BIOCHEMISTRY
including biophysical chemistry & molecular biology

Article

pubs.acs.org/biochemistry

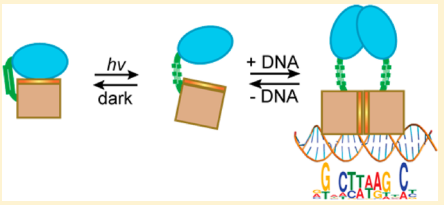
Identification of Natural and Artificial DNA Substrates for Light-Activated LOV–HTH Transcription Factor EL222

Giomar Rivera-Cancel, Laura B. Motta-Mena, and Kevin H. Gardner*

Departments of Biophysics and Biochemistry, University of Texas Southwestern Medical Center, Dallas, Texas 75390-8816, United States

Supporting Information

ABSTRACT: Light–oxygen–voltage (LOV) domains serve as the photosensory modules for a wide range of plant and bacterial proteins, conferring blue light-dependent regulation to effector activities as diverse as enzymes and DNA binding. LOV domains can also be engineered into a variety of exogenous targets, allowing similar regulation for new protein-based reagents. Common to these proteins is the ability for LOV domains to reversibly form a photochemical adduct between an internal flavin chromophore and the surrounding protein, using this to trigger conformational changes that affect output activity. Using the *Erythrobacter litoralis* protein EL222 model system that links LOV regulation to a helix–turn–helix (HTH) DNA binding domain, we demonstrated that the LOV domain binds and inhibits the HTH domain in the dark, releasing these interactions upon illumination [Nash, A. I., et al. (2011) *Proc. Natl. Acad. Sci. U.S.A. 108*, 9449–9454]. Here we combine genomic and in vitro selection approaches to identify optimal DNA binding sites for EL222. Within the bacterial host, we observe binding at several genomic sites using a 12 bp sequence consensus that is also found by in vitro selection methods. Sequence-specific alterations in the DNA consensus reduce EL222 binding affinity in a manner consistent with the expected binding mode, a protein dimer binding to two repeats. Finally, we demonstrate the light-dependent activation of transcription of two genes adjacent to an EL222 binding site. Taken together, these results shed light on the native function of EL222 and provide useful reagents for further basic and applications research of this versatile protein.



For cells to respond to changes in their environment, they rely on sensory proteins to perceive these changes and initiate appropriate responses at the biochemical level. Two critical aspects of this process, detecting the signal and transmitting this to downstream effectors, are elegantly combined within several types of small protein domains that bind environmentally sensitive cofactors, using these to trigger changes in protein structure that affect sensor–effector interactions. This principle has been demonstrated for several different types of sensory domains, including the PAS (Per-ARNT-Sim) domain family that includes sensors of oxygen, redox state, light, and other stimuli.[1]

The signaling mechanism of PAS domains is nicely exemplified by a subset that utilizes internally bound flavin chromophores to sense changes in blue light or redox state, known as LOV (light–oxygen–voltage) domains.[2] In the dark, LOV domains exist with a single noncovalently bound FMN or FAD molecule near a conserved set of residues within a mixed α/β fold common to all PAS domains. Upon illumination, a covalent adduct is formed between one of these residues, a cysteine, and the C4a position of the flavin isoalloxazine ring. This adduct formation triggers the rearrangement or dissociation of protein binding to the external surface of the β-sheet, controlling the activity of effector domains.[3,4] Originally demonstrated in studies of isolated LOV domains from phototropins,[2] a group of light-activated serine/threonine kinases from plants, this type of light-dependent regulation has since been found in a wide range of plant, algal, and bacterial proteins with very diverse effectors.[5] LOV regulation is portable enough to be engineered into a variety of downstream targets, allowing the successful design of fusion proteins conferring photoactivation to enzymatic and nonenzymatic targets.[6−8] As such, understanding the biophysical nature of this control is essential to understanding this type of natural photosensing and furthering engineering efforts.

In this vein, we have examined the generality of this signaling mechanism with studies of several bacterial LOV domain-containing proteins, which are members of the rapidly growing ensemble of photoreceptors that control diverse responses in phototrophic and nonphototropic bacteria (recently reviewed in ref 9). One such protein, EL222 from the alphaproteobacterium *Erythrobacter litoralis* HTCC2594,[10] provides one of the smallest complete LOV domain-containing proteins with both sensor and effector domains inside of a small framework (222 amino acids). An example of a "one-component" signaling protein,[11] EL222 contains both a LOV sensor and a helix–turn–helix (HTH) DNA binding domain. Combining this domain architecture and LOV signaling principles, we hypothesized that EL222 is a light-dependent DNA binding protein, which we tested with a combination of biophysical and biochemical approaches.[10] Structural studies indicated that the

Received: September 25, 2012
Revised: November 13, 2012
Published: December 3, 2012

LOV and HTH domains are tightly associated in the dark, with the LOV domain β-sheet docking to the HTH 4α helix, blocking the ability of this helix and protein to dimerize as is typically required for HTH domains to bind DNA.[12] Using nuclear magnetic resonance, limited proteolysis, and other approaches, we demonstrated that light-dependent conformational changes break this association. To survey the functional effects of these changes, we used a candidate-based approach to identify EL222-binding sequences from within the EL222 promoter. In vitro screening of more than 20 overlapping 45mer duplex DNA sites found several that bound specifically to EL222 in the light but not in the dark. However, the relatively low affinity of this interaction (5−10 μM) compared to the affinities of other HTH domain−DNA interactions reported to be between 0.1 and 1000 nM[13−16] suggested that higher-affinity DNA binding sites might exist. Such would be useful reagents for biochemical studies, to verify the binding site preference of EL222, and for engineering purposes.

To address this shortcoming, we have pursued two independent approaches to identify higher-affinity DNA binding sites. Our first approach was to use in vivo ChIP−Seq (chromatin immunoprecipitation−high-throughput sequencing[17,18]) to identify binding sites within the *E. litoralis* genome. We demonstrate that this approach was successful at providing 11 different sites that we validated with a variety of in vitro approaches. Further, we have found light-dependent gene activation near one of these sites, providing the first indication of EL222 serving as a light-dependent transcription factor. A complementary approach was provided by in vitro selection (SELEX[19]) of EL222 binding under lit state conditions. Both methods identified a converged sequence consensus between the natural and artificial sites that are sufficient to predict novel sites within the EL222 genome. Finally, we present results from mutagenesis studies of the EL222 DNA binding sites, showing position-dependent effects consistent with binding in the predicted dimeric, major groove binding mode expected of HTH proteins. Taken together, these data advance our understanding of this class of LOV-dependent proteins and lay the foundation for further functional work in the area.

EXPERIMENTAL PROCEDURES

ChIP−Seq (chromatin immunoprecipitation−high-throughput sequencing) Assay. *E. litoralis* HTCC2594 was grown in 1 L of ZoBell marine broth 2216 (HiMedia Laboratories) at 30 °C to a OD_{600} of 0.8. Cells were illuminated using a blue LED panel (14 W) and a white light flood lamp (150 W) for 20 min, cross-linked with 0.6% formaldehyde under illumination for an additional 20 min, and subsequently quenched with 0.5 mM glycine. Cells were harvested by centrifugation, washed twice with PBS, and stored at −80 °C until they were ready to be used. Pellets were thawed on ice, resuspended in 1 mL of lysis buffer [10 mM Tris (pH 8.0), 50 mM NaCl, 10 mM EDTA, and 20% sucrose], and incubated with lysozyme (10 mg/mL) for 30 min at 30 °C. The suspension was diluted via addition of 4 mL of PBS with 1% Triton X-100 and sonicated on ice at 60% power for four cycles of 15 pulses (1 s on, 1 s off) with 1 min between cycles. Samples were supplemented with 1 mM PMSF and clarified by centrifugation. Chromatin was precleared by incubation with 100 μL of protein A/G-Plus agarose beads (Santa Cruz Biotechnology) for 2 h at 4 °C in a rotator and subsequently divided into 800 μL aliquots. Each aliquot was incubated overnight at 4 °C with 20 μL of protein A/G agarose beads and 4 μg of anti-EL222 antibody (YenZym) (with the exception of the mock sample, which did not have the antibody). Beads were washed four times with PBS, after which protein and DNA were eluted in 50 μL of elution buffer by being heated (65 °C for 15 min). Eluted fractions were treated with RNase A and proteinase K, heated for 16 h at 65 °C, and processed with the Qiaquick polymerase chain reaction (PCR) clean-up kit (Qiagen).

Deep-Sequencing Analysis. Single end read libraries were constructed for IP and mock samples following standard protocols (Illumina). After we had checked the library quality using an Agilent Technologies 2100 Bionalyzer and measured the concentration with PicoGreen reagent (Invitrogen), approximately 0.1 pg of DNA from IP and mock samples was sequenced on an Illumina Genome Analyzer IIx at the University of Texas Southwest Next Generation Sequencing Core. Reads were aligned to the *E. litoralis* HTCC2594 genome[20] using the Burrows-Wheeler Aligner,[21] and the aligned reads were visualized using IGB.[22] Sequencing statistics are listed in Table S1 of the Supporting Information.

Gene Expression Analysis. *E. litoralis* HTCC2594 cultures were grown in triplicate in ZoBell marine broth 2216 (HiMedia Laboratories) to an OD_{600} of 0.8. Samples (1 mL) were collected at three time points: (1) preillumination, (2) after white and blue light illumination for 30 min, and (3) after incubation for 30 min in the dark following illumination. RNA was extracted using Trizol reagent (Invitrogen) and treated with DNase I before generation of cDNA with the iScript cDNA synthesis kit (Bio-Rad). Real-time PCR was performed in duplicate for each biological replicate using the iTaq SYBR Green Supermix (Bio-Rad). Changes in mRNA levels were analyzed using the comparative C_T method[23] using *rpoD* expression as a reference.

Overexpression and Purification of EL222 Protein. WT-EL222 and A79Q-EL222 proteins[10,24] were expressed in *Escherichia coli* BL21(DE3) cells, grown in LB-AMP at 37 °C in the dark, and induced with 0.5 mM IPTG until the OD_{600} reached ∼0.5−0.7. After being induced for 20 h at 18 °C, cells were centrifuged, and the resulting pellets were resuspended in buffer A [50 mM Tris-HCl (pH 8.0) and 100 mM NaCl] and subsequently lysed by sonication. For SELEX experiments, protein was purified in the dark at 4 °C by gravity-flow chromatography with Ni-NTA agarose (Qiagen) equilibrated in buffer A. Proteins were eluted in buffer A with 75 mM imidazole, exchanged into imidazole-free buffer A, and concentrated to 100−200 μM. For EMSA experiments conducted with ChIP−Seq-derived DNA, WT-EL222 was purified by FPLC using Ni-affinity chromatography as previously described,[10] followed by a final Superdex 75 size exclusion chromatography step.

SELEX Search for EL222-Binding Sequences. The initial single-stranded oligonucleotide library [5′-GGGAATGGATC-CACATCTACG-$(N)_{33}$-TTCAACTTGACGAAGCTTGCC-3′] was chemically synthesized (IDT). To amplify the DNA pool, six 50 μL PCR mixtures were set up using the synthetic pool oligonucleotide (0.1 μM) as the template, 2 μM primers (Fwd-L1, 5′-GGGAATGGATCCACATCTACG-3′; and Rev-L1, 5′-GGCAAGCTTCGTCAAGTTGAA-3′), 200 μM dNTPs, 20 mM Tris-HCl (pH 8.8), 10 mM $(NH_4)_2SO_4$, 10 mM KCl, 2 mM $MgSO_4$, 0.1% Triton X-100, 4 mM $MgCl_2$, and 0.04 unit of Vent (New England Biolabs). Amplified DNAs were purified using the QIAquick PCR purification kit (Qiagen). In a total volume of 500 μL, approximately 15.7

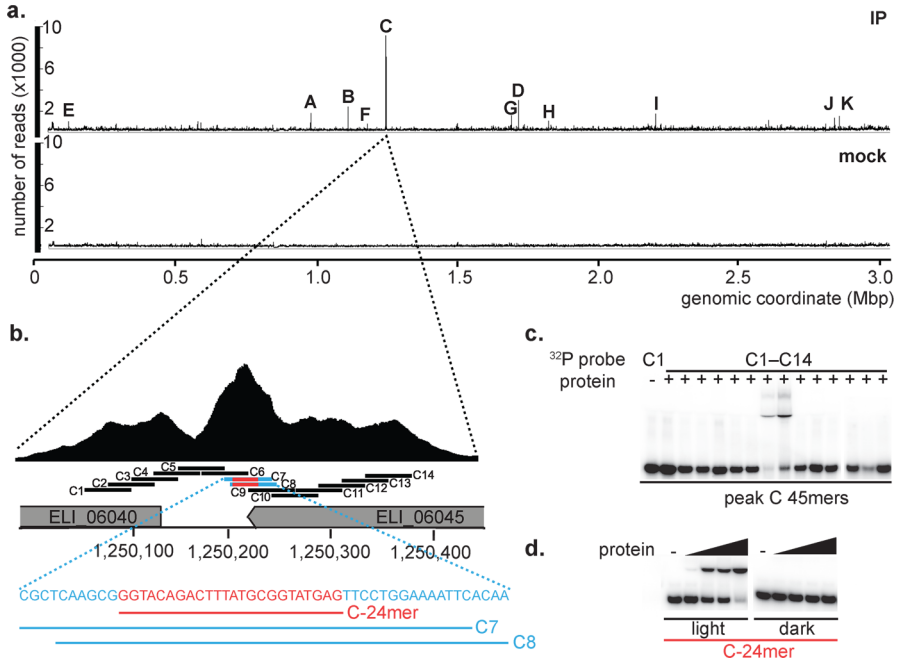


Figure 1. ChIP–Seq identification of EL222 binding sites within the *E. litoralis* genome. (a) Comparison of ChIP–Seq results from *E. litoralis* cultures grown and cross-linked in light; histograms show results of samples with an anti-EL222 antibody (IP, top) and a negative control without antibody (Mock, bottom). Numbers of reads (*y*-axis) are plotted against genomic coordinates (*x*-axis). (b) Expanded profile of peak C with an overview of the strategy for determining a specific binding sequence and validating binding by ChIP–Seq. DNA probes C1–C14, schematically represented by bars under the profile, were tested for binding by an EMSA (c). Additional information includes the location of two open reading frames within the peak area (gray arrows pointing from 5′ to 3′) and the combined sequence of probes C7 and C8 (blue, full sequence; red, 24 bp fragment used in panel d). (c) Lit state EMSA of probes C1–C14, with either no protein (−, negative control) or 0.5 μM EL222 (+). (d) Light dependence of EL222 binding of DNA demonstrated by an EMSA (24 bp sequence in red, panel b) in the presence of increasing concentrations of protein (0, 0.25, 0.50, and 0.75 μM).

μM DNA (5×10^{14} molecules) and 7.8 μM His_6-tagged EL222 protein (containing an A79Q point mutation to slow dark state recovery[24]) were incubated in binding buffer [10 mM Tris-HCl (pH 8.0), 80 mM NaCl, 3 mM $MgCl_2$, 10% glycerol, 0.025 mg/mL poly(dI-dC), and 0.01 mg/mL BSA]. The binding reaction mixture was mixed by rotation for 25 min at 4 °C and kept under continuous illumination with a fluorescent white light bulb (23 W). Ni-NTA agarose beads (Qiagen) were preblocked with 0.02 mg/mL poly(dI-dC), added to the binding reaction mixture, and incubated for an additional 25 min at 4 °C with mixing and continuous illumination. The bead–EL222–DNA complexes were pulled down by centrifugation at 400*g* for 1 min at 4 °C. Next, complexes were washed with 300 μL of wash buffer [50 mM Tris-HCl (pH 8.0) and 300 mM NaCl] and then centrifuged to pellet the beads; this step was repeated two or more times. After the last wash, the bead–EL222–DNA complexes were resuspended in 400 μL of binding buffer [without poly(dI-dC) and BSA] and incubated in the dark for 30 min at 4 °C to elute the DNAs. The bead–EL222 complexes were pulled down by centrifugation, and the supernatant (containing the DNAs) was transferred to a new tube. A phenol/chloroform/isoamyl alcohol mixture was added to the supernatant in a 1:1 ratio, and the sample was vortexed and centrifuged at 15800*g* for 5 min at 4 °C. The top aqueous phase, which contains the eluted DNAs, was transferred to a new tube containing 1 mL of 100% ethanol, 0.3 M NaOAc (pH 5.2), and 0.01 mg/mL glycogen. The DNAs were precipitated overnight at −20 °C and subsequently recovered by centrifugation at 15800*g* for 20 min at 4 °C. The pelleted DNAs were resuspended in 12 μL of DNase-free water and later used as the template DNA pool in a second PCR amplification step and round of selection. To sequence individual DNA sequences from the DNA pools obtained after each SELEX round, the DNA pools were cloned into the pBlueSkript+ vector (Stratagene) using BamHI and HindIII restriction sites. Computational analysis of the sequences identified with SELEX was conducted using the MEME algorithm.[25]

Electrophoretic Mobility Shift Assay (EMSA). All EMSA experiments conducted with ChIP–Seq-derived DNA sequences were performed as previously described.[10] For EMSAs conducted with DNAs derived from SELEX, the experiment was conducted as follows. The DNA pools from each cycle were first amplified via PCR as described for the SELEX procedure. For the individual SELEX-derived clones, complementary oligonucleotides were chemically synthesized (Sigma). The oligonucleotides were annealed by being heated to 95–100 °C for 5 min and then left to cool to room temperature. The DNA pools and the individual clones were 5′-end labeled in a 50 μL reaction mixture containing 68 nM DNA substrate, 70 mM Tris-HCl (pH 7.6), 10 mM $MgCl_2$, 5 mM DTT, 0.6 μCi of [γ-^{32}P]ATP (Perkin-Elmer), and 0.2 unit of PNK (New England Biolabs). The reaction mixture was incubated for 30 min at 37 °C, followed by a 20 min incubation at 65 °C to heat-inactivate the enzyme. The ^{32}P-labeled DNA was purified from the unincorporated [γ-^{32}P]ATP using Illustra ProbeQuant G-50 microcolumns (GE Healthcare). Approximately 13.6 nM radiolabeled DNA was incubated with varying concentrations of WT-EL222 in the same binding buffer used for the SELEX procedure for 25 min at 4 °C with continuous illumination with

dx.doi.org/10.1021/bi301306t | *Biochemistry* 2012, 51, 10024–10034

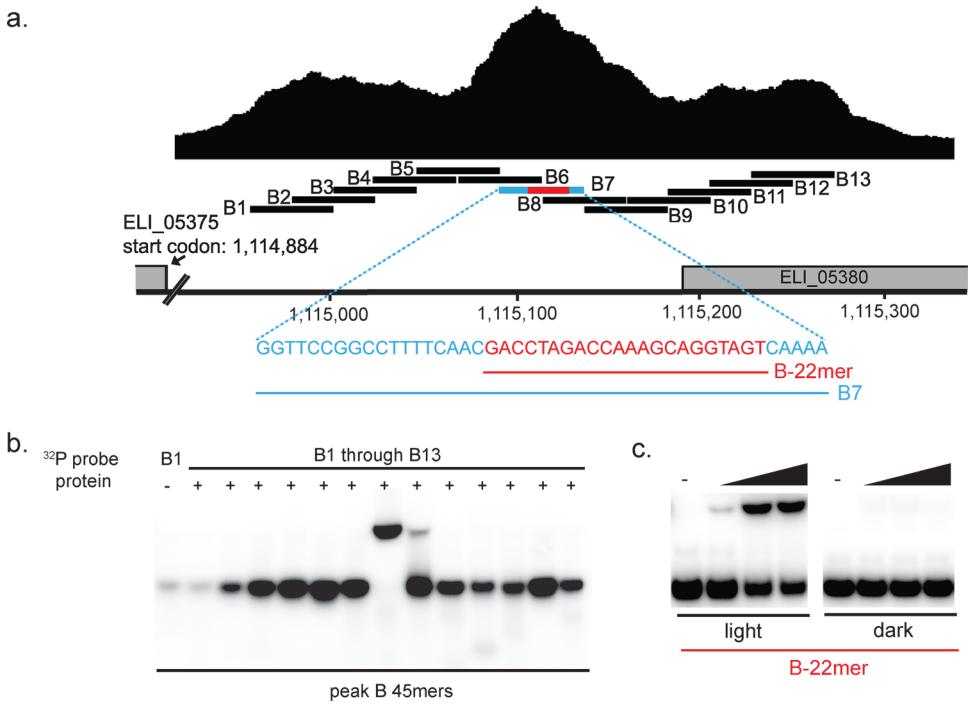


Figure 2. Refinement of the EL222-binding sequence within ChIP−Seq peak B. (a) Profile of ChIP−Seq peak B. The location of 45 bp DNA probes tested by an EMSA is shown as horizontal bars under the profile. Sequences for probe B7 and a derived 22mer are also indicated. (b) Probes from part a were tested for binding by an EMSA. Lanes containing 0.5 μM EL222 are indicated by +, and a negative control without protein is indicated by −. (c) Titration of the 22 bp probe based on probe B7 was used to confirm that binding of EL222 to this sequence was light-dependent. Protein concentrations of 0, 0.25, 0.50, and 0.75 μM were used (from left to right, respectively).

a fluorescent white light bulb or kept in the dark. Reaction mixtures were analyzed on a 5% native gel (acrylamide:bisacrylamide ratio of 29:1) and run in TBE buffer at 150 V for 1.5 h at 4 °C. The gel was exposed to a phosphorimaging plate and visualized using a FujiFilm FLA-5100 imaging system.

RESULTS

Identification of EL222 Binding Sites by ChIP−Seq. As noted above, we first demonstrated the light-dependent DNA binding properties of EL222 using a candidate-based approach to find putative DNA binding sites. Assuming EL222 might be autoregulatory and bind to sequences in its own promoter, we used an EMSA to survey twenty-one 45 bp sequences located within 350 bp of the EL222 gene itself.[10] Three sequences bound with low micromolar affinities, the best of which, AN-45, called oligomer 1 in ref 10, bound EL222 with an EC_{50} of 5−10 μM in the light with minimal binding in the dark, expected from the presence of the photosensitive LOV domain. However, while AN-45 bound to EL222 with the highest affinity of this limited set of oligonucleotides, the relatively low affinity compared to those of other HTH domain-containing proteins[13−16] led us to suspect that it was not an ideal binding site.

To search for EL222-binding sequences more broadly and in an unbiased manner, we used ChIP−Seq[17,18] (Figure 1a). We compared results from two data sets, both of which were generated using *E. litoralis* cultures grown and cross-linked under lit state conditions. One set used an anti-EL222 antibody for the immunoprecipitation step (IP in Figure 1a); an antibody-free parallel experiment (Mock) controlled for nonspecific binding. Reads for both IP and mock IP were mapped to the *E. litoralis* HTCC2594 genome sequence.[20]

From these data, we identified 11 putative EL222 binding sites (peaks A−K in Figure 1a) that were enriched up to >30-fold in the IP data set over the mock IP counterpart. For our initial characterization, we focused on peak C, which showed the greatest enrichment in the ChIP−Seq experiment. Using an EMSA to survey the ability of 45 bp probes covering this peak to bind 0.5 μM EL222 in vitro, we found two sequences (C7 and C8) that were substantially bound (Figure 1b,c); both are located at the center of the ChIP−Seq peak. We tested smaller 20−24 bp probes covering the sequence comprised by probes C7 and C8 and found a 24 bp probe (C-24mer) that bound with similar affinity as the 45 bp probe C8 (Figure S1 of the Supporting Information). Finally, we confirmed that the binding of EL222 to this 24 bp sequence was light-dependent by comparing EMSA data collected under dark and lit conditions (Figure 1d). Binding in the dark was negligible compared to binding under lit conditions at all concentrations tested, consistent with our prior findings.[10] Similar studies with peaks A, B, and D established 45 bp sequences near the centers of each peak that bound EL222 (Figure 2 and Figure S2 of the Supporting Information). In a manner analogous to that used for peak C, we refined probe B7 of peak B to a 22 bp sequence that bound in a light-dependent manner with similar affinity to the C-24mer (Figure 2c).

SELEX Identifies Additional EL222-Binding Sequences from Randomized Libraries. To complement the results we obtained from the ChIP experiment, we developed an in vitro selection strategy based on SELEX (systematic evolution of ligands by exponential enrichment[19]). To do so, we used a large library of double-stranded oligonucleotides (~5 × 10^{14} molecules) consisting of a random 33 bp central portion flanked by constant 5′ and 3′ ends (each 21 bp long) with primer binding sites (Figure 3a). Next, we incubated the

dx.doi.org/10.1021/bi301306t | *Biochemistry* 2012, 51, 10024−10034

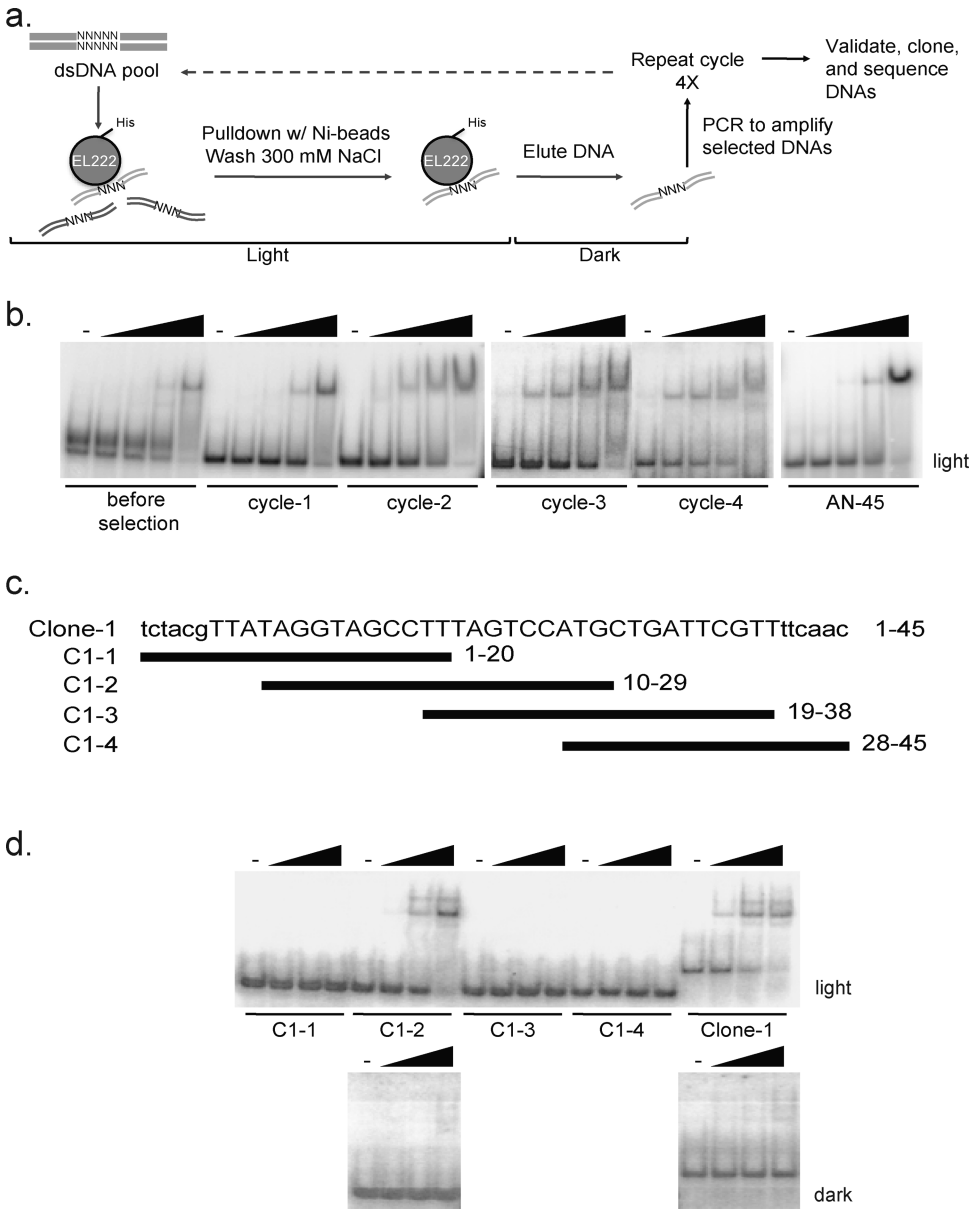


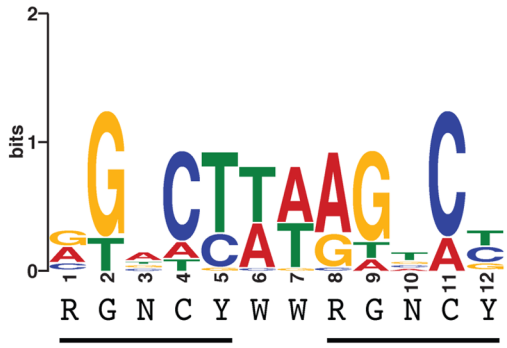
Figure 3. In vitro selection (or SELEX) identifies additional high-affinity EL222 target sequences. (a) A pool of dsDNA molecules, each 75 bp in length and containing a region of 33 random nucleotides flanked by primer binding sites, was incubated with recombinant His-tagged EL222 protein and exposed to light. Next, the protein−DNA complexes were isolated, washed, and incubated in the dark to elute the DNAs. Finally, PCR was conducted to enrich the eluted DNAs, which were used in subsequent rounds of selection. (b) Radiolabeled versions of the DNA pools from each selection round were incubated with increasing amounts (0.3, 1, 3, and 9 μM) of EL222, and the mixtures were exposed to light and electrophoresed on a native polyacrylamide gel. (c) Sequence of the Clone-1 DNA identified from cycle 4 of SELEX. The 45 bp Clone-1 DNA contains a 33 bp randomized region (uppercase letters) flanked by 6 bp of the primer binding sequence on each side (lowercase letters). Four 20 bp fragments (C1-1−C1-4) were used to map the sequence within the Clone-1 DNA to which EL222 binds. (d) Radiolabeled versions of 20 bp DNAs were incubated with increasing amounts (0.1, 0.3, and 1 μM) of EL222, and the mixtures were exposed to light or kept in the dark and electrophoresed on a native polyacrylamide gel.

oligonucleotide library with recombinant His_6-tagged EL222 protein and exposed the mixture to light to activate the protein. EL222−DNA complexes were purified by affinity chromatography using Ni-NTA beads under stringent binding conditions, and the DNAs were eluted from the bead-bound complexes by incubation in the dark. Eluted DNAs were amplified by PCR and used as a library for a subsequent round of selection; this entire cycle was repeated a total of four times. To verify that the binding affinity of the DNA pools increased with each successive round, EMSAs were conducted after each selection cycle. As expected, by the fourth cycle of SELEX, we had enriched a pool of DNAs that bound EL222 appreciably tighter than the AN-45 substrate (Figure 3b).

To characterize the features of these EL222-binding sequences, we cloned and sequenced 57 unique sequences from the DNA pool obtained after the fourth round of selection (Table S2 of the Supporting Information). On the basis of analyses conducted using the motif-based sequence analysis tool MEME (Multiple Em for Motif Elicitation[25]), the 10 highest-scoring DNAs were chosen for follow-up binding studies (Figure S3a,b of the Supporting Information). Gel shift assays showed that Clone-1 bound EL222 most tightly of the

dx.doi.org/10.1021/bi301306t | *Biochemistry* 2012, 51, 10024−10034

a.



b.

Name	Strand	p-value	Sites
Clone-3	-	7.23e-07	CTTGTGTACT **AGACTTAAGTCT** A
Clone-22	+	5.66e-06	TATA **GGGCTTTAGTCT** GTATGTGGGA
Clone-1	+	5.66e-06	TTATAGGT **AGCCTTTAGTCC** ATGCTGATTC
Peak-A	-	3.78e-05	TGTCCATCAT **GGGCCAAGGTCC** ATATTTTCTT
Clone-53	+	8.30e-05	TAGTACTTAC **AGTCCTTAGGCC** ATG
Peak-C	-	2.94e-04	GAACTCATAC **CGCATAAAGTCT** GTACCCGCTT
Clone-2	+	3.60e-04	AGCTA **GGCTTTTGGTCT** GTATAGCTGT
Peak-B	-	4.25e-04	TTTTGACTAC **CTGCTTTGGTCT** AGGTCGTTGA
Clone-77	-	4.73e-04	CAGATAC **ATGCTTTAGTAC** GTATATTCGA
Clone-74	+	1.29e-03	TTATTAAGTC **CGTATTAAGTAT** GGAGGGAGAG
Clone-55	-	1.59e-03	TCAGTAAACC **CTTATATAGTCC** TTAAGTA
Peak-D	+	4.65e-03	**TGCCTTGGGCCC** GCGCCGCAGC
Clone-78	+	4.90e-03	CCCTAAGACT **GTATCTTATTCG** TGAGTTTTAA
Clone-88	+	7.11e-03	CTTATGACCA **GGTTGTTGGACT** GATTGTAAC

Figure 4. Identification of a DNA consensus sequence for EL222. (a) Sequence logo representation of the EL222-specific motif obtained using the MEME algorithm.[25] The 12 bp consensus sequence consists of two binding sites (underlined); the IUPAC version for the sequence is shown. (b) MEME alignment results (http://meme.sdsc.edu/meme/cgi-bin/meme.cgi) for the EL222 target DNA sequences obtained from the ChIP–Seq and SELEX studies in which the consensus motif is shown in bold.

10 DNAs tested, with approximately 16–30-fold higher affinity than AN-45 [EC_{50} = 0.3 μM Clone-1 vs EC_{50} = 5–10 μM AN-45 (Figure S3d of the Supporting Information and ref 10)]. To further refine the sequence determinants for binding within the 45 bp Clone-1 DNA, we designed four overlapping 20 bp fragments derived from Clone-1 and assessed their binding to EL222 by a gel shift assay (Figure 3c,d). Interestingly, only the C1-2 fragment was bound by recombinant EL222 protein at protein concentrations between 0.3 and 1 μM, while the other three fragments showed no binding under these conditions. Thus, this result suggests that C1-2 contains all of the residues necessary for the binding of EL222.

ChIP and SELEX Reveal a Common Binding Motif. From these data, it is clear that both ChIP and SELEX were capable of independently identifying sequences with submicromolar affinities for EL222. To determine if any shared motifs existed within these sequences, we again used the motif-based sequence analysis tool MEME.[25] We aligned the 45 bp EL222-binding sequences from ChIP–Seq peaks A–D (probes B7 and D8, and the combination of probes A5 and A6 and probes C7 and C8) and 33 bp regions from the 10 SELEX sites validated by EMSA (Figure S3a,b and Table S2 of the Supporting Information) and searched for palindromic motifs. A search for a ≥12 bp motif revealed the consensus sequence RGNCYW-WRGNCY [Y = C or T, W = A or T, R = A or G, and N = any nucleotide (Figure 4)], while searches with the default minimal motif width of 6 returned only part of this motif, RGNCY. As observed for other HTH domains,[12] the 12 bp consensus contained two repeats separated by a short spacer. Here, we found 5 bp repeats, each of which contained a highly conserved GNC element flanked by a purine (5′) and pyrimidine (3′) residue, separated by an AT-rich 2 bp central spacer. Within the search sequences, we found examples of both inverted and direct arrangements of these repeats. Remarkably, we find that the MEME-predicted EL222 consensus motif maps specifically to the 20–24 bp fragments identified by truncation analyses of ChIP–Seq peaks B and C (Figures 1 and 2 Figure S1 of the Supporting Information) and Clone-1 (Figure 3d). We also confirmed that a 22 bp fragment of peak A containing the motif found in the MEME alignment was bound by EL222 (Figure S2c of the Supporting Information). These results provide independent support for the possibility that the MEME motif may serve as an EL222 binding site; however, they do not confirm that the motif predicted by MEME is the bona fide consensus sequence for EL222, nor do they directly establish the relative energetic importance of different residues within the motif.

To further explore the EL222 binding roles of residues within the MEME motif, we engineered mutations to disrupt sequence elements in the predicted consensus sequence. We used SELEX Clone-1 as the template for these mutations, as it both bound EL222 tightly (Figure 3d) and contains a copy of the consensus sequence (Figure 4B), and used an EMSA to analyze EL222 binding. Purified recombinant EL222 efficiently binds the wild-type Clone-1 sequence with submicromolar affinity (Figure 5b and Figure S3 of the Supporting Information); in contrast, sequences with mutations in the either of the two GNC elements substantially weakened

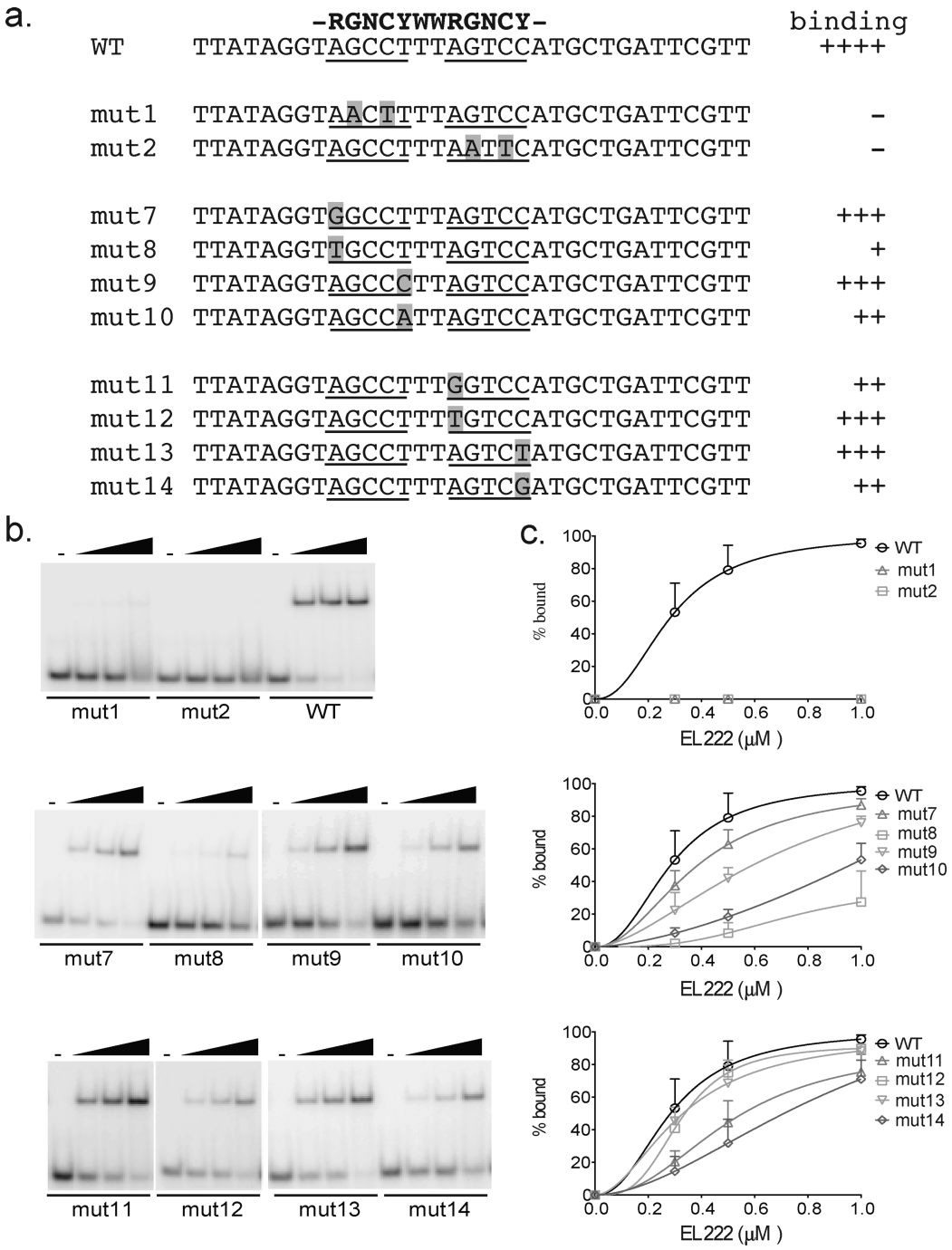


Figure 5. Mutagenesis of residues within the MEME-derived motif in Clone-1 disrupts the binding of EL222. (a) Schematic of the sequences of wild-type and mutant Clone-1 DNAs. Only the 33 bp randomized region is shown. Each mutation is boxed in gray, and the repeats are underlined. The binding data are summarized using an arbitrary scale from − (weakest) to ++++ (strongest). (b) Radiolabeled versions of the DNAs shown in panel a were incubated with increasing amounts (0.3, 0.5, and 1 μM) of EL222, and the mixtures were exposed to light and analyzed on a native polyacrylamide gel. (c) Quantification of the bound species from three independent gel shift experiments performed with each of the DNAs shown in panel a.

binding (Figure 5b and Figure S4 of the Supporting Information). Simultaneous mutation of both G and C residues in a single repeat was most detrimental to binding (Figure 5b, mut1 and mut2). We also saw lowered binding affinities when individual G or C residues were changed within either repeat, albeit not as dramatically as in the double mutant (Figure S4 of the Supporting Information, mut3−6). These results suggest that EL222 requires two intact GNC repeats to bind to the Clone-1 sequence with submicromolar affinity. The fact that disrupting only one of the repeats is enough to weaken the binding of EL222 in our gel shift assays is consistent with our signaling model that EL222 binds to DNA as a dimer,[10] as is commonly observed for other HTH proteins.[12]

Turning to less well-conserved residues, we examined the importance of the purine and pyrimidine residues that MEME predicted to flank the GNC element on the 5′ and 3′ sides, respectively. Consistent with this, swaps among the purines (A → G on 5′, mut7) or pyrimidines (T → C on 3′, mut9) within

dx.doi.org/10.1021/bi301306t | *Biochemistry* 2012, 51, 10024−10034

the left repeat were tolerated, while purine to pyrimidine changes and vice versa (A → T, mut8; T → A, mut10) more substantially lowered the binding affinity (Figure 5). Similar results were obtained when we made the equivalent mutations in the right repeat (Figure 5, mut11−mut14). These results confirm that the sequence conservation of these flanking residues reflects an inherent energetic preference for specific types of residues at these sites.

Finally, we also examined the importance of the least conserved positions, the N residues of the GNC repeats and the inter-repeat spacers. As shown in Figure S4 of the Supporting Information, binding of EL222 to DNAs with mutations in the middle N position of either repeat (mut15 or mut16) or both (mut17) was virtually identical to that of wild-type Clone-1. Turning to the spacer region separating the two repeats, we found changes to the spacer sequence had little to no effect on binding (mutating both T → A, mut18) (Figure S4 of the Supporting Information). Importantly, however, the length of the spacer itself was critical; insertion of two additional T residues to convert the spacer from 2 to 4 bp (mut19) completely eliminated binding. In conclusion, these data demonstrate that two properly spaced RGNCY repeats are the primary determinants of EL222 binding.

The Consensus Motif Can Be Detected near the Center of Additional ChIP−Seq Peaks. Using the EL222-binding consensus sequence identified from ChIP−Seq peaks A−D and SELEX data (Figure 4), we predicted binding sites for EL222 in sequences for the remaining ChIP−Seq peaks E−K. We searched for the consensus near the centers of these peaks, based on our observations that this is where the binding sequences were located in the peaks we investigated initially. We then tested if EL222 would bind to these sequences in an EMSA (Figure S5 of the Supporting Information); indeed, EL222 bound to 22 bp oligonucleotides containing the sequences predicted from the consensus at protein concentrations between 0.25 and 0.75 μM.

Genes Downstream of ChIP−Seq Peak B Exhibit Light-Dependent Regulation. The ChIP−Seq results suggested that EL222 bound to specific sites within the *E. litoralis* genome, indicating a possible light-dependent regulatory function. To test the specific possibility that this would affect transcription, we examined the expression levels of genes downstream of peaks with the most ChIP−Seq reads, peaks B (ELI_05375, protein of unknown function, and ELI_05380, radical SAM protein, putative pyrimidine dimer lyase), C (ELI_06040, putative indoleamine 2,3-dioxygenase), and D (ELI_08405, NAD synthetase). Using qRT-PCR, we quantitated expression of these four genes in *E. litoralis* cultures at three points: before exposure to light, immediately after a 30 min illumination, and after a 30 min postillumination dark state recovery period. Notably, we saw induction of both genes near peak B after illumination: 6-fold increase for ELI_05375 and 2-fold increase for ELI_05380 (Figure 6). Expression levels diminished after the cultures had been returned to the dark for 30 min, supporting the light dependence of this effect. Importantly, the timing of this decrease in mRNA levels correlates with the quick recovery of the dark state EL222 conformation postillumination, which is approximately 30 s at room temperature.[10] We did not observe significant changes in expression for genes from the other two peaks that were studied.

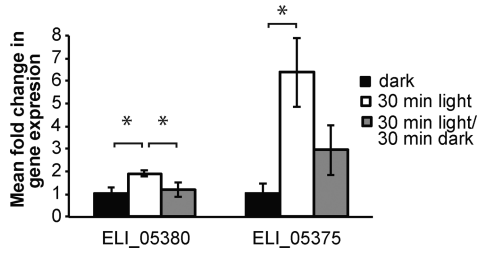


Figure 6. Light regulation of genes adjacent to ChIP−Seq peak B. Light-dependent changes in the expression of genes ELI_05380 and ELI_05375 were measured by RT-PCR and normalized to *rpoD* expression levels. Data were obtained from three independent experiments, and the error bars represent the standard deviation. Samples for which $p < 0.05$ are indicated by a bracket and asterisk.

DISCUSSION

Utilizing independent ChIP−Seq and SELEX approaches, we have identified a 12 bp consensus sequence recognized by the light-dependent LOV−HTH protein EL222. Several properties of this consensus are consistent with our model for EL222 signaling[10] and, more broadly, with general characteristics of HTH domain−DNA interactions. First, the consensus contains two 5 bp binding sites that can be oriented as direct or inverted repeats separated by an AT-rich 2 bp spacer ("5-2-5") for binding an EL222 dimer (Figure 4). Similar arrangements are commonly used by other HTH domains to bind DNA as dimers,[12] with minor differences in the number of base pairs involved: the related LuxR-type HTH domain proteins DosR,[26] NarL,[27] and TraR[28−30] utilize inverted repeats in 9-2-9, 7-2-7, and 8-2-8 configurations, respectively. Interestingly, several HTH proteins have been demonstrated to bind both inverted and direct repeats, including NarL[27] and the serine recombinase Sin.[31] Such binding is facilitated for these proteins by relatively long linkers (approximately 30 amino acids) between the HTH domain and the adjacent domain,[31,32] as we observed between the LOV and HTH domains in EL222 (22 amino acids[10]). Studies of several of these proteins show that the inter-repeat spacing is critical to positioning the two half-sites at ideal distances and phasing for interacting with the HTH domains. Alteration of these AT-rich spacers greatly diminishes the protein binding capability of these repeats, as demonstrated for 2 bp insertions for EL222 (Figure S4 of the Supporting Information, mut19) or TraR,[30] despite the lack of specific protein−DNA base contacts in this region.[28,29] This parallel supports a key feature of our model for EL222 activation by blue light, in which covalent adduct formation within the LOV domain triggers an allosteric process that releases inhibitory interactions between the LOV and HTH domains, converting a dark state monomer into a dimeric form that can bind DNA.[10]

Upon examination of each DNA half-site, HTH domains typically utilize sequence-specific contacts between residues from one of the HTH domain helices (predicted to be helix 3α for EL222) and groups located in the major groove from bases in these repeats,[12] making them particularly important for binding as we observed for EL222 (Figure 5). In the case of TraR, structural data predicted sequence-specific contacts between the protein and the *tra* box DNA that were subsequently tested by using synthetic *tra* boxes containing various base substitutions in in vitro binding assays.[30] Substitutions at positions near the center of each *tra* box half-site that make direct contacts with TraR[28,29] caused the

most severe defects in binding affinity. Comparably, our studies showed that binding of EL222 to Clone-1 DNA was most affected by mutations in the GNC element found in the middle of each repeat (Figure 5 and Figure S4 of the Supporting Information, mut1−mut6). Taken together, these data support parallels between EL222 and other HTH proteins, which will be quantitatively examined in future biophysical and structural studies of lit state EL222 in complex with its cognate DNA.

Regarding the predicted functional effects of blue light activation on EL222, we note that this is an example of a "one-component" bacterial signaling protein,[11] directly linking an environmental sensor domain with a DNA binding effector domain instead of separating these into distinct "two-component" kinase/response regulator proteins. EL222 links the most common sensor and effector functions (small molecule sensing and DNA binding) observed among one-component proteins,[11] leading to our proposal that EL222 is a light-dependent transcription factor. Here we provide several lines of data supporting this hypothesis. (1) The majority of the EL222 binding sites found by ChIP−Seq are in intergenic regions. (2) All ChIP−Seq binding sites but one are within 300 bp upstream of a predicted open reading frame (ORF) translational start (the two genes adjacent to the binding site in peak A, one of which is EL222 itself, have their 3′ ends pointing toward the binding site). (3) In vitro binding assays confirmed EL222 bound to sites within the ChIP−Seq peaks with submicromolar affinity. (4) Genes adjacent to ChIP−Seq peak B exhibited statistically significant upregulation upon illumination (Figure 6). Further studies are needed to determine the complete set of genes regulated by blue light illumination and to establish that EL222 is directly responsible for this control, particularly given that *E. litoralis* HTCC2594 contains two functional blue light-sensing histidine kinases that rely on comparable LOV photochemistry.[33] Strong circumstantial evidence in favor of a direct role for EL222 is provided by the fast decrease in transcript level after removal of light, which is consistent with the rapid rate of LOV dark state recovery ($\tau \sim 30$ s)[10] but is much faster than expected for the slowly cycling LOV-HK proteins in *E. litoralis* and other α-proteobacteria ($\tau > 30$ min).[33,34]

If EL222 indeed functions as an *E. litoralis* transcription factor, what purpose could it serve? We note the growing appreciation of light-dependent changes in fundamental aspects of microbial metabolism, some of which are mediated by other LOV domain-containing proteins.[9] In this vein, several *Erythrobacter* strains are facultative phototrophs, providing numerous candidates for light-dependent regulation. However, the lack of key genes for photosynthesis (e.g., bacteriochlorophyll synthesis, photosynthetic reaction centers, and CO_2 fixation)[20] and an obvious light-dependent growth phenotype (G. Rivera-Cancel and K. H. Gardner, unpublished results) for *E. litoralis* HTCC2594 strongly suggest that other biological responses are controlled by EL222. Some insight is provided by our ChIP−Seq analysis (Table S3 of the Supporting Information); however, limited homology and domain information hampers an in-depth analysis of many genes, including light-regulated ELI_05375 (Figure 6). Nevertheless, several broad themes can be observed, among them metabolism (including NAD biosynthesis), metabolite/ion transport, and DNA repair.

Further examining the last of these three areas, we see several lines of evidence supporting the role of EL222 in controlling DNA repair. Notably, one of our light-induced genes (ELI_05380) encodes a protein that could potentially be implicated in repair because of the radical *S*-adenosylmethionine (SAM) and helix−hairpin−helix (HHH) domains contained within this protein. Using SAM and [4Fe-4S] cluster cofactors, radical SAM domains catalyze a large variety of biochemical reactions,[35] including the spore product (SP) lyase repair of a type of UV-induced thymine dimer.[35,36] As shown in the recently determined structure of *Geobacillus thermodenitrificans* SP lyase,[36] this enzyme contains both a radical SAM domain and a C-terminal β-hairpin, the latter of which is thought to be important for SP lesion recognition and insertion into the enzymatic active site. ELI_05380 resembles SP lyase in that it is predicted to have both a radical SAM domain and a C-terminal HHH domain, a conserved non-sequence-specific DNA-binding domain.[37] This similarity in domain organization implies that ELI_05380 might also have a role in the repair of UV-induced DNA lesions such as SP. Further support for this assertion is provided by the presence of ELI_05385 adjacent to the putative pyrimidine lyase encoded by ELI_05380. This gene contains a uracil-DNA glycosylase domain, which initiates the removal of uracil from DNA.[38] Uracil in DNA often results from cytosine deamination, which is accelerated in UV-induced cytosine photodimers[39] and can lead to C → T transitions if left unrepaired. The pairing of these two types of proteins is quite common, having been observed in ∼20% of bacterial genomes as adjacent genes,[40] supporting the possible use in DNA repair.

At this time, EL222 homologues have been found in the genomes of eight α-proteobacteria. Intriguingly, in five of these strains (*Novosphingobium aromaticivorans* DSM12444, *Sphingopyxis alaskensis* RB2256, *Sphingomonas* sp. KC8, *Sphingomonas elodea* ATCC31461, and *Novosphingobium nitrogenifigens* DSM19370), these homologues are located adjacent to genes encoding DNA photolyases and enzymes needed for synthesis of the pterin cofactors,[41] GTP cyclohydrolases. Another homologue with an inverted domain composition (HTH−LOV, instead of LOV−HTH) can be found in the ε-proteobacterium *Sulfuromonas denitrificans* DSM1251 (previously *Thiomicrospira denitrificans*) and is also adjacent to a GTP cyclohydrolase. The clustering of these genes among among several strains suggests that LOV−HTH light-dependent regulation of DNA repair enzymes might be a conserved theme.

While the natural targets of EL222 regulation remain open, we note the ability to control DNA binding with light provides tremendous opportunities for engineering EL222 to flexibly control transcription and other DNA-modifying activities with unparalleled spatial and temporal resolution. A particular advantage of EL222 is that it directly regulates DNA binding with light, rather than relying on light-dependent activation of protein dimerization of separate DNA-binding and transactivation components[42,43] or second-messenger systems,[44] allowing its use as a single component in heterologous systems. Notably, preliminary data indicate that the DNA binding sites discovered here are sufficient to work in a variety of different eukaryotic environments (L. B. Motta-Mena and K. H. Gardner, unpublished results), permitting discoveries in areas far from their native marine bacterial environments.

■ ASSOCIATED CONTENT

Supporting Information

EL222 binding to a short 24 bp probe from ChIP−Seq peak B (Figure S1), EL222 binding to fragments of ChIP−Seq peaks A and D (Figure S2), differential binding affinity of EL222 for

different SELEX-derived sequences (Figure S3), mutational analyses of binding of EL222 to SELEX Clone-1 (Figure S4), verification of binding of EL222 to sequences within ChIP−Seq peaks E−K, statistics from high-throughput sequencing for ChIP−Seq (Table S1), a list of SELEX-derived EL222 binding sites (Table S2), and a list of *E. litoralis* genes near EL222 binding sites (Table S3). This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*Phone: (214) 645-6365. E-mail: kevin.gardner@utsouthwestern.edu.

Author Contributions

G.R.-C. and L.B.M.-M. contributed equally to this work.

Funding

We gratefully acknowledge funding from the National Institutes of Health (Grant R01 GM081875 to K.H.G.) and the Robert A. Welch Foundation (Grant I-1424 to K.H.G.) in support of this research.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We thank Fernando Correa and Victor Ocasio for their comments on the manuscript. The pBlueSkript+ vector was a kind gift from Dr. Michael Dellinger (University of Texas Southwestern Medical Center). *E. litoralis* HTCC2594 genome sequence data were provided in advance of publication 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.

ABBREVIATIONS

EMSA, electromobility shift assay; LOV, light−oxygen−voltage; PAS, Per-ARNT-Sim; HTH, helix−turn−helix; ChIP−Seq, chromatin immunoprecipitation−high-throughput sequencing; SELEX, systematic evolution of ligands by exponential enrichment; MEME, Multiple Em for Motif Elicitation.

REFERENCES

(1) Moglich, A., Ayers, R. A., and Moffat, K. (2009) Structure and signaling mechanism of Per-ARNT-Sim domains. *Structure 17*, 1282−1294.

(2) Huala, E., Oeller, P. W., Liscum, E., Han, I. S., Larsen, E., and Briggs, W. R. (1997) *Arabidopsis* NPH1: A protein kinase with a putative redox-sensing domain. *Science 278*, 2120−2123.

(3) Harper, S. M., Neil, L. C., and Gardner, K. H. (2003) Structural basis of a phototropin light switch. *Science 301*, 1541−1544.

(4) Harper, S. M., Christie, J. M., and Gardner, K. H. (2004) Disruption of the LOV-Jα helix interaction activates phototropin kinase activity. *Biochemistry 43*, 16184−16192.

(5) Krauss, U., Minh, B. Q., Losi, A., Gartner, W., Eggert, T., von Haeseler, A., and Jaeger, K. E. (2009) Distribution and phylogeny of light-oxygen-voltage-blue-light-signaling proteins in the three kingdoms of life. *J. Bacteriol. 191*, 7234−7242.

(6) Strickland, D., Moffat, K., and Sosnick, T. R. (2008) Light-activated DNA binding in a designed allosteric protein. *Proc. Natl. Acad. Sci. U.S.A. 105*, 10709−10714.

(7) Wu, Y. I., Frey, D., Lungu, O. I., Jaehrig, A., Schlichting, I., Kuhlman, B., and Hahn, K. M. (2009) A genetically encoded photoactivatable Rac controls the motility of living cells. *Nature 461*, 104−108.

(8) Strickland, D., Lin, Y., Wagner, E., Hope, C. M., Zayner, J., Antoniou, C., Sosnick, T. R., Weiss, E. L., and Glotzer, M. (2012) TULIPs: Tunable, light-controlled interacting protein tags for cell biology. *Nat. Methods 9*, 379−384.

(9) Gomelsky, M., and Hoff, W. D. (2011) Light helps bacteria make important lifestyle decisions. *Trends Microbiol. 19*, 441−448.

(10) Nash, A. I., McNulty, R., Shillito, M. E., Swartz, T. E., Bogomolni, R. A., Luecke, H., and Gardner, K. H. (2011) Structural basis of photosensitivity in a bacterial LOV-HTH DNA binding protein. *Proc. Natl. Acad. Sci. U.S.A. 108*, 9449−9454.

(11) Ulrich, L. E., Koonin, E. V., and Zhulin, I. B. (2005) One-component systems dominate signal transduction in prokaryotes. *Trends Microbiol. 13*, 52−56.

(12) Harrison, S. C., and Aggarwal, A. K. (1990) DNA recognition by proteins with the helix-turn-helix motif. *Annu. Rev. Biochem. 59*, 933−969.

(13) Eiting, M., Hageluken, G., Schubert, W. D., and Heinz, D. W. (2005) The mutation G145S in PrfA, a key virulence regulator of *Listeria monocytogenes*, increases DNA-binding affinity by stabilizing the HTH motif. *Mol. Microbiol. 56*, 433−446.

(14) Lee, D. H., Jeong, H. S., Jeong, H. G., Kim, K. M., Kim, H., and Choi, S. H. (2008) A consensus sequence for binding of SmcR, a *Vibrio vulnificus* LuxR homologue, and genome-wide identification of the SmcR regulon. *J. Biol. Chem. 283*, 23610−23618.

(15) Qin, N., Callahan, S. M., Dunlap, P. V., and Stevens, A. M. (2007) Analysis of LuxR regulon gene expression during quorum sensing in *Vibrio fischeri*. *J. Bacteriol. 189*, 4127−4134.

(16) Zhu, J., and Winans, S. C. (2001) The quorum-sensing transcriptional regulator TraR requires its cognate signaling ligand for protein folding, protease resistance, and dimerization. *Proc. Natl. Acad. Sci. U.S.A. 98*, 1507−1512.

(17) Barski, A., Cuddapah, S., Cui, K., Roh, T. Y., Schones, D. E., Wang, Z., Wei, G., Chepelev, I., and Zhao, K. (2007) High-resolution profiling of histone methylations in the human genome. *Cell 129*, 823−837.

(18) Johnson, D. S., Mortazavi, A., Myers, R. M., and Wold, B. (2007) Genome-wide mapping of in vivo protein-DNA interactions. *Science 316*, 1497−1502.

(19) Tuerk, C., and Gold, L. (1990) Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. *Science 249*, 505−510.

(20) Oh, H. M., Giovannoni, S. J., Ferriera, S., Johnson, J., and Cho, J. C. (2009) Complete genome sequence of *Erythrobacter litoralis* HTCC2594. *J. Bacteriol. 191*, 2419−2420.

(21) Li, H., and Durbin, R. (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics 25*, 1754−1760.

(22) Nicol, J. W., Helt, G. A., Blanchard, S. G., Jr., Raja, A., and Loraine, A. E. (2009) The Integrated Genome Browser: free software for distribution and exploration of genome-scale datasets. *Bioinformatics 25*, 2730−2731.

(23) Schmittgen, T. D., and Livak, K. J. (2008) Analyzing real-time PCR data by the comparative C(T) method. *Nat. Protoc. 3*, 1101−1108.

(24) Zoltowski, B. D., Nash, A. I., and Gardner, K. H. (2011) Variations in protein-flavin hydrogen bonding in a light, oxygen, voltage domain produce non-Arrhenius kinetics of adduct decay. *Biochemistry 50*, 8771−8779.

(25) Bailey, T. L., Boden, M., Buske, F. A., Frith, M., Grant, C. E., Clementi, L., Ren, J., Li, W. W., and Noble, W. S. (2009) MEME SUITE: Tools for motif discovery and searching. *Nucleic Acids Res. 37*, W202−W208.

(26) Wisedchaisri, G., Wu, M., Rice, A. E., Roberts, D. M., Sherman, D. R., and Hol, W. G. (2005) Structures of *Mycobacterium tuberculosis* DosR and DosR-DNA complex involved in gene activation during adaptation to hypoxic latency. *J. Mol. Biol. 354*, 630−641.

(27) Maris, A. E., Sawaya, M. R., Kaczor-Grzeskowiak, M., Jarvis, M. R., Bearson, S. M., Kopka, M. L., Schroder, I., Gunsalus, R. P., and Dickerson, R. E. (2002) Dimerization allows DNA target site recognition by the NarL response regulator. *Nat. Struct. Biol. 9*, 771−778.
(28) Zhang, R. G., Pappas, T., Brace, J. L., Miller, P. C., Oulmassov, T., Molyneaux, J. M., Anderson, J. C., Bashkin, J. K., Winans, S. C., and Joachimiak, A. (2002) Structure of a bacterial quorum-sensing transcription factor complexed with pheromone and DNA. *Nature 417*, 971−974.
(29) Vannini, A., Volpari, C., Gargioli, C., Muraglia, E., Cortese, R., De Francesco, R., Neddermann, P., and Marco, S. D. (2002) The crystal structure of the quorum sensing protein TraR bound to its autoinducer and target DNA. *EMBO J. 21*, 4393−4401.
(30) White, C. E., and Winans, S. C. (2007) The quorum-sensing transcription factor TraR decodes its DNA binding site by direct contacts with DNA bases and by detection of DNA flexibility. *Mol. Microbiol. 64*, 245−256.
(31) Mouw, K. W., Rowland, S. J., Gajjar, M. M., Boocock, M. R., Stark, W. M., and Rice, P. A. (2008) Architecture of a serine recombinase-DNA regulatory complex. *Mol. Cell 30*, 145−155.
(32) Baikalov, I., Schroder, I., Kaczor-Grzeskowiak, M., Grzeskowiak, K., Gunsalus, R. P., and Dickerson, R. E. (1996) Structure of the *Escherichia coli* response regulator NarL. *Biochemistry 35*, 11053−11061.
(33) Swartz, T. E., Tseng, T.-S., Frederickson, M. A., Paris, G., Comerci, D. J., Rajashekara, G., Kim, J.-G., Mudgett, M. B., Splitter, G. A., Ugalde, R. A., Goldbaum, F. A., Briggs, W. R., and Bogomolni, R. A. (2007) Blue-light-activated histidine kinases: Two-component sensors in bacteria. *Science 317*, 1090−1093.
(34) Purcell, E. B., Siegal-Gaskins, D., Rawling, D. C., Fiebig, A., and Crosson, S. (2007) A photosensory two-component system regulates bacterial cell attachment. *Proc. Natl. Acad. Sci. U.S.A. 104*, 18241−18246.
(35) Frey, P. A., Hegeman, A. D., and Ruzicka, F. J. (2008) The Radical SAM Superfamily. *Crit. Rev. Biochem. Mol. Biol. 43*, 63−88.
(36) Benjdia, A., Heil, K., Barends, T. R., Carell, T., and Schlichting, I. (2012) Structural insights into recognition and repair of UV-DNA damage by spore photoproduct lyase, a radical SAM enzyme. *Nucleic Acids Res. 40*, 9308−9318.
(37) Doherty, A. J., Serpell, L. C., and Ponting, C. P. (1996) The helix-hairpin-helix DNA-binding motif: A structural basis for non-sequence-specific recognition of DNA. *Nucleic Acids Res. 24*, 2488−2497.
(38) Pearl, L. H. (2000) Structure and function in the uracil-DNA glycosylase superfamily. *Mutat. Res. 460*, 165−181.
(39) Barak, Y., Cohen-Fix, O., and Livneh, Z. (1995) Deamination of cytosine-containing pyrimidine photodimers in UV-irradiated DNA. Significance for UV light mutagenesis. *J. Biol. Chem. 270*, 24174−24179.
(40) Marchler-Bauer, A., Lu, S., Anderson, J. B., Chitsaz, F., Derbyshire, M. K., DeWeese-Scott, C., Fong, J. H., Geer, L. Y., Geer, R. C., Gonzales, N. R., Gwadz, M., Hurwitz, D. I., Jackson, J. D., Ke, Z., Lanczycki, C. J., Lu, F., Marchler, G. H., Mullokandov, M., Omelchenko, M. V., Robertson, C. L., Song, J. S., Thanki, N., Yamashita, R. A., Zhang, D., Zhang, N., Zheng, C., and Bryant, S. H. (2011) CDD: A Conserved Domain Database for the functional annotation of proteins. *Nucleic Acids Res. 39*, D225−D229.
(41) Sancar, A. (2003) Structure and function of DNA photolyase and cryptochrome blue-light photoreceptors. *Chem. Rev. 103*, 2203−2237.
(42) Wang, X., Chen, X., and Yang, Y. (2012) Spatiotemporal control of gene expression by a light-switchable transgene system. *Nat. Methods 9*, 266−269.
(43) Shimizu-Sato, S., Huq, E., Tepperman, J. M., and Quail, P. H. (2002) A light-switchable gene promoter system. *Nat. Biotechnol. 20*, 1041−1044.
(44) Ye, H., Daoud-El Baba, M., Peng, R. W., and Fussenegger, M. (2011) A synthetic optogenetic transcription device enhances blood-glucose homeostasis in mice. *Science 332*, 1565−1568.