-dependent DNA-binding activity. Although the DNA sequence used in these gel shift experiments bound with the highest affinity of all sequences tested, we suspect that this is not an optimal binding sequence for EL222 based on the affinities of similar HTH-containing proteins for their cognate DNA sequences (34, 35). Nevertheless, these data suggest that this DNA sequence retains its utility for assaying protein activity in future structural and/or functional experiments.

Taken together, our data demonstrate that conformational changes propagate through the LOV domain upon illumination, disrupting inhibitory LOV-HTH interactions mediated by the LOV β-sheet. To test this, we mutated several sites to constitu-

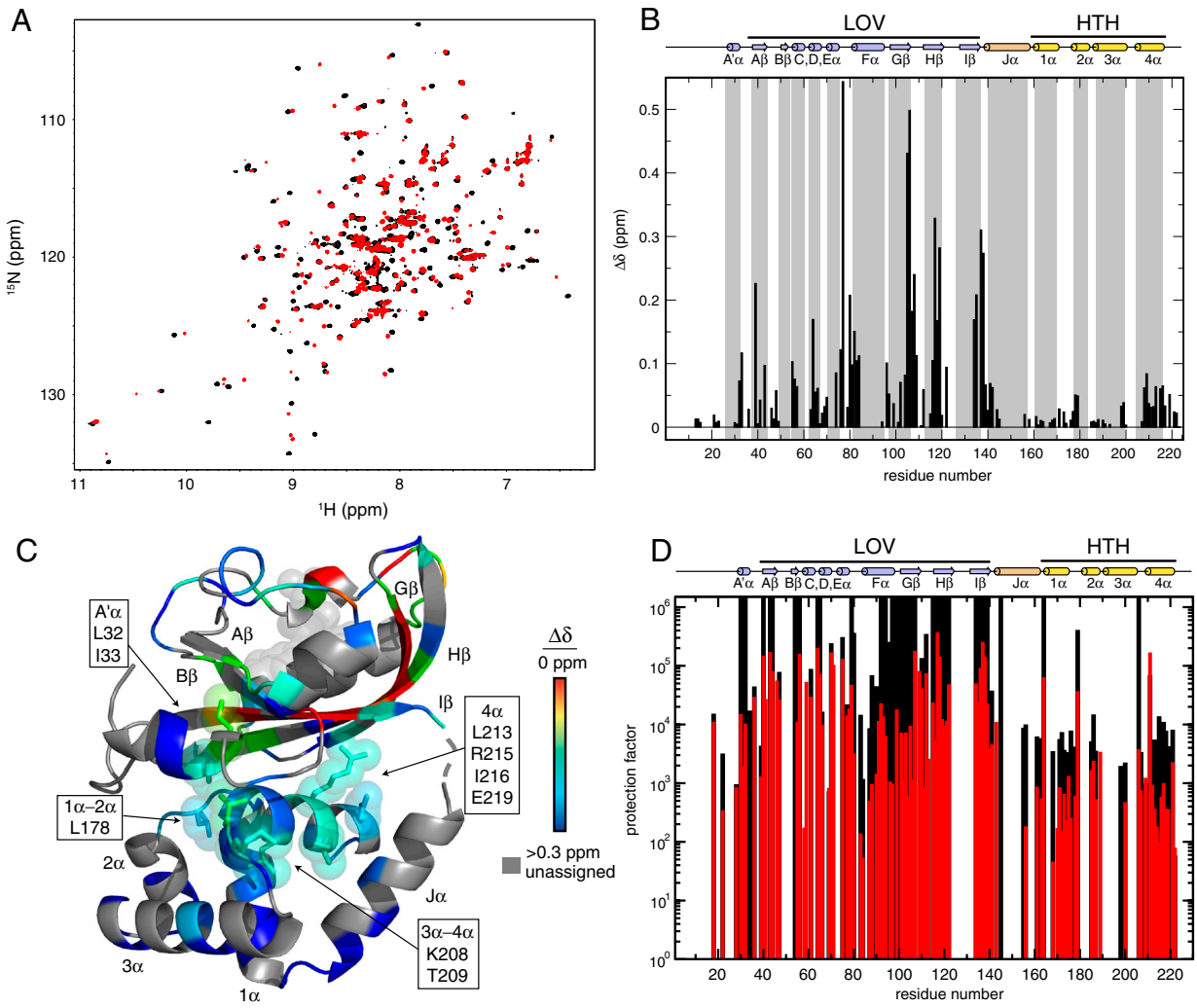

Fig. 3. Solution NMR data suggests EL222 undergoes light-induced rearrangement of two ordered domains. (*A*) Superposition of ^{15}N/^{1}H HSQC spectra of EL222 acquired under dark (black) or lit (red) conditions show light-induced changes in peak location and intensity. (*B*) Chemical shift difference analysis of ^{15}N/^{1}H HSQC spectra shown in Fig. 3*A* indicate significant changes occurring in both domains, including the HTH 1α-2α loop, 3α-4α loop, and 4α-helix located at the interface with the LOV domain. Secondary structure elements as indicated by the NMR data and X-ray structure are indicated. (*C*) Mapping values from Fig. 3*B* onto the dark-state crystal structure illustrates the pattern of chemical shift perturbations at the interdomain interface. Side chains are indicated for 1α-2α loop, 3α-4α loop, and 4α-helix residues in the HTH domain with ^{15}N/^{1}H chemical shift changes upon illumination. (*D*) ^{2}H exchange protection factor analyses (32) of EL222 conducted in the dark (black) and lit (red) states show similar protection, but to a lower overall degree upon illumination, consistent with reorganization of two ordered domains. Protection factors >10^6 are lower bound estimates because these sites did not sufficiently exchange for robust fitting of the time-dependent peak intensity changes.

tively break the LOV/HTH dark-state interaction and generate proteins locked in the DNA-binding conformation. One of these mutations, L120K, targeted a hydrophobic patch between the β-sheet surface of the LOV domain and the HTH 4α-helix (Fig. S6*A*). Gel filtration chromatography established that this mutant is a monomer in solution, as is wild-type EL222 (Fig. S6*B*). Gel shift assays conducted under dark-state conditions demonstrated that EL222 L120K bound DNA with similar affinity to wild type under lit-state conditions (Fig. S6*C*). Limited proteolysis of L120K using chymotrypsin showed little difference between the protein in the dark or lit state (Fig. S6*D*), with both resembling the lit state of wild-type protein. These results suggest that the L120K mutation forces EL222 into a lit-state-like structure that constitutively binds DNA.

Discussion

Within the context of regulation of HTH-containing proteins, our data are consistent with the EL222 LOV domain inhibiting DNA binding in the dark state via interactions with the HTH 4α-helix and several interhelical loops (Figs. 1, 3, and 5). Disruption of these interdomain contacts by light-induced conformational changes in the LOV domain (or mutagenesis of residues at the LOV/HTH interface) induces DNA-binding activity. A similar regulatory model is used by other two-domain response regulator proteins, including the *Escherichia coli* nitrite/nitrate response protein NarL. In this case, transfer of a phosphate group to the N-terminal receiver domain disrupts inhibitory contacts of this domain with the C-terminal LuxR-type HTH domain, allowing dimerization and DNA binding (27, 36, 37). Studies of response regulator proteins from NarL and other LuxR family members indicate that their regulatory domains also contact the HTH 1α-2α loop and 4α-helix (28, 36, 37), similar to EL222. Although this aspect of regulation shows strong parallels between NarL and EL222, we note that they are activated quite differently. In contrast with the intramolecular mechanism that we describe for EL222, NarL activation is entirely dependent on a separate sensor protein (NarQ or NarX) that detects an environmental signal (nitrate or nitrite) (38, 39) and initiates an intermolecular phosphotransfer to NarL. Finally, although the combination of N-terminal sensory and C-terminal HTH DNA-

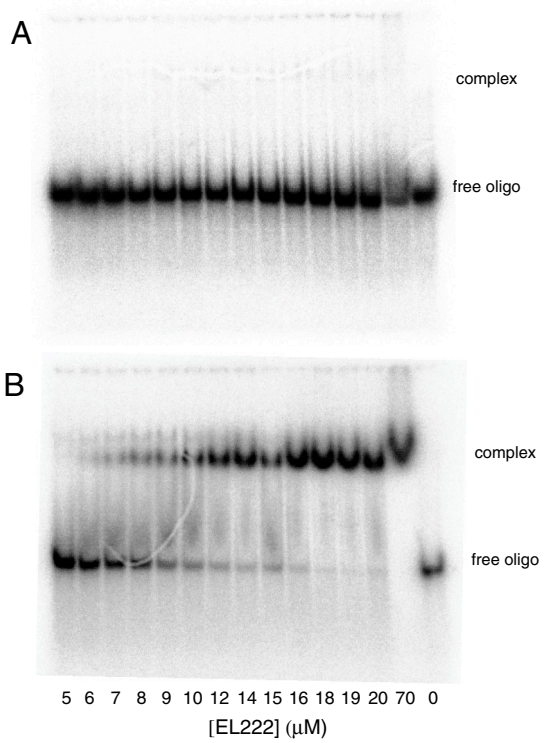

Fig. 4. EL222 is a light-activated DNA-binding protein. (*A*) EL222 demonstrates no observable DNA binding to the 45-bp dsDNA oligomer 1 under dark-state conditions at protein concentrations up to 70 μM. (*B*) Following illumination, cooperative DNA binding is observed to the same 45-bp dsDNA oligomer used for dark-state conditions.

binding domains may suggest that EL222 resembles response regulators that directly detect diffusible small ligands in the cell (35, 40), we note that some of these proteins may likely be controlled through ligand-induced protein folding (35) rather than covalent bond formation as seen in NarL and EL222.

Our results also further validate a conserved aspect of LOV domain and, more generally, PAS domain signaling via the β-sheet surface. Many PAS and LOV domains use this surface for hetero- or homodimerization, whereas others bind different N- and C-terminal segments that are essential to signaling (15, 18, 25). Some of these interactions can be modulated by cofactors within the PAS/LOV domain, providing a ligand-regulated

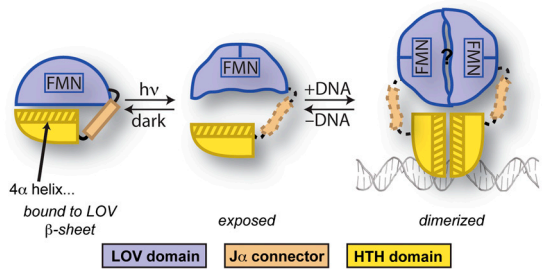

Fig. 5. Model for EL222 activation by blue light. In the dark, EL222 is incapable of binding DNA as the LOV domain sequesters the HTH 4α-helix and has steric conflicts with DNA if it could bind in monomeric form. The photochemical formation of a cysteinyl/flavin adduct in the LOV domain generates conformational changes that release inhibitory LOV/HTH interactions and expose the 4α helix, likely with concomitant changes in the interdomain LOV/HTH linker. The freed 4α-helix is then free to participate in HTH homodimerization upon binding DNA substrates, as observed in other HTH/DNA complexes, potentially also involving LOV/LOV interactions between EL222 monomers.

environmental switch. EL222 extends this paradigm by demonstrating that fully folded effector domains can bind to this surface, harnessing conformational changes within the LOV domain to rearrange the LOV-effector complex (without unfolding the effectors, as seen with the isolated Jα-helix in AsLOV2; ref. 15). Notably, these effector domains have different structures but appear to work through a common mechanism involving the β-sheet, potentially explaining how a single type of sensory domain can regulate a diverse group of effectors (7). Such information is particularly useful for both understanding naturally occurring LOV-regulated proteins and engineering light-regulated systems. These currently include LOV fusions to small GTPases, metabolic enzymes, DNA-binding domains, and other enzymes (8–10). All of these designed proteins have taken advantage of the well-characterized signaling mechanism of AsLOV2, including the PA-Rac1 light-activated GTPase (8). This fusion protein tethers the photosensory LOV domain closely to the effector GTPase when the Jα-helix is bound by the LOV domain, inhibiting enzymatic activity. With the knowledge of the broader principles provided here by EL222, such engineering may well be extended to an even larger range of target effectors as part of the rapidly growing toolbox of "optogenetic" tools (41) that offer precise spatial and temporal control of protein activity in vitro and in vivo.

Methods

Protein Expression and Solution Characterization. EL222 protein samples were obtained using standard *E. coli* heterologous expression and affinity purification methods as detailed in *SI Methods*. Thin layer chromatography established that EL222 bound FMN, not FAD or riboflavin. Additional solution characterization included UV-visible absorbance spectroscopy (60 μM sample; Varian Cary 50 spectrophotometer), CD spectroscopy (15 μM sample; AVIV 62DS), and limited proteolysis (1:43 wt:wt ratio of chymotrypsin:EL222). Photoexcited adduct-containing states were generated using a photographic flash (UV-vis absorbance, CD) or filtered mercury lamp (limited proteolysis).

Crystallographic Structure Determination. Crystals of EL222 were grown using the hanging drop method, using equal volumes of 8 mg/mL EL222 (1–222) and a reservoir of 20% (wt/vol) PEG 8K, 0.1 M MOPS (pH 7.5), 0.1 M ammonium acetate. X-ray diffraction data were collected from a single crystal on beam line 7-1 at Stanford Synchrotron Radiation Laboratory. The structure was solved by four-step molecular replacement using PHASER (42), with independent search models for the LOV and HTH domains (without a Jα-helix for the LOV domain). The structure of the Jα-interdomain helix was built manually as supported by difference density. The initial model of EL222 was subjected to iterative cycles of model building with COOT (43) and subsequent refinement with REFMAC5 (44) and PHENIX (45). Final R and R_{free} values were 26.3% and 32.9%, respectively, with further statistics of the refinement available in Table S1.

Solution NMR Studies. Solution NMR data were collected at University of Texas Southwestern using Varian 600 and 800 MHz spectrometers equipped with cryogenically cooled probes and laser illumination as previously described (15), with samples between 250–650 μM. NMR data were processed using NMRPipe (46) and analyzed with NMRView (47). Backbone and limited side-chain chemical shift assignments of dark-state EL222 were obtained using ^{1}H-CH$_3$ (V/I/L), U-^{2}H, ^{13}C, ^{15}N-labeled protein and a combination of ^{2}H-modified triple resonance and NOESY experiments. Lit-state chemical shift differences were determined using ^{15}N/^{1}H Scotch data to correlate dark- and lit-state chemical shifts (31), whereas lit-state ^{2}H exchange rates were determined using interleaved dark/lit acquisition (15).

DNA-Binding Studies. DNA-binding activity was assessed using gel shift assays using ^{32}P-labeled dsDNA 45-bp oligonucleotide fragments of DNA located to the 5′ end of the EL222 gene, as detailed in *SI Methods*. Gel shift results presented in Fig. 4 used one of these fragments (oligomer 1, genomic position 983532–983577), using a photographic flash to generate the photoexcited adduct state.

ACKNOWLEDGMENTS. We thank Brian Zoltowski, Giomar Rivera-Cancel, and Laura Motta-Mena for assistance with data collection and analysis, and further thank all members of the Gardner laboratory for constructive comments provided on this manuscript. This work was supported by grants

from the National Institutes of Health (R01 GM081875 to K.H.G., R01 AI07000 to H.L., T32 GM008297 supporting A.I.N.), National Science Foundation (0843662 to R.A.B.), The Robert A. Welch Foundation (I-1424 to K.H.G.), and a University of California, Irvine Chancellor's Fellowship to H.L. The *E. litoralis* genome sequence data was provided by Stephen Giovannoni's laboratory (Oregon State University, Corvallis, OR) and The J. Craig Venter Institute with grant support from The Gordon and Betty Moore Foundation Microbial Genome Sequencing Project.

1. Taylor BL, Zhulin IB (1999) PAS domains: Internal sensors of oxygen, redox potential, and light. *Microbiol Mol Biol Rev* 63:479–506.
2. Huala E, et al. (1997) Arabidopsis NPH1: A protein kinase with a putative redox-sensing domain. *Science* 278:2120–2123.
3. Liscum E, Briggs WR (1995) Mutations in the NPH1 locus of Arabidopsis disrupt the perception of phototropic stimuli. *Plant Cell* 7:473–485.
4. Imaizumi T, Tran HG, Swartz TE, Briggs WR, Kay SA (2003) FKF1 is essential for photoperiodic-specific light signalling in Arabidopsis. *Nature* 426:302–306.
5. Avila-Perez M, Hellingwerf KJ, Kort R (2006) Blue light activates the sigmaB-dependent stress response of Bacillus subtilis via YtvA. *J Bacteriol* 188:6411–6414.
6. Suzuki N, Takaya N, Hoshino T, Nakamura A (2007) Enhancement of a sigma(B)-dependent stress response in Bacillus subtilis by light via YtvA photoreceptor. *J Gen Appl Microbiol* 53:81–88.
7. Crosson S, Rajagopal S, Moffat K (2003) The LOV domain family: Photoresponsive signaling modules coupled to diverse output domains. *Biochemistry* 42:2–10.
8. Wu YI, et al. (2009) A genetically encoded photoactivatable Rac controls the motility of living cells. *Nature* 461:104–108.
9. Strickland D, Moffat K, Sosnick TR (2008) Light-activated DNA binding in a designed allosteric protein. *Proc Natl Acad Sci USA* 105:10709–10714.
10. Lee J, et al. (2008) Surface sites for engineering allosteric control in proteins. *Science* 322:438–442.
11. Möglich A, Ayers RA, Moffat K (2009) Design and signaling mechanism of light-regulated histidine kinases. *J Mol Biol* 385:1433–1444.
12. Salomon M, Christie JM, Knieb E, Lempert U, Briggs WR (2000) Photochemical and mutational analysis of the FMN-binding domains of the plant blue light receptor, phototropin. *Biochemistry* 39:9401–9410.
13. Swartz TE, et al. (2001) The photocycle of a flavin-binding domain of the blue light photoreceptor phototropin. *J Biol Chem* 276:36493–36500.
14. Halavaty A, 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* 46:14001–14009.
15. Harper SM, Neil LC, Gardner KH (2003) Structural basis of a phototropin light switch. *Science* 301:1541–1544.
16. Losi A, Kottke T, Hegemann P (2004) Recording of blue light-induced energy and volume changes within the wild-type and mutated phot-LOV1 domain from Chlamydomonas reinhardtii. *Biophys J* 86:1051–1060.
17. Nakasako M, Matsuoka D, Zikihara K, Tokutomi S (2005) Quaternary structure of LOV-domain containing polypeptide of Arabidopsis FKF1 protein. *FEBS Lett* 579:1067–1071.
18. Zoltowski BD, et al. (2007) Conformational switching in the fungal light sensor Vivid. *Science* 316:1054–1057.
19. Harper SM, Christie JM, Gardner KH (2004) Disruption of the LOV-Jalpha helix interaction activates phototropin kinase activity. *Biochemistry* 43:16184–16192.
20. Malzahn E, Ciprianidis S, Kaldi K, Schafmeier T, Brunner M (2010) Photoadaptation in Neurospora by competitive interaction of activating and inhibitory LOV domains. *Cell* 142:762–772.
21. Takahashi F, et al. (2007) AUREOCHROME, a photoreceptor required for photomorphogenesis in stramenopiles. *Proc Natl Acad Sci USA* 104:19625–19630.
22. Henikoff S, Wallace JC, Brown JP (1990) Finding protein similarities with nucleotide sequence databases. *Methods Enzymol* 183:111–132.
23. Choi SH, Greenberg EP (1991) The C-terminal region of the Vibrio fischeri LuxR protein contains an inducer-independent lux gene activating domain. *Proc Natl Acad Sci USA* 88:11115–11119.
24. Möglich A, Moffat K (2007) Structural basis for light-dependent signaling in the dimeric LOV domain of the photosensor YtvA. *J Mol Biol* 373:112–126.
25. Möglich A, Ayers RA, Moffat K (2009) Structure and signaling mechanism of Per-ARNT-Sim domains. *Structure* 17:1282–1294.
26. Yao X, Rosen MK, Gardner KH (2008) Estimation of the available free energy in a LOV2-J alpha photoswitch. *Nat Chem Biol* 4:491–497.
27. Baikalov I, et al. (1996) Structure of the *Escherichia coli* response regulator NarL. *Biochemistry* 35:11053–11061.
28. Wisedchaisri G, Wu M, Sherman DR, Hol WG (2008) Crystal structures of the response regulator DosR from Mycobacterium tuberculosis suggest a helix rearrangement mechanism for phosphorylation activation. *J Mol Biol* 378:227–242.
29. Christie JM, Salomon M, Nozue K, Wada M, Briggs WR (1999) LOV (light, oxygen, or voltage) domains of the blue-light photoreceptor phototropin (nph1): Binding sites for the chromophore flavin mononucleotide. *Proc Natl Acad Sci USA* 96:8779–8783.
30. Cornilescu G, Delaglio F, Bax A (1999) Protein backbone angle restraints from searching a database for chemical shift and sequence homology. *J Biomol NMR* 13:289–302.
31. Rubinstenn G, et al. (1999) NMR experiments for the study of photointermediates: Application to the photoactive yellow protein. *J Magn Reson* 137:443–447.
32. Bai Y, Milne JS, Mayne L, Englander SW (1993) Primary structure effects on peptide group hydrogen exchange. *Proteins* 17:75–86.
33. Brudler R, et al. (2006) PAS domain allostery and light-induced conformational changes in photoactive yellow protein upon I2 intermediate formation, probed with enhanced hydrogen/deuterium exchange mass spectrometry. *J Mol Biol* 363:148–160.
34. Da Re S, et al. (1999) Phosphorylation-induced dimerization of the FixJ receiver domain. *Mol Microbiol* 34:504–511.
35. Zhu J, Winans SC (2001) The quorum-sensing transcriptional regulator TraR requires its cognate signaling ligand for protein folding, protease resistance, and dimerization. *Proc Natl Acad Sci USA* 98:1507–1512.
36. Zhang JH, Xiao G, Gunsalus RP, Hubbell WL (2003) Phosphorylation triggers domain separation in the DNA binding response regulator NarL. *Biochemistry* 42:2552–2559.
37. Eldridge AM, Kang HS, Johnson E, Gunsalus R, Dahlquist FW (2002) Effect of phosphorylation on the interdomain interaction of the response regulator, NarL. *Biochemistry* 41:15173–15180.
38. Walker MS, DeMoss JA (1993) Phosphorylation and dephosphorylation catalyzed in vitro by purified components of the nitrate sensing system, NarX and NarL. *J Biol Chem* 268:8391–8393.
39. Schroder I, Wolin CD, Cavicchioli R, Gunsalus RP (1994) Phosphorylation and dephosphorylation of the NarQ, NarX, and NarL proteins of the nitrate-dependent two-component regulatory system of *Escherichia coli*. *J Bacteriol* 176:4985–4992.
40. Urbanowski ML, Lostroh CP, Greenberg EP (2004) Reversible acyl-homoserine lactone binding to purified Vibrio fischeri LuxR protein. *J Bacteriol* 186:631–637.
41. Moglich A, Moffat K (2010) Engineered photoreceptors as novel optogenetic tools. *Photochem Photobiol Sci* 9:1286–1300.
42. McCoy AJ, et al. (2007) Phaser crystallographic software. *J Appl Crystallogr* 40:658–674.
43. Emsley P, Cowtan K (2004) Coot: Model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* 60:2126–2132.
44. Murshudov GN, Vagin AA, Dodson EJ (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr* 53:240–255.
45. Adams PD, et al. (2010) PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* 66:213–221.
46. Delaglio F, et al. (1995) NMRPipe: A multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR* 6:277–293.
47. Johnson BA, Blevins RA (1994) NMRView: A computer program for the visualization and analysis of NMR data. *J Biomol NMR* 4:603–614.

Supporting Information

Nash et al. 10.1073/pnas.1100262108

SI Methods

Cloning, Expression, and Purification of EL222 for Crystallographic Studies. DNA containing full-length EL222, a light-regulated DNA-binding protein, was cloned into pET/100D TOPO (Invitrogen) and expressed in *Escherichia coli* BL21(DE3) cells. Cells were grown in LB-AMP at 37 °C in the dark. Upon reaching A_{600} 0.6–0.8, expression was induced with 1 mM IPTG. After 2 h induction, cells were centrifuged at 4,400 × *g* for 15 min and subsequently lysed with a French Pressure Cell containing 20 mM Tris (pH 7.5), 1 mM PMSF, 300 mM NaCl, and one tablet of Complete™ EDTA-free (Roche) protease inhibitor. Lysates were centrifuged at 39,800 × *g* for 1 h before purifying the protein using FPLC at 4 °C in the dark. First, supernatant was loaded onto a CO^{2+}-Talon column with buffer A [20 mM (Tris pH) 7.5, 30 mM NaCl, 10% glycerol]. Elution was performed with buffer B [500 mM imidazole, 20 mM Tris (pH 7.5), 30 mM NaCl, 10% glycerol]. Next, protein was loaded onto a Hitrap Q FF anion exchange column equilibrated with buffer C [20 mM Tris (pH 7.5), 30 mM NaCl, 10% glycerol, 1 mM DTT] and eluted with a linear gradient mixing in buffer D [20 mM Tris (pH 7.5), 1 M NaCl, 10% glycerol, 1 mM DTT]. Protein was then dialyzed into 20 mM Tris (pH 7.5), 30 mM NaCl, 10% glycerol and concentrated to 16 mg/mL.

Crystallization. Crystallization of purified protein was first achieved with the Hauptman-Woodward Institute (HWI) robot (1). We are grateful to Tina Veatch from the HWI for manually checking that the crystals in the black and white images were actually colored yellow, consistent with flavin-bound protein. Subsequently, the protocol was adjusted for hanging drop vapor diffusion. Drops contained 2 μL of 20% (wt/vol) PEG 8 K, 0.1 M MOPS (pH 7.5), 0.1 M ammonium chloride from the reservoir, and 2 μL of 8 mg/mL protein in 20 mM Tris (pH 7.5), 300 mM NaCl. Drops also contained 0.1 M guanidine HCl and 1% glycerol as an additive. Trays were kept in the dark at 16 °C. Crystals took up to 2 mo to grow. Crystals were harvested under red light and transferred stepwise to a reservoir solution containing 5%, 10%, and 15% glycerol as a cryoprotectant prior to flash-cooling with liquid nitrogen.

X-Ray Data Collection. X-ray diffraction data were collected remotely at beam line 7-1 at the Stanford Synchrotron Radiation Laboratory using a Q315 CCD area detector from Area Detector Systems Corporation. Two separate datasets, 180° and 360°, were collected on a single crystal with an oscillation width of 1°. The exposure time was 20 s per image. The data from both runs were indexed and integrated using the program d*trek (2) and subsequently scaled and merged to 2.1 Å. For data reduction and refinement statistics, see Table S1.

Structure Solution and Refinement. Search models were generated with the program SWISS-MODEL (3) based on Protein Data Bank (PDB) entries 2PR5 and 1YIO for the light-oxygen-voltage (LOV) domain and helix-turn-helix (HTH) domain, respectively. The amino acid sequence of each model was substituted with that of EL222. A four-step molecular replacement search was performed with the program PHASER (4). Space groups *P*2 and $P2_1$ were considered for these sequential six-dimensional searches, the latter being the correct choice. First, two copies of the LOV domain were placed, followed by a search for two copies of the LuxR domain, resulting in an initial *R* factor of 49.1%. To improve the quality of the phases, solvent flattening was performed with the program DM (5). The resulting map was used to build a more accurate model from scratch using the program BUCCANEER (6). Subsequently, density not modeled by Buccaneer was built manually using the program COOT (7). Positional and isotropic B-factor refinement was performed iteratively using the program REFMAC5 (8). A final round of refinement was done with PHENIX (9). The final *R* factor and R_{free} were 26.3% and 32.9%, respectively. Model geometry was analyzed using the program PROCHECK (10). Coordinates have been deposited with the PDB/Research Collaboratory for Structural Bioinformatics (accession code 3P7N).

Differences Between the Two Molecules in the Asymmetric Unit. The main-chain heavy atom rmsd, computed with the program SSM (11), between the two independent copies in the asymmetric unit (molecules A and B) is 1.23 Å, significantly larger than the values between their individual LOV domains (0.52 Å) and their individual HTH domains (0.81 Å). This indicates a rigid-body motion of LOV vs. HTH domains between the two copies of the protein in the asymmetric unit (Fig. S1*E*), specifically a relative translation of one HTH domain by about half a helix pitch (2.5 Å) parallel to the axis of helix 4α. Although this results in some differences in interdomain interactions (Table S2), both molecules still involve the same residues in making a mix of hydrophobic and polar contacts at the LOV/HTH interface (Fig. 1*B*). Using PISA (12) to analyze the interfaces between the LOV and HTH domains (disregarding the two Jα-helices bridging pairs of domains), we found that the LOV/HTH interface buries 714 and 612 $Å^2$ of surface area for molecules A and B, respectively (Table S2). Consistent with the smaller surface area, molecule B also has a smaller number of hydrogen bonds at the LOV/HTH interface (eight for molecule A, four for B) and is predicted to have a slightly smaller solvation free energy gain for complexing the LOV and HTH domains (−6.8 kcal/mol for A, −6.6 kcal/mol for B) (Table S2) (12). Finally, we also noted differences between the main-chain temperature factors of the two molecules (37 $Å^2$ for A, 46 $Å^2$ for B), albeit with the same pattern of lower and higher values along the length of the protein (Fig. S7).

Cloning, Expression, and Purification of EL222 for NMR and Functional Studies. DNA encoding full-length EL222 and a 13-residue N-terminal truncation (residues 1–222 and 14–222, respectively) were subcloned into the expression vector pHis-Gβ1-Parallel1, a derivative of the pHis-Parallel1 vector (13, 14). Mutagenesis was carried out using QuikChange II XL from Stratagene according to manufacturer's instructions. *E. coli* were transformed and grown in either LB broth or M9 minimal media as described for individual experiments below. Cultures were grown at 37 °C to an A_{600} of 0.6–0.9 and then induced at 20 °C in the dark by addition of 0.5 mM IPTG. After 16 h induction, cells were centrifuged and the resulting pellets resuspended in 50 mM Tris, 100 mM NaCl, pH 8.0 buffer at 4 °C and lysed by sonication. Lysates were clarified by centrifugation at 48,000 × *g* for 30 min. The resulting supernatant was loaded on a Ni^{2+}-nitrilotriacetate (NTA) column, allowing for rapid affinity purification of His-Gβ1 tagged proteins by gradient elution, mixing in the same buffer containing additional 500 mM imidazole. After exchanging the protein-containing fractions into 50 mM Tris, pH 8.0 buffer, the His-Gβ1 tag was cleaved using 1 mg His_6-Tobacco Etch Virus (TEV) protease (15) per 30 mg of fusion protein. Proteolysis reactions were incubated overnight at 4 °C and stopped using a Ni^{2+}-NTA column to remove the His_6-Gβ1 tag and His_6-TEV

protease. Postcleavage, the resulting protein contains only three vector-derived residues, GEF at the N-terminus. TEV-cleaved protein was loaded onto a MonoQ column to separate small remaining impurities. Highly pure protein was eluted with a linear gradient mixing in 50 mM Tris, 1 M NaCl pH 8.0 buffer, exchanged into 50 mM sodium phosphate, 100 mM NaCl pH 6.0 buffer, and concentrated to a final protein concentration of 100–250 μM.

Thin Layer Chromatography Analysis of Flavin Content in EL222. Given that LOV domains are commonly associated with either FMN or FAD as their flavin cofactor, we experimentally determined that EL222 preferentially associates with FMN using thin layer chromatography (TLC) analysis. Samples of full-length EL222 (1–222) were exchanged into H_2O and concentrated to approximately 250 μM. A 2× volume of ethanol was added to the sample and incubated for 2 min in a boiling water bath. The denatured protein sample was blotted onto a silica gel-based TLC plate, with standard samples of FMN, FAD, and riboflavin (250 μM each) in separate lanes. Separation proceeded in a 12:3:5 *n*-butanol: acetic acid: H_2O solvent system, and the resulting bands were visualized with a UV lamp. This analysis showed that recombinantly expressed EL222 only associates with FMN, with no significant incorporation of FAD or riboflavin.

UV-Visible Absorbance Spectroscopy and Photocycle Kinetics. UV-visible absorbance spectra were measured on a Varian Cary Series 50 spectrophotometer from 250–550 nm. Dark-state spectra were obtained on 60-μM samples exposed only to red light for the preceding 24 h, whereas lit-state spectra were obtained immediately after exposing the sample to illumination from a photographic flash. Kinetic experiments monitored the return of the A_{450} signal following illumination. Data points were fitted using a first-order exponential to obtain the reported time constant (τ).

Limited Proteolysis. Full-length (1–222) EL222 samples which retained an N-terminal His_6 tag were used for limited proteolysis, providing an internal control for proteolytic activity between the His_6 tag and the LOV domain. A 1:43 ratio (wt/wt) of chymotrypsin to protein was used in a single volume, with subsequent samples collected from a common stock. Samples collected for each time point were stopped by the addition of SDS loading buffer containing 25% glycerol and visualized on 20% SDS-PAGE gel. For mass spectrometry analysis, reactions were stopped with 1% trifluoroacetic acid. Dark-state experiments were conducted under dim red light, whereas lit-state experiments were performed under constant irradiation produced by an Oriel mercury lamp (model number 66902) at 50 mW power with a broadband blue-green filter (Oriel no. 51970).

Circular Dichroism Spectroscopy. A total of 500 μL of 15 μM full-length EL222 was used for each CD experiment, using a 1-mm path length cell. Dark-state spectra were collected under complete darkness, whereas lit-state spectra were recorded following exposure to photographic flash. CD data were collected using a wavelength range from 195 to 260 nm at 12 °C with 1.5-nm bandwidth and 3 s averaging time. For lit-state spectra, data collection was paused every 9 s and the sample reexposed to a photographic flash to ensure lit-state protein. Final data were generated from an average of three repeats.

Solution NMR Spectroscopy. All of the following NMR experiments utilized uniformly ^{15}N, ^{13}C-labeled EL222(14–222) protein, generated from bacteria grown in M9 minimal media containing 1 g/L of $^{15}NH_4Cl$ and 3 g/L of $^{13}C_6$ glucose. Solution NMR experiments were performed on Varian Inova 600 and 800 MHz spectrometers equipped with cryogenically cooled probes at 25 °C, using NMRPipe (16) for data processing and NMRview (17) for analysis. Lit-state $^{15}N/^{1}H$ and $^{13}C/^{1}H$ heteronuclear single quantum coherence (HSQC) spectra were acquired with a 488-nm, 200-mW Coherent Sapphire laser. The output from this laser was focused into a 10-m long, 0.6-mm diameter quartz fiber optic. The other end of the fiber was placed into the bottom of a coaxial insert tube (Wilmad) inside a 5-mm NMR sample tube, allowing the illuminated tip to be immersed in protein solution without contamination. Power level measurements were conducted prior to every experiment to establish the efficiency of coupling the laser output to the fiber optic, and all power levels reported here are those measured at the end of the fiber. Lit-state $^{15}N/^{1}H$ and $^{13}C/^{1}H$ HSQC spectra were recorded by preceding each transient in the experiment with a 50-mW, 120-ms laser pulse during the 1.06-s delay between transients (14).

Backbone chemical shift assignments were acquired with ^{2}H-modified triple resonance data on a uniformly ^{15}N, ^{13}C, ^{2}H-labeled sample. This sample was prepared with M9 media for protein growth composed of D_2O, 1 g/L $^{15}NH_4Cl$, and 3 g/L $^{13}C_6$ glucose as the sole carbon source for U-$^{2}H/^{15}N/^{13}C$ protein (14–222). The purified protein was buffer exchanged into 50 mM MES, 100 mM NaCl pH 6.0 buffer and concentrated to 250 μM. Spectra were taken at 25 °C on a Varian Inova 600 MHz spectrometer fitted with a triple resonance cryoprobe. Assignment of the backbone and Cβ NMR resonances was achieved with the following 3D NMR experiments modified for use on highly deuterated samples: HNCA (18), HN(CO)CA (19), HNCACB (19), HN(CO)CACB (19). Side-chain ^{13}C and ^{1}H resonances were assigned with $^{15}N/^{13}C$-edited NOESY spectra on protein grown in M9 minimal media for U-$^{15}N/^{13}C$ protein. Additional assignments were obtained using (H)C(CO)NH-total correlation spectroscopy (20) on perdeuterated protein with protonated methyl groups at the valine γ, leucine δ, and isoleucine δ_1 positions. This protein was produced by growing *E. coli* in D_2O based M9 minimal media containing both $^{15}NH_4Cl$ and $^{13}C_6$ glucose, supplemented 30 min prior to induction with 80 mg/L α-ketovalerate and 50 mg/L α-ketobutyrate (21).

To measure deuterium exchange rates on EL222 using solution NMR methods, a 550-μL sample of 600 μM ^{15}N-labeled protein (14–222) in 50 mM sodium phosphate (pH 6.0), 100 mM NaCl, pH 6.0 buffer was lyophilized. Immediately prior to acquiring spectra, the lyophilized sample was rehydrated with 550 μL D_2O and placed in the magnet. $^{15}N/^{1}H$ HSQC spectra were acquired as an automated series of 79–90 datasets, each taking approximately 20 min to complete. Lit-state spectra were obtained using a 50-mW laser pulse as described above. The lit-state spectra were recorded interleaved with dark-state spectra, letting us take advantage of the chemical shift assignments in the dark state. Protection factors were then calculated based on standard methods (22), with a correction for the lit-state rates given the interconversion between the dark and lit states.

DNA Binding as Assessed by Electrophoretic Mobility Shift Assay. We designated 45-bp lengths of DNA within 350 bp of the gene start site to use in a gel shift assay. Oligonucleotides (Integrated DNA Technologies) were staggered to cover all possible binding sites in this region. Lyophilized DNA was resuspended in 50 mM Tris, 100 mM NaCl, 2 mM $MgCl_2$, 12% glycerol, pH 8.0 buffer. Reverse complementary pairs were annealed by heating 40 μL of 100 pmol/μL DNA to 95 °C in a heat block for 5 min., followed by slow cooling to room temperature over 1 h. Following annealing, 100 ng of DNA were labeled with ^{32}P ATP using T4 polynucleotide kinase in a 20 μL reaction volume. Unincorporated ^{32}P was purified away using ProbeQuant G50 Microcolumns (GE Healthcare). Concentrations were estimated assuming 80% recovery from column. Unlabeled protein (both 1–222 and 14–222) was purified and concentrated to 250 μM. Reaction conditions included 0.01 mg/mL BSA, 0.02 mg/mL deoxyinosine/deoxycytosine, 0.04 ng/μL ^{32}P-labeled DNA (1.45 nM),

and varying concentrations of protein in 50 mM Tris, 100 mM NaCl, 2 mM $MgCl_2$, 12% glycerol, pH 8.0 buffer. Protein was the final addition to each reaction followed by brief centrifugation and incubation in ice slush for 30 min. For dark-state reactions, samples were prepared under dim red light and incubated in a covered ice bucket, whereas lit-state reactions were conducted under bright white light illumination in a clear glass beaker with periodic exposure to camera flash. After incubation, samples were separated on a 5–10% TAE gels at 4 °C for 2 h at 100 V. Samples allowed to recover after illumination were flashed in the absence of DNA, then stored in the dark at 4 °C for 24 h to allow complete recovery of the dark-state conformation. Again, this was conducted under either dim red light or bright white light. Gels were dried with a heated slab gel dryer fitted with vacuum pump for 1 h and developed using FujiFilm FLA-5100 imaging system following exposure to PhosphorImaging plates. Quantification of percent DNA bound was performed using FujiFilm MultiGauge v2.3 software (FUJIFILM Medical Systems USA, Inc.).

These surveys of EL222 binding to sequences within the 350 bp located 5′ of the start of EL222 translation indicated two sequences that exhibited higher affinity binding: oligomer 1 (genomic base pairs 983532–983577), GGTAGGATCCATCGGGCAGTGCGGTCAGCGGCATGCCGGCAGCAG; oligomer 2 (genomic base pairs 983647–983692), GGCCCCGAGGTCCAGCACCAACGCAGTCCCCTTTGGTACGCCGAC.

Experimental data presented in the main text and Fig. S5 utilized oligomer 1; Fig. S5 also utilized a lower affinity oligomer from elsewhere in the EL222 promoter (base pairs 983469–983513, ACAGCAATTGCAATGGTGCCGCGAGGGCTGTGAACTACCTGTTGC).

1. Luft JR, et al. (2003) A deliberate approach to screening for initial crystallization conditions of biological macromolecules. *J Struct Biol* 142:170–179.
2. Pflugrath JW (1999) The finer things in X-ray diffraction data collection. *Acta Crystallogr D Biol Crystallogr* 55:1718–1725.
3. Arnold K, Bordoli L, Kopp J, Schwede T (2006) The SWISS-MODEL workspace: A web-based environment for protein structure homology modelling. *Bioinformatics* 22:195–201.
4. McCoy AJ, et al. (2007) Phaser crystallographic software. *J Appl Crystallogr* 40:658–674.
5. Cowtan K (1994) dm: An automated procedure for phase improvement by density modification. *Joint CCP4 and ESF-EACBM Newsletter on Protein Crystallography* 31:34–38.
6. Cowtan K (2006) The Buccaneer software for automated model building. 1. Tracing protein chains. *Acta Crystallogr D Biol Crystallogr* 62:1002–1011.
7. Emsley P, Cowtan K (2004) Coot: Model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* 60:2126–2132.
8. Murshudov GN, Vagin AA, Dodson EJ (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr* 53:240–255.
9. Adams P, et al. (2010) PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* 66:213–221.
10. Laskowski RA, Macarther MW, Moss DS, Thornton JM (1993) Procheck—a program to check the stereochemical quality of protein structures. *J Appl Crystallog* 26:283–291.
11. Krissinel E, Henrick K (2004) Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallogr D Biol Crystallogr* 60:2256–2268.
12. Krissinel E, Henrick K (2007) Inference of macromolecular assemblies from crystalline state. *J Mol Biol* 372:774–797.
13. Sheffield P, Garrard S, Derewenda Z (1999) Overcoming expression and purification problems of RhoGDI using a family of "parallel" expression vectors. *Protein Expr Purif* 15:34–39.
14. Harper SM, Neil LC, Gardner KH (2003) Structural basis of a phototropin light switch. *Science* 301:1541–1544.
15. Blommel PG, Fox BG (2007) A combined approach to improving large-scale production of tobacco etch virus protease. *Protein Expr Purif* 55:53–68.
16. Delaglio F, et al. (1995) NMRPipe: A multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR* 6:227–339.
17. Johnson BA, Blevins, RA (1994) NMRView: A computer program for the visualization and analysis of NMR data. *J Biomol NMR* 4:595–740.
18. Yamazaki T, et al. (1994) An HNCA pulse scheme for the backbone assignment of ^{15}N, ^{13}C, ^{2}H-labeled proteins: Application to a 37 kDa trp repressor-DNA complex. *J Am Chem Soc* 116:6464–6465.
19. Yamazaki T, Lee W, Arrowsmith CH, Muhandiram DR, Kay LE (1994) A suite of triple resonance NMR experiments for the backbone assignment of ^{15}N, ^{13}C, ^{2}H labeled proteins with high sensitivity. *J Am Chem Soc* 116:11655–11666.
20. Gardner KH, Konrat R, Rosen MK, Kay LE (1996) An (H)C(CO)NH-TOCSY pulse scheme for sequential assignment of protonated methyl groups in otherwise deuterated ^{15}N, ^{13}C-labeled proteins. *J Biomol NMR* 8:351–356.
21. Goto NK, Gardner KH, Mueller GA, Willis RC, Kay LE (1999) A robust and cost-effective method for the production of Val, Leu, Ile (delta 1) methyl-protonated ^{15}N-, ^{13}C-, ^{2}H-labeled proteins. *J Biomol NMR* 13:369–374.
22. Bai Y, Milne JS, Mayne L, Englander SW (1993) Primary structure effects on peptide group hydrogen exchange. *Proteins* 17:75–86.

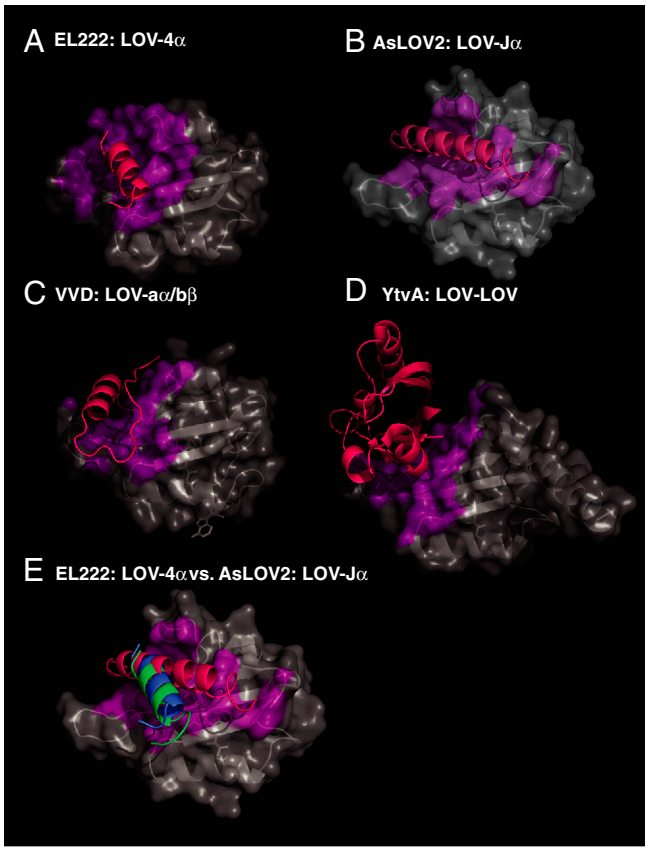

Fig. S1. LOV domains utilize a common β-sheet interface for interacting with a wide range of effectors. Structures of four LOV-containing proteins have been aligned to a common view, presenting the β-sheet surface toward the reader. Effector regions located outside the canonical LOV domain are shown in bright pink, with regions of the LOV domain surface located within 5 Å of those effectors colored in magenta. This alignment shows that all of these structures share the use of the LOV β-sheet surface to interact with their effectors, despite the diversity of surfaces they provide, including predominantly α-helical surfaces from (*A*) 4α-helix of EL222 HTH domain and (*B*) Jα-helix of *Avena sativa* phototropin 1 LOV2 (AsLOV2) (1, 2), (*C*) mixed α/β surface from Vivid (VVD) (3), and (*D*) predominantly β-sheet surface from the YtvA LOV–LOV homodimer (4). We emphasize that this same interface is used by both of the EL222 molecules observed in the crystallographic asymmetric unit, despite the translation of the HTH domains by approximately 2 Å along the direction of the C-terminal 4α-helix. This is demonstrated in *E* by a superposition of LOV domains from EL222 and AsLOV2 (gray) and indicating the 4α-helices from EL222 molecules A and B (green and blue, respectively) with the Jα-helix of AsLOV2 (magenta). This demonstrates that all of these helices associate with similar positions on the LOV domain surface, but with slightly different orientations (relative approximately 35° rotation between EL222's 4α and AsLOV2's Jα-helices). Consistent with this movement, the main-chain atom rmsd between all residues in the two EL222 molecules (1.23 Å) is higher than that between the isolated LOV (0.52 Å) and the HTH domains (0.81 Å). Interactions within 5 Å of the C-terminal helix of molecule A with its LOV domain are highlighted magenta.

1 Harper SM, Neil LC, Gardner KH (2003) Structural basis of a phototropin light switch. *Science* 301:1541–1544.
2 Halavaty A, 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* 46:14001–14009.
3 Zoltowski BD, et al. (2007) Conformational switching in the fungal light sensor Vivid. *Science* 316:1054–1057.
4 Möglich A, Moffat K (2007) Structural basis for light-dependent signaling in the dimeric LOV domain of the photosensor YtvA. *J Mol Biol* 373:112–126.

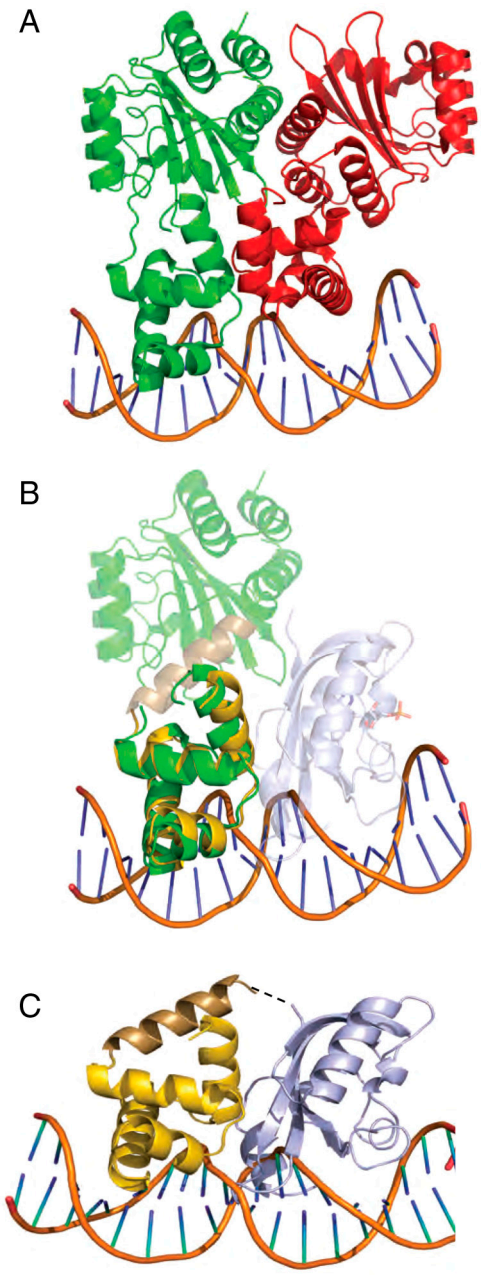

Fig. S2. LOV-HTH interactions observed in the EL222 crystal structure are predicted to interfere with HTH dimerization and DNA binding. (*A*) Crystal structure of TraR/DNA complex (1), with TraR dimer shown as green and red monomers. (*B*) Superposition of one TraR molecule (green) and EL222 (gold, blue) by least-squares fitting of the HTH domains of both proteins (shown with solid ribbons, whereas the regulatory domains are partially transparent). (*C*) Ribbon diagram of EL222 alone (in orientation shown in *B*) demonstrates that the LOV and HTH domains are oriented in the dark-state crystal structure of EL222 in a mode inconsistent with DNA binding without major domain rearrangement. In particular, the LOV domain clashes with DNA and occludes the HTH 4α-helical dimerization interface.

1 Vannini A, et al. (2002) The crystal structure of the quorum sensing protein TraR bound to its autoinducer and target DNA. *EMBO J* 21:4393–4401.

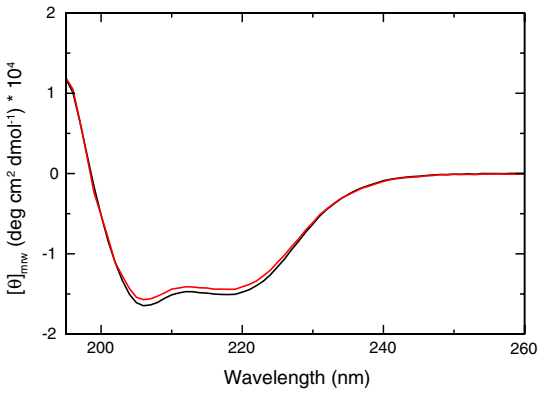

Fig. S3. Circular dichroism spectra of dark- and lit-state EL222 indicate minimal change in secondary structure content upon illumination. The double minima at 208 and 222 nm suggest well-folded protein with significant α-helical character in the dark-state (black), consistent with the crystal structure of this protein. Following illumination, little change was observed in the secondary structure characteristics of the protein (red).

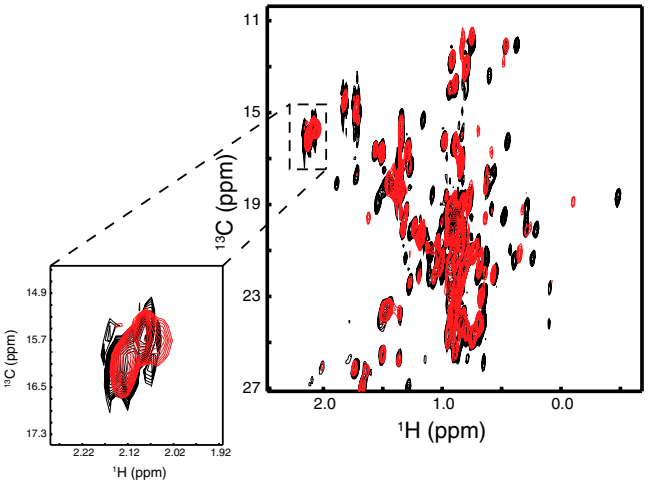

Fig. S4. Light-induced chemical shift changes in the methyl region of ^{13}C/^{1}H HSQC spectra of EL222. Overlaid spectra were acquired under dark (black) or lit (red) conditions. Analysis of the methionine Cε region located near 2 ppm ^{1}H, 15 ppm ^{13}C (and validated by changes in peak sign in constant time ^{13}C/^{1}H HSQC spectra) indicated a perturbation in one Met Cε methyl group upon illumination. Because all the methionine residues in EL222 are located outside the canonical LOV domain, this supports the long-range propagation of light-induced chemical shift and structural changes.

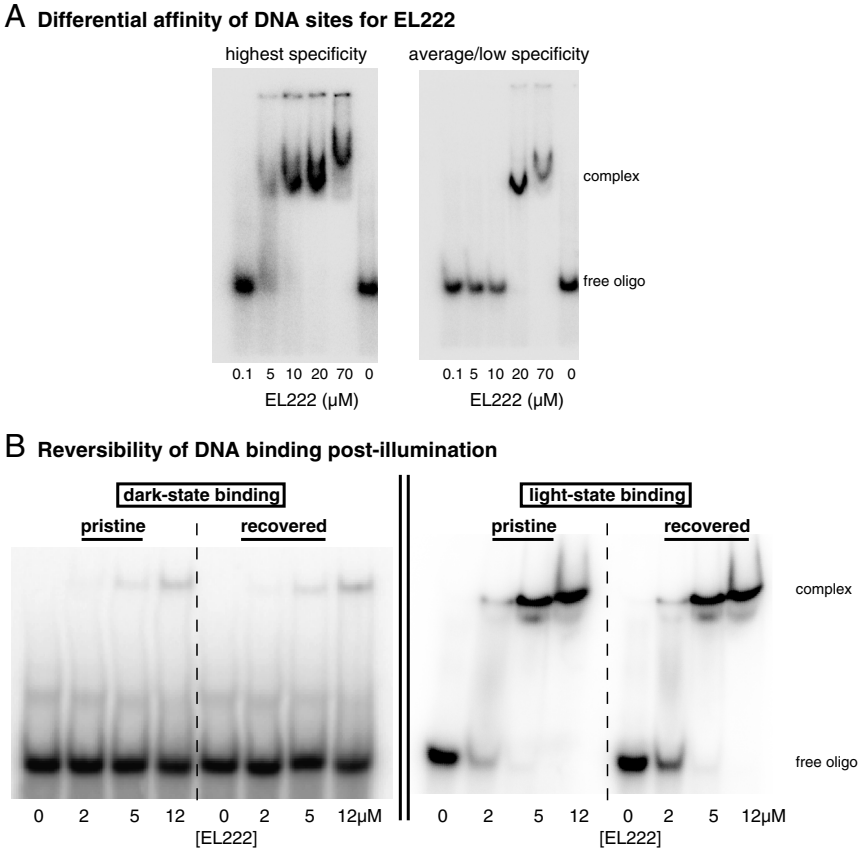

Fig. S5. EL222 binds with differential affinities to different DNA sequences in the EL222 promoter and does so in a light-dependent manner that is not permanently triggered by illumination. (*A*) Comparison of the EL222 concentration dependence of EMSA band shifts for a higher affinity site (oligomer 1, GGTAGGATCCATCGGGCAGTGCGGTCAGCGGCATGCCGGCAGCAG, genomic base pairs 983532–983577) with a lower affinity one located elsewhere in the EL222 promoter (ACAGCAATTGCAATGGTGCCGCGAGGGCTGTGAACTACCTGTTGC, genomic base pairs 983469–983513) reveals an approximate fivefold difference in binding affinity under these conditions. (*B*) DNA binding requires illumination and is not permanently induced by exposure to light. Comparison of EMSA assays of samples that have not been exposed to significant blue light ("pristine") and those that have been illuminated then allowed to recover in the dark for 24 h at 4 °C ("recovered") show that neither bind the high affinity oligomer 1 significantly in the dark (maximum DNA binding = 2–4% at 12 μM EL222). Both proteins bind DNA avidly and equivalently when illuminated, with approximately 100% DNA bound under the same condition.

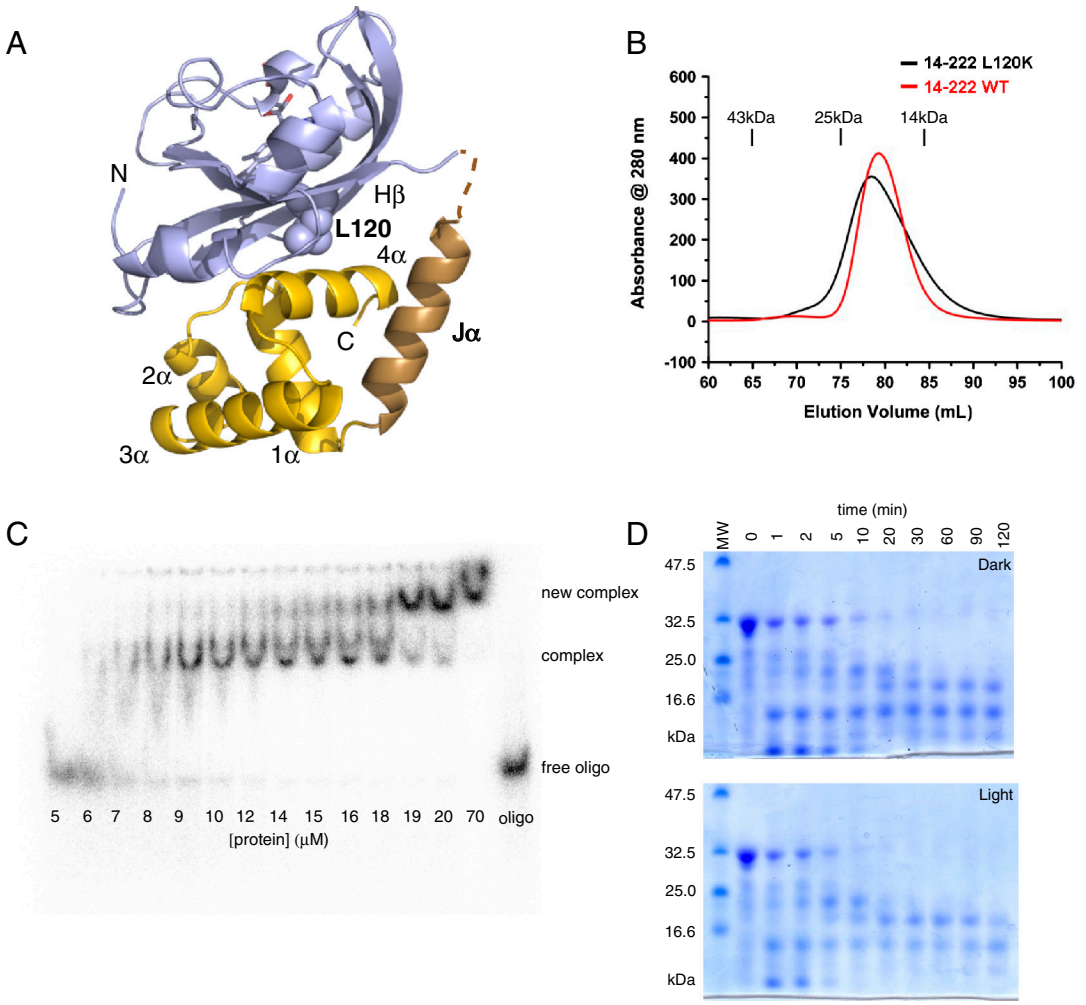

Fig. S6. A point mutation predicted to disrupt LOV-HTH interactions in EL222 leads to constitutive DNA binding. (*A*) L120K is located on the Hβ-strand of the LOV domain (blue) and points down toward 4α of the HTH (gold). (*B*) The L120K mutant appears to be monomeric (black), similar to wild type (red), as shown by comparable retention times in Superdex 75 gel filtration chromatography. The elution volumes for three standards are shown: ovalbumin, 43 kDa; chymotrypsin A, 25 kDa; RNase A, 13.7 kDa. (*C*) Electrophoretic mobility shift assays show L120K binds DNA (oligomer 1) in the dark with comparable affinity as wild-type EL222 in the light. However, L120K appears to have higher order binding at higher protein concentrations (labeled new complex) that wild type does not demonstrate. (*D*) Limited proteolysis with chymotrypsin in the dark and lit states reveals little change in enzyme accessibility with both states behaving similarly to the wild-type lit-state protein.

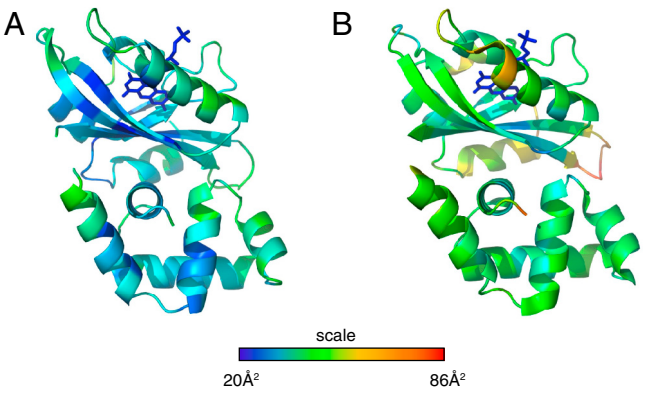

Fig. S7. Differences in temperature factors between the two EL222 molecules in the asymmetric unit. B factors of molecule A and B (*A* and *B*, respectively) are represented as a heat map (low, blue; high, red, ranging from 20 to 86 Å^2). The average B factor for molecule B (46.3 Å^2) is considerably higher than that of molecule A (37.1 Å^2). In both molecules, the HTH domain has a higher average B factor than the corresponding LOV domain and, in turn, the Jα-linker helix has a higher B factor than either the LOV or HTH domains.

Table S1. Crystallographic data reduction and refinement statistics for the EL222 dark-state structure

Data reduction	
Space group	$P2_1$
Unit cell parameters	
a, Å	37.04
b, Å	81.60
c, Å	65.48
β, °	94.57
Matthews coefficient, $Å^3$/Da	1.98
No. of molecules in the asymmetric unit	2
Solvent content, %	37.8
Resolution, Å	36.93 - 2.10 (2.18 - 2.10)
Total observations	174,625
Unique reflections	22,731
Completeness, %	99.8 (99.8)
R_{merge}	0.102 (0.473)
Avg I/sig(I)	10.1 (2.5)
Mosaicity, °	2.21
Avg redundancy	7.68 (4.14)
Refinement	
Resolution range, Å	36.93 - 2.10
No. of reflections	
Working set	22,135
Test set	1,092
R_{work}, %	26.3
R_{free}, %	32.9
Avg B factor, $Å^2$	41.7
Rmsd bond lengths (Å)/angles, °	0.003/0.779
Ramachandran plot, %	
Most favorable regions	90.8
Allowed regions	9.2
Disallowed regions	0
Model details	
No. of protein atoms	3,026
No. of water molecules	98

Note, values in parentheses are for the highest resolution shell (2.18–2.10 Å).

Table S2. Molecular interactions are different at the LOV-HTH domain interface for the two molecules in the asymmetric unit

	Buried interface area, $Å^2$	ΔG_i, kcal/mol	ΔG_i P value	HB	SB
EL222 molecule A without Jα	714	−6.8	0.249	8	0
EL222 molecule B without Jα	612	−6.6	0.233	4	0
AsLOV2	811	−15.3	0.076	9	2

There is a difference in the number of hydrogen bonds (HB) for molecules A (8) and B (4) and no salt bridges (SB) are present at the interface for EL222. The computed solvation free energy gains (ΔG_i) of the domain/domain interactions are favorable for molecules A (−6.8 kcal/mol) and B (−6.6 kcal/mol), both of which are significantly lower than that of *Avena sativa* phototropin 1 LOV2 (AsLOV2) with its Jα-helix (−15.3 kcal/mol). The ΔG_i P values ($P < 0.5$) of all interactions support that the interface surfaces are interaction specific. All information in this table was computed using the PISA (Protein Interfaces, Surfaces, and Assemblies) service at the European Bioinformatics Institute (http://www.ebi.ac.uk/msd-srv/prot_int/pistart.html) (1).

1 Krissinel E, Henrick K (2007) Inference of macromolecular assemblies from crystalline state. *J Mol Biol* 372:774–797.