STRUCTURAL BASIS FOR SIGNAL TRANSDUCTION IN LOV BLUE LIGHT PHOTORECEPTORS

APPROVED BY SUPERVISORY COMMITTEE

Kevin H. Gardner, Ph.D.

Margaret A. Phillips, Ph.D.

Michael K. Rosen, Ph.D.

Philip J. Thomas, Ph.D.

DEDICATION

I would like to thank Kevin Gardner, who provided excellent mentorship, guidance, and understanding throughout the years. His willingness to take on new tasks and experiment with new techniques is something I hope to maintain in my own career. Thank you members of my committee, Dr. Margaret Phillips, Dr. Michael Rosen and Dr. Philip Thomas, for much help and guidance throughout this process. I'd also like to thank all the members of the Gardner lab that helped make my experience fantastic. This thesis reflects not only my work but also the insights, suggestions and interactions I had with everyone during my time in the lab. This includes Shannon Harper who took me under her wing as a rotation student and taught me all about LOV. Carlos Amezcua has not only maintained the NMR facility but offered support and friendship for which I extend my thanks. Charles Dann III and Cecelia Sieling were fantastic lab mates that provided invaluable advice and friendship. I miss you both! Finally, I want to thank my family and friends, without whom, I never would have made it this far.

STRUCTURAL BASIS FOR SIGNAL TRANSDUCTION IN LOV BLUE LIGHT PHOTORECEPTORS

by

ABIGAIL I. NASH

DISSERTATION

Presented to the Faculty of the Graduate School of Biomedical Sciences

The University of Texas Southwestern Medical Center at Dallas

In Partial Fulfillment of the Requirements

For the Degree of

DOCTOR OF PHILOSOPHY

The University of Texas Southwestern Medical Center at Dallas

Dallas, Texas

June, 2011

STRUCTURAL BASIS FOR SIGNAL TRANSDUCTION IN LOV BLUE LIGHT PHOTORECEPTORS

ABIGAIL I. NASH

The University of Texas Southwestern Medical Center at Dallas, 2011

KEVIN H. GARDNER, Ph.D.

My graduate research focused on studying the mechanisms of signal transduction within Light-Oxygen-Voltage domains, a subset of the PAS domain family. The first of two projects addressed intradomain signaling from the hydrophobic core to the domain surface. In this study, we addressed the role of a specific conserved residue in transmitting activating signal from the domain core to surface structural elements. Through biophysical and biochemical studies of LOV proteins containing point mutations at key residues, we determined that structural strain placed on the domain following light-induced covalent adduct formation regulates both structural based signal transduction as well as dark state recovery kinetics. In the second project, I characterized a novel LOV

containing protein comprised of an N-terminal LOV domain and a C-terminal DNA binding helix-turn-helix (HTH) domain. Following initial characterization of this protein, I was able to determine how light-induced covalent adduct formation in the N-terminus leads to interdomain separation through release of inhibitory contacts with the HTH domain, allowing for DNA binding. Comparisons of this protein with other known HTH proteins highlight the conserved signal transduction pathways of both the LOV domain and the HTH domain.

TABLE OF CONTENTS

CHAPTER 1. Cellular Signaling and the Role of PAS Domains 15

PAS DOMAINS 17

PAS Domain Structure 18

Mechanism of PAS Domain Signaling 20

LOV DOMAINS 21

LOV Domain Photochemistry 22

LOV Domain Signal Transmission 24

KEY QUESTIONS IN THE FIELD 25

CHAPTER 2. Investigating Intradomain Signal Transduction at the Amino Acid Level 27

LOV DOMAINS 29

Photochemistry of the LOV Domain 30

Phototropin-based Studies of LOV Domain Signaling 31

MATERIALS AND METHODS 34

EFFECTS OF Q513 MUTATIONS ON FMN SPECTRAL PROPERTIES 39

STRUCTURAL EFFECTS OF Q513 MUTATIONS 42

UNDERSTANDING THE MECHANISM OF INTRADOMAIN COMMUNICATION 52

Literature Review of the Role of Glutamine 513 52

Structural Effects of Q513L and Q513N Point Mutations 52

Dark Recovery Kinetic Effects of Q513L and Q513N Point Mutations 54

Models of Light Induced Movement in Q513 56

Role of Q513 in Full-Length LOV-Containing Proteins 56

CHAPTER 3. Introduction to DNA Binding Proteins 62
DNA BINDING PROTEIN STRUCTURAL VARIABILITY 62
REGULATION OF DNA BINDING ABILITY 65
CHAPTER 4. Investigating Interdomain Signal Transduction in a Novel LOV-HTH Protein 70
A NOVEL LOV-CONTAINING PROTEIN 71
LOV Domains in Environmental Signaling 71
LOV-containing Transcription Factors 72
HTH-Type DNA Binding Proteins 72
EL222 as a Model for LOV and HTH Regulation 73
MATERIALS AND METHODS 74
SPECTROSCOPIC ANALYSIS OF EL222 81
LOW RESOLUTION STUDIES OF LIGHT-INDUCED CONFORMATIONAL CHANGES 83
CRYSTAL STRUCTURE OF DARK STATE EL222 88
SOLUTION STRUCTURAL STUDIES OF EL222 89
TESTING THE ABILITY OF EL222 TO BIND DNA 95
DISCUSSION: EL222 IS A LIGHT-REGULATED DNA BINDING PROTEIN 100
Evidence for Light-Induced Conformational Changes 100
Model for EL222 Activation 102
Generality of LOV Domain Signaling 105
Interdomain Signal Propagation Pathways 105
Role of EL222 in *Erythrobacter litoralis HTCC2594* 107

CHAPTER 5. Conclusions 109
WHAT WE ALREADY KNEW 109
LOV Domains 109
HTH Domains 110
IMPACT OF THIS WORK 112
LOV Domain Intradomain Signaling 112
LOV Domain Interdomain Signaling 114
HTH Interdomain Signaling 114

PRIOR PUBLICATIONS

Ko, W.-H., Nash, A. I., Gardner, K. H. (2007) A LOVely view of blue light photosensing. *Nat. Chem. Biol.* 3, 372-374.

Nash, A. I., Ko, W.-H., Harper, S. M., Gardner, K. H. (2008) A conserved glutamine plays a central role in LOV domain signal transmission and duration. *Biochemistry* 47, 13842-9.

LIST OF FIGURES

CHAPTER 1

FIGURE 1-1 Basic mechanism of cell signaling cascade..... 16

FIGURE 1-2 Overall PAS domain fold and associated ligands..... 19

FIGURE 1-3 Additional structural elements are bound N- and C-terminal to the PAS core and play a role in sensing and signaling 20

FIGURE 1-4 LOV domain photochemistry..... 23

FIGURE 1-5 LOV domains are coupled to diverse signal output domains 25

CHAPTER 2

FIGURE 2-1 Proposed signal transduction pathway in the *As*LOV2 domain..... 31

FIGURE 2-2 UV-visible absorbance spectra of wildtype, Q513L and Q513N mutants 40

FIGURE 2-3 Q513 mutants demonstrate typical LOV domain spectra 41

FIGURE 2-4 Structural effects of Q513 mutations in the dark and lit states monitored by circular dichroism..... 44

FIGURE 2-5 Structural effects of Q513 mutations in the dark and lit states monitored by limited proteolysis..... 46

FIGURE 2-6 Overlay of dark and lit state $^{15}N/^{1}H$ HSQC correlation spectra for wildtype *As*LOV2 domain 47

FIGURE 2-7 $^{15}N/^{1}H$ HSQC spectra of the tryptophan $H_{\varepsilon}1$-$N_{\varepsilon}1$ region 49

FIGURE 2-8 Overlay of dark and lit state $^{15}N/^{1}H$ HSQC correlation spectra for Q514L and Q513N 51

FIGURE 2-9 Sequence alignment of LOV domains..... 60-61

CHAPTER 3

FIGURE 3-1 Single helix-turn-helix structure and function 64

FIGURE 3-2 Representative HTH-containing proteins 64

FIGURE 3-3 Model for activity in two-component response regulator 66

CHAPTER 4

FIGURE 4-1 UV-visible absorbance spectroscopy and Arrhenius analysis of EL222 82

FIGURE 4-2 Dark vs. light state circular dichroism spectroscopy 84

FIGURE 4-3 Limited proteolysis in both the dark and lit states 86

FIGURE 4-4 Studies of C75M mutant to mimic lit state of protein 87

FIGURE 4-5 Crystal and solution structures of dark state EL222 89

FIGURE 4-6 Solution NMR secondary structure predictions 90

FIGURE 4-7 Solution NMR HSQCs of dark versus lit state EL222 91

FIGURE 4-8 Light induced changes monitored at the amino acid level 94

FIGURE 4-9 EL222 binds DNA cooperatively under lit state conditions 97

FIGURE 4-10 L120K mutation successfully disrupts the interdomain interaction . 100

CHAPTER 5

FIGURE 5-1 Interdomain interfaces of LuxR-type response regulators 112

LIST OF TABLES

TABLE 2-1 Comparisons of the kinetic time constants (τ) and the absorption coefficients (ε) at the isosbestic points for wildtype and Q513 mutanats in the dark state 42

LIST OF ABBREVIATIONS

PAS - Period-ARNT-Single Minded (Per-ARNT-Sim)

LOV - Light-Oxygen-Voltage

FMN - flavin mononucleotide

FAD - flavin adenine dinucleotide

*As*LOV2 - *Avena sativa* phototropin 1 LOV2 domain

VVD - *Neurospora crassa* Vivid protein

WC-1 - *Neurospora crassa* white collar-1 protein

HTH - helix-turn-helix

wHTH - winged helix-turn-helix

PER - period protein

ARNT - aryl hydrocarbon receptor nuclear translocator

Sim - single-minded protein

HIF-α - hypoxia inducible factor protein

AHR - aryl hydrocarbon receptor protein

CHAPTER 1

CELLULAR SIGNALING AND THE ROLE OF PAS DOMAINS

Cells must regulate their functions in response to both the intracellular and extracellular environments. For example, single-celled organisms must be able to detect the presence of nearby nutrient sources and regulate their metabolic processes accordingly. Within multicellular organisms, cells must be aware of their neighbors and/or chemical modifiers to regulate the processes of growth and development in a uniform fashion. This is accomplished using signal transduction pathways – combinations of proteins designed to interact with each other to communicate a signal and cause some type of response, whether it be behavioral or biochemical. Disruption or deregulation of signal transduction can lead to cell death or disease.

Signaling cascades frequently begin with a protein that is able to directly respond to a stimulus. This protein, called a receptor, is often, but not always, located at the cellular membrane. Upon activation, the receptor undergoes a structural alteration that leads to changes in its activity and/or modified interactions between the receptor and downstream interaction partner proteins (*13*)(Fig.1-1). There are two principal ways a receptor senses the presence of its activating signal: small molecule/ligand binding and post-translational modification. The first of these, small molecule binding, is exemplified in bacterial quorum sensing. One class of quorum sensor proteins, exemplified by

TraR, binds N-acyl homoserine lactone when a sufficient bacterial population has been established. Upon binding this small molecule, the receptors undergo a structural reorganization that allows for DNA binding and transcriptional regulation (*14-16*). Post-translational modification, such as phosphorylation, is a common activation mechanism utilized by bacterial two-component response regulators. Upon sensing an external stimulus, the histidine kinase component autophosphorylates on a histidine residue, then transfers this phosphate group to an aspartate residue on the response regulator (*17, 18*). Both small molecule binding and post-translational modification are essential elements of signal transduction cascades and can often be simultaneously employed to activate a wide range of responses to different elements present in an environment.

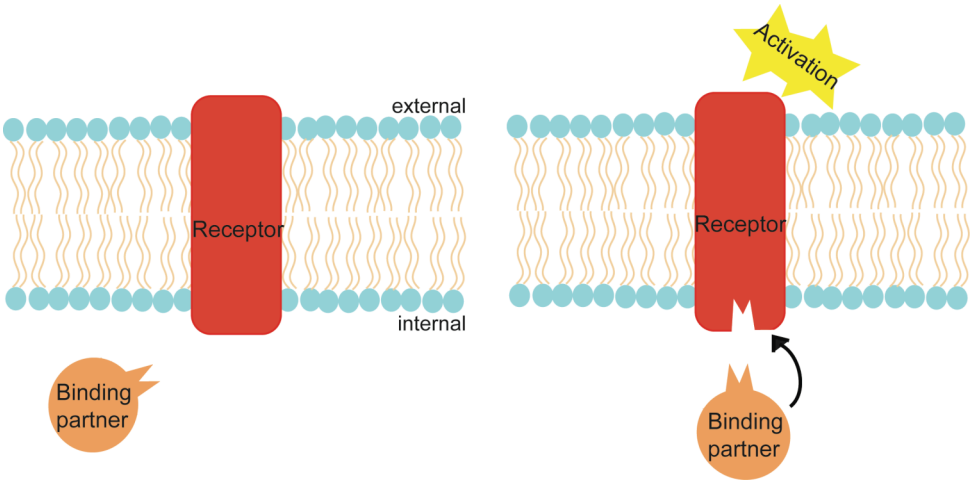

Fig. 1-1. Basic mechanism of cell signaling cascade
Receptor proteins in the cell membrane are able to sense environmental stimuli and trigger signaling cascades to modify the cell's behavior accordingly. Following activation by a stimulus, the receptor undergoes a conformational change that may modify its own activity and/or alter its ability to interact with other proteins.

Because an organism's environment is neither simple nor static, signal transduction pathways are often convoluted and multi-branched, allowing for a single protein to have varied effects on a wide range of cellular behaviors. This requires that certain elements of the signaling cascade interact with multiple proteins or ligands (*19*). To retain specificity among interaction partners, many signaling proteins are composed of distinct structural elements called domains. These domains may be characterized by a common fold or specific function. A protein may be composed of a single domain or multiple domains, each imparting different functional and chemical characteristics (*13*). Understanding how signal transduction proteins regulate their activity and specificity lies in understanding the multiple roles of the component domains from which they are constructed and how these modulate each other in the context of full-length proteins.

PAS Domains

One type of domain, commonly found in signal transduction pathways is the PAS domain. PAS stands for Per-ARNT-Sim, the three eukaryotic transcription factors where an imperfect repeat sequence was first recognized (*20, 21*). The *Drosophila* period protein (PER) is involved in circadian rhythms (*22*), aryl hydrocarbon receptor nuclear translocator protein (ARNT) is responsible for hypoxic and xenobiotic responses (*23, 24*) and *Drosophila* single-minded (Sim) is important for CNS development (*25*). Members of this family are small (~110 aa)

domains capable of sensing the environment through protein-protein interactions, and in some cases, protein-ligand relationships (*20*). They are capable of sensing diverse stimuli such as oxygen, ligands, light, and redox potential and are found in all kingdoms of life (*20*). Given the functionality of PAS domains, they are found in a wide variety of proteins and thus are integral components of numerous signal transduction pathways. They are involved with processes ranging from hypoxia response (*24*) to phototropism (*26*) to circadian clock dependent protein degradation (*27*).

PAS Domain Structure

Despite vastly different functions *in vivo* and low sequence similarity (<20%), PAS domains share a common mixed α/β fold comprised of three or four helical segments flanking one side of a five-stranded antiparallel β-sheet (*28*) (Fig. 1-2a). Additional structural elements located N- and C-terminal to this PAS core are present in some PAS domains and may play a role in regulating their function (Fig. 1-3a-b) (*7, 29-31*). Finally, in a number of PAS domains, the hydrophobic core serves as a ligand binding site as in FixL (heme) (*32, 33*), NifL (*9, 34*) and PYP (*28, 35*) (Fig. 1-2b-d).

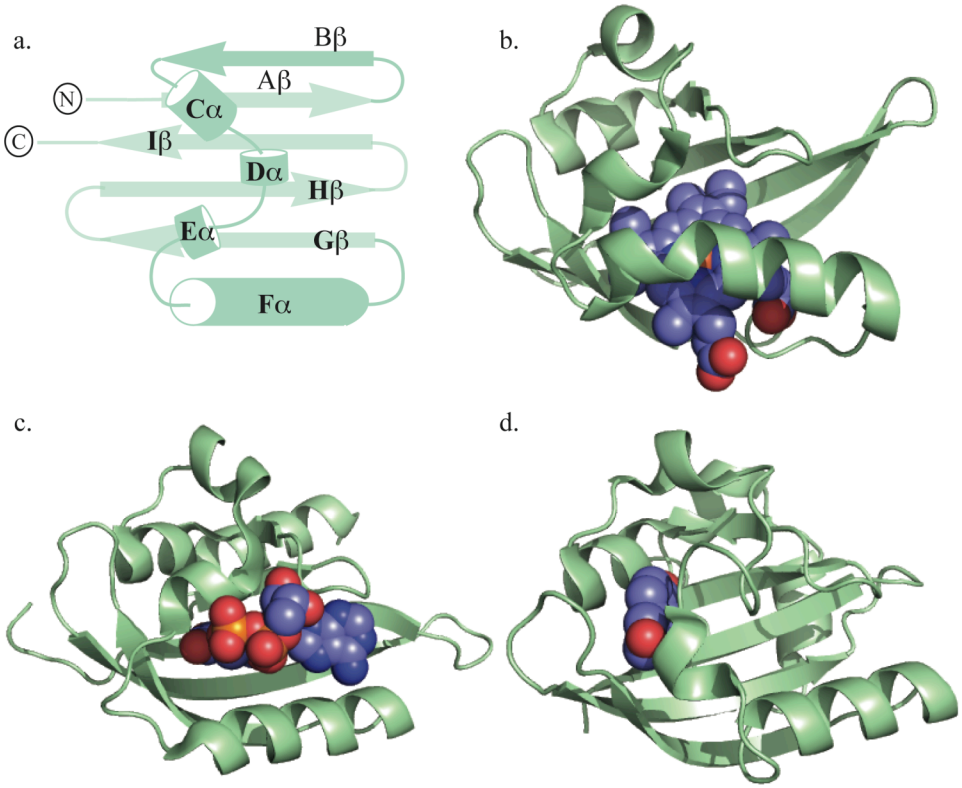

Fig. 1-2. Overall PAS domain fold and associated ligands
a.) PAS domains share a common mixed α/β fold consisting of three or four helical segments flanking one side of a five-stranded antiparallel β-sheet. b.-d.) Examples of PAS domains containing ligands in their hydrophobic cores b.) Bacterial oxygen sensor FixL contains a heme ligand PDB code: 1XJ4 (*5*). c.) Bacterial oxygen sensor NifL has an FAD ligand within its core PDB code: 2GJ3 (*9*).
d.) Bacterial blue light sensor PYP uses hydroxycinnamic acid in its core to sense light PDB code: 2ZOH (*12*).

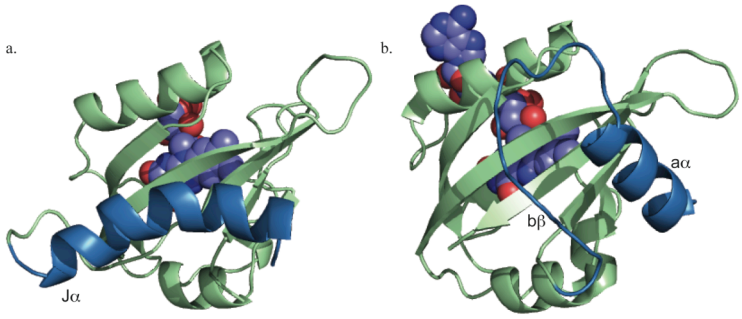

Fig. 1-3. Additional structural elements are bound N- and C-terminal to the PAS core and play a role in sensing and signaling
The PAS core of each domain is colored in green, while the additional structural elements beyond this core are in blue. a.) Phototropin 1 LOV2 from Avena sativa. The C-terminal Jα helix has been shown to dissociate from the PAS core and partially unfold following activation. b.) VIVID protein from Neurospora crassa PDB code: 2PD7 (*7*). The N-terminal aα and bβ units appear to repack against the surface of the protein following illumination, resulting in an overall increase in hydrodynamic radius.

Mechanism of PAS Domain Signaling

The mechanism through which PAS domains transduce their sensory input into changes in protein activity is not well characterized. Given the diversity among PAS domains in their ability to bind ligand as well as the structural changes known to occur post-activation, a single, universal mechanism for signal transduction has been difficult to identify. One functional element that appears to be common among the PAS domains is the ability to interact with a variety of protein partners. Unfortunately, it is difficult to identify potential binding partners due to the vast number of interactions characterized thus far. For example, the canonical PAS-containing protein ARNT is known to heterodimerize with several other PAS domain proteins including hypoxia inducible factor (HIF-α), aryl hydrocarbon receptor (AHR) and single-minded

protein (Sim) (*36-39*). Dimerization specifically through the PAS-B domain of ARNT has also been demonstrated (*40*), however, how PAS-B dimerization influences full-length protein dimerization in the context of the other two domains is not known.

What has become apparent is that PAS domains are capable of participating in a wide variety of protein-protein interactions, facilitating their presence among signal transduction proteins (*20*). Active research is being conducted into how these domains are able to achieve such flexibility in their binding partners while maintaining the specificity necessary to discriminate among signaling pathways. The following dissertation focuses on how a subset of these ubiquitous domains transduces an environmental signal into protein activation both through intradomain structural perturbation as well as interdomain communication.

LOV Domains

As described earlier, some PAS domains have associated ligands that impart the ability to sense different environmental factors. A specific subset of PAS proteins, the LOV (light-oxygen-voltage) domains (*41*), is capable of sensing light, oxygen or voltage. My work will focus on LOV domains that sense blue light in the environment and convert it into a biochemical signal in a wide variety of proteins.

LOV Domain Photochemistry

LOV domains contain a series of highly conserved residues (GXNCRFL) (*42*) surrounding an internally bound flavin mononucleotide (FMN) or flavin adenine dinucleotide (FAD) chromophore (Fig. 1-3a-b, 1-4) that converts blue light into protein structural changes. Spectroscopic studies on LOV-FMN and LOV-FAD complexes showed that blue light induces the formation of a covalent adduct between the isoalloxazine C4a position and a conserved cysteine residue within the LOV domain (Fig. 1-4b) (*43, 44*). The chemistry behind this covalent adduct formation remains a debated topic in the field. An ionic mechanism for adduct formation involving a flavin cation attacked by a thiolate anion at the conserved cysteine residue was originally proposed based on pH-dependent fluorescence experiments (*44*). Later, however, infrared spectroscopy data showed this cysteine to be protonated in the dark, ground state (*45, 46*). A second, radical-based mechanism, first proposed based on *ab initio* quantum mechanical calculations and model compounds (*47*), was later supported by EPR studies of a mutant LOV domain in which the conserved cysteine was substituted with a methionine residue (*48*). The covalent adduct formed by this mutant is between the methionine and the N(5) position of the flavin ring, not the C(4a) as is seen in wildtype LOV domains (*48*). Bond formation with the N(5) position suggests this atom is likely not involved with stabilizing an ion pair by accepting a proton, as proposed for the ionic mechanism. In addition to having an EPR spectrum consistent with a neutral radical species, the methionine possesses a

methyl group at the position of a proton in cysteine. This rules out the possibility of a two-step ionic mechanism since a methyl group is unlikely to be extracted from the residue to create a thiolate. Radical pair formation has also been identified directly by EPR spectroscopy in wildtype domains at low temperature (*49, 50*)(Fig. 1-4c).

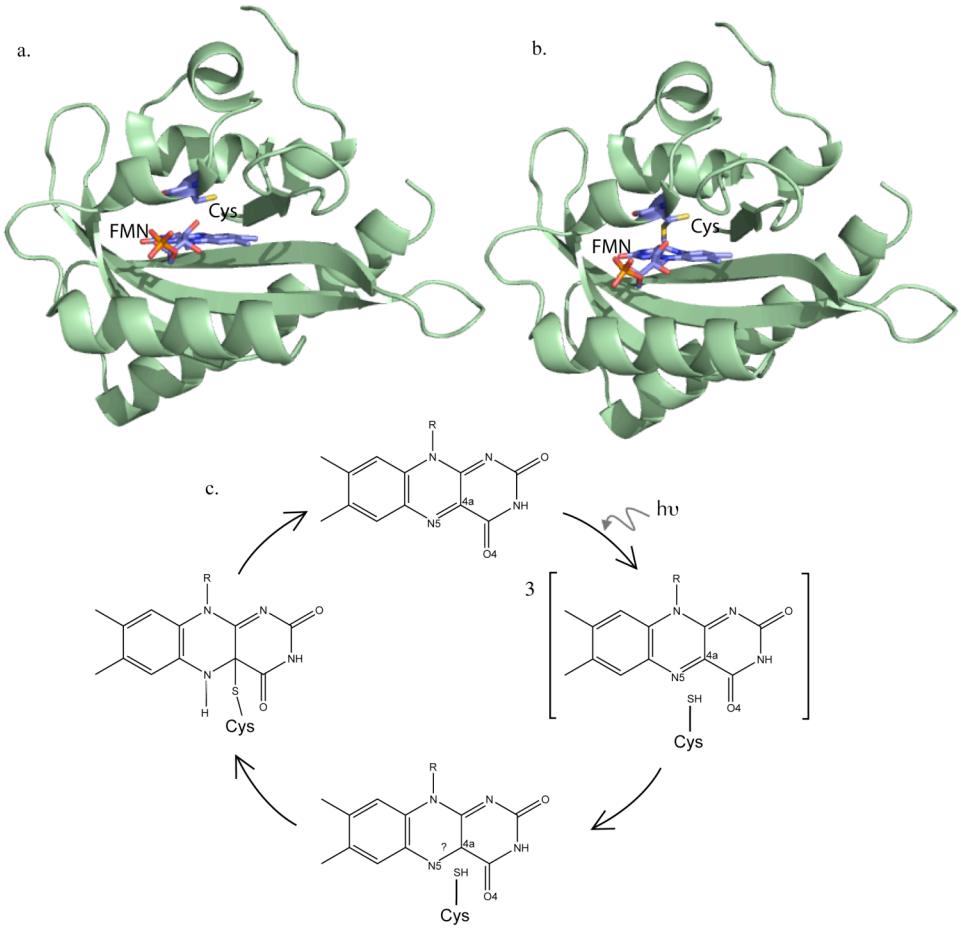

Fig. 1-4. LOV domain photochemistry
a.) LOV domains have an internally associated FMN or FAD molecule. b.) In response to exposure to blue light, an absolutely conserved cysteine residue forms a covalent adduct with the isoalloxazine ring of the flavin. c.) The photochemistry behind the covalent adduct formation and decay is still under debate. Following illumination, the flavin is excited to a triple state, then a reaction takes place to form the covalent adduct. The covalent adduct exists for a finite period of time ranging from seconds to days in vitro.

The stability of this photoadduct is variable among LOV domains; most spontaneously relax back to the noncovalent dark state within several seconds to many hours (*51, 52*), but several systems appear to be effectively irreversible *in vitro* on a biological time scale (*50, 53*). The mechanism of the back reaction remains unclear at this time, however there is evidence that many different factors can influence the rate of dark state recovery. These include the presence of basic compounds (*e.g.* imidazole) (*54*), alkaline pH (*55, 56*) and ionic strength (*55, 56*).

LOV Domain Signal Transmission

LOV domains are coupled to a wide variety of output domains, including serine/threonine kinases, STAS domains, GGDEF domains, and various DNA binding domains (*41*) (Fig. 1-5). Given this diversity, and the range of domains with which the LOV must interact to transduce light activation into protein activity, many questions exist as to how LOV domains communicate to their downstream effector domains or protein interaction partners. Like PAS domains in general, a common mechanism for signal transduction has been difficult to observe. Is there a common mechanism among all PAS domains or will LOV domains exhibit novel features due to their associated ligands and/or the nature of their sensory activity? Studies on isolated LOV domains have suggested that structural elements outside the canonical LOV core may be involved with signal transduction (*7, 29, 30*), however, not all LOV domains contain these additional elements and some conflicting data exists (*57*). There is also evidence of

dimerization or higher order multimerization in some isolated LOV domains (*9, 58-64*) though these may be artifactual due to truncation of secondary structural elements in the constructs used in these studies. These LOV-LOV interactions may be constitutive or light dependent, however, the role of oligomerization for signal communication is unknown. Clearly, more studies on multi-domain LOV-containing proteins and higher resolution structural information of proteins in both the dark and lit states are required to address the nature of LOV interactions.

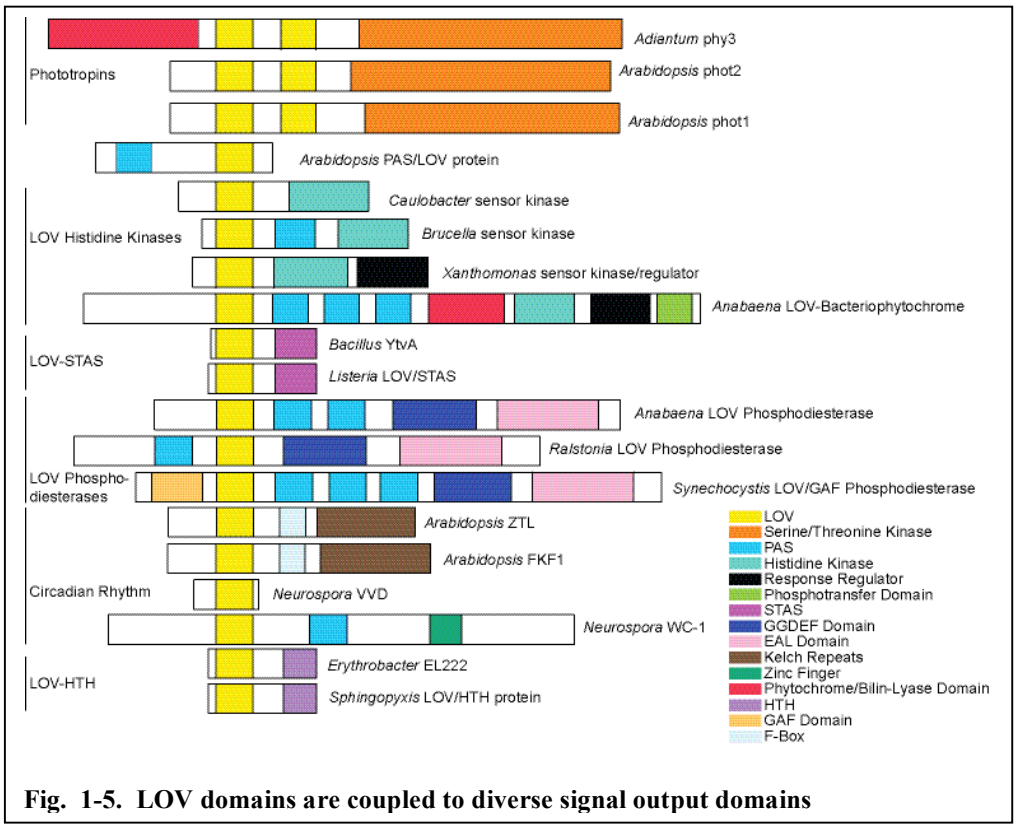

Fig. 1-5. LOV domains are coupled to diverse signal output domains

Key Questions in the Field

From my perspective, there are two key questions in the field of PAS/LOV

domains. The first of these is: How is signal-induced activation transduced throughout the domain? Activation of the domain mainly occurs either through post-translational modification at a specific residue or though ligand binding at a site on the domain. In either case, the residues directly involved in the sensory aspect of the domain function may not be the same residues required for signal transduction in the pathway, thus necessitating signal transfer from one area of the domain to another.

This brings us to a second major question in the field: How do PAS domains communicate signal-induced activation to downstream effector domains and/or binding partners? As stated earlier, PAS domains appear to interact with various binding partners. Following activation of the domain, signal must be transferred to either an effector domain, such as a kinase domain, or to another protein, as in the case of the single domain PAS protein Vivid (VVD) (*7*). How this signal is communicated between domains is unclear.

This thesis addresses both of these questions. Initially, I investigated the role of a specific amino acid in intradomain signal transmission following activation. In chapter 4, I will focus on interdomain communication in a protein whose two domains appear to significantly reorient with respect to each other following activation.

CHAPTER 2

INVESTIGATING INTRADOMAIN SIGNAL TRANSDUCTION AT THE AMINO ACID LEVEL

LOV domains are able to respond to illumination with blue light via an associated flavin moiety within their hydrophobic core. Interestingly, while all light-activated LOV domains form a covalent adduct between a conserved cysteine residue and the flavin isoalloxazine ring, we do not have a clear understanding of how this adduct formation leads to protein activation. Several lines of evidence have shown light induced structural changes in some LOV domains (*7, 29, 51, 65-69*) as well as light dependent activity for some LOV-containing proteins (*70, 71*). Somehow, adduct formation at the flavin must be transduced to the surface of the protein to account for these findings. We investigated the role of a specific and conserved glutamine residue in the Iβ strand that hydrogen bonds to the flavin ring in the dark state. Conflicting data exists regarding whether this amino acid side chain rotates following illumination to form a new hydrogen bond at another point in the flavin ring, but it is well accepted that the dark state hydrogen bond is lost. The following is excerpted from our published manuscript (*72*) describing our methods and findings.

A Conserved Glutamine Plays a Central Role in LOV Domain Signal Transmission and Its Duration

Abigail I. Nash[‡], Wen-Huang Ko[‡], Shannon M. Harper and Kevin H. Gardner

Biochemistry, **2008**, 47: 13842-13849

[‡]These authors contributed equally to the work

Abstract

Light is a key stimulus for plant biological functions, several of which are controlled by light-activated kinases known as phototropins, a group of kinases that contain two light-sensing domains (LOV, Light-Oxygen-Voltage domains) and a C-terminal serine/threonine kinase domain. The second sensory domain, LOV2, plays a key role in regulating kinase enzymatic activity via the photochemical formation of a covalent adduct between a LOV2 cysteine residue and an internally-bound flavin mononucleotide (FMN) chromophore. Subsequent conformational changes in LOV2 lead to the unfolding of a peripheral Jα helix, and ultimately, phototropin kinase activation. To date, the mechanism coupling bond formation and helix dissociation has remained unclear. Previous studies found that a conserved glutamine residue (Q513 in the *Avena sativa* phototropin 1 LOV2 (*As*LOV2) domain) switches its hydrogen-bonding pattern with FMN upon light stimulation. Located in the immediate vicinity of the FMN binding site, this Gln residue is provided by the Iβ strand that interacts with the Jα helix, suggesting a route for signal propagation from the core of the LOV domain to its

peripheral Jα helix. To test whether Q513 plays a key role in tuning the photochemical and transduction properties of *As*LOV2, we designed two point mutations, Q513L and Q513N, and monitored the effects on the chromophore and protein using a combination of UV-visible absorbance and circular dichroism spectroscopy, limited proteolysis, and solution NMR. The results show that these mutations significantly dampen the changes between the dark and lit state *As*LOV2 structures, leaving the protein in a pseudo-dark state (Q513L) or a pseudo-lit state (Q513N) conformation. Further, both mutations changed the photochemical properties of this receptor, particularly the lifetime of the photoexcited signaling states. Together, these data establish that this residue plays a central role in both spectral tuning and signal propagation from the core of the LOV domain through the Iβ strand to the peripheral Jα helix.

LOV Domains

Protein signaling cascades are central to organismal growth, adaptation, and communication; therefore, the regulation of these cascades is key to survival. PAS (Per-ARNT-Sim) domain-containing proteins are well characterized as vital members of many such regulatory paths, including adaptation to hypoxia (*73*), circadian rhythm-dependent gene transcription (*74*), and phototropism and chloroplast organization in plants (*75*). A specific subset of these types of PAS proteins, the LOV (light-oxygen-voltage) domains (*41*), is capable of sensing blue light as an environmental signal and converting it into a biochemical signal in a wide variety of proteins.

Photochemistry of the LOV Domain

LOV domains sense blue light through an internally bound flavin molecule (FMN or FAD). Following illumination, a covalent adduct between the isoalloxazine C4a position and an absolutely conserved cysteine residue is formed (Fig. 2-1a-b) (*43, 44*). This photoadduct is stable over a range of time scales from seconds to days before spontaneously relaxing back to the noncovalent state, depending on the specific LOV domain (*51, 53*). At the limiting case, some LOV domains generate adducts which are irreversible *in vitro* on a biological time scale (*50, 52*).

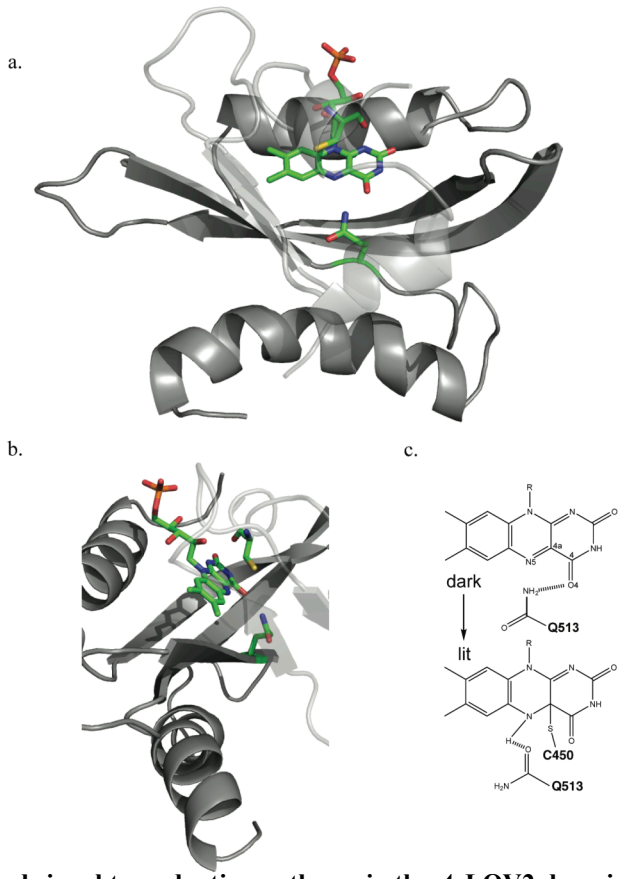

Fig. 2-1. Proposed signal transduction pathway in the *As*LOV2 domain.
Front (a) and side view (b) of AsLOV2 domain structure including Jα helix is shown in grey. FMN, Q513 and C450 residues are shown as stick figures with carbon (green), oxygen (red), nitrogen (blue), sulfur (yellow), and phosphorous (orange). (a) The hydrogen bond between Q513 side chain amide proton and FMN C4 carbonyl in the dark state are shown with a yellow dashed line. The side view (b) shows the relative orientations between C450, FMN, Q513, and the Jα helix. Bond formation between C450 and FMN leads to signal propagation through Q513 and ultimately to the dissociation of Jα helix from the Iβ strand. (c) Proposed side chain rotation and hydrogen bond switch by Q513 residue. Light-induced rotation of the Q513 side chain leads to breakage of a hydrogen bond between the Q513 amide and the C4 carbonyl of FMN and possibly formation of a new hydrogen bond between the Q513 carboxyl group and N5 of FMN.

Phototropin-based Studies of LOV Domain Signaling

The sensory role played by LOV domains is characterized in a variety of proteins, including transcription factors, ubiquitin ligases and kinases. Previous studies on phototropins, a group of plant photoreceptors that contain two LOV

domains and a C-terminal serine/threonine kinase, demonstrated that they form the expected covalent cysteinyl adducts and exhibit a corresponding robust increase in autophosphorylation activity upon illumination (*76*). While the role of the N-terminal LOV (LOV1) domain remains poorly understood, light-induced changes in the C-terminal LOV (LOV2) domain structure are both necessary and sufficient for kinase activation (*77*). Despite our knowledge of LOV2 in the context of the full-length protein, the molecular mechanism by which the blue light signal is communicated to the kinase domain remains unclear. Harper *et al.* (*29*) proposed a mechanism for signal transduction in the *Avena sativa* phototropin 1 LOV2 domain (*As*LOV2) that involves light-induced unfolding of a helix, termed Jα, that is C-terminal to the conserved LOV core domain. In the dark state, the Jα helix interacts with the β-sheet of the LOV domain, particularly the Gβ, Hβ and Iβ strands. Disruption of the Jα helix interaction with the Iβ strand by site-directed mutagenesis was sufficient to induce a pseudo-lit state structure of the LOV domain and constitutively activate kinase function in the absence of illumination (Fig. 2-1a-b) (*71*). Recent crystallographic data on *As*LOV2 containing the Jα helix also support a role for Jα helix in signal transduction (*67*). While these studies clearly implicate the C-terminal Jα helix in communicating photodetection events to a downstream effector domain, it remains unclear how covalent adduct formation in the core leads to α-helical unfolding on the surface of the domain.

Insight into this question was provided by X-ray crystallography and

molecular dynamics simulations that show a reorganization of the protein/FMN hydrogen-bonding network upon covalent adduct formation (*78, 79*). In the dark state, crystallography shows that the side chain amide of a conserved glutamine residue (Q1029) in the *A. capillus-veneris* neochrome1 LOV2 domain donates a hydrogen bond to the O4 atom of FMN (Fig. 2-1c) (*80*). Upon illumination, the glutamine side chain rotates, breaking this bond to O4 and instead allowing Q513 to accept a hydrogen bond from the newly-protonated N5 atom of FMN (*78, 79*). Additionally, FTIR studies demonstrated that mutation of Q1029 to leucine alters the electronic state of FMN and reduces the magnitude of light-induced protein structural changes (*81*). As a result, Nozaki *et al*. proposed that these changes are due to the absence of the glutamine carbonyl hydrogen bond to FMN in the lit state. Subsequent studies of the mutation at the corresponding glutamine in full-length *Arabidopsis thaliana* phototropin1 demonstrated attenuated autophosphorylation activity in the light versus wildtype protein (*82*). In addition, studies of the corresponding glutamine residue in the Vivid fungal photoreceptor also show a leucine mutant diminishes light-induced activity *in vivo* (*7*).

To better understand the role of this conserved glutamine residue in LOV domain signaling, we used a variety of biochemical and biophysical techniques to characterize how two mutations of this essential residue affect photochemistry and structural perturbations upon blue light illumination. Specifically, we made the corresponding glutamine to leucine (Q513L) mutation as well as a glutamine to asparagine (Q513N) mutation in the *As*LOV2 domain. These mutations

allowed us to probe how subtle perturbations of this side chain affect photochemistry and signal transmission. We observed significant changes in the electronic and structural properties of these mutants in comparison to wildtype *As*LOV2 using a combination of UV-visible spectroscopy, limited proteolysis, circular dichroism, and NMR spectroscopy. While both mutant domains maintained photocycling capabilities and demonstrated light-induced structural changes, they appeared to dampen the degree of light-induced structural change, leaving one mutant (Q513L) in a pseudo-dark state while the other (Q513N) is more pseudo-lit state compared to wildtype *As*LOV2. These data underscore the importance of hydrogen bond networks between FMN and the protein β-sheet in tuning properties of the chromophore and communicating light-induced structural changes throughout the domain.

Materials and Methods

Cloning, Expression, and Purification of AsLOV2

Plasmid DNA encoding the *As*LOV2 domain plus the Jα helix (residues 404-560 (*29*)was used to generate Q513N and Q513L mutants. Mutagenesis was carried out according to the Quick Change II site-directed mutagenesis kit (Stratagene) following manufacturer's instructions and verified by DNA sequencing. Proteins were expressed in *E. coli* BL21(DE3) cells grown in M9 minimal medium supplemented with $^{15}NH_4Cl$ (1 g/L) at 37^{o}C to an A_{600} of 0.6-0.8 and then induced with IPTG (0.12 g/L). After 16 hr induction at 20^{o}C, cells were

centrifuged and pellets resuspended in 50 mM Tris, 100 mM NaCl, pH 8 buffer. Cells were lysed using sonication and clarified with centrifugation at 10,000 g for 40 min. The soluble fraction was loaded onto a Ni^{+2}-NTA column, allowing for rapid affinity purification of His-Gβ1 tagged (*29*) LOV fusions by eluting with 250 mM imidazole. After exchanging the LOV-containing fractions into 50 mM Tris, 100 mM NaCl pH 8.0 buffer, the His-Gβ1 tag was cleaved by adding 1 mg His_6-TEV protease per 30 mg of fusion protein. Proteolysis reactions were allowed to proceed overnight at 4°C and stopped using a Ni^{+2}-NTA column to remove the His- Gβ1 and His_6-TEV protease. Post-cleavage, the resulting proteins contain only GEF (N-terminal) and G (C-terminal) residues as cloning artifacts.

Obtaining Absorption coefficients for wildtype AsLOV2, Q513N, and Q513L

To obtain absorption coefficients we employed trichloroacetic acid (TCA) precipitation of our proteins to isolate FMN (*26*). Following addition of 10% TCA, each protein sample was incubated in the dark at room temperature 5 min., then centrifuged at 20,000 x g, 4°C for 10 min. to clarify supernatant. An FMN standard curve was prepared using the A_{446} measurements of FMN at concentrations of 1 μM, 10 μM, 50 μM, 100 μM, and 250 μM in 50 mM sodium phosphate, 100 mM NaCl, pH 6.0 with 10% TCA. The UV-visible spectrum of each sample's supernatant was recorded and the concentrations of FMN determined using the standard curve. Assuming a 1:1 protein/FMN

stoichiometry, we assume the concentration of FMN is equal to total concentration of protein and calculate an absorption coefficient for each protein using the following formula:

$$\varepsilon_{446,TCA} = A_{446,TCA} / [C_{TCA} * 1 \text{ nm}]$$

Where ε is the absorption coefficient, A is the absorption, and C is the concentration of total protein.

To convert the $\varepsilon_{446,TCA}$ to the an ε_{446} under non-acidic buffer conditions, we measured the ratio between the A_{446} measured under non-acidic versus acidic conditions ($A_{446,non\text{-}acidic}/A_{446,TCA}$) and multiplied the $\varepsilon_{446,TCA}$ by this ratio. The resulting absorption coefficient (ε_{446}) for wildtype *As*LOV2, Q513N, and Q513L are reported in Table 2-1.

Protein-Flavin Stoichiometry Calculation

UV-visible absorbance spectra (from 250 nm to 550 nm) were recorded for all three freshly purified proteins following buffer exchange into 50 mM sodium phosphate, 100 mM NaCl, pH 6.0. During buffer exchange the flow through fraction was monitored by UV-visible spectroscopy for the presence of free FMN. Using the A_{280}/A_{446} ratio for wildtype (2.60) as a reference for 1:1 protein/FMN stoichiometry (*26*), this same ratio was calculated for each of the mutant domains (2.62 for Q513N and 2.76 for Q513L), revealing an approximately 1:1 protein/FMN stoichiometry for both Q513N and Q513L, suggesting the mutations do not significantly affect flavin incorporation.

UV-visible Absorbance Spectroscopy and Photocycle Kinetics

All proteins were concentrated to <70 μM in buffer containing 50 mM sodium phosphate, 100 mM NaCl, pH 6.0. UV-visible absorbance spectra were measured on a Varian Cary Series 50 spectrophotometer from 250-550 nm. Dark state spectra were obtained on samples exposed only to red light for the past 24 hr, while lit state spectra were obtained immediately after exposing sample to illumination from a photographic flash. Kinetic experiments monitored the return of the A_{446} signal following illumination. Data points were fitted using a first order rate equation to obtained the time constant (τ).

Limited Proteolysis

Proteins were buffer exchanged to 50 mM sodium phosphate, 100 mM NaCl, pH 7.5 buffer. A 1:90 ratio (w/w) of chymotrypsin to protein was used in a single volume with subsequent samples collected from this larger quantity. 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. Dark state experiments were done under dim red light while lit state experiments were performed under constant irradiation with 488 nm laser light at 50 mW power.

Circular Dichroism Spectroscopy

Proteins were buffer exchanged into buffer containing 50 mM sodium phosphate and 100 mM NaCl at pH 6.0. A total of 500 μl of 15 μM sample was used for each CD experiment. Dark state spectra were collected under dim red light; while lit state spectra were recorded following exposure to photographic flash every 10 s during the course of the experiment. CD data were collected using a wavelength range from 195 to 260 nm at 10ºC with 1.5 nm bandwidth and 3 s averaging time. Final data were generated from an average of 3 repeats.

Nuclear Magnetic Resonance spectroscopy

Proteins were concentrated to 1 mM in pH 6.0 buffer containing 50 mM sodium phosphate, 50 μM FMN, and 100 mM NaCl, with 10% (v/v) D_2O added to all samples prior to all NMR experiments. NMR experiments were performed on Varian Inova 500 and 600 MHz spectrometers at 25^{o}C, using nmrPipe (*83*) for data processing and NMRview (*84*) for analysis. Lit state HSQC spectra were acquired with a 488 nm 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 designed to hold external chemical shift standards inside a 5 mm NMR sample tube. This allowed 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. Each $^{15}N/^{1}H$ HSQC spectrum was recorded by preceding each transient in the experiment with a 50 mW 200 ms laser pulse during the 1.06 sec delay between transients (*29*).

Sequence Alignment

A multiple sequence alignment of LOV domains was generated using CLUSTAL W (*85*) and sequences were displayed using ESPript.cgi Version 3.06 (*86*).

Effects of Q513 mutations on FMN spectral properties

The electronic state of FMN within the LOV protein core is easily observed by UV-visible absorbance spectroscopy. Quantitative measurements of the flavin:protein stoichiometry indicated that both mutants retain FMN at approximately a 1:1 ratio with protein as previously described for wildtype *As*LOV2(*26*) (Fig. 2-2). As with wildtype, both the Q513L and Q513N mutants demonstrate typical LOV domain spectra with three characteristic absorbance peaks between 400 and 500 nm in the dark state and three isosbestic points (Fig. 2-3), with similar absorption coefficients to known flavoproteins at 446 nm (Table 2-1). Both mutants also display similar loss of this fine structure upon illumination with blue light, indicating the formation of covalent adduct. However, the Q513L absorbance profile is blue-shifted 9 nm in the dark state, as shown before (*81, 82*), indicating a change in the electronic environment surrounding the FMN. This contrasts with the Q513N absorbance spectra which

do not significantly deviate from wildtype spectra, indicating comparatively little change in the electronic environment surrounding FMN in the dark state. Similarly, we observed that Q513L caused greater changes in the locations of the three isosbestic points compared to Q513N (Table 2-1). Overall, these spectroscopic data indicate that the electronic environment of the FMN exhibits a larger change in the Q513L mutant than the Q513N mutant, most likely due to loss of hydrogen bond contacts between FMN and the altered side chain.

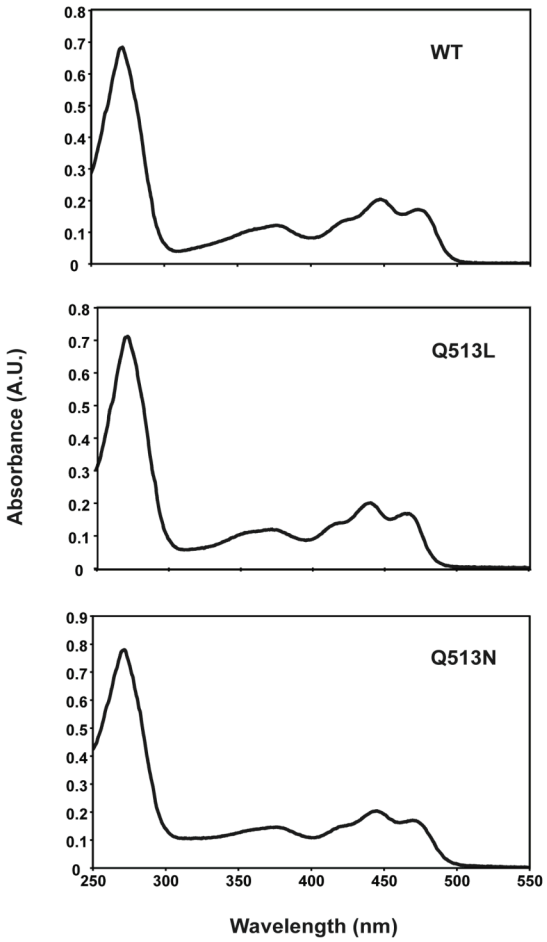

Fig. 2-2. UV-visible absorbance spectra of wildtype, Q513L and Q513N mutants UV-visible absorbance spectra shown here were recorded from 250nm to 550nm for each protein at concentrations between 50 μM and 70 μM in buffer containing 50 mM sodium phosphate (pH 6.0) and 100 mM NaCl. Calculated A_{280}/A_{446} ratios are 2.60 for wildtype, 2.63 for Q513N and 2.76 for Q513L (Q513L ratio taken at A_{438} due to blue shift)

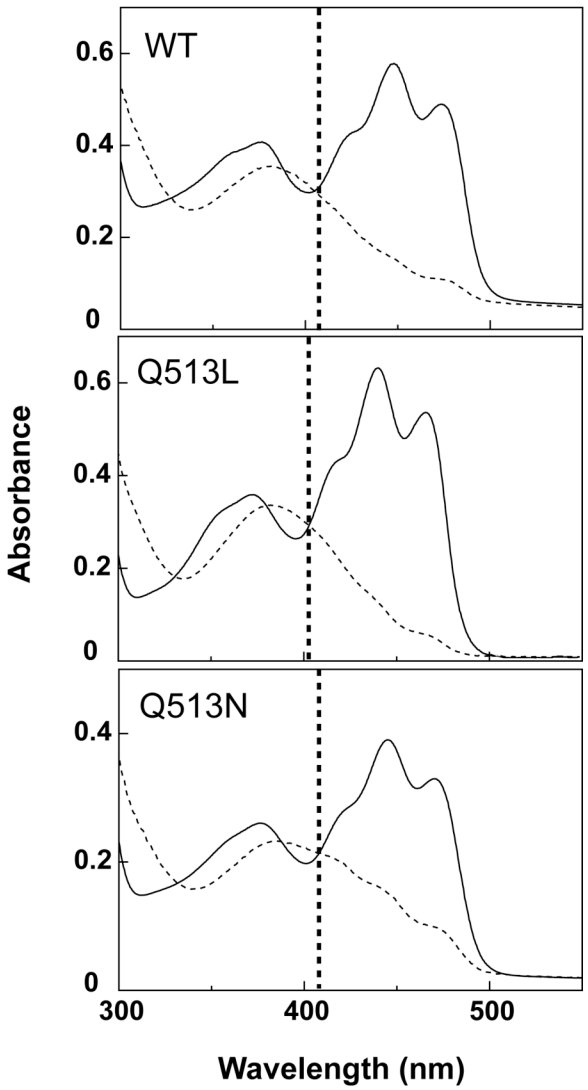

Fig. 2-3. Q513 mutants demonstrate typical LOV domain spectra
Solid traces represent the dark state spectra and dashed traces represent the lit state spectra. The mutants all display the same characteristic dark state absorbance profile for typical LOV domains, with three distinctive maxima between 400 nm to 500 nm. These maxima diminish in the lit state in all three cases. The vertical dashed line is aligned with the largest wavelength isosbestic points of the LOV domains (406 nm for AsLOV2 and Q513N; 403 nm for Q513L).

Construct	$\tau_{dark\ recovery}[A_{446}]$ (s)	$\tau_{dark\ recovery}[\theta_{222}]$ (s)	ε_{446} (M^{-1} cm^{-1})	Isosbestic points (nm)
WT	68.3	72.4	11700	327, 388, and 406
Q513N	37.3	40.6	12400	330, 388, and 406
Q513L	1080	>1000	10700	330, 380, and 403

Table 2-1. Comparisons of the kinetic time constants (τ) and the absorption coefficients (ε) at the isosbestic points for wildtype and Q513 mutants in the dark state.
The kinetics experiments are recorded at room temperature (22°C) for 200 s and the data points were fitted using a first order exponential rate equation to obtain the time constant (τ). The dark recovery time constant at A_{446} is measured by UV-visible absorbance spectroscopy while the dark recovery time constant at θ_{222} is measured via CD spectroscopy.

Given the alterations in the environment surrounding the FMN cofactor, we sought to determine if these mutations would affect the photocycle of the LOV domain. We found the dark recovery time constant of the Q513L mutant followed by illumination is 1080 s, approximately 15-fold longer than wildtype (68 s, Table 2-1). In contrast, the Q513N mutation has a shorter recovery time constant (37 s, Table 2-1). These kinetic data indicate that the Q513L point mutation has a more significant effect on the relative energetics of *As*LOV2 lit state and transition state it visits during the recovery process, complementing the steady state absorbance data which show a similarly larger effect of this mutation.

Structural Effects of Q513 Mutations

Circular Dichroism

Prior solution NMR studies of wildtype *As*LOV2 show that the C-terminal

Jα helix dissociates from the core LOV domain and unfolds upon illumination, while the LOV domain itself remains intact and folded (*29, 87*). Circular dichroism reflects global secondary structure so the spectra presented represent the total mixed α/β fold of the LOV domains. The double minima at 208 nm and 222 nm and maximum at 195 nm are features of helical secondary structure (Fig. 2-4), which demonstrate a clear decrease in molar residue ellipticity in wildtype *As*LOV2 following illumination with white light (*66*). In contrast, both the Q513L and Q513N mutants show reduced light-induced changes by CD (Fig. 2-4). This is consistent with the limited proteolysis and NMR results (*vide infra*) suggesting mutation of Q513 leads to some decoupling of covalent adduct formation from conformational changes.

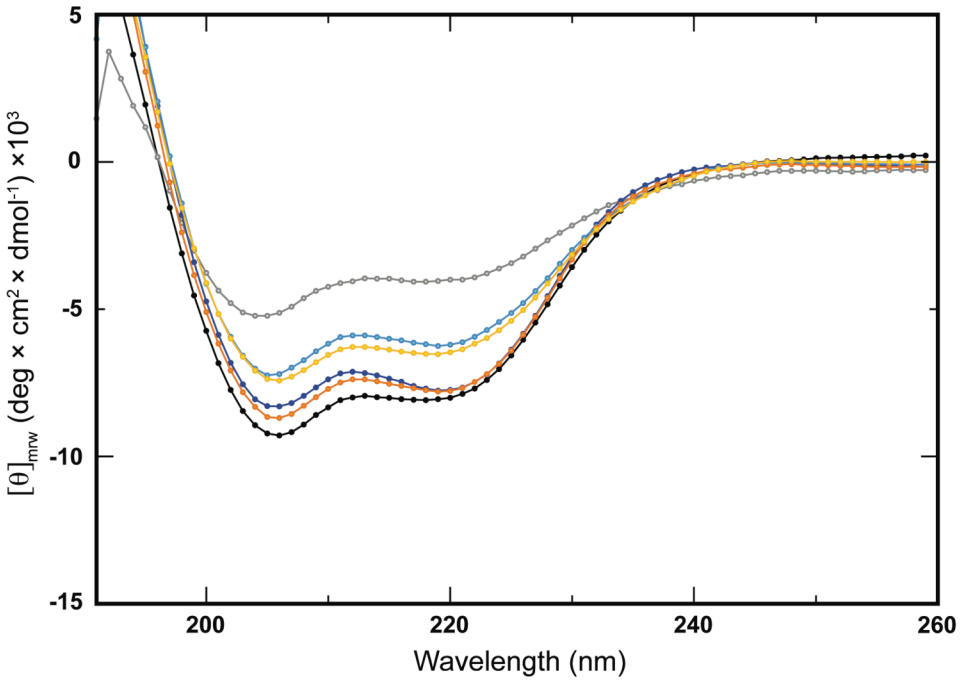

Fig. 2-4. Structural effects of Q513 mutations in the dark and lit states monitored by circular dichroism
The room-temperature far UV region (190 nm to 250 nm) CD spectra of the wildtype dark (black) and lit (grey) states, Q513L dark (orange) and lit (yellow) states, and the Q513N dark (dark blue) and lit (light blue) states are overlaid here for comparison. The differences between the dark and lit states are more pronounced in the wildtype compared to the Q513 mutants.

Despite the reduced amplitude of light-induced changes in the CD signals of Q513L and Q513N, we were still able to monitor the kinetics of dark state recovery via changes in the secondary structure during this process. As observed with UV-visible absorbance spectroscopy, both mutants undergo a normal photocycle with complete recovery to the dark state following illumination. Also consistent with the UV-visible absorbance results, Q513L shows significantly slower recovery kinetics while Q513N is slightly accelerated (Table 2-1). We observed similar recovery kinetics regardless of whether we monitored the change via protein (CD) or chromophore (UV-visible absorbance), suggesting that these

processes have a common rate-limiting step (*51, 66*).

Limited Proteolysis

To further document how the Q513 mutations affected the overall stability of the LOV domain, we used limited proteolysis. Wildtype *As*LOV2 becomes more susceptible to proteolytic cleavage by chymotrypsin upon illumination, specifically at Met530 located in the middle of the Jα helix (*29*). This is reflected in the light-induced acceleration in the appearance of a lower molecular weight species in SDS-PAGE (Fig. 2-5). Notably, neither mutant domain demonstrates as dramatic an increase in proteolysis after covalent adduct formation. The Q513L mutant is less susceptible to proteolysis from chymotrypsin in the lit state than wildtype *As*LOV2, while the dark state demonstrates resistance similar to wildtype. In contrast, the Q513N mutant displays the opposite effect, with a protease-susceptible dark state and a lit state that is as easily proteolyzed as wildtype. The primary species formed upon cleavage of both mutants is consistent with that formed by similar treatment of wildtype *As*LOV2, suggesting that Met530 is the likely cleavage site. In addition, chymotrypsin treatment of Q513N produces a lower molecular weight species that is formed very quickly upon addition of protease to the lit state. This additional band suggests that Q513N may adopt another domain conformation or have increased dynamics that allow protease accessibility to an otherwise inaccessible residue.

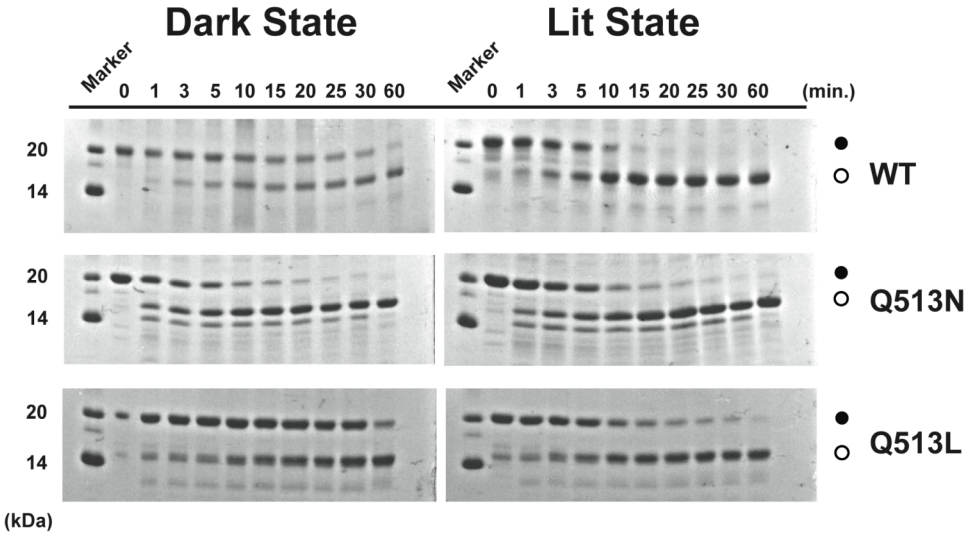

Fig. 2-5. Structural effects of Q513 mutations in the dark and lit states monitored by limited proteolysis
Limited proteolysis by chymotrypsin digestion was performed at room temperature for the wildtype protein and Q513 mutants in both the dark (left panel) and lit states (right panel). Time points ranging from 0 to 60 min are shown above the gels. The molecular weight marker indicating 14 kDa and 20 kDa are also shown on the left side of the gels. Black filled circles represent the full-length undigested domain and open circles represent the largest stable digested product.

NMR Spectroscopy

The low-resolution structural information provided by CD spectroscopy and limited proteolysis clearly show that both Q513L and Q513N mutants have fewer conformational changes upon illumination compared to wildtype *As*LOV2. To examine this in further detail, we used two-dimensional ^{15}N-^{1}H HSQC spectra to monitor the environments of the pairs of J-coupled ^{15}N-^{1}H nuclei within the domain in both the dark and lit state. The ^{15}N-^{1}H HSQC spectrum of wildtype *As*LOV2 in the dark state shows well-dispersed peaks consistent with our previous NMR results (Fig. 2-6) (*29*). Upon illumination, we observe a general loss of amide proton chemical shift dispersion and the appearance of several new

intense peaks in the center of the spectrum, indicative of increased dynamics in the LOV domain and dissociation of the Jα helix.

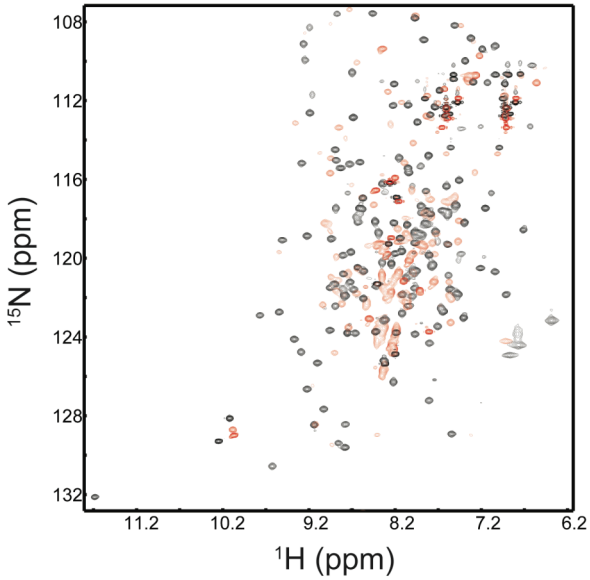

Fig. 2-6. Overlay of dark and lit state ^{15}N/^{1}H HSQC correlation spectra for wildtype *As*LOV2 domain
Spectra were recorded at 25°C at 500 MHz. Black spectrum represents the dark state and red spectrum represents the lit state.

Specific analysis of two tryptophan side chain indole ($H_{\varepsilon 1}$-$N_{\varepsilon 1}$) crosspeaks highlights the light-induced structural changes surrounding residues W491 and W556, located near the LOV-Jα helix hinge region and C-terminal to the Jα helix respectively (*29*). In the dark state of wildtype *As*LOV2, these indole crosspeaks are clearly separated (Fig. 2-7a). Upon illumination, they collapse towards a central position that is near the average location for protein tryptophans in general (Fig. 2-7b) (*88*), also consistent with the Jα helix unfolding and the tryptophans adopting less distinctive chemical environments after the covalent adduct is formed. Comparison with this same region in the Q513L spectrum again shows

two clearly resolved crosspeaks that overlay with the wildtype dark state (Fig. 2-7a). After light induction, there is a small chemical shift perturbation of the Q513L crosspeaks (Fig. 2-7b), indicating that a significant majority of the Q513L population still remains closer to the dark state conformation after light induction. Conversely, the Q513N mutant displays very different NMR spectra in the tryptophan indole region from Q513L or wildtype. Prior to light irradiation, the tryptophan signals are closer to their counterparts in the wildtype lit state (Fig. 2-7a). Additionally, the crosspeaks do not collapse toward each other to the same extent as wildtype upon illumination (Fig. 2-7b), suggesting that the Q513N mutant adopts a pseudo-lit structure that resembles the wildtype lit state and hence undergoes relatively few light-induced structural changes. While we discuss Q513L and Q513N as pseudo-dark and –lit structures here, we suggest that both are significantly more dynamic than wildtype given the increased linebroadening present in both spectra (Fig. 2-7). Further, signs of peak doubling can be observed in several of these spectra, suggestive of slow (τ~ms or longer) interconversion between states. Overall, these data support and extend both the proteolysis and CD data, showing that both Q513L and Q513N undergo more limited structural changes with illumination and appear to be poised more towards either the dark- or lit-state structure of the wildtype.

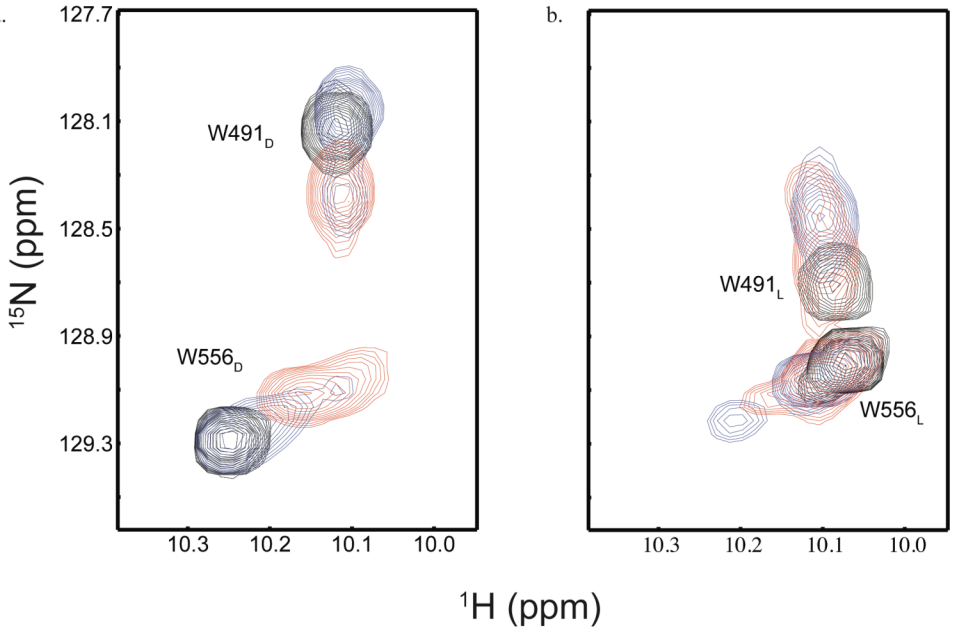

Fig. 2-7. ^{15}N/^{1}H HSQC spectra of the tryptophan $H_\varepsilon 1$-$N_\varepsilon 1$ region
The overlaid spectra of the dark (a) and lit (b) states of the wildtype AsLOV2 (black), Q513L (blue), and Q513N (red) are shown. The tryptophan indole assignments are indicated in both the dark and lit state panels with subscript D or L respectively. In the lit state, there is a significant shift in this region in wildtype protein (b).

An analysis of the full ^{15}N/^{1}H HSQC spectra of wildtype, Q513L and Q513N (Fig. 2-8a-b) supports the assignment of Q513L as a pseudo-dark state structure and Q513N as pseudo-lit state. The dark state spectrum of Q513L (Fig. 2-8a) is quite similar to wildtype, consistent with limited proteolysis. In contrast, the Q513N dark state shows significant chemical shift changes and/or linebroadening, and is reminiscent of the spectra of the wildtype lit state (Fig. 2-8b). These data clearly indicate that the Q513N point mutation causes a greater structural perturbation of the wildtype structure than Q513L. As with the wildtype protein, illumination causes significant spectral changes for both the Q513L and Q513N mutants (Fig. 2-8a-b). Unfortunately, these perturbations

cannot be unambiguously interpreted to provide independent confirmation of the reduced conformational changes in the two Q513 mutants as reported above by proteolysis and CD. This is due to the fact that chemical shift changes originate from two interrelated sources: *bona fide* protein conformational changes and the significant alteration in the electronic structure of the FMN isoalloxazine ring upon adduct formation. Given that the adduct forms successfully in all three *As*LOV2 variants tested here, we expect significant chemical shift changes in these proteins regardless of their ability to couple this photochemical event with protein conformational changes. Despite this caveat, these NMR spectra support the assignment of Q513L and Q513N domains adopting pseudo-dark and pseudo-lit state structures in the dark.

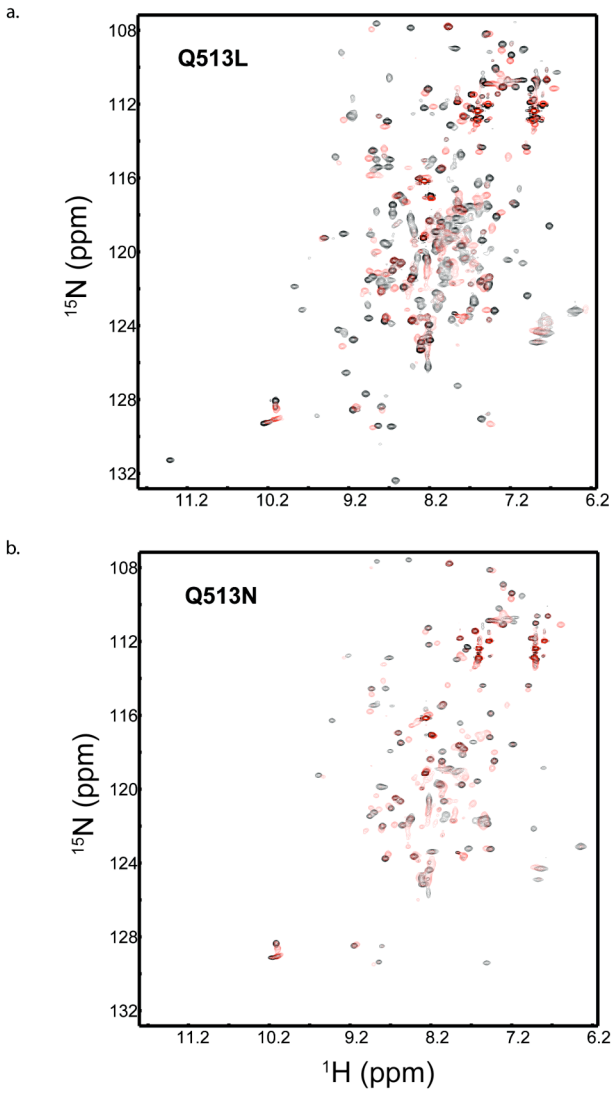

Fig. 2-8. Overlay of dark and lit state ^{15}N/^{1}H HSQC correlation spectra for Q513L and Q513N

Spectra were recorded at 25°C at 500 MHz. Black spectra represent the dark state and red spectra represent the lit state. Mutations lead to chemical shift changes as well as differential broadening across each spectrum.

Understanding the Mechanism of Intradomain Communication

Literature Review of the Role of Glutamine 513

While the connection between light-induced covalent adduct formation and protein conformational changes in LOV domains is well established (*7, 29, 51, 67, 71, 81, 89*), the mechanism through which this occurs remains unclear. A highly conserved glutamine residue (Q513 in *As*LOV2) in the core of the LOV domain was previously suggested to be crucial for this signaling process (*7, 81*). While FTIR studies show that illumination breaks a hydrogen bond between this residue and the FMN O4 position upon adduct formation (*81*), formation of a proposed new hydrogen bond between Q513 and the FMN N5 position has been more difficult to demonstrate. Some crystal structures show that the side chain of this glutamine rotates with illumination, consistent with formation of this new hydrogen bond (*7, 78, 79, 89*), while other structures argue against it (*67*). In light of this ambiguity, we targeted our point mutations to test the importance of the Q513/FMN interaction for intradomain signal communication. Ground state structures, light-induced structural changes and dark state recovery rates are all altered by mutations at this position, demonstrating an important role for Q513 in *As*LOV2 signaling.

Structural Effects of Q513L and Q513N Point Mutations

Residue Q513 is located on the Iβ strand, on the opposite side as the Jα-helix binding surface, thus suggesting a direct path from the internally bound

FMN to the Jα-helix on the surface (Fig. 2-1a). As such, it seemed reasonable that changing hydrogen-bonding patterns between Q513 and FMN would alter the structure of the anchoring Iβ strand in such a way as to interfere with this pathway. The Q513L mutation was designed to disrupt all hydrogen bonding with FMN while only slightly increasing the volume of a glutamine residue. The UV-visible absorbance profile of Q513L is blue-shifted in the dark state, consistent with the loss of this hydrogen bond to the O4 carbonyl oxygen of FMN. Without this hydrogen bonding capability, we anticipated that the Q513 side chain would not rotate in an organized fashion upon covalent adduct formation. With loss of this rotation, the Iβ strand structure and dynamics would likely remain unchanged with illumination. Our data bear out these predictions, as the Q513L mutant demonstrated similar structural properties to the wildtype dark state and also had reduced light-induced conformational changes. Previous findings of the analogous Q1029L mutation in neo1 LOV2 (*81, 90*) support this view.

In comparison, the Q513N mutant also had reduced amplitude of light-induced structural changes but appeared to adopt a pseudo-lit state structure in the dark. The Q513N mutation was designed to maintain hydrogen-bonding contacts, and the similarity of the Q513N and wildtype UV-visible absorbance spectra is consistent with hydrogen bonds being maintained between this residue and the FMN O4 and N5 atoms (if we assume this interaction occurs in the lit state). To maintain these bonds, the Iβ strand may be distorted to allow for FMN interaction

with the shorter asparagine side chain. This type of stress may somehow be similar to the type of movement or changes that normally induce Jα release, giving rise to a pseudo-lit state type structure. Given that Q513N is already in this pseudo-lit state in the dark, we suggest that illumination and cysteinyl/C4a adduct formation cannot induce further conformational changes, consistent with our results.

Dark State Recovery Kinetic Effects of Q513L and Q513N Point Mutations

While the Q513L and Q513N mutants were created to study their roles in the structural changes accompanying signal transduction, we found that both also affected dark state recovery rates. Other work has identified several solution parameters that perturb these rates (*54-56*); however, the most relevant factors for our results are related to the conformations of the lit state and the transition state between the dark and lit structures. Chemical transition state theory establishes that the energetics of these two states influence the kinetics of the return rate. In *As*LOV2, the difference between these two states is approximately $\Delta G^{\ddagger}$~14.5 kcal mol^{-1} based on the temperature dependence of dark state recovery (*51*). In parallel, the spontaneous relaxation of the lit state establishes that it is energetically less favorable than the dark state, leading to the suggestion that the lit state is somehow conformationally strained (*68*), which is experimentally supported by a light-dependent increase in ^{2}H exchange rates (*51*). As such, changes that lower $\Delta G^{\ddagger}$ by either destabilizing the lit state or stabilizing the

transition state are predicted to accelerate dark state recovery, and vice-versa. This is supported by the accelerated recovery rates of an *As*LOV2 I427V point mutant, which removes a methyl group that is predicted to stabilize the lit state Cys-C4a adduct (*91*).

Our data on the recovery rates of Q513L and Q513N are consistent with this model. Q513L exhibits a fifteen-fold slowing of the dark state return rate, as well as data indicating that this protein undergoes much smaller light-induced structural changes than the wildtype domain. These data are consistent with stabilization of the lit state and a corresponding increase in $\Delta G^{\ddagger}$, possibly by relieving tension that would otherwise be maintained by the Q513:FMN hydrogen bonds in the lit state. In agreement with this, an analogous Q1029L point mutation in *A. capillus-veneris* neo1 LOV2 slowed dark state recovery significantly (7-fold, (*90*)), as does a F1010L mutation at an adjacent position on the neighboring Hβ strand (10-fold, (*92*)). In contrast, we find that the Q513N point mutant accelerates dark state recovery two-fold. We suggest that this protein retains dark state hydrogen bonding between the N513 amide and the O4 carbonyl oxygen based on visible absorbance spectroscopy, and may retain similar interactions with the flavin cofactor in the light. The maintenance of these interactions despite the loss of a methylene group would likely lead to a more destabilized lit state, consistent with the rate acceleration we observe. While a detailed understanding of this process remains to be established, we suggest that these and other rate-perturbing mutations are providing useful evidence for the

features of *As*LOV2 that establish the lifetime of the signaling state.

Models of Light-Induced Movement in Q513

The combination of molecular dynamics simulations and x-ray crystallography has led to models for light-induced Q513 movement with somewhat opposing conclusions. Simulations of *As*LOV2 minus the Jα helix identified breakage of the dark state hydrogen bond between Q513 and FMN and further demonstrated light-induced hydrogen bond formation at the FMN N5 position (*79*). In addition, these simulations also suggested a second conformation in which Q513 interacted with neighboring residues on the Iβ strand, thereby increasing dynamics in this region. Such an alteration at the LOV-Jα interface may contribute towards Jα release and signal transduction. In contrast, recent crystal structures of *As*LOV2 containing the Jα (*67*) demonstrate neither any rotation of the Q513 side chain upon illumination nor a bent conformation. Consistently, crystal structures that fail to demonstrate side chain rotation also fail to demonstrate the previously described loss of hydrogen bonding to FMN O4. These inconsistencies among computational models, solution studies, and crystallographic structural methods suggest the role of Q513 side chain rotation in signal transduction merits further investigation.

Role of Q513 in Full-Length LOV-Containing Proteins

Studies on a single isolated domain, such as those discussed here, provide

interesting results from which we can postulate the role of Q513 in signal transduction, but the behavior in full-length proteins requires further investigation. Limited proteolysis, circular dichroism and NMR data all demonstrate fewer light/dark conformational changes for the mutant domains. Specifically, Q513L appears locked into a dark state-like conformation and Q513N retains lit state-like characteristics regardless of FMN electronic state. *In vitro* biochemical experiments with full-length phototropin containing a mutation analogous to Q513L support our findings, showing that this mutation attenuates light-activated autophosphorylation activity (*82*). These data are consistent with the Q513L mutant *As*LOV2 domain maintaining a dark state-like, inactive conformation, as we have found. This residue also plays a central role in the FAD-bound LOV domain photoreceptor, Vivid (*7*). A comparison of dark and lit state crystal structures of this protein shows a network of light-induced rearrangements in hydrogen bond contacts between the protein and FAD. In the wildtype protein, these lead to a series of side chain reorientations that ultimately alter the protein surface. Introduction of a leucine mutation at the equivalent glutamine position in Vivid (Q182 in Vivid) disrupts these changes, as shown by differential elution times in size exclusion chromatography. While this work further extends the results of the domain studies to full-length proteins, it also suggests that this conserved glutamine is important for signal communication in non-phototropin related LOV domains.

A large-scale sequence alignment of LOV domain sequences suggests that

Q513 is highly, but not absolutely, conserved (Fig. 2-9a-b). In particular, we see that several proteins contain naturally occurring leucine substitutions at this critical site. While most of the LOV domains with leucine substitutions have not been studied to the extent of determining structural information or dark state recovery time constants, a small cohort of *A. thaliana* proteins are of particular interest. These three proteins, FKF1, LKP2 and ZTL have extremely stable cysteinyl-flavin adducts: FKF1 demonstrates a dark-state recovery half-life of 62.5 hr (*50*), while the other two have been described as effectively irreversible (*74*). Intriguingly, LKP2 contains a leucine at the position equivalent to Q513 in *As*LOV2, suggesting it may play a role in extending the dark state recovery of this protein. While this is an enticing hypothesis, neither FKF1 nor ZTL have a leucine at this position, indicating there must be other factors influencing dark state recovery rates. Mutational studies have determined several other positions that contribute to dark state recovery kinetics (*43, 48, 91-93*). Of these, a phenylalanine to leucine mutation at the position equivalent to *As*LOV2 F494 led to a 10-fold increase in half-life of the excited state (*92*). FKF1, ZTL and LKP2 all contain this naturally-occurring leucine substitution, which occurs on the Hβ strand immediately adjacent to Q513, positing an important role for this residue in tuning photocycle kinetics. These data, combined with the studies on Q513 presented here, indicate that several residues of the chromophore-binding pocket of LOV domains collectively play roles in critical aspects of signaling, including signal transmission and regulation of signaling state lifetimes. A combination of

further biochemical and biophysical measurements are needed to characterize the detailed basis of this control, perhaps allowing artificial control of these features (*87*).

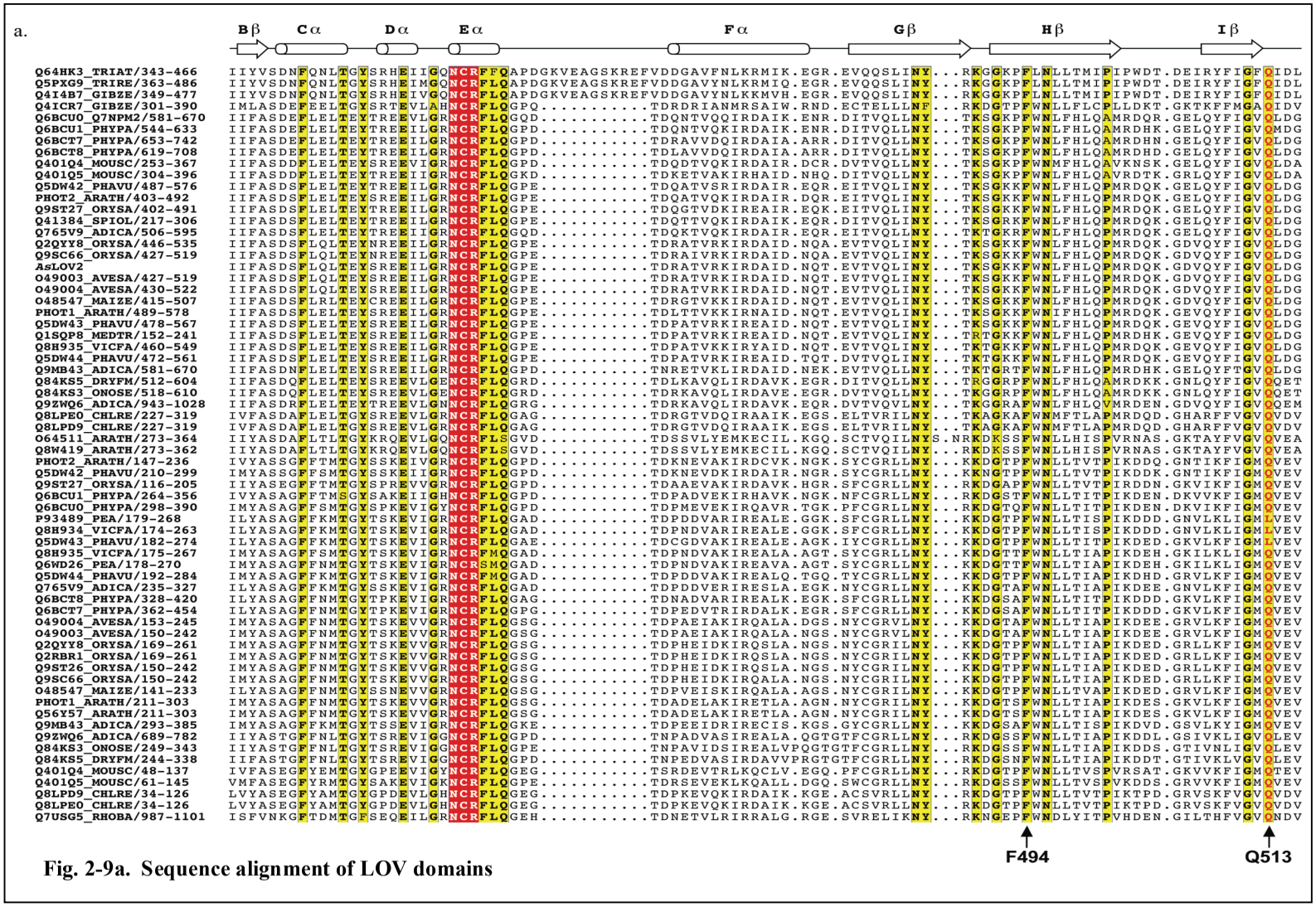

Fig. 2-9a. Sequence alignment of LOV domains

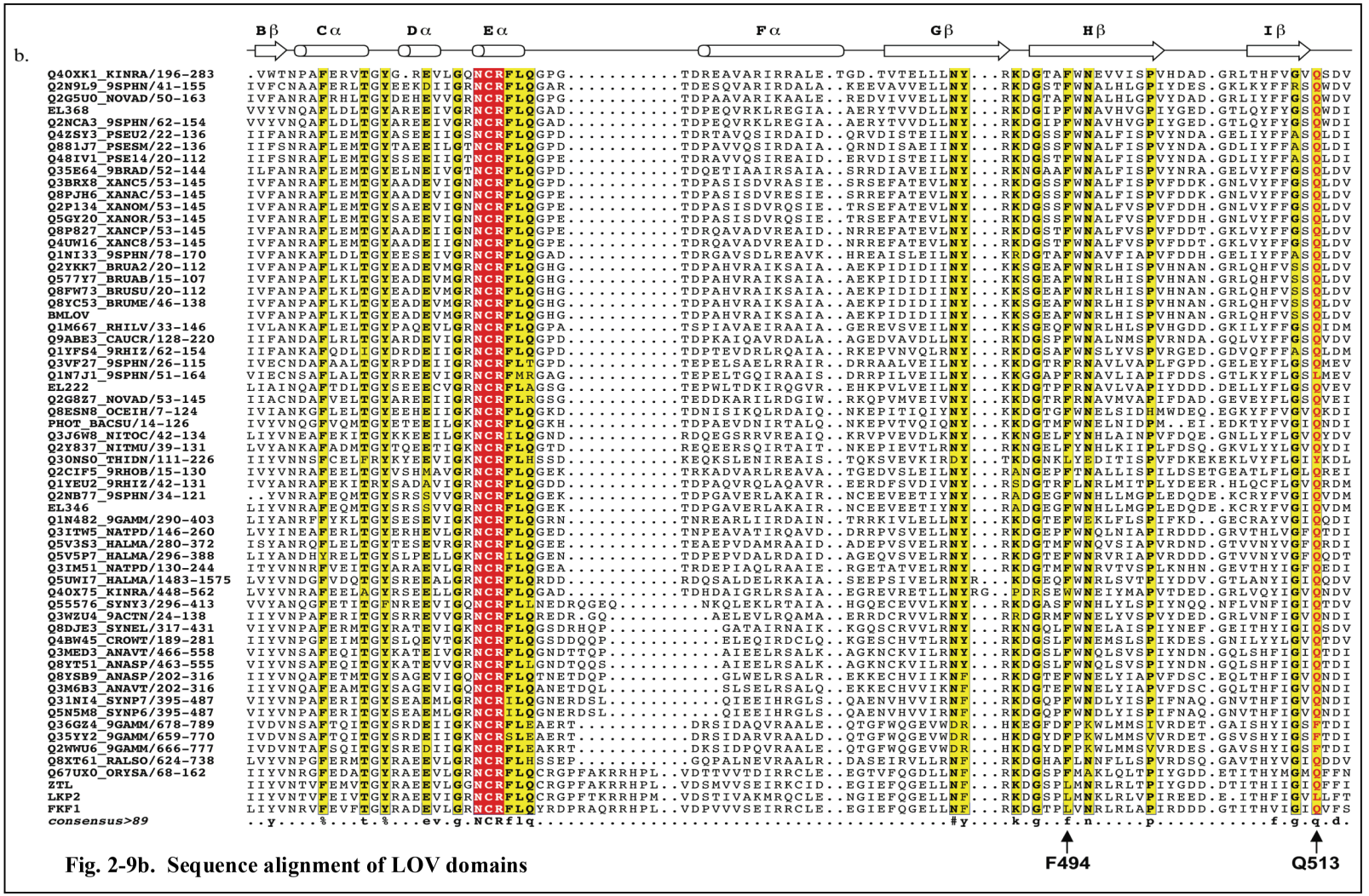

Fig. 2-9b. Sequence alignment of LOV domains

CHAPTER 3

INTRODUCTION TO DNA BINDING PROTEINS

Signal transduction pathways responding to environmental stimuli often terminate in a DNA binding protein that regulates gene transcription to alter the current behavior of a cell (*13*). These DNA binding proteins demonstrate great diversity in their domain components and overall structure; however, all must respond to their upstream signaling partner to regulate the ability to bind DNA.

DNA Binding Protein Structural Variability

There are several families of DNA binding proteins including, but not limited to, zinc-finger proteins, homeodomains, helix-loop-helix proteins, and helix-turn-helix proteins (*94*). These families are delineated according to their DNA recognition motif architecture. The last family, the helix-turn-helix (HTH) proteins, is the first DNA binding motif discovered and consists of an α-helix, a turn and a second α-helix (*95*). A complete HTH domain consists of three core helices in a right-handed helical bundle (*52*)(Fig. 3-1a). Several conserved hydrophobic motifs and residues localize to the interior, forming a hydrophobic core that stabilizes the domain (*52*). The third helix has been identified as the recognition helix, the major site of interaction between the protein and DNA (*94*)(Fig. 3-1a-b); however, additional elements beyond the HTH may also contact DNA (*52*). While the HTH motif defines the DNA binding protein family,

distinctive features allow members to be divided into two major structural classes and several other highly derived classes with extreme alterations to the HTH core (*52*). The simplest of the structural classes is the simple tri-helical bundle, represented by the bacterial proteins NtrC (*96*) or the quorum sensing protein HapR (*2*)(Fig. 3-2a). Extensions on this tri-helical bundle commonly include a fourth helix (Fig. 3-2b) (referred to as the LuxR class of HTH domains herein, given the characteristics of this protein), or in some cases, multiple additional helices (*52*). The other major class of HTH architecture is the winged HTH (wHTH). These domains are distinguished by an addition C-terminal β-strand hairpin that packs against the shallow cleft in the partially open tri-helical core (*52*). This variant, along with the tetra-helical architecture, is commonly found in prokaryotic transcription factors (*52*) like DrrB (*11*) (Fig. 3-2c). In the following chapter I will describe my investigations into a protein possessing a tetra-helical HTH. This architecture is commonly found in bacterial transcription factors belonging to the FixJ/NarL family as well as bacterial quorum sensing proteins in the LuxR family.

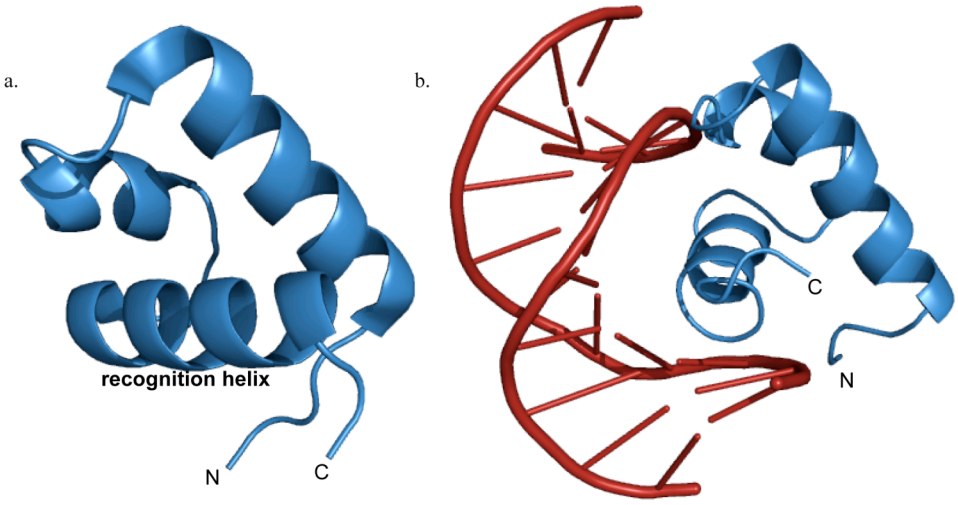

Fig. 3-1. Single helix-turn-helix structure and function
a.) The crystal structure of C-Myb R1 (1GUU) (*1*) demonstrates the most basic HTH structure. It consists of three α-helices. The third helix is the recognition helix that associates with the major groove of DNA. b.) Myb R2 associated with DNA (1MSE) (*3*). Notice the position of the recognition helix in the major groove.

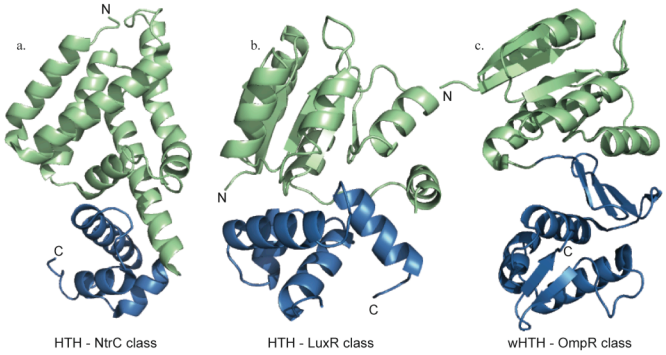

Fig. 3-2. Representative HTH-containing proteins
Proteins representing three common HTH folds. a.) HapR contains the basic three helix bundle of an HTH domain (2PBX) (*2*). b.) Bacterial response regulator NarL contains a common variant of the all helical HTH fold which includes an additional, fourth α-helix (1RNL) (*6*). c.) A second bacterial response regulator, DrrB, contains the other major class of HTH folding which includes a small antiparallel β-sheet in addition to the three α-helices of the HTH (1P2F) (*11*). This fold is designated a winged helix-turn-helix.

Helix-turn-helix domains are found in combination with a wide variety of domains, imparting different modes of regulation and functional purposes to the proteins. At the simplest level, the HTH may be a stand-alone protein with simple extensions to the N-terminus to aid in dimerization or transmembrane regions (*52*). Combinations of HTH with other nucleic acid binding domains and/or protein-protein interaction domains suggest some proteins may be involved with bridging macromolecular complexes, especially those involved with DNA replication, repair and restriction endonucleases (*52*). Finally, HTH domains in combination with catalytic domains have revealed several functional trends. The first of these is the use of the HTH as a substrate recognition or localization domain (*97-99*). The second trend is related to feedback regulation of metabolic pathways, where the HTH is combined with a domain catalyzing a key step in a biosynthetic pathway, such as biotin synthesis (*100*).

Regulation of DNA Binding Activity

A critical aspect of HTH activity is how their DNA binding activity can be regulated with external stimuli. Since this motif is often involved with gene transcription, it is important that its activity is modulated appropriately. Both of the major architectural groups of HTH have a significant number of family members involved in bacterial two-component, or one-component, response regulators. These signaling pathways are popular research topics given their

likely role in bacterial pathogenesis; therefore there is a wide range of knowledge regarding the activation and regulation of these proteins. Within these two-component systems, the HTH-containing protein often serves as the response regulator. In this role, the receiver domain will accept a phosphotransfer from its associated histidine kinase, thus altering the ability of the HTH domain to either activate or inhibit transcription (*101*) (Fig. 3-3). Understanding the mechanism of signal transduction from the receiver domain to the HTH domain has been hampered by the difficulty in obtaining structural information on both active and inactive forms of full-length response regulators.

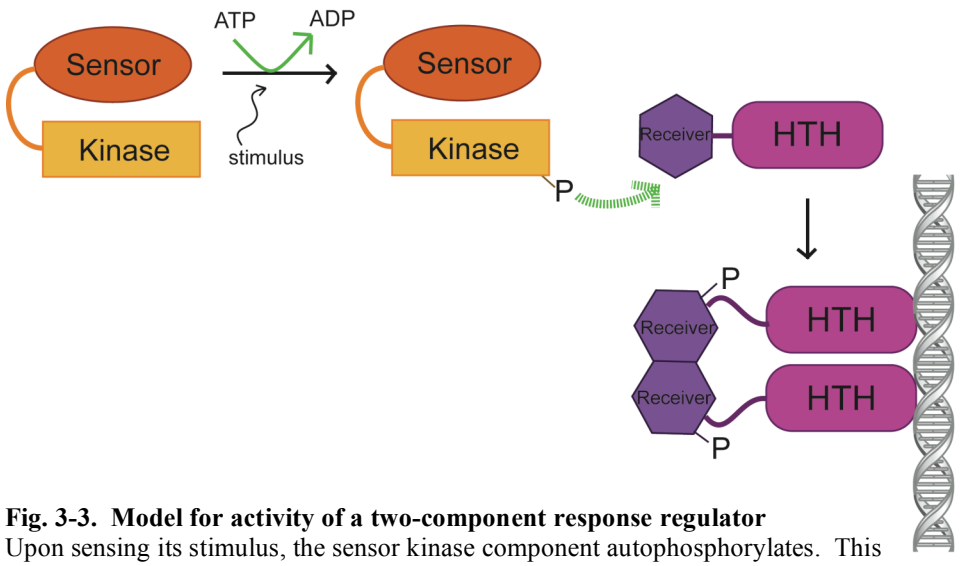

Fig. 3-3. Model for activity of a two-component response regulator
Upon sensing its stimulus, the sensor kinase component autophosphorylates. This phosphate is then transfered to the receiver domain of the response regulator. Upon phosphorylation, the response regulator undergoes some change, likely conformational, that affects the HTH-DNA binding interaction. Phosphorylation may activate binding or inhibit otherwise constitutive binding, depending on the protein itself. Response regulators generally dimerize following activation in the presence of their cognate binding sequence.

Studies on select proteins have given us some insight into what may occur following phosphorylation. The two most common results of phosphorylation are conformational change and dimerization, usually in that order. Structural information on the full-length response regulator NarL revealed the interaction between a response regulator and its LuxR-type HTH domain in atomic detail. Interestingly, in the unphosphorylated (inactive) state the DNA recognition helix was in direct contact with the receiver domain such that it could not interact with the DNA major groove (*6*)(Fig.3-2b). This suggested NarL must undergo a structural rearrangement upon activation to release the recognition helix from these inhibitory contacts (*6*). In addition, crystal structures of the *isolated* HTH domain show DNA binding via the recognition helix of dimeric domains (*102*). Again, the crystal structure of unphosphorylated NarL indicates the dimerization helix of the HTH is in contact with the receiver domain, further necessitating a conformational change following activation (Fig. 3-2b, Fig. 5-1a). Solution studies on the unphosphorylated receiver domain of NarL showed interaction with the HTH domain. However, these contacts could be disrupted following phosphorylation, consistent with an interdomain separation (*103*). EPR studies of spin labeled full-length protein revealed this same interdomain separation following phosphorylation (*104*).

Phosphorylation-induced conformational change to reduce inhibitory contacts and expose a dimerization interface is also observed in other response regulators with LuxR-type HTH domains. Crystal structures of the isolated N-

terminal domain of the response regulator FixJ, in both the unphosphorylated and phosphorylated states, revealed a key interaction between a conserved threonine residue and the phosphoryl group (*105*). This leads to significant changes in the positions of secondary structural elements ultimately resulting in formation of the dimerization interface necessary for DNA binding of active protein (*105*). The FixJ receiver domain also demonstrates inhibitory control on the HTH domain that can be disrupted with phosphorylation in the full-length protein (*106*). This theme of conformational change and dimerization is even found in the wHTH family of response regulators. Phosphorylation of the wHTH-containing PhoP appears to shift the monomer-dimer equilibrium in favor of dimeric protein and increases the protein's affinity for DNA (*107*). Similar results are also seen for OmpR (*108*) and CovR (*109*).

As presented in chapter 1 of this dissertation, PAS domains are found in myriad proteins and are able to regulate protein function through ligand binding, protein-protein interactions and post-translational modification. These domains coupled with a DNA binding domain allow us to explore two different fields of interest that share a common question: How is one domain of a protein regulated by the other? This work presents a two-domain protein with an N-terminal PAS domain that absorbs blue light, leading to activation. The C-terminal domain is a DNA binding domain that I will show binds DNA only in the active, light-induced, state. Through this work we gain insight into how PAS domains communicate to their downstream effector domains, as well as how HTH motifs

within a DNA binding domain are regulated

CHAPTER 4

INVESTIGATING INTERDOMAIN SIGNAL TRANSDUCTION IN A NOVEL LOV-HTH PROTEIN

LOV domain signaling originates in the core of the domain following covalent adduct formation between the conserved cysteine residue and the flavin isoalloxazine ring. It then propagates to the surface in order to alter the activity of the protein. Studies of other LOV-containing proteins have shown that structural elements beyond the canonical LOV core undergo conformational changes following illumination, and that these alterations are responsible for protein activation (*7, 29, 67, 71*). In addition, these auxiliary structural elements appear to flank the β-sheet surface of the domain, suggesting a direct pathway from the hydrophobic core, through the β-sheet, and to these structures. The previous chapter described an in-depth investigation into the role of a particular amino acid side chain, located on the β-sheet surface, in propagating light-induced structural changes. This study, however, is somewhat limited due to the inability to use the full-length, multi-domain protein for structural studies (although functional studies have been conducted (*82*)). In the following chapter, I will be studying interdomain signal propagation using a full-length, two-domain protein. With this protein, EL222, I can study the structural changes occurring within the LOV domain following illumination, as well as how these changes affect the C-

terminal effector domain and the likely path the signal passes through.

A Novel LOV-Containing Protein

LOV Domains in environmental signaling

As discussed in the previous chapters of this dissertation, PAS domains are often employed for cellular sensing of the environment. A subset of PAS domains, the Light-Oxygen-Voltage (LOV) domains, are able to directly sense blue light using a noncovalently associated flavin molecule in their hydrophobic cores. Following illumination, a covalent adduct forms between an essential cysteine residue and the flavin isoalloxazine ring (Fig. 1-4c). This adduct exists for a variable time period, ranging from seconds to days, (specific to the domain), before spontaneously decaying to the noncovalent state.

Following covalent adduct formation, the LOV domain must transmit this signal to its cognate effector domain within the protein or to an interacting partner protein. Investigations into this mechanism have been hampered in the past by the difficulty in working with full-length, multi-domain LOV proteins in the quantities, concentration, and purity required for high-resolution structural studies. Biophysical studies on isolated domains have demonstrated structural changes occurring within secondary structural elements beyond the LOV core and associated with the domain surface (*7, 29*). Studies on full-length, multi-domain LOV-containing proteins remain elusive.

LOV-containing Transcription Factors

LOV domains are associated with a wide variety of physiological responses. As such, they are coupled with diverse effector domains such as kinases, F-boxes, and zinc fingers (*41*). Among these LOV proteins, several known LOV-containing transcription factors have been described. These transcription factors include *Neurospora crassa* white collar-1 (WC-1) (*110*), and the AUREOCHROMEs from stramenophile algae (*111*). While WC-1 utilizes a zinc finger DNA binding domain, the AUREOCHROMES use a basic region/leucine zipper domain. These different DNA binding domains highlight the ability of the LOV domain to pair with different domains to regulate the same activity, in this case, DNA binding. Both proteins have been shown to bind their cognate DNA sequences only under lit state conditions (*110, 111*), suggesting a role for the LOV domain in this activity. Unfortunately, neither of these proteins has been amenable to structural studies. The mechanisms of both light-sensing and DNA binding have been described for the component domains of both proteins (*43, 44, 112-114*), however the signal transduction pathway between LOV domain and DNA binding domain remains unknown.

HTH-type DNA Binding Proteins

LOV domains have been described coupled to zinc finger domains and leucine zipper motifs, enabling light-dependent DNA binding. Here we present

an example of a bacterial protein, EL222, isolated from *Erythrobacter litoralis HTCC 2594*, that contains an N-terminal LOV domain and a C-terminal helix-turn-helix (HTH) DNA binding domain, representative of the LuxR-type DNA binding domains (named for the protein used as the template sequence for the group). These domains are well represented among bacterial response regulators (*115*) including the *Escherichia coli* protein NarL (*6*). Based on biochemical studies on truncated derivatives of LuxR-type proteins (*106, 116, 117*) as well as crystallographic and solution NMR studies (*6, 10, 103*), the model for transcriptional regulation by NarL proposes the N-terminal domain masks the C-terminal HTH in such a way as to prevent DNA interaction. Upon phosphorylation of the N-terminal domain, this inhibition is released.

EL222 as a Model for LOV and HTH Regulation

EL222 is a protein containing two domains, each type of which has been independently characterized in other proteins, but with questions surrounding the generality of signaling or regulation. In the case of LOV domains, we have a clear understanding of the mode of activation, but no single model for how this signal is communicated to downstream effector domains. Similarly, LuxR-type DNA binding domains have demonstrated clear inhibition of activity when in association with the N-terminal regulatory domains (*103, 104, 106, 118, 119*), but the mechanism for phosphorylation-induced release of this inhibition remains unclear. In this study, we focus on testing the generality of the existing models

for LOV signaling as well as HTH regulation in EL222. We present evidence that this protein can sense blue light using canonical LOV domain photochemistry and undergo a reversible photocycle that is coupled to conformational changes extending throughout both domains of the full-length protein. In addition to light-coupled conformational changes, gel shift experiments demonstrate a corresponding set of changes in the expected function of DNA binding. This activity is abolished when EL222 decays back to the dark/ground state. With additional information provided by a dark state crystal structure and point mutations along the interdomain interface, we propose that EL222 functions similarly to the NarL family of response regulators through interdomain inhibitory contacts that are released to allow DNA binding following activation.

Materials and Methods

Cloning, Expression, and Purification of EL222

Plasmid encoding full-length EL222 was obtained from Dr. Roberto Bogomolni (UC Santa Cruz). DNA encoding full-length EL222 (residues 1-222), and a 13-residue N-terminal truncation form (residues 14-222) were subcloned into the expression vector pHis-Gβ1-Parallel1, a derivative of the pHis-Parallel1 vector (*29, 120*). Mutagenesis was carried out using QuikChange II XL from Stratagene according to manufacturer's instructions. *E. coli* were transformed and grown in either Luria-Bertani broth for unlabeled protein or M9 minimal media as described for individual experiments below. Cultures were grown at 37^{o}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 hours induction, cells were centrifuged and the resulting pellets resuspended in 50 mM Tris, 100 mM NaCl pH 8.0 at 4°C buffer and lysed by sonication. Lysates were clarified by centrifugation at 48,000 x g for 30 min. The resulting supernatant was loaded on a Ni^{2+}-NTA column, allowing for rapid affinity purification of His-Gβ1 tagged proteins by eluting with the same buffer containing additional 250 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-TEV protease (*121*) per 30 mg of fusion protein. Proteolysis reactions were allowed to proceed overnight at 4°C and stopped using a Ni^{2+}-NTA column to remove the His-Gβ1 tag and His_6-TEV protease. Post-cleavage, the resulting protein contains only three vector-derived residues, GEF at the N-terminus. TEV-cleaved protein was loaded onto a MonoQ column to allow for charge-based separation of small impurities. Highly pure protein was eluted with 50 mM Tris, 250 mM NaCl pH 8.0 buffer and exchanged into 50 mM sodium phosphate, 100 mM NaCl pH 6.0 buffer and concentrated to a final protein concentration of 100-250 μM.

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 unlabeled samples exposed only to red light for the preceding 24 hr, while lit state

spectra were obtained immediately after exposing sample to illumination from a photographic flash. Kinetic experiments monitored the return of the A_{450} signal (acquisition rate of 0.1 min^{-1} or 0.1 sec^{-1} yields same values) following illumination. Data points were fitted using a first order exponential rate equation to obtained the time constant (τ).

Flavin Cofactor Determination

Full length EL222 (1-222) was buffer exchanged into H_2O and concentrated to ~250 μM. A 2x volume of ethanol was added to the sample and boiled for 2 minutes in a boiling water bath. The denatured protein sample was blotted, along with 250 μM FMN and FAD standards, onto a silica gel-based thin layer chromatography plate and separation proceeded in a 12:3:5 n-butanol:acetic acid:H_2O solvent system. The resulting bands were visualized with a UV-lamp.

Limited Proteolysis

Unlabeled full length (1-222) EL222 retaining an N-terminal His_6 tag was used for limited proteolysis to allow for an internal control for proteolytic activity under both dark and lit state conditions. A 1:43 ratio (w/w) of chymotrypsin to protein was used in a single volume with subsequent samples collected from this larger quantity. 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 while lit state experiments were performed under constant irradiation produced by an ORIEL mercury lamp (model number 66902) at 50 mW power with a broad-band blue-green filter (Oriel no. 51970). There was no significant heating of the sample under these light conditions compared to dark state as measured by digital thermometer.

Circular Dichroism Spectroscopy

A total of 500 μl of 15 μM unlabeled full length EL222 (1-222) was used for each CD experiment. Dark state spectra were collected under complete darkness, while lit state spectra were recorded following exposure to photographic flash. CD data were collected using a wavelength range from 195 to 260 nm at 12ºC with 1.5 nm bandwidth and 3 s averaging time. For the lit state spectrum, data collection was paused every 9 s and the sample exposed to a photographic flash to ensure lit state protein. Final data were generated from an average of 3 repeats.

Solution Nuclear Magnetic Resonance Spectroscopy

$^{15}N/^{1}H$ HSQC Dark and Lit NMR Spectroscopy

All of the following NMR experiments utilize protein grown in M9 minimal media with 1 g/L $^{15}NH_4Cl$ as the sole nitrogen source. Subsequent additions to the M9 will be noted under specific experimental methods. Solution

NMR experiments were performed on Varian Inova 600 MHz spectrometers at 25°C, using nmrPipe (*83*) for data processing and NMRview (*84*) for analysis. Lit state HSQC spectra were acquired with a 488 nm 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 designed to hold external chemical shift standards inside a 5 mm NMR sample tube. This allowed 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. Each $^{15}N/^{1}H$ HSQC spectrum was recorded by preceding each transient in the experiment with a 50 mW 120 ms laser pulse during the 1.06 sec delay between transients (*29*).

Chemical Shift Assignments

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 was 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: HNCA (*122*), HN(CO)CA (*123*), HNCACB (*123*), HN(CO)CACB (*123*). 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 and confirmations were obtained using (H)C(CO)NH-TOCSY(*124*) 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 supplemented with both $^{15}NH_4Cl$ and $^{13}C_6$ glucose. One-half hour prior to induction 0.08 g/L α-ketovalerate and 0.05 g/L α-ketobutyrate were added to cultures.

Deuterium Exchange

A 550 μL sample of 615 μM protein (14-222) was lyophilized in 50 mM sodium phosphate, 100 mM NaCl, pH 6.0 buffer. Immediately prior to acquiring spectra, the lyophilized sample was rehydrated with 550 μL D_2O and placed in the magnet. Spectra were acquired as an automated series of 79-90 spectra taking ~20 min. each 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 due to the increased peak broadening observed upon illumination independent of deuterium exchange. Protection factors were then calculated based on standard methods (*125*).

Electrophoretic Mobility Shift Assay

We designated 45 base-pair lengths of DNA within 350 base-pairs of the gene start site to use in a gel shift assay. Oligos were staggered to cover all possible binding sites and were obtained from Integrated DNA Technologies. 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^{o}C in a heat block for 5 min., followed by slow cooling to room temperature over 1 hr. 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 GE Healthcare ProbeQuant G50 Microcolumns following manufacturer's instructions. Concentration was 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 dI-dC, 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, while 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 native 10% TAE gel at 4^{o}C for 2 hr at 100 V. 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 one hour and developed using FujiFilm FLA-5100 imaging system following exposure to phosphor imaging plates. Quantification of percent DNA bound was performed using FujiFilm MultiGauge v2.3 software (FUJIFILM Medical Systems USA, Inc.).

Spectroscopic Analysis of EL222

UV-visible Absorbance Spectroscopy, Kinetics and Flavin Determination

Secondary structure predictions combined with the conserved sequence motif GXNCRFL predicted EL222 would contain a LOV domain (*42*). Interaction of the flavin chromophore with protein leads to three absorbance peaks centered around 450 nm (*26*). As expected, EL222 demonstrates this fine structure as observed for other LOV-flavin complexes. Following illumination, this fine structure diminishes revealing isosbestic points at 330 nm, 384 nm, and 407 nm (Fig.4-1a). Notably, we were unable to obtain a UV-visible absorption spectrum with completely photoconverted protein, likely due to the fast rate of recovery. Dark state recovery is observed as the characteristic absorbance profile returns with first order exponential kinetics. The τ for dark state recovery was determined to be 25.5 s at 25^{o}C. This rate decreased slightly with lower pH (28.7 s at pH 8.0) and increased with higher temperature (12.4 s at 32^{o}C) (data not shown). Given that LOV domains are commonly associated with either FMN or

FAD as their flavin cofactors, we wanted to experimentally determine this, using TLC analysis to establish that EL222 preferentially associates with FMN. No residual FAD or riboflavin was discernable on the plate following chromatography, suggesting little to no incorporation of these other cofactors.

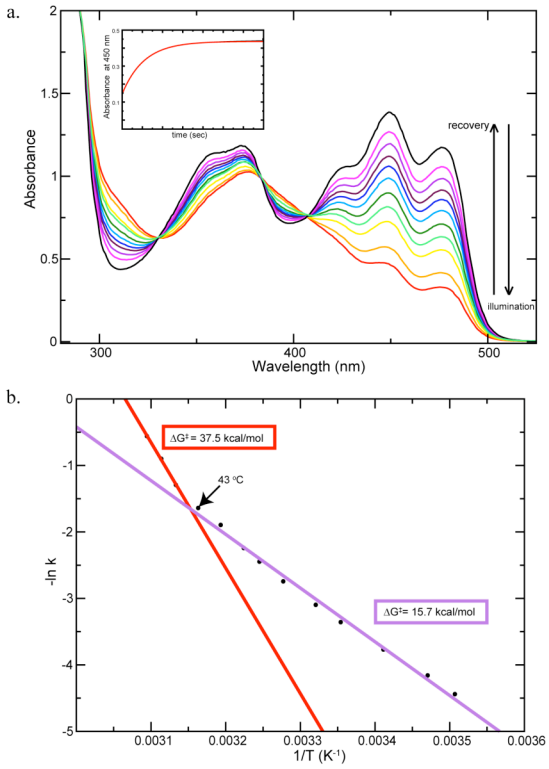

Fig. 4-1. UV-visible absorbance spectroscopy and Arrhenius analysis of EL222 a.) UV-visible absorbance spectrum of EL222. In the dark state (black) EL222 demonstrates the expected fine structure centered around 450 nm from protein-FMN interactions. Illumination induces covalent adduct formation (red). Return to the non-covalent state is observed as fine structure returns. Rate of dark state recovery is determined through curve fitting the absorbance recovery at 450 nm following illumination (inset). b.) Arrhenius plot of EL222. Kinetics of dark state recovery were measured as described for temperatures from 12°C to 50°C. A plot of the inverse ln(τ) versus inverse temperature yields an Arrhenius plot. From this we calculated the $\Delta G^{\ddagger}$ of the photochemistry using the formula: $\Delta G^{\ddagger} = -mR$ where m is the slope of the line and R is the gas constant. Fitting the line at temperatures 43°C and below yields a $\Delta G^{\ddagger}$ of 15.7 kcal/mol. This value doubles to 37.5 kcal/mol fitting temperatures above 43°C. The nonlinear behavior is suggestive of a conformational change at higher temperatures.

Arrhenius Analysis

To establish the temperature dependence of the photocycle, we performed kinetics measurements at temperatures ranging from 12^{o}C to 50^{o}C. From this data we were then able to construct an Arrhenius plot to determine the activation energy ($\Delta G^{\ddagger}$) of the photocycling chemistry. Interestingly, the Arrhenius plot is not linear over all temperatures, but instead, appears to become much steeper at higher temperatures (Fig. 4-1b) resulting in higher $\Delta G^{\ddagger}$ at these higher temperatures. Fitting the curve at all points less than 43^{o}C we can calculate a $\Delta G^{\ddagger}$ of 15.7 kcal/mol using the equation: $\Delta G^{\ddagger} = -mR$ where m is the slope of the line and R is the gas constant in units of cal K^{-1} mol^{-1}. This compares with fitting the final three points of the curve (temperatures 46^{o}C, 48^{o}C, and 50^{o}C) and obtaining a $\Delta G^{\ddagger}$ of 37.5 kCal/mol. Nonlinear Arrhenius behavior is suggestive of solvent exposure of hydrophobic surfaces resulting in a change in the heat capacity (*126, 127*). This would seem to indicate that higher temperatures (> 43^{o}C) cause a conformational change in the protein.

Low Resolution Studies of Light-Induced Conformational Changes

Circular Dichroism

As described above, many LOV domain proteins undergo structural changes following light-induced conformational changes. We employed two low-resolution methods to explore this conformational change from different perspectives. The first technique, circular dichroism spectroscopy, correlates

photocycle with global changes in secondary structure. EL222 demonstrated double minima at 208 nm and 222 nm, attributes of helical secondary structure (Fig. 4-2). Absorbance at these wavelengths is expected considering the helical content of the LOV domain in addition to the all-helical HTH domain. Unlike *As*LOV2, where the lit state CD spectrum reflects unfolding of the C-terminal α-helix (*51, 66, 72*), EL222 exhibited little difference between the dark and lit state CD spectra. Analysis with Dichroweb software (*128, 129*) indicated no significant change in the relative proportions of secondary structure elements between dark and lit states. This is consistent with the majority of the protein remaining intact with little to no global unfolding following illumination.

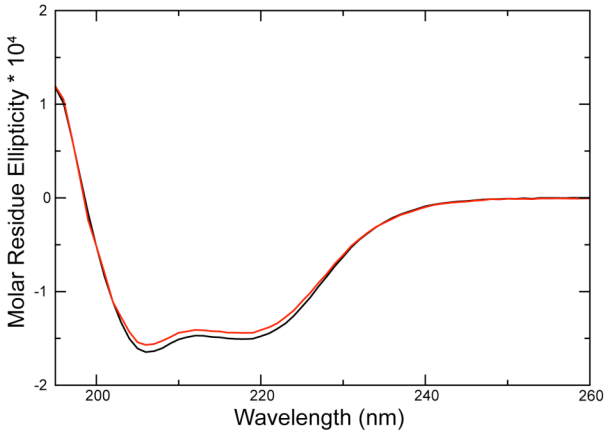

Fig. 4-2. Dark vs. light state circular dichroism spectroscopy
The double minima at 208 nm and 222 nm suggest well-folded protein with significant alpha helical character (black). Following illumination, little change was observed in the secondary structure characteristics of the protein (red).

Limited Proteolysis

Complementing this global analysis, we used limited proteolysis as a site-specific probe for conformational change. This experiment reports not only changes in secondary structure elements, but also any structural change that allows greater enzyme accessibility to a susceptible region of the protein. Both dark and lit state EL222 demonstrated an initial cleavage removing the N-terminal His_6-tag within the first five minutes (Fig. 4-3 species B) of chymotrypsin proteolysis, indicating chymotrypsin is not affected by light irradiation. Following this initial cut, there was little progression of proteolysis in the dark state, indicating a well-folded, structurally stable protein. Upon illumination there was much faster and more complete proteolysis, with little full-length protein remaining after 60 min. Notably, following chymotrypsin treatment in the lit state, three major species predominate. The higher molecular weight proteolytically stable fragment is consistent with an intact LOV domain (Fig. 4-3 species C). To ascertain the site at which chymotrypsin is cleaving, we submitted the 120 min. sample for mass spectrometry analysis and determined the proteolytic site to be Met159. Intriguingly, this residue is located in the linker between the LOV domain and C-terminal HTH domain (Fig. 4-3, 4-5a). Combined with the CD data, these results suggest light induced conformational changes are likely from structural reorientation of the two domains with respect to each other rather than increased disorder or unfolding within the protein.

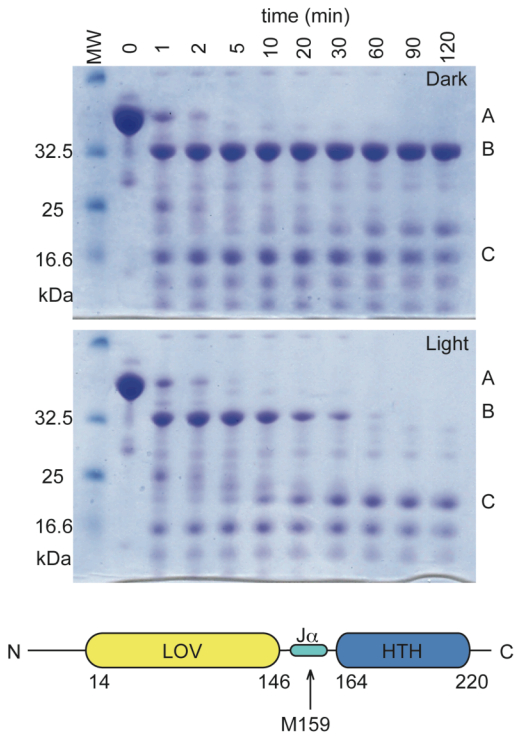

Fig. 4-3. Limited proteolysis in both the dark and lit states
Proteolytic digestion of EL222 by chymotrypsin is affected by the illumination state of the protein. Dark state protein is only partially cleaved after 120 min (top panel). In the illumnation state (bottom panels), EL222 is completely proteolyzed to a lower molecular weight species after 90 min. Mass spectroscopy determined the site of cleavage in the lit state is Met 159, located in the linker region between the two domains.

Movement between two domains, as suggested by limited proteolysis may, lead to an increased radius of the protein. Solution studies with other LOV domains have shown light-induced increases in volume (*7, 68*); therefore, we sought to compare hydrodynamic radii of both dark and lit state EL222. Unfortunately, the dark state recovery rate is too fast (τ at 12^{o}C is 85 sec) to maintain fully lit state protein during chromatography. To overcome this, we constructed a C75M mutation in the core of the LOV domain that allows covalent adduct formation, but does not spontaneously decay to a non-covalent state (*48*).

We characterized this mutant using the functional (DNA binding) assay described later in the document and confirmed it behaves as lit state wildtype EL222 (Fig. 4-4b). Gel filtration of this lit-state mutant indicates the hydrodynamic radius is not changed significantly following illumination (Fig. 4-4d). This data suggests that, while light produces increased solvent accessibility to the linker region of the protein, there is not sufficient interdomain separation to lead to an increased hydrodynamic radius.

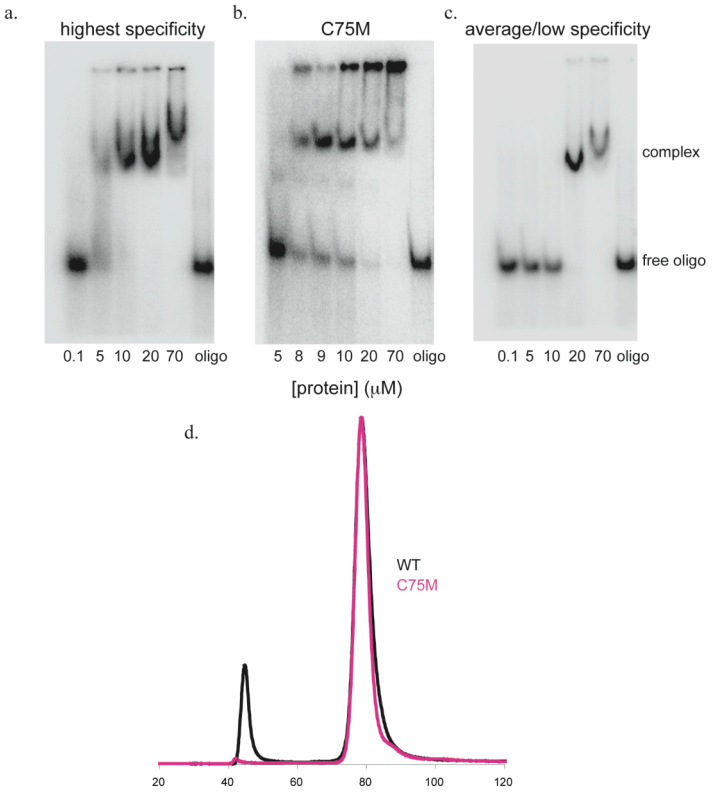

Fig. 4-4. Studies of C75M mutant to mimic lit state of protein
Mutation of the conserved cysteine involved in LOV photochemistry to a methionine prevents the covalent adduct from relaxing back to the dark state. To ensure this is true in EL222 I used an electrophoretic mobility shift assay to show that the protein bound DNA in the dark if exposed to light previously (b). This mutant also appeared to bind DNA with the same affinity as wildtype (a, c). (d) Gel filtration of the C75M mutant following illumination shows the protein has the samehydrodynamic radius as wildtype, indicating no dimerization.

Crystal Structure of Dark State EL222

Our collaborators obtained a 2.1 Å crystal structure of the dark state of the protein (Luecke et al., data unpublished) (Fig.4-5a). The LOV domain has the expected mixed α/β fold of a PAS domain and the HTH is a four-helix bundle at the C-terminus. The linker between the two domains is helical and interestingly, appears closely associated with the HTH domain (Fig. 4-5a). This is in contrast with some known LOV structures where helical extensions contact the β-sheet surface of the LOV domain (*7, 29*)(Fig. 1-3a-b). In these structures, the helices appear to be directly involved with light-induced activation of protein activity, suggesting perhaps the Jα of EL222 is not crucial for activation or signal transduction. The domain-domain interface is composed of a mix of hydrophobic residues and a hydrogen bond pair between the 4α helix of the HTH (R215) and the Hβ strand of the LOV domain (S137) (Fig. 4-5b). The N-terminal helical extension predicted by secondary structure prediction software (*130*) is not seen in the crystal structure. This is likely due to the addition of 32 residues to the N-terminus of the crystal construct in the form of a His_6 purification tag and other elements introduced through the DNA plasmid encoding the protein.

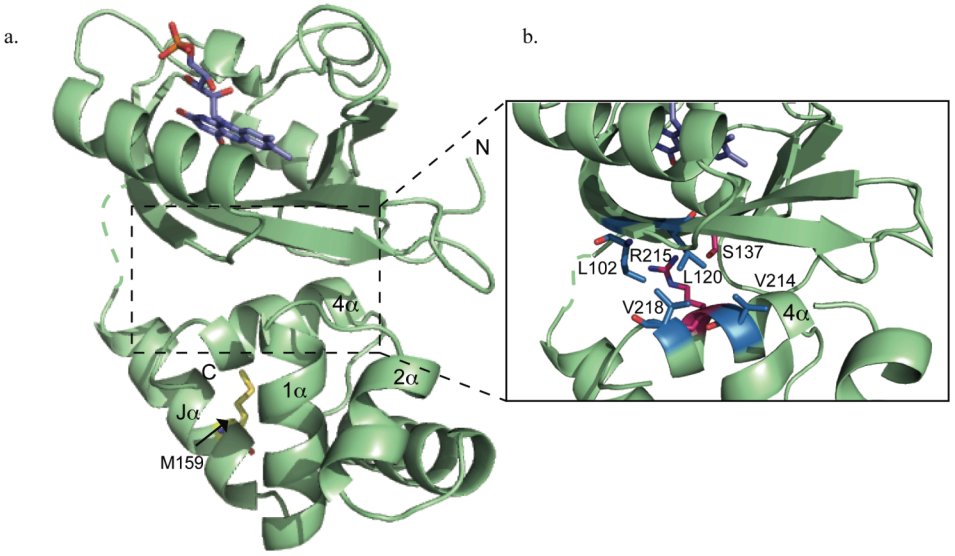

Fig. 4-5. Crystal and solution structures of dark state EL222
a.) Crystal structure of dark state EL222. Both the PAS and HTH domains contain the expected structural elements predicted from the primary sequence. The FMN is located in the hydrophobic pocket of the LOV domain on the β-sheet surface. The interface between the two domains is between the β-sheet surface of the LOV domain and helix 4α of the HTH. Met159 is shown in yellow b.) Zoom of the interdomain interface of EL222. Hydrophobic residues are shown in blue and the S137-R215 hydrogen bonding pair in magenta.

Solution Structural Studies of EL222

Circular dichroism and limited proteolysis indicate EL222 undergoes some form of light activated conformational change; however, these experiments give us only low-resolution detail. In order to probe this conformational change at the atomic level we undertook solution NMR spectroscopy. Using a combination of triple resonance and NOESY data, we assigned 81% of the protein backbone and 40% of residue side chains. These assignments, combined with through-space NOE data and TALOS prediction software (*4*), enabled us to determine the locations of secondary structure elements throughout the protein. From these analyses, the solution structure and crystal structure of dark state

EL222 appear to be similar (Fig. 4-6a-b). In addition to the canonical domain elements, we identified an N-terminal helical extension on the LOV domain termed the A'α helix.

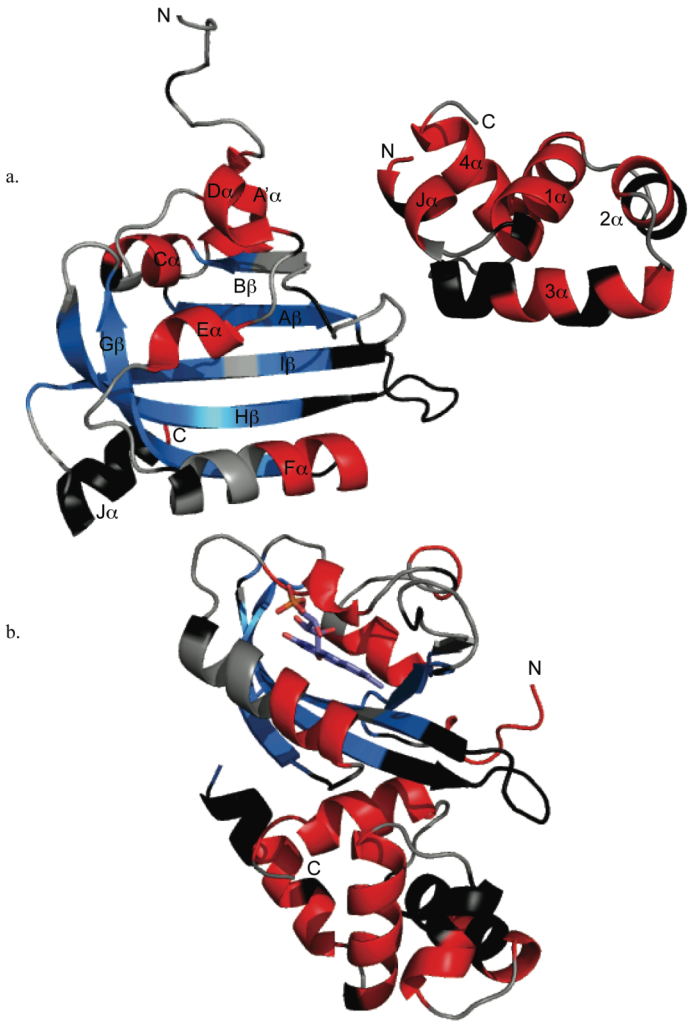

Fig. 4-6. Solution NMR secondary structure predictions
a.) The secondary structure elements predicted by solution NMR, TALOS (*4*) and modeled using the program MODELLER (*8*) are consistent with the crystal structure. TALOS predictions from NMR assignment are colored such that α-helical predictions are red, β-strand predictions are blue, and unassigned residues are black. Residues with no clear structural prediction are grey.
b.) TALOS predictions modeled onto the crystal structure.

Light Induced Structural Changes by NMR

Comparisons of dark versus light ^{15}N /^{1}H HSQC spectra show many chemical shift differences throughout EL222 (Fig. 4-7a) as well as some diminished peak intensities caused by differential line broadening. Chemical shift changes indicate the observed nuclei are experiencing new electronic environments, while line broadening is indicative of dynamic behavior. Complementing this, we observed chemical shift changes in the ^{13}C/^{1}H HSQC methyl region, (Fig. 4-7b) reflecting changes in the environment surrounding side chain methyl groups. Notably, in depth analysis of the methionine Cε region around 2 ppm ^{1}H, 15 ppm ^{13}C demonstrated one residue undergoes a light-induced change in chemical shift. Interestingly, all the methionine residues in EL222 are located outside the canonical LOV domain.

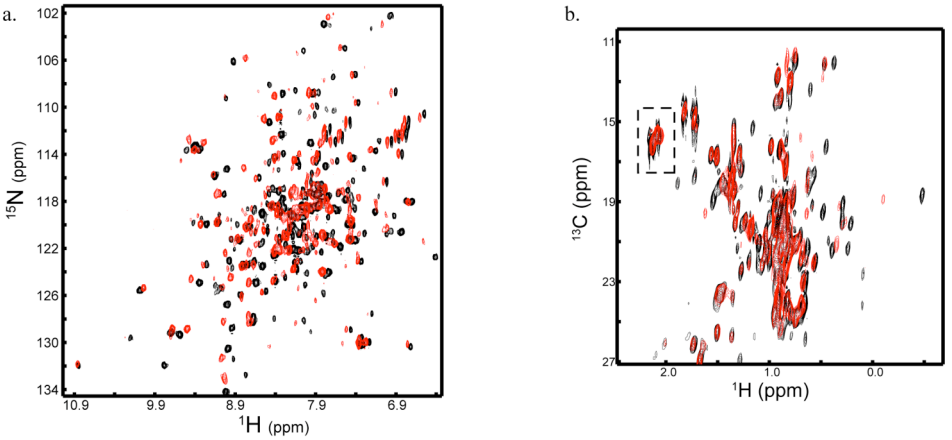

Fig. 4-7. Solution NMR HSQCs of dark versus lit state EL222
a.) ^{1}H/^{15}N HSQC shows many chemical shift changes between the dark (black) and lit (red) states of the protein. b.) While ^{15}N/^{1}H HSQC spectra report mainly on the backbone of the protein, the ^{13}C/^{1}H HSQC methyl region demonstrates chemical shift changes occuring within the side chains as well. Note the methionine chemical shift change in the lit state (dashed box).

While some chemical shift changes are expected due to light induced configurational changes of the FMN isoalloxazine ring and surrounding residues, EL222 exhibited far more changes than we would attribute to local effects near the cofactor, including sites > 16 Å away. Signal loss from differential line broadening indicated regions of the protein are undergoing light-induced movements on the μs-ms time scale. This type of behavior may be exhibited in areas that are released from otherwise constraining arrangements or are newly solvent exposed. Taken together, we interpret the chemical shift changes and differential line broadening of the lit state spectrum as indications of changes in structure and regional dynamics upon covalent adduct formation.

To obtain more site-specific information on the locations of these changes, we used a minimal chemical shift analysis of $^{15}N/^{1}H$ HSQC spectra acquired under dark and lit state conditions (*131*). From this analysis, we determined that chemical shift changes occur throughout the length of the protein (Fig. 4-8a). The clusters of changes occurring in the Eα helix are likely local effects due to changes in the FMN isoalloxazine ring and cysteinyl bond formation upon illumination. Residues undergoing significant chemical shift changes on the Gβ strand may be experiencing effects from the FMN cofactor, but given that the residues experiencing greatest change are on the solvent face of the sheet, it is likely these changes derive from another source. Interestingly, two clusters in the DNA binding domain appear significantly shifted. The first of these, in the loop region between helices 1α and 2α consists of Ala175, Gly177, and Arg179 (Fig.

4-8b). The second cluster locates to one face of helix 4α (Fig. 4-8b). These residues, Thr209, Asp212, and Arg215, as well as those clustered between helices 1α and 2α, likely point into the solvent in an isolated domain, refractory to significant chemical shift changes. The demonstrated chemical shift perturbations between dark and light, however, are more consistent with residues interacting with other areas of the protein. From the crystal structure, we can see that these residues are located at the domain-domain interface (Fig. 4-8b). The extent of chemical shift changes across the protein, beyond the LOV core, support our limited proteolysis findings that localized light-driven cysteinyl adduct formation triggers structural alterations throughout EL222. In addition, the significant perturbation of residues at the interface between the domains further supports an interdomain separation or other type of alteration in structure at this location, as suggested by limited proteolysis.

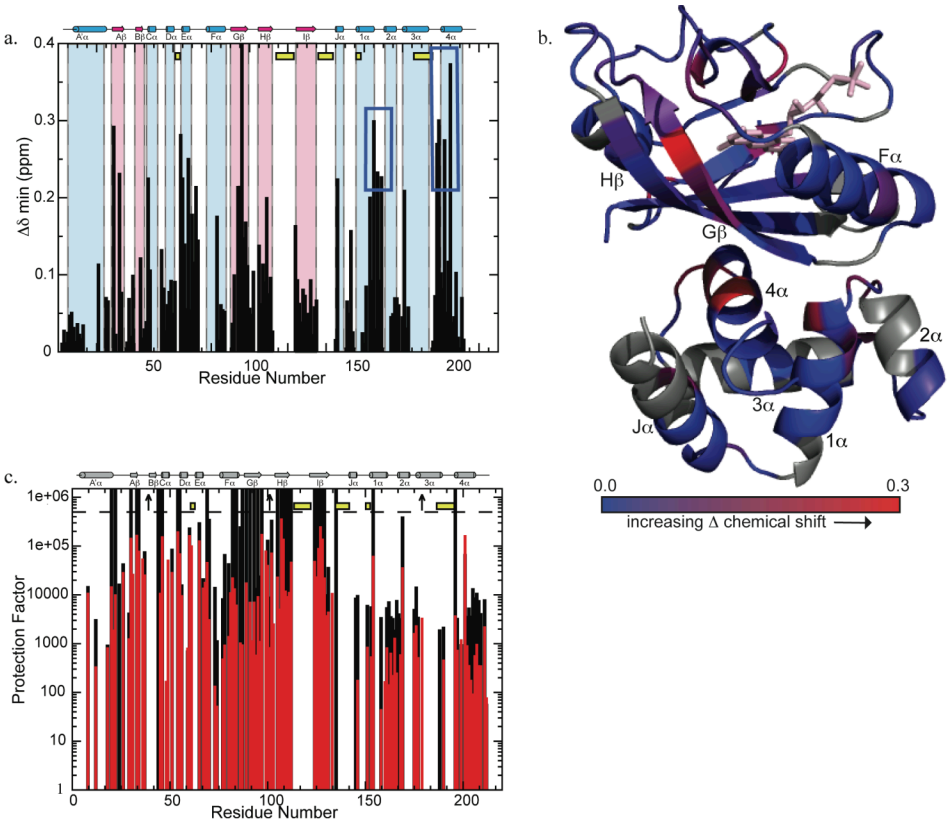

Fig. 4-8. Light induced changes monitored at the amino acid level
a.) Minimum chemical shift difference map demonstrates changes throughout full-length EL222. Blue boxed regions highlight the loop + α-helical face of the HTH that undergoes significant light induced changes. Yellow boxes denote unassigned residues b.) Minimal chemical shift difference mapped onto the crystal structure. Degree of chemical shift is minimal at dark blue and maximum at red regions. c.) Deuterium exchange shows significant loss in protection factor for several key areas across the protein including the β-sheet surface following illumination (black vs. red).

Deuterium Exchange

From limited proteolysis data combined with mass spectrometry analysis, we suspect some increased enzyme/solvent accessibilty in the linker region between the two domains following illumination. A higher resolution method to observe solvent accessibility is deuterium exchange. As protons exchange with deuterons in the buffer, we see loss of signal at these positions by $^{15}N/^{1}H$ HSQC.

Calculating the rate of exchange allows us to differentiate sites that exchange quickly, such as those in solvent exposed areas or unstructured regions of the protein, from those that exchange on a slower time scale due to sequestering from the solvent or their involvement in stabilizing interactions such as hydrogen bonds. We calculated protection factors (*125*) based on these kinetics such that the protection factor is inversely proportional to the rate of exchange. In the dark state, we can see there are numerous sites across the protein that appeared to exchange very slowly (Fig. 4-8c). Notably, the β-sheet surface of the LOV domain is very well protected as expected for PAS domains (*29, 65*) 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 demonstrate a decrease in protection factor, consistent with previous findings in *As*LOV2 (*29*). These findings indicate light-induced loss of protective interactions preventing solvent exchange at residues on these regions. An interdomain separation leading to solvent exposure at these sites, as suggested by limited proteolysis would be consistent with the observed increased deuterium exchange at these residues.

Testing the Ability of EL222 to Bind DNA

Limited proteolysis and NMR data are consistent with light-dependent structural changes in EL222, specifically, loss of interdomain contacts. As described above, many LOV proteins couple their activity levels with these types

of conformational changes (*7, 71*). Based on this precedent, we sought a method for monitoring protein activity in relation to illumination status. Unfortunately, little is known about this protein, therefore, while DNA binding is likely, the sequence to which EL222 binds is unknown. As an initial method to find a DNA binding sequence, we assumed EL222 might be autoregulatory through a feedback loop. Based on this assumption, we focused our efforts on evaluating the binding ability of the protein to DNA sequences located up to 350 bp upstream of the EL222 gene. Of 21 overlapping 45 bp candidate sequences tested, none appeared to bind EL222 in dark state conditions. Interestingly, all demonstrated light-dependent protein binding at or above 70 μM protein (Fig. 4-9a), suggesting that this protein is capable of inherent non-specific binding at this concentration. Using this value as a positive control for titrations, two sequences appeared to bind EL222 at protein concentrations of 5 μM. The first sequence spans base pairs 983532-983577 of the genome and is: GGTAGGATCCATCGGGCAGTGCGGTCAGCGGCATGCCGGCAGCAG. The second sequence occurs between base pairs 983647-983692 and is: GGCCCCGAGGTCCAGCACCAACGCAGTCCCCTTTGGTACGCCGAC. In both instances DNA binding only occurred when the protein:DNA mix was incubated under white light, ensuring lit state protein (Fig. 4-9a). No binding occurred under dark state conditions, even at protein concentrations capable of non-specific DNA binding (Fig. 4-9b). Protein previously exposed to bright light for ~ten minutes, then allowed to recover to dark state overnight at 4^{o}C also

demonstrated no residual DNA binding activity, indicating the activity is completely reversible and light dependent (Fig. 4-9b). Notably, protein exposed to bright light for 30 min., allowed to recover, then exposed to light for another 30 min. precipitates out of solution.

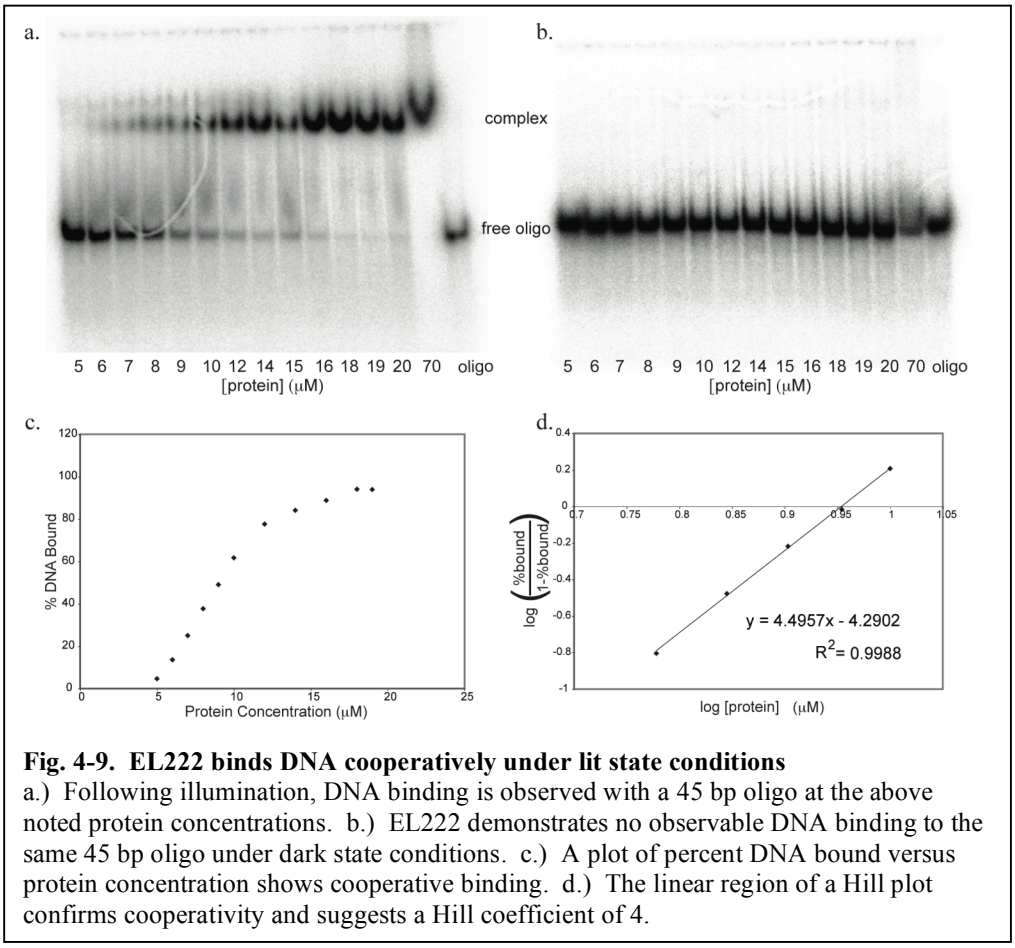

Fig. 4-9. EL222 binds DNA cooperatively under lit state conditions
a.) Following illumination, DNA binding is observed with a 45 bp oligo at the above noted protein concentrations. b.) EL222 demonstrates no observable DNA binding to the same 45 bp oligo under dark state conditions. c.) A plot of percent DNA bound versus protein concentration shows cooperative binding. d.) The linear region of a Hill plot confirms cooperativity and suggests a Hill coefficient of 4.

Given that our protein titrations demonstrated total conversion of DNA from 0 to 100% bound over a narrow concentration range, we hypothesized that EL222 may cooperatively bind DNA. The sigmoidal nature of a graph of percent DNA bound versus protein concentration (Fig. 4-9c) suggests EL222 may bind

cooperatively to the DNA; therefore we constructed a Hill plot and determined a Hill coefficient of approximately 4 (Fig. 4-9d). Four molecules could bind to a 45 bp oligo as two dimers given one dimer molecule would require ~25 bp if the DNA is not significantly bent (this accounts for both monomeric units binding the same side of the DNA, in the major groove). If dimers were binding on either side of the DNA, there should be sufficient space available. This result suggests that, although EL222 exists as a monomer in solution (Fig. 4-4d), upon binding DNA, oligomerization occurs.

From this data, we can conclude that EL222 demonstrates light-dependent DNA binding activity. While the DNA sequence used in these gel shift experiments bound with the highest affinity of all sequences tested, we strongly 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 (*14, 119, 132*). That being said, these data suggest that this DNA sequence retains its utility for assaying protein activity in future structural and/or functional experiments.

Effects of Point Mutations on DNA binding in Dark State

Based on the x-ray crystallographic data obtained from our collaborator, several prospective sites were designated for mutagenesis with the purpose of disrupting the interdomain interaction between LOV and HTH. The first of these mutations, L120K, breaks up a hydrophobic patch between the β-sheet surface of

the LOV domain and the fourth helix of the HTH domain (Fig. 4-10a). Initial studies to determine the effects on the protein showed no evidence for aggregration by gel filtration (Fig. 4-10b). NMR spectroscopy in both the dark and lit states suggested the protein was not as stable as wildtype under NMR conditions, precipitating during the course of the experiment, therefore little information regarding structural changes between the dark and lit states was obtained (Fig. 4-10d). Gel shift assay of the protein conducted under dark state conditions demonstrated DNA binding similar to wildtype under lit state conditions (Fig. 4-10c). Interestingly, this mutant appears to bind DNA with a slightly higher affinity than wildtype and also demonstrates an additional bound state at higher protein concentrations. The higher affinity may be due to more efficient and complete disruption of the interdomain inhibitory interaction by mutation than by maintaining a lit-state majority of wildtype protein with illumination. Due to the second protein-DNA complex, cooperativity was not assessed as with wildtype. Limited proteolysis of L120K using chymotrypsin showed little difference between the protein in the dark or lit state (Fig. 4-10e), both resembling the lit state of wildtype protein. This is consistent with stabilizing a lit state-like structure that maintains the ability to bind DNA in the dark state as shown by gel shift.

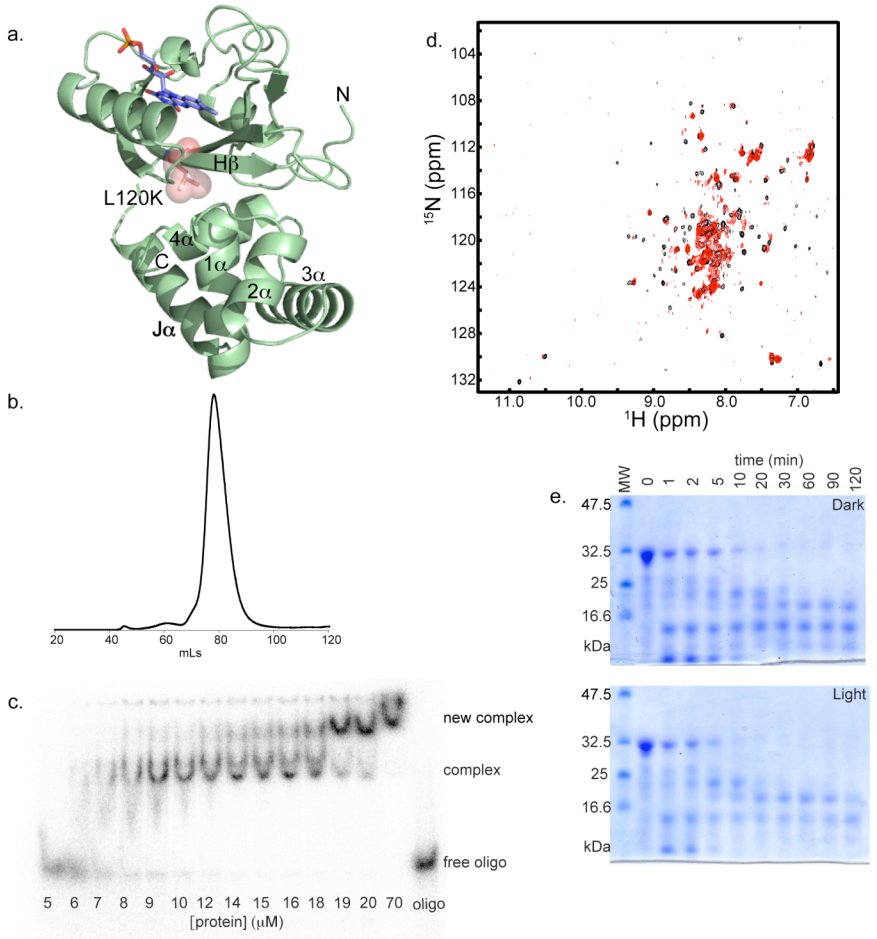

Fig. 4-10. L120K mutation successfully disrupts the interdomain interaction
a.) L120K is located on the Hb strand of the LOV domain and points down to the HTH. b.) L120K mutant appears to be monomeric, like wildtype, by gel filtration. c.) Electrophoretic mobility shift assay shows L120K binds DNA with the same affinity as wildtype EL222, however, it appears to have higher order binding at higher protein concentrations (labeled new complex) that wildtype does not demonstrate. d.) $^{15}N/^{1}H$ HSQC shows some dark (black) versus light (red) changes however due to the inherent instability of the protein the signal to noise is very low and little information can be gathered. e.) Limited proteolysis with chymotrypsin in the dark and lit states reveals little change in enzyme accessibility with both states looking similar to wildtype lit state protein.

Discussion

EL222 is a Light-Activated DNA Binding Protein

Evidence for Light-Induced Conformational Changes

Secondary structure predictions and sequence homology suggested EL222

is a two-domain protein consisting of a light-sensing LOV domain and a C-terminal LuxR-type DNA binding domain. We originally sought to confirm this prediction as well as to establish EL222 as a light-activated DNA binding protein. Initial characterization by UV-visible absorbance spectroscopy demonstrated reversible light induced covalent adduct formation between the associated FMN moiety and LOV domain. Following this result, we sought evidence for covalent adduct formation coupling to conformational changes using a combination of biophysical and biochemical approaches. Limited proteolysis and NMR spectroscopy demonstrated light-induced conformational changes at the tertiary structure and atomic levels respectively. Finally, gel shift experiments revealed these light-induced conformational changes lead to activation of an anticipated DNA binding activity, though we emphasize that the DNA sequence used in these experiments is likely not the highest affinity interaction site for EL222 *in vivo*.

We have further characterized the types of conformational changes occurring upon illumination through analysis and comparison of the data above. Limited proteolysis data suggest the linker region between the LOV and HTH becomes more proteolytically susceptible following covalent adduct formation. NMR data further support this observation with a combination of chemical shift data and deuterium exchange protection factors. Residues that were in close proximity in the dark state may become separated and/or more solvent exposed in the lit state leading to changes in their chemical shifts or solvent exchangeability. Minimal chemical shift analysis and deuterium exchange data highlight several

areas on both domains that appear to experience more significant changes in their environments following illumination. Crystallographic evidence shows these residues, the LOV β-sheet surface and between HTH helices 1α and 2α and along helix 4α (Fig. 4-8a-c), at the interdomain surface in the dark state. Light-induced disruption of these interdomain contacts would lead to both changes in these residues' environments (*inter*protein contacts become surface residues) as well as increased solvent accessibility. If light induces an interdomain separation, this is likely a more flexible region of the protein and may be more susceptible to temperature-dependent changes. The interdomain interface is composed of a series of hydrophobic residues (Fig. 4-5b) that may account for the curvature in the Arrhenius plot if exposed to solvent. Based on these data, we propose a model for EL222 activity similar to the NarL-type bacterial two-component response regulators.

Model for EL222 Activation

To develop a model for EL222 activation, it is helpful to compare it with other systems with a HTH domain. Specifically, the bacterial response regulators have been well studied and much data exists as to the mechanism for DNA binding regulation. Unfortunately, most of this data must be assembled from a mixture of experimental approaches including full-length and isolated domains. Those caveats being said, excellent work has been conducted that suggests common elements regulating DNA binding in various HTH-containing response

regulators.

A prime example of this work is with the two-domain response regulator protein NarL, from *Escherichia coli*. Transfer of a phosphate group to the N-terminal receiver domain likely leads to release of inhibitory contacts with the C-terminal LuxR-type HTH domain, thus allowing dimerization and DNA binding (*6, 102-104*). Similarly to NarL, EL222 adduct formation triggers conformational change that activates the protein, allowing for DNA binding (Fig. 4-9). Based on this model, we would expect some loss of interdomain contacts upon activation. As described above, limited proteolysis and NMR spectroscopic data support this model. Comparative studies of isolated HTH domains and full-length response regulator proteins using NMR and EPR suggest the LuxR subfamily may share a common interaction surface between the N- and C-terminal domains (*103, 104, 119*). Significant changes in chemical shift are consistently observed in the loop region between the first and second helices of the HTH domain as well as in the N-terminal portion of the final helix of the structure (*103, 118*). The crystal structure confirms a common face on the HTH domain for interdomain contacts (Fig. 4-5a-b). The orientation of the HTH with respect to the N-terminal domain in the inactive form of the NarL-type response regulators prevents DNA binding and oligomerization (*6*). Given the similar interface, we would expect EL222 to bind DNA and oligomerize only upon activation by illumination. Our gel shift data confirm that, indeed, EL222 binds DNA in a cooperative manner upon illumination.

An interesting similarity between EL222 and NarL is that both proteins appear to oligomerize only in the presence of *both* activating signal and DNA (*102*). Both proteins are monomers in solution without DNA (*6, 102*). This may be due to the interaction between the dimerization helix of the HTH (helix 4α) with the N-terminal domains in the inactive forms of the proteins (*6*)(Fig. 4-5a and Fig. 5-1a). This interaction is released following activation, allowing the helix to reorient and provide a dimerization interface. (*102-104, 118*) While we believe NarL represents an excellent model system for EL222 activity, the mode of activation between the two proteins is significantly different. Conformational changes in NarL rely on phosphotransfer from a separate sensor protein (NarQ or NarX) that detects the environmental signal (nitrate or nitrite) (*133, 134*). Unlike NarL, EL222's N-terminal domain is able to directly sense the environmental signal leading to conformational changes and protein activation. The combination of a N-terminal sensor domain and C-terminal HTH DNA binding domain suggest EL222 more closely resembles many of the bacterial quorum sensing proteins which directly detect a diffusible small ligand in the cell to activate or repress gene transcription (*14-16, 135-138*). We prefer the NarL model however, because the DNA binding activity of the quorum sensing proteins may likely be controlled through ligand-induced protein folding (*14*) rather than covalent bond formation as seen in NarL and EL222.

Generality of LOV domain signaling

A common element of LOV domain, and more generally PAS domain, signaling is the β-sheet surface. This surface has been identified as the site for both hetero- and homodimerization in many PAS and LOV domains (*40, 61, 63*) as well as a site for interdomain signal transfer. As demonstrated through work on *As*LOV2, YtvA and Vivid (*7, 29, 61*), light induced adduct formation in the core of the LOV domain is transduced to the β-sheet surface resulting in signal transduction to the effector domain. Our data show the β-sheet surface of the LOV domain as the site of interdomain interaction. The residues identified through deuterium exchange information and minimal chemical shift data that undergo the most significant perturbation are present at the domain-domain interface, suggesting the interdomain contacts are lost upon illumination. In addition to this information, truncation of the C-terminal HTH domain results in a LOV homodimer (data not shown), suggesting the HTH domain interacts with the same surface through which dimerization occurs. From a combination of biochemical and biophysical data, EL222 appears to transduce covalent adduct formation in the LOV core to interdomain movement through the LOV β-sheet surface.

Interdomain Signal Propagation Pathways

Secondary structural elements either N-terminal or C-terminal to the canonical LOV domain have recently been described and proposed as important

mediators of protein activation (*7, 67, 71*). Displacement of a C-terminal Jα helix upon illumination of *As*LOV2 appears to be directly responsible for kinase activation *in vitro* (*71*). In full-length phototropin, this Jα helix may act as a helical linker that helps to regulate LOV2 inhibition of kinase activity. While EL222 fails to demonstrate complete unfolding of any secondary structural elements, as observed by CD, it does contain an α-helix C-terminal to the LOV domain that functions as an interdomain linker. Interestingly, light-enhanced chymotrypsin proteolysis occurs within this linker helix, suggesting some structural alteration to allow greater enzyme accessibility in this region. Rather than displacement followed by unfolding, perhaps EL222's linker helix is displaced to a lesser degree such that unfolding does not occur, while still activating the binding activity of the HTH domain. The methionine identified as the chymotrypsin cut site is near the C-terminal end of the linker helix (Fig. 4-5a). Interestingly, this helix appears to be closely associated with the HTH domain rather than the N-terminal domain as see in the *As*LOV2 structure. Methionine 159 may be protected from the solvent through its interactions with the HTH domain. Perhaps movement of the HTH upon illumination leads to increased solvent exposure of this residue.

The kind of behavior demonstrated by the helical linker of EL222 more closely resembles the VVD and YtvA proteins. The single domain, fungal protein VVD contains an N-terminal α–helix/β-strand unit that packs against the β-sheet surface in a manner similar to Jα in *As*LOV2 (*7*)(Fig. 1-3b). Through a

combination of mutagenesis and *in vitro* work, it was shown that these elements repack against the surface of the protein following light-induced covalent adduct formation (*7*). YtvA is a bacterial protein containing an N-terminal LOV domain with a light-stable Jα helix. Unlike *As*LOV2, this Jα does not unfold following illumination as monitored by CD and tryptophan fluorescence (*69*). Structural studies of YtvA suggest this helical linker may, in fact, not be directly involved in signal propagation beyond simply tethering two domains together (*61*). While the role of the helical linker in EL222 remains unknown, it is clear from our studies that it is, at the very least, located in a region of the protein that undergoes changes in its environment upon illumination. Whether it acts as an otherwise unremarkable interdomain tether, or is more directly involved in signal propagation is yet to be determined.

Role of EL222 in Erythrobacter litoralis HTCC2594

LOV domain-containing proteins have been characterized in a wide variety of plants, fungus and bacteria. Within these organisms, they are implicated in phototropism, stomatal opening (*139*), chloroplast movement (*140, 141*), circadian clock function (*27, 142*), stress responses (*70*), branch development (*111*) and bacterial virulence (*143*). What then is the role for a blue-light receptor in a non-pathogenic marine bacterium? An initial assumption would be that this protein could be used for photosynthesis, however there is no evidence of photosynthetic machinery in this bacterium (*144, 145*).

While blue light is used directly as an energy source in photosynthetic organisms, it may also be employed as a method to identify other energy sources. Nutrients tend to be more abundant in coastal waters (*146*) and different wavelengths of light penetrate to varying depths (*147*). We postulate that EL222 may be used for depth determination, thereby giving information regarding possible nutrient content. This may be especially important for this species given its native habitat in the relatively nutrient-poor Sargasso Sea (*146*). As more information becomes available for *E. litoralis* HTCC2594, we will learn more about how this organism gleans energy from its environment and, hopefully, the role of blue light in this process.

CHAPTER 5

CONCLUSIONS

What We Already Knew

LOV Domains

Previous work on isolated LOV domains has focused on determining how signaling is communicated from the hydrophobic core, at the site of covalent adduct formation, to the surface of the domain. A likely path for this signal transduction is from the flavin ring to amino acids on the β-sheet immediately beneath the ring, to the domain surface. A mutational study of the LOV domain oxygen sensor Aer found that residues in contact with the FAD ring appeared to be involved in signal transduction (*148*). High-resolution crystal structures of both dark and lit LOV domain showed a highly conserved glutamine residue, located on the Iβ strand of the domain, maintains hydrogen bonds with the FMN isoalloxazine ring by rotating its side-chain between dark and lit states (*78, 80*). A follow-up infrared resonance spectroscopy study mutated this residue to a leucine to investigate how loss of the hydrogen bond contacts might affect light-dependent conformational changes (*81*). They found that light-induced structure changes were significantly decreased with this mutant, suggesting that this really was a key residue in intradomain signal transduction.

Studies on LOV domain communication to downstream effector domains has been hampered by difficulties in working with multidomain LOV-containing proteins. No structures exist of a full-length multi-domain LOV protein (until this

EL222 crystal structure), however, some information regarding interdomain interactions has been inferred through low-resolution biochemical and biophysical studies. Studies of phototropin1 activity have shown the LOV2 domain inhibits kinase activity *in trans in vitro* in the dark state; this inhibition is released following illumination (*57*). This is consistent with studies on full-length phototropins where kinase activity relies on disruption of a C-terminal Jα helix following illumination or by mutation (*71, 82*). These studies do not isolate the interaction surface of the LOV domain however. Mutation of a glutamic acid residue on the β-sheet surface of the bacterial LOV-containing protein YtvA leads to increased dark versus lit changes in the CD spectra of the full-length protein (*69*). These changes are also not of the same nature as those seen in wildtype YtvA. This implies the β-sheet surface is important for transmitting light-induced conformational change through the full-length protein. To fully investigate how LOV domains communicate with effector domains and the possible role of the β-sheet surface in this communication, more high-resolution studies must be conducted on multi-domain proteins.

HTH Domains

The four-helix HTH is a common domain element of bacterial response regulators. As such, much work has been done on these domains, however, obtaining structural information on full-length response regulator proteins has proven difficult. Biochemical evidence demonstrated phosphorylation is necessary for DNA binding activity in full-length proteins, (*101, 149, 150*).

However, HTH domains demonstrate constitutive DNA binding when isolated, often with higher affinity than the full-length proteins (*106, 116, 151, 152*). Solution studies demonstrated *in trans* interactions between isolated unphosphorylated receiver and HTH domains (*103, 104, 118*), as reflected in the NarL structure. These interactions are disrupted following phosphorylation of the receiver domain, corresponding to the activated, DNA binding, state of the protein (*103, 104, 118*). Low-resolution studies of the response regulator FixJ showed an inhibitory effect of the two domains with respect to each other that prevents DNA binding in the inactive state (*106, 119*). To date, only three structures of LuxR-type HTH response regulators have been solved (*6, 10, 153*). The high-resolution structures of full-length NarL and DosR have given us some insight into the exact nature of this receiver domain-based inhibition. Both proteins appear to make contact with the recognition and/or dimerization helices of the HTH (Fig. 5-1), thereby impeding the ability of the domain to interact with the DNA major groove or dimerize for higher affinity binding (*6, 10*). From this evidence, it is clear that phosphorylation must induce a structural change to release the HTH from inhibitory contacts and allow interaction with DNA.

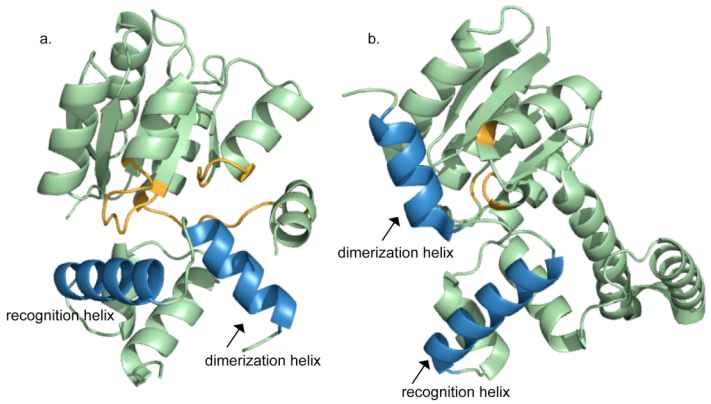

Fig. 5-1. Interdomain interfaces of LuxR-type response regulators
Response regulators make inhibitory contacts on their HTH DNA binding domains through loops and terminal regions of some helices (orange) in the receiver domain. The LuxR-type HTH domains interact with receiver domains through the recognition and dimerization helices (blue). a.) NarL crystal structure (1RNL) (*6*) b.) DosR crystal structure (3C3W) (*10*)

Impact of This Work

LOV domain intradomain signaling

My studies on the conserved glutamine residue on the β-sheet surface of LOV domains have corroborated the previous findings regarding decreased light-induced conformational changes with loss of hydrogen bonding contacts with the FMN ring. When leucine was substituted for glutamine, we were able to show that the domain maintained a dark-like structure regardless of covalent adduct formation, effectively decoupling cysteinyl bond formation from conformational changes. Our work led to studies of this mutant in full-length phototropin that showed the protein activity is significantly attenuated and resembles that of dark state wildtype protein regardless of illumination (*82*). We went on to study the effect of an asparagine substitution at this position. As described in a previous chapter, this shorter side chain may alter the structure of the LOV domain β-sheet

to maintain hydrogen bonds with the flavin ring, thus leading to the lit-like structure and reduced dark versus lit changes of the mutant domain.

In the course of the experiments, we also had the surprising finding that both mutations at the glutamine residue altered the length of the dark state recovery time. The determinants of dark state recovery are poorly understood so the glutamine studies have given some insight into this question. I believe the likely explanation for this relationship between the glutamine and dark state recovery kinetics is from decreased or increased structural strain placed on the domain in the lit state with each of the mutations. The leucine mutant may stabilize the lit state of the domain by decoupling adduct formation from conformational changes elsewhere in the domain. Conversely, the shorter side chain of the asparagine mutant may place greater strain on the domain in the lit state, leading to a destabilized structure and faster recovery to the noncovalent state.

Our studies of this glutamine residue were conducted in a LOV domain with a C-terminal helical extension shown to be important for interdomain signaling and protein activity (*71*). As described previously, the leucine mutant attenuates the kinase activity of full-length phototropin, thereby demonstrating this mutant affects both intradomain signal transduction and interdomain communication. This implicates the β-sheet surface in interdomain signaling, as expected from other known PAS domain proteins (*36-39, 69*). This study was able to extend our knowledge of LOV domain signaling beyond the isolated domain, a feat that has proven difficult in the past.

LOV domain interdomain signaling

While the glutamine study provided insight into how LOV domains may communicate to effector domains, the most appropriate way to study these interactions is with a full-length protein. EL222 is an ideal candidate for this research because it is relatively small, soluble and generally practical for use in structural studies. My studies, combined with the crystallographic data from collaborators, demonstrated the β-sheet surface is the site for interdomain contacts. This again supports the idea of a common interface for LOV domain communication. Unlike other LOV domains that have demonstrated significant light-induced structural changes (*7, 51*), this protein appears not to rely on a structural element beyond the LOV core. This makes it more similar to other isolated LOV domains that either do not have these types of additional structural elements or do not demonstrate significant changes in these elements following illumination (*61, 89*), suggesting the β-sheet face as a signaling surface more generally in the LOV field.

HTH interdomain signaling

Previous studies on LuxR-type HTH domains have been conducted in bacterial response regulator or quorum sensor proteins. From the few structures that exist, it has been suggested that the HTH activity is regulated through inhibitory contacts with the receiver domain. These contacts are lost following activation. One structure, StyR, appears to be in a conformation capable of DNA binding even in the unphosphorylated state. These conflicting structures

necessitate more structural information on LuxR-type proteins in general. EL222 adopts a conformation similar to NarL and DosR, where contacts between the N- and C-terminal domains likely inhibit DNA binding. Biophysical and biochemical data presented here support the idea of activation induced interdomain separation, as has been proposed for NarL and other response regulators. Similar to NarL the fourth helix of the HTH, the dimerization helix, makes contacts with the N-terminal domain in the inactive state. Dimerization is known to significantly increase the protein's affinity for DNA (*119, 154*) or be necessary for DNA binding (*102*) in these types of response regulators so it seems likely that without disruption of these contacts, DNA binding does not occur. Consistent with this, EL222 does not bind DNA in the inactive, dark state, where the dimerization helix is clearly inhibited by the LOV domain.

Interestingly, unlike NarL or DosR, the interdomain contact surface is the β-sheet face of the LOV domain. In both NarL and DosR, the inhibitory contacts are made through loops in the receiver domain (Fig. 5-1a-b). Notably, the receiver domains of these proteins are not PAS domains. My study of a novel LOV-HTH demonstrates the interdomain interface for LuxR-type HTH domains is the dimerization helix and recognition helices. Regardless of the interaction partner, this is the face making contacts in the inactive state of the proteins. The LOV or receiver domains modify the type of contacts made between the domains to maintain transduction through these HTH helices and their respective signaling surfaces. In the case of the receiver domains, phosphorylation induces structural changes in the loop regions making contact with the HTH (*105, 155-158*). For

LOV domains, the β-sheet surface makes the inhibitory contacts with the HTH. This study emphasizes the modular nature of protein domains. Regardless of interaction partners, each individual domain functions in a similar way to transduce activating signal and regulate protein function.

Bibliography

1. Tahirov, T. H., Sato, K., Ichikawa-Iwata, E., Sasaki, M., Inoue-Bungo, T., Shiina, M., Kimura, K., Takata, S., Fujikawa, A., Morii, H., Kumasaka, T., Yamamoto, M., Ishii, S., and Ogata, K. (2002) Mechanism of c-Myb-C/EBPβ cooperation from separated sites on a promoter, *Cell 108*, 57-70.
2. De Silva, R. S., Kovacikova, G., Lin, W., Taylor, R. K., Skorupski, K., and Kull, F. J. (2007) Crystal structure of the Vibrio cholerae quorum-sensing regulatory protein HapR, *J Bacteriol 189*, 5683-5691.
3. Ogata, K., Morikawa, S., Nakamura, H., Sekikawa, A., Inoue, T., Kanai, H., Sarai, A., Ishii, S., and Nishimura, Y. (1994) Solution structure of a specific DNA complex of the Myb DNA-binding domain with cooperative recognition helices, *Cell 79*, 639-648.
4. Cornilescu, G., Delaglio, F., and Bax, A. (1999) Protein backbone angle restraints from searching a database for chemical shift and sequence homology, *Journal of Biomolecular NMR 13*, 289-302.
5. Key, J., and Moffat, K. (2005) Crystal structures of deoxy and CO-bound *bj*FixLH reveal details of ligand recognition and signaling, *Biochemistry 44*, 4627-4635.
6. 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.
7. Zoltowski, B. D., Schwerdtfeger, C., Widom, J., Loros, J. J., Bilwes, A. M., Dunlap, J. C., and Crane, B. R. (2007) Conformational switching in the fungal light sensor Vivid, *Science 316*, 1054-1057.
8. Sali, A., and Blundell, T. L. (1993) Comparative protein modelling by satisfaction of spatial restraints, *J Mol Biol 234*, 779-815.
9. Key, J., Hefti, M., Purcell, E. B., and Moffat, K. (2007) Structure of the redox sensor domain of *Azotobacter vinelandii* NifL at atomic resolution: signaling, dimerization, and mechanism, *Biochemistry 46*, 3614-3623.
10. Wisedchaisri, G., Wu, M., Sherman, D. R., and Hol, W. G. (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.

11. Robinson, V. L., Wu, T., and Stock, A. M. (2003) Structural analysis of the domain interface in DrrB, a response regulator of the OmpR/PhoB subfamily, *J Bacteriol 185*, 4186-4194.
12. Yamaguchi, S., Kamikubo, H., Kurihara, K., Kuroki, R., Niimura, N., Shimizu, N., Yamazaki, Y., and Kataoka, M. (2009) Low-barrier hydrogen bond in photoactive yellow protein, *Proc Natl Acad Sci U S A 106*, 440-444.
13. Downward, J. (2001) The ins and outs of signalling, *Nature 411*, 759-762.
14. 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.
15. Urbanowski, M. L., Lostroh, C. P., and Greenberg, E. P. (2004) Reversible acyl-homoserine lactone binding to purified *Vibrio fischeri* LuxR protein, *J Bacteriol 186*, 631-637.
16. Castang, S., Reverchon, S., Gouet, P., and Nasser, W. (2006) Direct evidence for the modulation of the activity of the *Erwinia chrysanthemi* quorum-sensing regulator ExpR by acylhomoserine lactone pheromone, *J Biol Chem 281*, 29972-29987.
17. Posas, F., Wurgler-Murphy, S. M., Maeda, T., Witten, E. A., Thai, T. C., and Saito, H. (1996) Yeast HOG1 MAP kinase cascade is regulated by a multistep phosphorelay mechanism in the SLN1-YPD1-SSK1 "two-component" osmosensor, *Cell 86*, 865-875.
18. Wolanin, P. M., Thomason, P. A., and Stock, J. B. (2002) Histidine protein kinases: key signal transducers outside the animal kingdom, *Genome Biol 3*, REVIEWS3013.
19. Pawson, T., and Nash, P. (2003) Assembly of cell regulatory systems

through protein interaction domains, *Science 300*, 445-452.

20. Taylor, B. L., and Zhulin, I. B. (1999) PAS domains: internal sensors of oxygen, redox potential, and light, *Microbiol Mol Biol Rev 63*, 479-506.
21. Crews, S. T., Thomas, J. B., and Goodman, C. S. (1988) The Drosophila *single-minded* gene encodes a nuclear protein with sequence similarity to the *per* gene product, *Cell 52*, 143-151.
22. Reddy, P., Zehring, W. A., Wheeler, D. A., Pirrotta, V., Hadfield, C., Hall, J. C., and Rosbash, M. (1984) Molecular analysis of the *period* locus in Drosophila melanogaster and identification of a transcript involved in biological rhythms, *Cell 38*, 701-710.
23. Hoffman, E. C., Reyes, H., Chu, F. F., Sander, F., Conley, L. H., Brooks, B. A., and Hankinson, O. (1991) Cloning of a factor required for activity of the Ah (dioxin) receptor, *Science 252*, 954-958.
24. Wood, S. M., Gleadle, J. M., Pugh, C. W., Hankinson, O., and Ratcliffe, P. J. (1996) The role of the aryl hydrocarbon receptor nuclear translocator (ARNT) in hypoxic induction of gene expression. Studies in ARNT-deficient cells, *J Biol Chem 271*, 15117-15123.
25. Nambu, J. R., Lewis, J. O., Wharton, K. A., Jr., and Crews, S. T. (1991) The Drosophila *single-minded* gene encodes a helix-loop-helix protein that acts as a master regulator of CNS midline development, *Cell 67*, 1157-1167.
26. Christie, J. M., Salomon, M., Nozue, K., Wada, M., and Briggs, W. R. (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 U S A 96*, 8779-8783.
27. Imaizumi, T., Schultz, T. F., Harmon, F. G., Ho, L. A., and Kay, S. A. (2005) FKF1 F-box protein mediates cyclic degradation of a repressor of *CONSTANS* in *Arabidopsis*, *Science 309*, 293-297.
28. Borgstahl, G. E., Williams, D. R., and Getzoff, E. D. (1995) 1.4 A structure of photoactive yellow protein, a cytosolic photoreceptor: unusual

fold, active site, and chromophore, *Biochemistry 34*, 6278-6287.

29. Harper, S. M., Neil, L. C., and Gardner, K. H. (2003) Structural basis of a phototropin light switch, *Science 301*, 1541-1544.
30. Watts, K. J., Sommer, K., Fry, S. L., Johnson, M. S., and Taylor, B. L. (2006) Function of the N-terminal cap of the PAS domain in signaling by the aerotaxis receptor Aer, *J Bacteriol 188*, 2154-2162.
31. Hoersch, D., Otto, H., Joshi, C. P., Borucki, B., Cusanovich, M. A., and Heyn, M. P. (2007) Role of a conserved salt bridge between the PAS core and the N-terminal domain in the activation of the photoreceptor photoactive yellow protein, *Biophys J 93*, 1687-1699.
32. Gilles-Gonzalez, M. A., Ditta, G. S., and Helinski, D. R. (1991) A haemoprotein with kinase activity encoded by the oxygen sensor of *Rhizobium meliloti*, *Nature 350*, 170-172.
33. Gong, W., Hao, B., Mansy, S. S., Gonzalez, G., Gilles-Gonzalez, M. A., and Chan, M. K. (1998) Structure of a biological oxygen sensor: a new mechanism for heme-driven signal transduction, *Proc Natl Acad Sci U S A 95*, 15177-15182.
34. Hill, S., Austin, S., Eydmann, T., Jones, T., and Dixon, R. (1996) *Azotobacter vinelandii* NIFL is a flavoprotein that modulates transcriptional activation of nitrogen-fixation genes via a redox-sensitive switch, *Proc Natl Acad Sci U S A 93*, 2143-2148.
35. Meyer, T. E. (1985) Isolation and characterization of soluble cytochromes, ferredoxins and other chromophoric proteins from the halophilic phototrophic bacterium *Ectothiorhodospira halophila*, *Biochim Biophys Acta 806*, 175-183.
36. Hirose, K., Morita, M., Ema, M., Mimura, J., Hamada, H., Fujii, H., Saijo, Y., Gotoh, O., Sogawa, K., and Fujii-Kuriyama, Y. (1996) cDNA cloning and tissue-specific expression of a novel basic helix-loop-helix/PAS factor (Arnt2) with close sequence similarity to the aryl hydrocarbon receptor nuclear translocator (Arnt), *Mol Cell Biol 16*, 1706-1713.

37. Wang, G. L., Jiang, B. H., Rue, E. A., and Semenza, G. L. (1995) Hypoxia-inducible factor 1 is a basic-helix-loop-helix-PAS heterodimer regulated by cellular O2 tension, *Proc Natl Acad Sci U S A 92*, 5510-5514.
38. Lees, M. J., and Whitelaw, M. L. (1999) Multiple roles of ligand in transforming the dioxin receptor to an active basic helix-loop-helix/PAS transcription factor complex with the nuclear protein Arnt, *Mol Cell Biol 19*, 5811-5822.
39. Ema, M., Morita, M., Ikawa, S., Tanaka, M., Matsuda, Y., Gotoh, O., Saijoh, Y., Fujii, H., Hamada, H., Kikuchi, Y., and Fujii-Kuriyama, Y. (1996) Two new members of the murine *Sim* gene family are transcriptional repressors and show different expression patterns during mouse embryogenesis, *Mol Cell Biol 16*, 5865-5875.
40. Card, P. B., Erbel, P. J., and Gardner, K. H. (2005) Structural basis of ARNT PAS-B dimerization: use of a common beta-sheet interface for hetero- and homodimerization, *J Mol Biol 353*, 664-677.
41. Crosson, S., Rajagopal, S., and Moffat, K. (2003) The LOV domain family: photoresponsive signaling modules coupled to diverse output domains, *Biochemistry 42*, 2-10.
42. Briggs, W. R. (2007) The LOV domain: a chromophore module servicing multiple photoreceptors, *J Biomed Sci 14*, 499-504.
43. Salomon, M., Christie, J. M., Knieb, E., Lempert, U., and Briggs, W. R. (2000) Photochemical and mutational analysis of the FMN-binding domains of the plant blue light receptor, phototropin, *Biochemistry 39*, 9401-9410.
44. Swartz, T. E., Corchnoy, S. B., Christie, J. M., Lewis, J. W., Szundi, I., Briggs, W. R., and Bogomolni, R. A. (2001) The photocycle of a flavin-binding domain of the blue light photoreceptor phototropin, *J Biol Chem 276*, 36493-36500.
45. Iwata, T., Tokutomi, S., and Kandori, H. (2002) Photoreaction of the cysteine S-H group in the LOV2 domain of *Adiantum* phytochrome3, *J*

Am Chem Soc 124, 11840-11841.

46. Ataka, K., Hegemann, P., and Heberle, J. (2003) Vibrational spectroscopy of an algal Phot-LOV1 domain probes the molecular changes associated with blue-light reception, *Biophys J 84*, 466-474.
47. Neiss, C., and Saalfrank, P. (2003) *Ab initio* quantum chemical investigation of the first steps of the photocycle of phototropin: a model study, *Photochem Photobiol 77*, 101-109.
48. Kottke, T., Dick, B., Fedorov, R., Schlichting, I., Deutzmann, R., and Hegemann, P. (2003) Irreversible photoreduction of flavin in a mutated Phot-LOV1 domain, *Biochemistry 42*, 9854-9862.
49. Schleicher, E., Kowalczyk, R. M., Kay, C. W., Hegemann, P., Bacher, A., Fischer, M., Bittl, R., Richter, G., and Weber, S. (2004) On the reaction mechanism of adduct formation in LOV domains of the plant blue-light receptor phototropin, *J Am Chem Soc 126*, 11067-11076.
50. Zikihara, K., Iwata, T., Matsuoka, D., Kandori, H., Todo, T., and Tokutomi, S. (2006) Photoreaction cycle of the light, oxygen, and voltage domain in FKF1 determined by low-temperature absorption spectroscopy, *Biochemistry 45*, 10828-10837.
51. Harper, S. M., Neil, L. C., Day, I. J., Hore, P. J., and Gardner, K. H. (2004) Conformational changes in a photosensory LOV domain monitored by time-resolved NMR spectroscopy, *J Am Chem Soc 126*, 3390-3391.
52. Aravind, L., Anantharaman, V., Balaji, S., Babu, M. M., and Iyer, L. M. (2005) The many faces of the helix-turn-helix domain: transcription regulation and beyond, *FEMS Microbiol Rev 29*, 231-262.
53. He, Q., and Liu, Y. (2005) Molecular mechanism of light responses in *Neurospora*: from light-induced transcription to photoadaptation, *Genes Dev 19*, 2888-2899.
54. Alexandre, M. T., Arents, J. C., van Grondelle, R., Hellingwerf, K. J., and Kennis, J. T. (2007) A base-catalyzed mechanism for dark state recovery in the *Avena sativa* phototropin-1 LOV2 domain, *Biochemistry 46*, 3129-

3137.

55. Kottke, T., Heberle, J., Hehn, D., Dick, B., and Hegemann, P. (2003) Phot-LOV1: photocycle of a blue-light receptor domain from the green alga *Chlamydomonas reinhardtii*, *Biophys J 84*, 1192-1201.
56. Guo, H., Kottke, T., Hegemann, P., and Dick, B. (2005) The phot LOV2 domain and its interaction with LOV1, *Biophys J 89*, 402-412.
57. Matsuoka, D., and Tokutomi, S. (2005) Blue light-regulated molecular switch of Ser/Thr kinase in phototropin, *Proc Natl Acad Sci U S A 102*, 13337-13342.
58. Ballario, P., Talora, C., Galli, D., Linden, H., and Macino, G. (1998) Roles in dimerization and blue light photoresponse of the PAS and LOV domains of *Neurospora crassa* white collar proteins, *Mol Microbiol 29*, 719-729.
59. Buttani, V., Losi, A., Eggert, T., Krauss, U., Jaeger, K. E., Cao, Z., and Gartner, W. (2007) Conformational analysis of the blue-light sensing protein YtvA reveals a competitive interface for LOV-LOV dimerization and interdomain interactions, *Photochem Photobiol Sci 6*, 41-49.
60. Lamb, J. S., Zoltowski, B. D., Pabit, S. A., Crane, B. R., and Pollack, L. (2008) Time-resolved dimerization of a PAS-LOV protein measured with photocoupled small angle X-ray scattering, *J Am Chem Soc 130*, 12226-12227.
61. Moglich, A., and Moffat, K. (2007) Structural basis for light-dependent signaling in the dimeric LOV domain of the photosensor YtvA, *J Mol Biol 373*, 112-126.
62. Nakasako, M., Matsuoka, D., Zikihara, K., and Tokutomi, S. (2005) Quaternary structure of LOV-domain containing polypeptide of *Arabidopsis* FKF1 protein, *FEBS Lett 579*, 1067-1071.
63. Nakasako, M., Zikihara, K., Matsuoka, D., Katsura, H., and Tokutomi, S. (2008) Structural basis of the LOV1 dimerization of *Arabidopsis* phototropins 1 and 2, *J Mol Biol 381*, 718-733.

64. Salomon, M., Lempert, U., and Rudiger, W. (2004) Dimerization of the plant photoreceptor phototropin is probably mediated by the LOV1 domain, *FEBS Lett 572*, 8-10.

65. Brudler, R., Gessner, C. R., Li, S., Tyndall, S., Getzoff, E. D., and Woods, V. L., Jr. (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.

66. Corchnoy, S. B., Swartz, T. E., Lewis, J. W., Szundi, I., Briggs, W. R., and Bogomolni, R. A. (2003) Intramolecular proton transfers and structural changes during the photocycle of the LOV2 domain of phototropin 1, *J Biol Chem 278*, 724-731.

67. Halavaty, A. S., and 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.

68. Losi, A., Kottke, T., and 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.

69. Losi, A., Ghiraldelli, E., Jansen, S., and Gartner, W. (2005) Mutational effects on protein structural changes and interdomain interactions in the blue-light sensing LOV protein YtvA, *Photochem Photobiol 81*, 1145-1152.

70. Avila-Perez, M., Hellingwerf, K. J., and Kort, R. (2006) Blue light activates the σB-dependent stress response of *Bacillus subtilis* via YtvA, *J Bacteriol 188*, 6411-6414.

71. 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.

72. Nash, A. I., Ko, W. H., Harper, S. M., and Gardner, K. H. (2008) A conserved glutamine plays a central role in LOV domain signal transmission and its duration, *Biochemistry 47*, 13842-13849.
73. Semenza, G. L. (1999) Regulation of mammalian O_2 homeostasis by hypoxia-inducible factor 1, *Annu Rev Cell Dev Biol 15*, 551-578.
74. Imaizumi, T., Tran, H. G., Swartz, T. E., Briggs, W. R., and Kay, S. A. (2003) FKF1 is essential for photoperiodic-specific light signalling in *Arabidopsis*, *Nature 426*, 302-306.
75. Liscum, E., and Briggs, W. R. (1995) Mutations in the *NPH1* locus of *Arabidopsis* disrupt the perception of phototropic stimuli, *Plant Cell 7*, 473-485.
76. Christie, J. M., Reymond, P., Powell, G. K., Bernasconi, P., Raibekas, A. A., Liscum, E., and Briggs, W. R. (1998) *Arabidopsis NPH1*: a flavoprotein with the properties of a photoreceptor for phototropism, *Science 282*, 1698-1701.
77. Christie, J. M., Swartz, T. E., Bogomolni, R. A., and Briggs, W. R. (2002) Phototropin LOV domains exhibit distinct roles in regulating photoreceptor function, *Plant J 32*, 205-219.
78. Crosson, S., and Moffat, K. (2002) Photoexcited structure of a plant photoreceptor domain reveals a light-driven molecular switch, *Plant Cell 14*, 1067-1075.
79. Freddolino, P. L., Dittrich, M., and Schulten, K. (2006) Dynamic switching mechanisms in LOV1 and LOV2 domains of plant phototropins, *Biophys J 91*, 3630-3639.
80. Crosson, S., and Moffat, K. (2001) Structure of a flavin-binding plant photoreceptor domain: insights into light-mediated signal transduction, *Proc Natl Acad Sci U S A 98*, 2995-3000.
81. Nozaki, D., Iwata, T., Ishikawa, T., Todo, T., Tokutomi, S., and Kandori, H. (2004) Role of Gln1029 in the photoactivation processes of the LOV2 domain in *adiantum* phytochrome3, *Biochemistry 43*, 8373-8379.

82. Jones, M. A., Feeney, K. A., Kelly, S. M., and Christie, J. M. (2007) Mutational analysis of phototropin 1 provides insights into the mechanism underlying LOV2 signal transmission, *J Biol Chem 282*, 6405-6414.
83. Delaglio, F., Grzesiek, S., Vuister, G.W., Zhu, G., Pfeifer, J., Bax, A. (1995) NMRPipe: A multidimensional spectral processing system based on UNIX pipes, *Journal of Biomolecular NMR 6*, 227-339.
84. Johnson, B. A., Blevins, R.A. (1994) NMR View: A computer program for the visualization and analysis of NMR data, *Journal of Biomolecular NMR 4*, 595-740.
85. Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice, *Nucleic Acids Res 22*, 4673-4680.
86. Gouet, P., Courcelle, E., Stuart, D. I., and Metoz, F. (1999) ESPript: analysis of multiple sequence alignments in PostScript, *Bioinformatics 15*, 305-308.
87. Yao, X., Rosen, M. K., and Gardner, K. H. (2008) Estimation of the available free energy in a LOV2-Jα photoswitch, *Nat Chem Biol*.
88. Ulrich, E. L., Akutsu, H., Doreleijers, J. F., Harano, Y., Ioannidis, Y. E., Lin, J., Livny, M., Mading, S., Maziuk, D., Miller, Z., Nakatani, E., Schulte, C. F., Tolmie, D. E., Kent Wenger, R., Yao, H., and Markley, J. L. (2008) BioMagResBank, *Nucleic Acids Res 36*, D402-408.
89. Fedorov, R., Schlichting, I., Hartmann, E., Domratcheva, T., Fuhrmann, M., and Hegemann, P. (2003) Crystal structures and molecular mechanism of a light-induced signaling switch: The Phot-LOV1 domain from *Chlamydomonas reinhardtii*, *Biophys J 84*, 2474-2482.
90. Iwata, T., Nozaki, D., Tokutomi, S., and Kandori, H. (2005) Comparative investigation of the LOV1 and LOV2 domains in *Adiantum* phytochrome3, *Biochemistry 44*, 7427-7434.
91. Christie, J. M., Corchnoy, S. B., Swartz, T. E., Hokenson, M., Han, I. S.,

Briggs, W. R., and Bogomolni, R. A. (2007) Steric interactions stabilize the signaling state of the LOV2 domain of phototropin 1, *Biochemistry 46*, 9310-9319.

92. Yamamoto, A., Iwata, T., Tokutomi, S., and Kandori, H. (2008) Role of Phe1010 in light-induced structural changes of the neo1-LOV2 domain of *Adiantum*, *Biochemistry 47*, 922-928.
93. Song, S. H., Dick, B., Penzkofer, A., and Hegemann, P. (2007) Photo-reduction of flavin mononucleotide to semiquinone form in LOV domain mutants of blue-light receptor phot from *Chlamydomonas reinhardtii*, *J Photochem Photobiol B 87*, 37-48.
94. Pabo, C. O., and Sauer, R. T. (1992) Transcription factors: structural families and principles of DNA recognition, *Annu Rev Biochem 61*, 1053-1095.
95. Steitz, T. A., Ohlendorf, D. H., McKay, D. B., Anderson, W. F., and Matthews, B. W. (1982) Structural similarity in the DNA-binding domains of catabolite gene activator and cro repressor proteins, *Proc Natl Acad Sci U S A 79*, 3097-3100.
96. Pelton, J. G., Kustu, S., and Wemmer, D. E. (1999) Solution structure of the DNA-binding domain of NtrC with three alanine substitutions, *J Mol Biol 292*, 1095-1110.
97. Wah, D. A., Hirsch, J. A., Dorner, L. F., Schildkraut, I., and Aggarwal, A. K. (1997) Structure of the multimodular endonuclease FokI bound to DNA, *Nature 388*, 97-100.
98. Giraldo, R., and Fernandez-Tresguerres, M. E. (2004) Twenty years of the pPS10 replicon: insights on the molecular mechanism for the activation of DNA replication in iteron-containing bacterial plasmids, *Plasmid 52*, 69-83.
99. Peat, T. S., Frank, E. G., McDonald, J. P., Levine, A. S., Woodgate, R., and Hendrickson, W. A. (1996) Structure of the UmuD' protein and its regulation in response to DNA damage, *Nature 380*, 727-730.

100. Wilson, K. P., Shewchuk, L. M., Brennan, R. G., Otsuka, A. J., and Matthews, B. W. (1992) *Escherichia coli* biotin holoenzyme synthetase/bio repressor crystal structure delineates the biotin- and DNA-binding domains, *Proc Natl Acad Sci U S A 89*, 9257-9261.
101. Stock, A. M., Robinson, V. L., and Goudreau, P. N. (2000) Two-component signal transduction, *Annu Rev Biochem 69*, 183-215.
102. 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.
103. Eldridge, A. M., Kang, H. S., Johnson, E., Gunsalus, R., and Dahlquist, F. W. (2002) Effect of phosphorylation on the interdomain interaction of the response regulator, NarL, *Biochemistry 41*, 15173-15180.
104. Zhang, J. H., Xiao, G., Gunsalus, R. P., and Hubbell, W. L. (2003) Phosphorylation triggers domain separation in the DNA binding response regulator NarL, *Biochemistry 42*, 2552-2559.
105. Birck, C., Mourey, L., Gouet, P., Fabry, B., Schumacher, J., Rousseau, P., Kahn, D., and Samama, J. P. (1999) Conformational changes induced by phosphorylation of the FixJ receiver domain, *Structure 7*, 1505-1515.
106. Da Re, S., Bertagnoli, S., Fourment, J., Reyrat, J. M., and Kahn, D. (1994) Intramolecular signal transduction within the FixJ transcriptional activator: in vitro evidence for the inhibitory effect of the phosphorylatable regulatory domain, *Nucleic Acids Res 22*, 1555-1561.
107. Sinha, A., Gupta, S., Bhutani, S., Pathak, A., and Sarkar, D. (2008) PhoP-PhoP interaction at adjacent PhoP binding sites is influenced by protein phosphorylation, *J Bacteriol 190*, 1317-1328.
108. Maris, A. E., Walthers, D., Mattison, K., Byers, N., and Kenney, L. J. (2005) The response regulator OmpR oligomerizes via beta-sheets to form head-to-head dimers, *J Mol Biol 350*, 843-856.
109. Gusa, A. A., Gao, J., Stringer, V., Churchward, G., and Scott, J. R. (2006)

Phosphorylation of the group A Streptococcal CovR response regulator causes dimerization and promoter-specific recruitment by RNA polymerase, *J Bacteriol 188*, 4620-4626.

110. Ballario, P., Vittorioso, P., Magrelli, A., Talora, C., Cabibbo, A., and Macino, G. (1996) White collar-1, a central regulator of blue light responses in *Neurospora*, is a zinc finger protein, *EMBO J 15*, 1650-1657.
111. Takahashi, F., Yamagata, D., Ishikawa, M., Fukamatsu, Y., Ogura, Y., Kasahara, M., Kiyosue, T., Kikuyama, M., Wada, M., and Kataoka, H. (2007) AUREOCHROME, a photoreceptor required for photomorphogenesis in stramenopiles, *Proc Natl Acad Sci U S A 104*, 19625-19630.
112. Vinson, C. R., Sigler, P. B., and McKnight, S. L. (1989) Scissors-grip model for DNA recognition by a family of leucine zipper proteins, *Science 246*, 911-916.
113. Berg, J. M. (1988) Proposed structure for the zinc-binding domains from transcription factor IIIA and related proteins, *Proc Natl Acad Sci U S A 85*, 99-102.
114. Miller, J., McLachlan, A. D., and Klug, A. (1985) Repetitive zinc-binding domains in the protein transcription factor IIIA from *Xenopus* oocytes, *EMBO J 4*, 1609-1614.
115. Henikoff, S., Wallace, J. C., and Brown, J. P. (1990) Finding protein similarities with nucleotide sequence databases, *Methods Enzymol 183*, 111-132.
116. Choi, S. H., and Greenberg, E. P. (1991) The C-terminal region of the *Vibrio fischeri* LuxR protein contains an inducer-independent lux gene activating domain, *Proc Natl Acad Sci U S A 88*, 11115-11119.
117. Kahn, D., and Ditta, G. (1991) Modular structure of FixJ: homology of the transcriptional activator domain with the -35 binding domain of sigma factors, *Mol Microbiol 5*, 987-997.
118. Donaldson, L. W. (2008) The NMR structure of the *Staphylococcus*

aureus response regulator VraR DNA binding domain reveals a dynamic relationship between it and its associated receiver domain, *Biochemistry 47*, 3379-3388.

119. Da Re, S., Schumacher, J., Rousseau, P., Fourment, J., Ebel, C., and Kahn, D. (1999) Phosphorylation-induced dimerization of the FixJ receiver domain, *Mol Microbiol 34*, 504-511.
120. Sheffield, P., Garrard, S., and Derewenda, Z. (1999) Overcoming expression and purification problems of RhoGDI using a family of "parallel" expression vectors, *Protein Expr Purif 15*, 34-39.
121. Blommel, P. G., and Fox, B. G. (2007) A combined approach to improving large-scale production of tobacco etch virus protease, *Protein Expr Purif 55*, 53-68.
122. Yamazaki, T., Lee, W., Revington, M., Mattiello, D. L., Dahlquist, F. W., Arrowsmith, C. H., and Kay, L. E. (1994) An HNCA pulse scheme for the backbone assignment of 15N, 13C, 2H-labeled proteins: application to a 37 kDa *trp* repressor-DNA complex., *Journal of the American Chemical Society 116*, 6464-6465.
123. Yamazaki, T., Lee, W., Arrowsmith, C. H., Muhandiram, D. R., and Kay, L. E. (1994) A suite of triple resonance NMR experiments for the backbone assignment of 15N, 13C, 2H labeled proteins with high sensitivity, *Journal of the American Chemical Society 116*, 11655-11666.
124. Gardner, K. H., Konrat, R., Rosen, M. K., and Kay, L. E. (1996) An (H)C(CO)NH-TOCSY pulse scheme for sequential assignment of protonated methyl groups in otherwise deuterated 15N, 13C-labeled proteins, *J. Biomol. NMR 8*, 351-356.
125. Bai, Y., Milne, J. S., Mayne, L., and Englander, S. W. (1993) Primary structure effects on peptide group hydrogen exchange, *Proteins 17*, 75-86.
126. Makhatadze, G. I., and Privalov, P. L. (1990) Heat capacity of proteins. I. Partial molar heat capacity of individual amino acid residues in aqueous solution: hydration effect, *J Mol Biol 213*, 375-384.

127. Privalov, P. L., and Makhatadze, G. I. (1992) Contribution of hydration and non-covalent interactions to the heat capacity effect on protein unfolding, *J Mol Biol 224*, 715-723.
128. Whitmore, L., and Wallace, B. A. (2008) Protein secondary structure analyses from circular dichroism spectroscopy: methods and reference databases, *Biopolymers 89*, 392-400.
129. Whitmore, L., and Wallace, B. A. (2004) DICHROWEB, an online server for protein secondary structure analyses from circular dichroism spectroscopic data, *Nucleic Acids Res 32*, W668-673.
130. Cole, C., Barber, J. D., and Barton, G. J. (2008) The Jpred 3 secondary structure prediction server, *Nucleic Acids Res 36*, W197-201.
131. Farmer, B. T., 2nd, Constantine, K. L., Goldfarb, V., Friedrichs, M. S., Wittekind, M., Yanchunas, J., Jr., Robertson, J. G., and Mueller, L. (1996) Localizing the NADP+ binding site on the MurB enzyme by NMR, *Nat Struct Biol 3*, 995-997.
132. Stevens, A. M., Dolan, K. M., and Greenberg, E. P. (1994) Synergistic binding of the *Vibrio fischeri* LuxR transcriptional activator domain and RNA polymerase to the *lux* promoter region, *Proc Natl Acad Sci U S A 91*, 12619-12623.
133. Walker, M. S., and DeMoss, J. A. (1993) Phosphorylation and dephosphorylation catalyzed in vitro by purified components of the nitrate sensing system, NarX and NarL, *J Biol Chem 268*, 8391-8393.
134. Schroder, I., Wolin, C. D., Cavicchioli, R., and Gunsalus, R. P. (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.
135. Schuster, M., Urbanowski, M. L., and Greenberg, E. P. (2004) Promoter specificity in *Pseudomonas aeruginosa* quorum sensing revealed by DNA binding of purified LasR, *Proc Natl Acad Sci U S A 101*, 15833-15839.
136. Weingart, C. L., White, C. E., Liu, S., Chai, Y., Cho, H., Tsai, C. S., Wei,

Y., Delay, N. R., Gronquist, M. R., Eberhard, A., and Winans, S. C. (2005) Direct binding of the quorum sensing regulator CepR of *Burkholderia cenocepacia* to two target promoters in vitro, *Mol Microbiol 57*, 452-467.

137. Lee, J. H., Lequette, Y., and Greenberg, E. P. (2006) Activity of purified QscR, a *Pseudomonas aeruginosa* orphan quorum-sensing transcription factor, *Mol Microbiol 59*, 602-609.
138. Minogue, T. D., Carlier, A. L., Koutsoudis, M. D., and von Bodman, S. B. (2005) The cell density-dependent expression of stewartan exopolysaccharide in *Pantoea stewartii* ssp. *stewartii* is a function of EsaR-mediated repression of the *rcsA* gene, *Mol Microbiol 56*, 189-203.
139. Kinoshita, T., Doi, M., Suetsugu, N., Kagawa, T., Wada, M., and Shimazaki, K. (2001) Phot1 and phot2 mediate blue light regulation of stomatal opening, *Nature 414*, 656-660.
140. Jarillo, J. A., Gabrys, H., Capel, J., Alonso, J. M., Ecker, J. R., and Cashmore, A. R. (2001) Phototropin-related *NPL1* controls chloroplast relocation induced by blue light, *Nature 410*, 952-954.
141. Kagawa, T., Sakai, T., Suetsugu, N., Oikawa, K., Ishiguro, S., Kato, T., Tabata, S., Okada, K., and Wada, M. (2001) *Arabidopsis NPL1*: a phototropin homolog controlling the chloroplast high-light avoidance response, *Science 291*, 2138-2141.
142. He, Q., Cheng, P., Yang, Y., Wang, L., Gardner, K. H., and Liu, Y. (2002) White collar-1, a DNA binding transcription factor and a light sensor, *Science 297*, 840-843.
143. 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.
144. Raymond, J., and Swingley, W. D. (2008) Phototroph genomics ten years

on, *Photosynth Res*.

145. 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.
146. Moran, M. A., and Miller, W. L. (2007) Resourceful heterotrophs make the most of light in the coastal ocean, *Nat Rev Microbiol 5*, 792-800.
147. Administration, N. O. a. A. (2006).
148. Repik, A., Rebbapragada, A., Johnson, M. S., Haznedar, J. O., Zhulin, I. B., and Taylor, B. L. (2000) PAS domain residues involved in signal transduction by the Aer redox sensor of *Escherichia coli*, *Mol Microbiol 36*, 806-816.
149. Parkinson, J. A. (1995) *Two-component Signal Transduction* ASM Press, Washington, DC.
150. Stock, J. B., Ninfa, A. J., and Stock, A. M. (1989) Protein phosphorylation and regulation of adaptive responses in bacteria, *Microbiol Rev 53*, 450-490.
151. 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.
152. Belcheva, A., and Golemi-Kotra, D. (2008) A close-up view of the VraSR two-component system. A mediator of *Staphylococcus aureus* response to cell wall damage, *J Biol Chem 283*, 12354-12364.
153. Milani, M., Leoni, L., Rampioni, G., Zennaro, E., Ascenzi, P., and Bolognesi, M. (2005) An active-like structure in the unphosphorylated StyR response regulator suggests a phosphorylation- dependent allosteric activation mechanism, *Structure 13*, 1289-1297.
154. Leoni, L., Ascenzi, P., Bocedi, A., Rampioni, G., Castellini, L., and Zennaro, E. (2003) Styrene-catabolism regulation in *Pseudomonas fluorescens* ST: phosphorylation of StyR induces dimerization and

cooperative DNA-binding, *Biochem Biophys Res Commun 303*, 926-931.

155. Halkides, C. J., McEvoy, M. M., Casper, E., Matsumura, P., Volz, K., and Dahlquist, F. W. (2000) The 1.9 Å resolution crystal structure of phosphono-CheY, an analogue of the active form of the response regulator, CheY, *Biochemistry 39*, 5280-5286.

156. Kern, D., Volkman, B. F., Luginbuhl, P., Nohaile, M. J., Kustu, S., and Wemmer, D. E. (1999) Structure of a transiently phosphorylated switch in bacterial signal transduction, *Nature 402*, 894-898.

157. Lee, S. Y., Cho, H. S., Pelton, J. G., Yan, D., Berry, E. A., and Wemmer, D. E. (2001) Crystal structure of activated CheY. Comparison with other activated receiver domains, *J Biol Chem 276*, 16425-16431.

158. Lewis, R. J., Brannigan, J. A., Muchova, K., Barak, I., and Wilkinson, A. J. (1999) Phosphorylated aspartate in the structure of a response regulator protein, *J Mol Biol 294*, 9-15.