PHOTOSENSING WITH LOV: MOLECULAR INSIGHTS INTO THE SHARED MECHANISM OF LIGHT REGULATED KINASE SIGNALING BETWEEN PLANT PHOTOTROPINS AND BACTERIAL TWO–COMPONENT SYSTEMS

APPROVED BY SUPERVISORY COMMITTEE

Kevin Gardner, Ph.D., Advisor

Elizabeth Goldsmith, Ph.D., Committee chair

Jose Rizo–Rey, Ph.D.

Melanie Cobb, Ph.D.

This thesis is dedicated to my parents

Pao–Hsien Liu & Sun–Mu Ko

ACKNOWLEDGEMENTS

First of all, I would like to thank my mentor Dr. Kevin Gardner. He always supports me with endless patient and encouragement regardless what mistakes I made. He keeps me motivated about science during my "darkest" years when I have to grow, purify, and perform all protein experiments without light. I would not have accomplished all these works without his guidance and unique sense of humor. I am very grateful of him dedicating so much time training me and teaching me how to be a better scientist.

I also would like to send out my gratitude to all my committee members, Dr. Elizabeth Goldsmith, Dr. Jose Rizo–Rey, and Dr. Melanie Cobb for their guidance and encouragement during my graduate career. In addition, I would also like to thank other UT Southwestern faculties, Dr. Michael Rosen, Dr. Carlos Amezcua, Dr. Diana Tomchick, Dr. Marie–Alda Gonzales, Dr. Xuewu Zhang, and Dr. David Chuang and Dr. Chung-I Chang for their advice and suggestion over these years.

I would like to give my thanks to all the current and previous members in the Gardner lab. I would like to thank Lori Neil and Gardner lab alumni Dr. Shannon Harper, who initiated the work and provided the framework for me to continue on the phototropin project. I would like to thank Abby Nash, who worked with Shannon during her rotation and later with me on the phototropin Q513 project. I would like to thank Dr. Qiong Wu for her countless help and

support during the time in the Gardner lab. I want to thank Dr. Fernando Corrêa who works closely with me during the EL346/EL368 project and I wish him and Dr. Giomar Rivera–Cancel the best for continuing these two–component system projects. I would also like to thank Jo Appleton and Darlene Reynolds for their administrative assistance and support, which have made my graduate life worry free. Additionally, I want to thank members in the Rosen lab and the Rizo–Rey lab for generating such a great and friendly environment for work and providing me with generous help and suggestions.

A special thanks goes to my pre–graduate mentor Dr. Daniel O'Leary. I would not have come to UTSW and met so many wonderful people without him. I want to thank him for his guidance and support, and for encouraging and believing in me.

Finally, I would like to thank my dear friends and family, especially Christine Yao, Anthony Chi, Kyle Brewer, my two sisters Wen–Min Leila Ko, Wen–Huei Vivian Ko, my mother Pao–Hsien Liu, and my father Sun–Mu Ko. They are the cheerleaders behind the one–woman team Ko, I would not have make it through without them.

PHOTOSENSING WITH LOV: MOLECULAR INSIGHTS INTO THE SHARED MECHANISM OF LIGHT REGULATED KINASE SIGNALING BETWEEN PLANT PHOTOTROPINS AND BACTERIAL TWO–COMPONENT SYSTEMS

by

WEN–HUANG KO

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

November, 2009

PHOTOSENSING WITH LOV: MOLECULAR INSIGHTS INTO THE SHARED MECHANISM OF LIGHT REGULATED KINASE SIGNALING BETWEEN PLANT PHOTOTROPINS AND BACTERIAL TWO–COMPONENT SYSTEMS

Publication No. _____

Wen–Huang Ko, Ph.D.

The University of Texas Southwestern Medical Center at Dallas, 2009

Supervising Professor: Kevin H. Gardner, Ph.D.

With the overall goal of understanding the basis of Light–Oxygen–Voltage (LOV) domain mediated signaling transduction pathways, my graduate research focused on characterizing the role of LOV domains in kinase signaling and the mechanism of signal propagation from LOV domains to a diverse set of downstream effectors. I started from the studies of the LOV domain regulated

serine/threonine kinase activity from plant phototropins, later comparing this with LOV domains light–regulated histidine kinases in marine bacteria. Though the plant and bacterial kinases have diverse protein folds and functions, they both share the common photosensing module. Using a combination of biophysical and biochemical techniques, I examined the properties of these light mediated proteins and provided molecular insights in which light signal is transmitted from the core to the surface of the LOV domain. This work also presented detailed biochemical and biophysical characterization of the full–length LOV domain–containing histidine kinase, providing a solid foundation for studying the regulation of the LOV domain in the context of full–length protein. Taken together, I hope to address both the mechanisms that LOV domains use to regulate their effectors, and further how histidine kinases are regulated by their sensory domains. These are the central questions in both the LOV domain and the histidine kinase domain fields, and are of central importance to signal transduction in general.

TABLE OF CONTENTS

PRIOR PUBLICATIONS

Jung, A., Domratcheva, T., Tarutina, M., Wu, Q., **Ko, W.H.**, Shoeman, R.L., Gomelsky, M., Gardner, K.H., Schlichting, I. (2005) Structure of a bacterial BLUF photoreceptor: insights into blue light-mediated signal transduction. *Proc. Natl. Acad. Sci.* 102(35): 12350-12355.

Ko, W.H., Nash, A.I., Gardner, K.H. (2007) A LOVely view of blue light photosensing. *Nat. Chem. Biol.* 3(7): 372-374.

Wu, Q., **Ko, W.H.**, Gardner, K.H. (2008) Structural requirements for key residue and auxiliary portions of a BLUF domain. *Biochmistry* 47(39): 10271-10280.

Nash, A.I.*, **Ko, W.H.***, Gardner, K.H. (2008)
A conserved glutamine plays a central role in LOV domain signal transmission and its duration. *Biochemistry* 47(52): 13842-13849.
*Co-first authors

LIST OF FIGURES

Chapter 1: General introduction

Chapter 2: Signal transduction mechanism in the *As*LOV2 domain

LIST OF TABLES

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

DHp, dimerization and histidine phosphotransfer domain

CA, catalytic and ATP-binding domain

HKs, histidine kinases

HPK, histidine protein kinase

Chapter 1 General introduction

1.1 Blue light sensing module: the LOV domain

From small prokaryotes to large complex organisms such as plants and animals, cells rely on their abilities to sense and response to a variety of stimuli to survive and adapt to the changes in the environments. Light is an example of such essential stimuli. It governs many cellular processes, such as cell attachment and virulence in bacteria (*1, 2*), phototropism in plants (*3*), and circadian rhythm in fungi (*4*). Photodetection is dependent on specialized photoreceptors, which can be classified into seven general families depending on the chromophore and protein fold (*5*). The specific responses listed above, bacteria infectivity, plant phototropism, and entrainment of the fungal circadian clock, are all controlled by a common photoreceptor domain called the Light–Oxygen–Voltage (LOV) domain (*3*). LOV domains utilize a flavin cofactor (FAD or FMN) to specifically sense blue light ($\lambda \sim 400$–450 nm) and undergo a self–contained photocycle that is dependent on a highly conserved cysteine residue located in the core of the LOV domain (*6*). Although it was first identified in the higher plant phototropins, LOV domains have also been found in bacteria and fungi, but have not been confirmed in animals at this point (*7*). Although the LOV domain is known to act as a biological light switch, the exact mechanism of how this small modular domain regulates a very wide range of downstream effector domains remains unclear. The focus of my research lies on the understanding of how a light signal

is propagated through the LOV domain, what is the role of light in the signaling transduction, and the details of the interdomain interaction between the LOV domains and their effectors.

1.1.1 The LOV domain is a subset of the environmental sensor PAS domain superfamily

Per–ARNT–Sim (PAS) domains are important signaling modules found in all kingdoms of life. They are capable of monitoring internal or environmental stimuli such as oxygen, redox potential, light, small ligands, and overall energy level of a cell via protein–protein interactions (*8*). Ligand binding or structural changes control protein–protein interactions. The specificity in sensing is determined by the cofactor that the PAS domains associated. PAS domains have been found to associate with heme, FMN, FAD, or in some cases, orphan ligands (Fig. 1.1B) (*8*). For example, the photoactive yellow protein (PYP) PAS domain binds to a 4'–hydroxycinnamic acid in the core and uses it to sense blue light (*9*). The PAS domain of an oxygen sensor, FixL, contains a heme cofactor (*10*). Another blue light sensing photoreceptor from plants, named phototropin1, contains two tandem PAS domains that binds and use a single flavin mononucleotide (FMN) to sense light (*11*). The PAS domain in the redox sensor, NifL, binds to flavin adenine dinucleotide (FAD) (*12*). Last but not least, the cardiac potassium ion channel, human–ether–a–go–go related gene (HERG), was

found with no ligand bound in the PAS domain. However, people speculate the existence of an orphan ligand that is yet to be identified (*13*).

The average sequence identity among PAS domains is less than 20%. but all three–dimensional structures determined to date are similar (*14*). The typical PAS domain fold consists of a five stranded, antiparallel β sheet, flanked to one or both sides with several α helices (*15, 16*). Interestingly, a similar combination of low sequence homology and conserved structure has also been observed in other protein modules such as the Pleckstrin homology (PH) domain superfamily (*17*). This suggests that it may be a preferred mechanism for evolution to structurally fine tune a common protein module instead of creating new protein folds to handle multiple cellular functions.

The LOV domain is defined as a subset of the PAS domain because the core of the LOV domain resembles the PAS domain core structure (Fig. 1.1A) (*16*). The major factor that distinguishes the LOV domain from other PAS domains is the conserved cysteine in the sequence, NCRFL (*6*). Although multiple repeats of PAS domains have been observed in many of the bacterial histidine kinases, the only proteins that containing two tandem LOV domains are the plant phototropins (*7*). Instead, most LOV–containing proteins have only a single LOV domain, which can couple to other sensory domains such as phytochrome or other type of PAS domains (*7*).

1.1.2 Diversity of the LOV domain

In addition to coupling to other sensory domains, LOV domains are also found to associate with a wild variety of other effector domains (e.g. histidine and serine/threonine kinase domains, STAS domain, GGDEF and EAL domains), demonstrating their ability to regulate diverse sets of biological functions (Fig. 1.2) (*7*). In general, the effector domains dictate the function of these photoreceptors. For instance, coupling to a DNA binding domain usually impliesthe protein is involved in the regulation of gene transcription, while coupling to the GGDEF and EAL effector domains implies that the protein regulates bacterial second messenger c–di–GMP levels (*7*).

1.1.3 LOV domain photocycle

UV–visible absorbance spectroscopy has been the standard method for easy initial characterization of the flavin–containing proteins. The FMN or FAD cofactor is non–covalently bound to the LOV domain in the dark state. Proteins of this type have a characteristic absorption profile, where three maxima are observed between 400 and 500 nm. The light reaction induces changes in the electronic state of the flavin isoalloxazine ring generate a covalent adduct between the flavin C4a carbon and the sulfur from the cysteine side chain on the LOV domain (Fig. 1.3) (*18*). A significant loss of the three–maxima peaks observed in the blue light region of the absorption spectrum is consistent with

flavin–cysteinyl covalent adduct formation (*19*). Although the forward photochemistry of the LOV photocycle remains controversial, recent EPR studies proposed a radical–pair reaction mechanism, which may served as the primary model for the LOV domains photoreaction (Fig. 1.3) (*18*).

1.1.4 The role of LOV domains in plant phototropins.

LOV domains were first identified in the plant phototropins as a blue light photoreceptor involved in the phototropism (3). Unlike other LOV domain containing proteins, phototropins contain two tandem LOV domains (Fig. 1.4). The importance of the role of each LOV domain have been tested with *in vitro* and in plants studies of phototropins containing mutations at the cysteine of either or both LOV domains. These studies have shown that the second LOV domain (LOV2) is the primary photosensory domain that is predominantly responsible for light–dependent control of kinase activity and subsequent phototropic function, while the function of the LOV1 domain remains unclear (*20*). Unlike the LOV2 domain, which functions as a monomeric entity, isolated LOV1 domains from several phototropins have been found to be dimers. This has led to the proposal that LOV1 domains mediate the dimerization of the plant photoreceptor phototropin (*21*).

1.1.5 The LOV domain structure and signaling mechanism

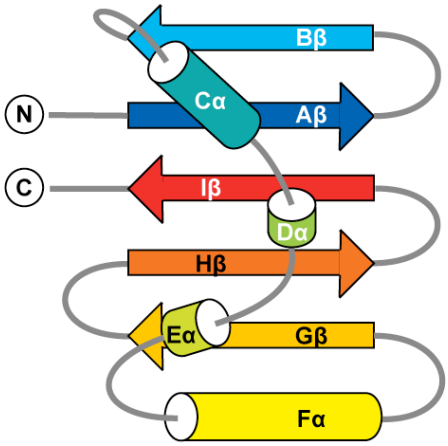
The 3D structures of several LOV domains have been determined. In all cases, they have a common PAS core consisting of an antiparallel β sheet with five β strands (denoted Aβ, Bβ, Gβ, Hβ, and Iβ) flanking with several α helices (denoted Cα, Dα, Eα, and Fα) (*22*) (Fig. 1.1A and Fig. 1.5). In general, the β sheet surface is identified to be an important site for protein–protein interaction or homodimerization in the PAS domains (*23*).

In addition to the PAS core, there are N– or C–terminal structural elements, which vary greatly in different LOV domains, outside of the PAS domain core (Fig. 1.5). In fact, these structural elements have been proposed to play an important role in the LOV domain regulation. As illustrated in Figure 1.5A for the phototropin1 LOV2 domain of *Avena sativa* (*24*), the canonical PAS core is followed immediately by a C–terminal amphipathic helix, called Jα. It associates with the β sheet surface in the dark state but unfolds and dissociates from the β sheet upon light illumination (*25*). This light induced conformational change has been shown to affect light enhanced phototropin kinase activity and an *in vivo* Cys to Ala mutation in the *At*Phot1 LOV2 domain that blocks the ability of the LOV domain to sense light also blocks phototropic and other phototropic–mediated responses (*20*) (Fig. 1.4). In the fungal LOV protein (Vivid), shown in Figure 1.5B, the N–terminal helices fold around to interact with the β sheet. It also undergoes conformational changes upon blue

light illumination, and these light activated conformational changes may be used to regulate light–activated gene expression (*26, 27*). However, not all additional structural elements interact with the β sheet surface. For example, in *Bs*YtvA LOV domain, the C–terminal Jα helix is pointed away from the protein, although this may be a function of crystallization (Fig. 1.5C). Together with the β sheet surface, they are involved in the protein dimerization instead (*28*).

Figure 1.1 Typical PAS domain fold and various types of PAS domains. Nomenclatures of the PAS domain showed in A. are adapted from Gong *et. al.* (*22, 29*). The PAS domain can bind to a wide variety of ligands as shown in B. Stimuli recognized by PAS domains include light, oxygen, redox potential, and small molecules. For instance, 4–hydroxycinnamic acid in photoactive yellow protein (PYP)(*9*), FAD in NifL (*12*), and FMN in the plant Phot1 photoreceptor (*30*) are used to sense light. Some PAS domains do not know their cofactor: the crystal structure of the PAS domain from the HERG cardiac K^+ channel was found without a cofactor (*13*).

A. PAS fold



B. Various types of PAS domains

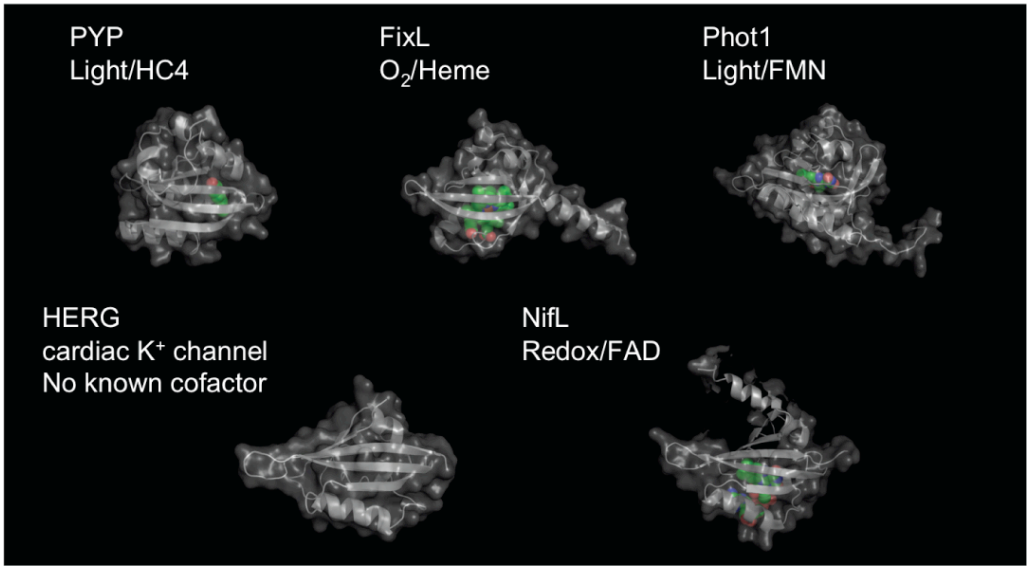


Figure 1.2 Diversity of the LOV domains adapted from Crosson *et. al.* (*7*).

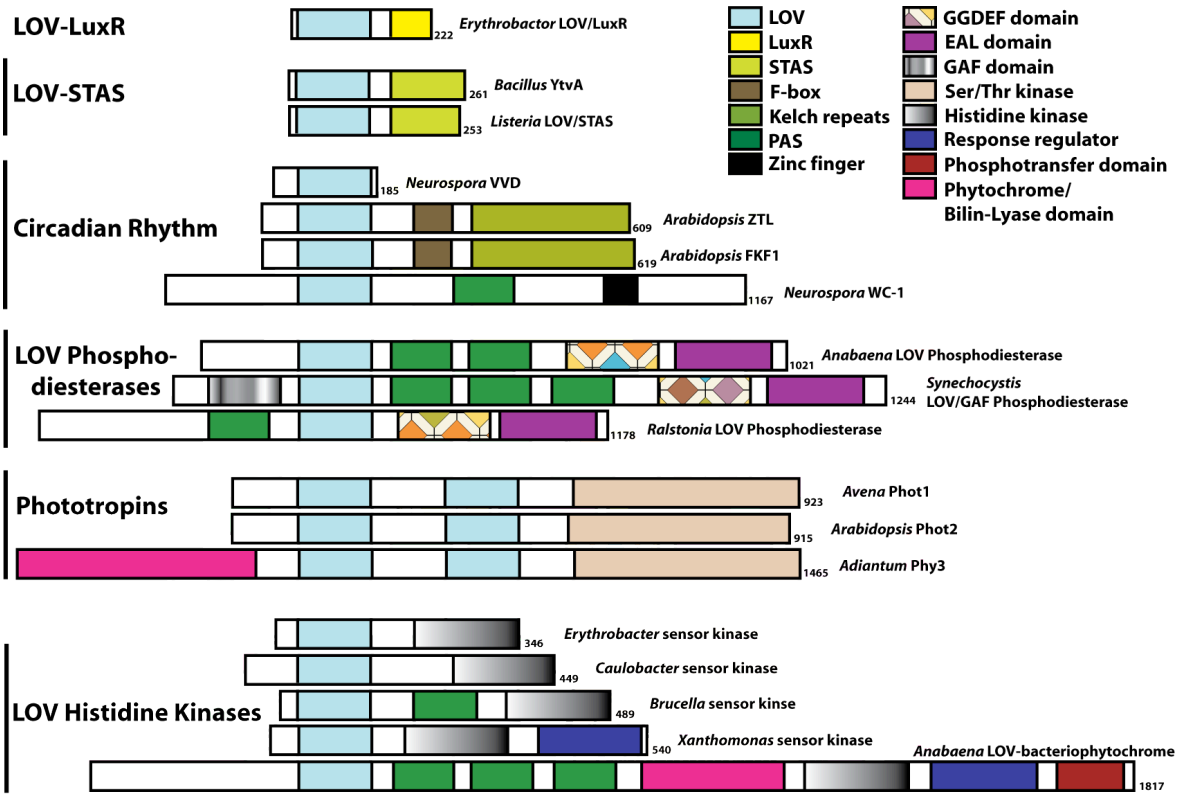


Figure 1.3 The mechanism of the LOV domain photocycle proposed by Schleicher E. *et. al.* (*18*). Upon light initiation, the noncovalently associated FMN undergoes some radical transition, which ultimately leads to a covalent adduct in the kinetically stable light state. The light state will spontaneously return to the dark state, ranging from seconds to hours and even days depend on the proteins that the LOV domain is coupled to.

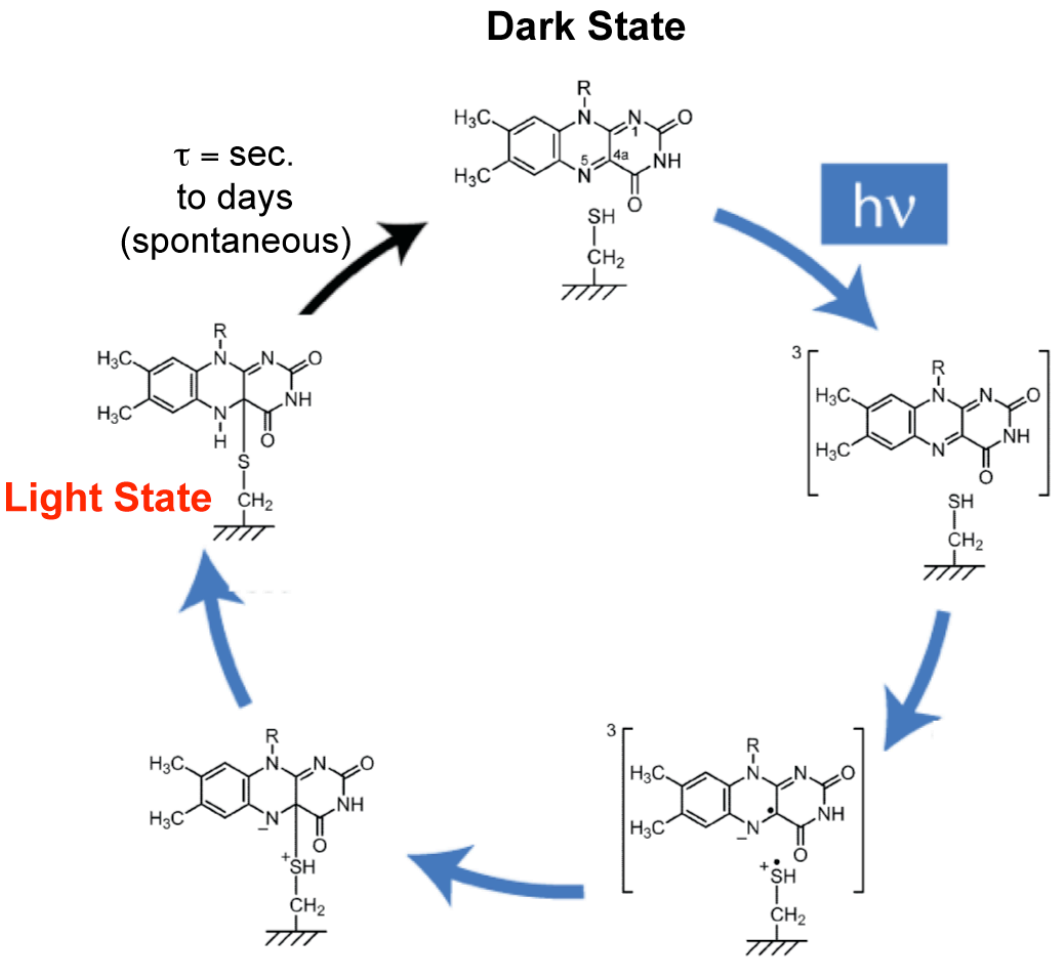


Figure 1.4 Important roles of LOV domains in the plant phototropin1. *In vitro* kinase assays and *in vivo* phototropic growth experiments from *A. thaliana* demonstrate that mutation of cysteine residue alone abolishes the LOV2 domain light sensing ability can affect both kinase phosphorylation and phototropism, data adapted from Christie *et. al*. (*20*). Grey bars across each of the LOV domain indicate a Cys to Aln mutation that abolishes the ability of the LOV to sense blue light.

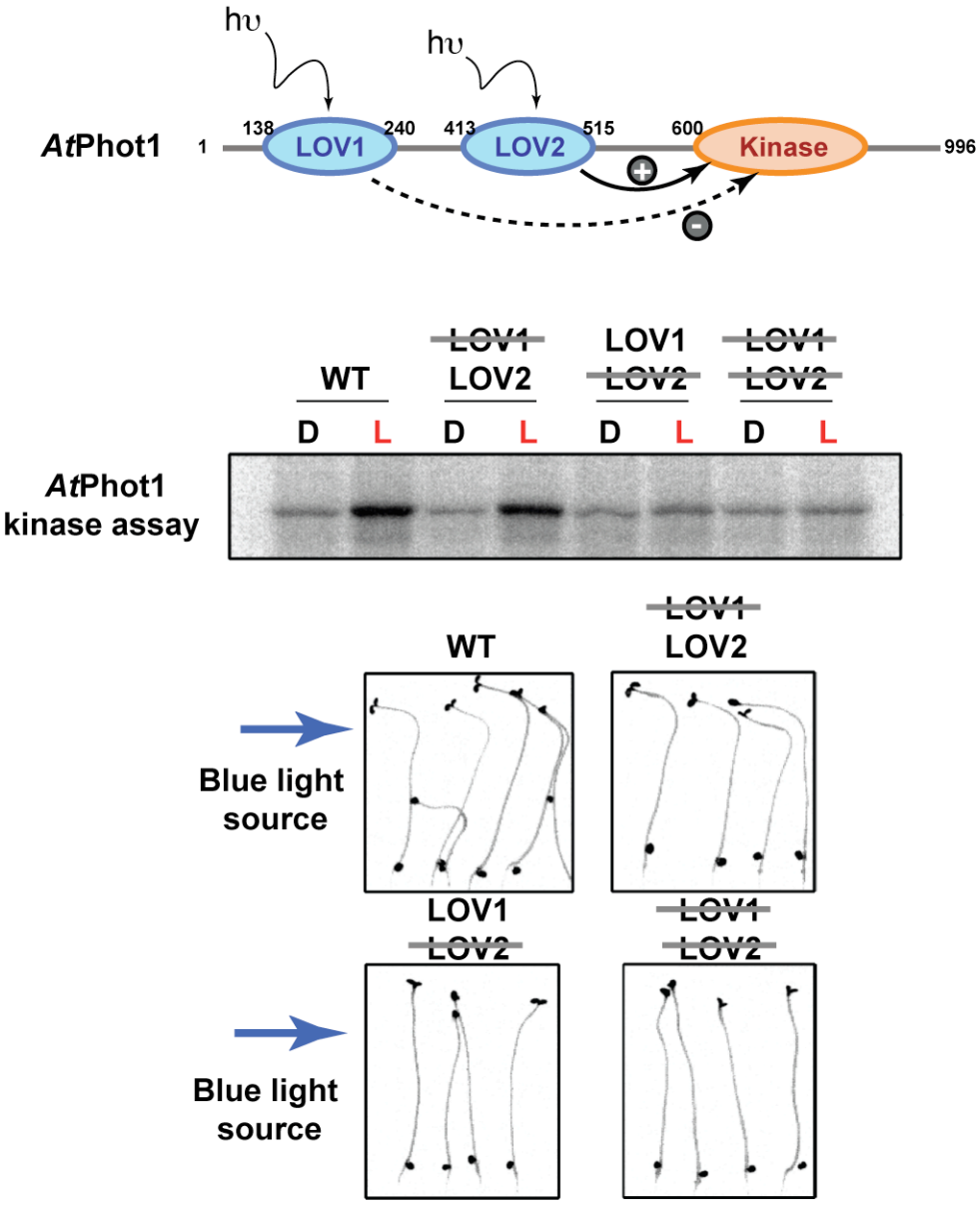
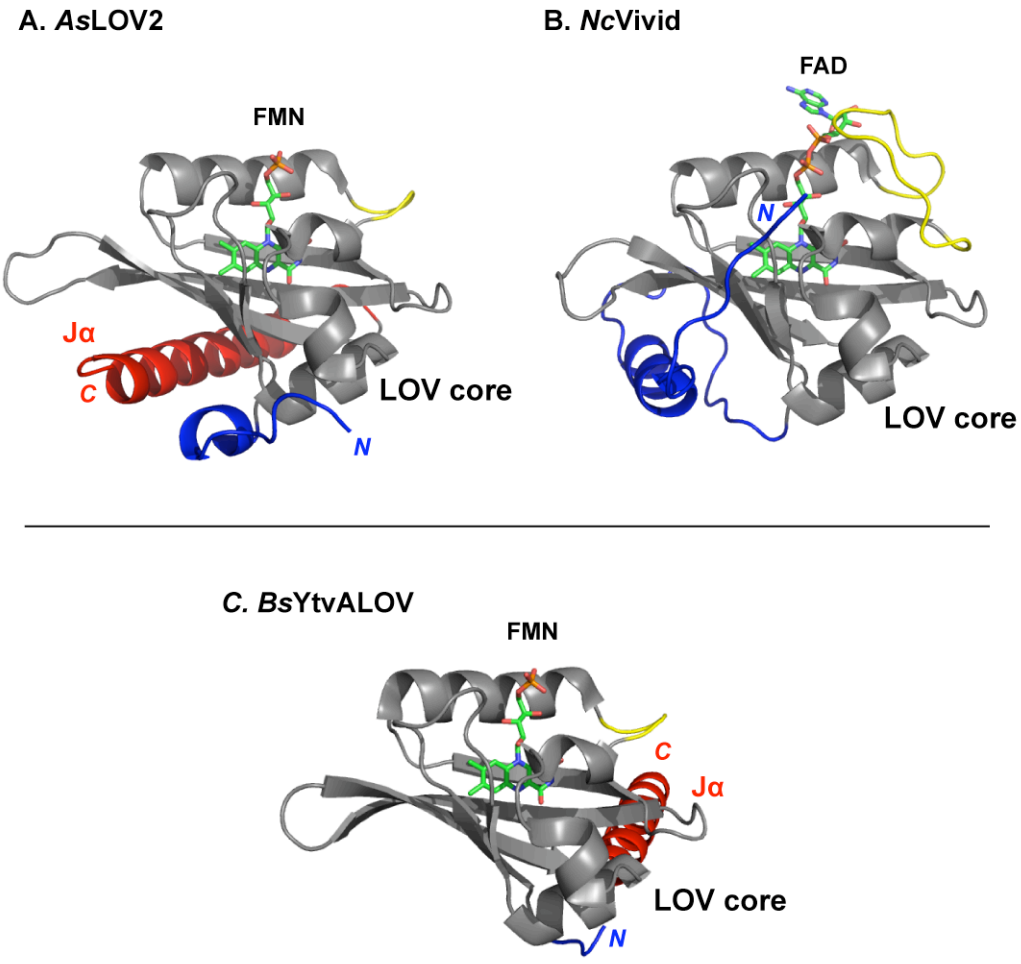


Figure 1.5 Structural comparisons of various LOV domains from A) *As*LOV2 domain (PDB: 2V0U), B) *Nc*Vivid LOV domain (PDB: 2PD7), and C) *Bs*YtvA LOV domain (2PR5). The N– and C–terminal structural elements outside of the canonical LOV core (shown in grey) were colored as blue and red, respectively. The C–terminal helix (Jα, in red) in the *As*LOV2 docks onto the β sheet surface while it points away from the β sheet surface in YtvA LOV domain. In *Nc*Vivid, although it is the N–terminal elements (helix and loop, in blue) also associate with the β sheet surface, the exact location is different from the other two proteins. The overall structures of the core are very similar except the loop region between the Eα and Fα (shown in yellow). The *As*LOV2 and YtvALOV contain a shorter loop, while *Nc*Vivid contains a much longer loop that interacts with its chromophore, FAD.



1.2 Two–component regulatory system in bacteria

Bacteria sense, respond, and adapt to a wide array of environmental signals via a general mechanism called two–component regulatory systems (TCSs) (*31*). TCSs fundamentally utilize two conserved protein superfamilies of His–Asp phosphorelay systems: a sensory histidine kinase (HK) that contains a highly conserved kinase (HK) domain and an output response regulator (RR) that contains a highly conserved receiver (REC) domain. Environmental signals are detected and transduced within the bacteria by these highly conserved domains from both proteins through fine–tuned processes such as autophosphorylation, phosphotransfer, and phosphatase activities (*32*). In these transduction pathways, environmental signals are initially detected by HKs, which then autophosphorylate and transfer this phosphoryl group to RRs. Protein phosphorylation is an important mechanism in the regulation of enzymatic activities in cells. Compared to protein phosphorylation on serine, threonine, or tyrosine, the discovery of the protein phosphorylation on histidine residues in proteins was delayed because most routine protein purification procedures called for acidic conditions, and the phosphoramidate bond of phosphohistidines are very prone to acid hydrolysis (*33*). Although they were originally discovered in bacteria, TCSs have been found in all kingdoms of life except for animals (*34*).

TCSs are present in both Gram negative and Gram positive, pathogenic and non–pathogenic bacteria. In pathogenic bacteria, they control expression of

toxins and other proteins important for pathogenesis, in addition to regulating other basic housekeeping functions (*35*). Two–component proteins typically constitute ~1% of encoded proteins in most eubacteria, with some bacteria containing up to as many as 200 two–component proteins, necessitating tight regulation to prevent potentially undesirable crosstalk (*36*). In the marine bacteria *Erythrobacter litoralis* HTCC2594, which is the focus of much work in Chapters 4 and 5, 50 two–component proteins can be found (out of a total of 3011 proteins) (*37, 38*). These pathways are potentially involved in regulating a variety of processes such as chemotaxis, osmoregulation, metabolism, or stress response based on sequence prediction and/or structure homology of the output domain (*39*).

The signaling mechanism in TSCs involves several distinct steps. First, the histidine kinase (HK) binds and uses ATP as a phosphate donor, undergoing autophosphorylation at a conserved histidine residue upon detection of an incoming stimulus. Second, the histidyl phosphate group is transferred to a conserved aspartate residue on a cognate response regulator (RR). Third, the phosphorylated RR (RR~P) interacts with target genes or other protein targets to generate the output response and activate corresponding cellular responses. Lastly, the signal is terminated by dephosphorylation of the activated response regulator, either by HK or by RR itself (*32*). The details of each component of

TCSs will be discussed independently in the next few sections, followed by a summary of a special type of light regulated histidine kinases.

1.2.1 First component: Histidine kinases (HKs)

A typical HK is composed of a diverse sensory domain and a conserved histidine kinase domain that contains conserved sequence motifs called the H, N, D, F, and G boxes (*35, 40, 41*). The H box contains a conserved histidine that is phosphorylated by the HK itself (Fig. 1.6B and Fig. 1.7A). The N, D, F, and G boxes are involved in ATP binding and constitute the *c*atalytic and *A*TP–binding (CA) domain in HKs (Fig. 1.6B and Fig. 1.7B). In the first step of the TCS signaling cascade, most HKs form homodimer and undergo transphosphorylation, where the CA domain of one subunit transfers the γ–phosphate from the ATP to the histidine residue in the H box of the other subunit (*32*). In more complex cases, HKs may contain multiple repeats of the sensory domains or additional accessory domains that are linked to other specific functions. Moreover, the second component (RR) may be connected to the kinase in sequence, making it a hybrid HK. HKs recognize one or an explicit set of RRs through the DHp domain (*36*). The specificity in the protein–protein interaction has been found to rely on a subset of coevolving residues between the DHp domain in the HKs and RRs, and mutagenesis studies have shown to successfully alter substrate specificity in some HKs (*42*).

1.2.1.1 Classification of histidine kinases

Apart from the conserved CA domain, HKs vary substantially in design. Essentially, they can be classified in three ways:

First, HKs are divided into three groups based on the localization of the sensory domain in the cell (Fig. 1.6A) (*43*). The first and largest group is the periplasmic–sensing HKs. It contains a signal peptide and transmembrane region(s) that separate the protein into a periplasmic N–terminal sensing domain and a cytoplasmic C–terminal kinase domain (Fig. 1.6A) (*43*). An example of this group is EnvZ that plays a central role in osmoregulation in *Escherichia coli* (*44*). The second group is called the intramembrane–sensing HKs, where stimulus perception can occur either by a combination of the transmembrane helices that is linked by short extracellular loops or just with the helices alone (see Fig. 1.6A) (*43*). An example of this group is a quorum–sensing HK, QseC, which is a bacterial adrenergic receptor found in the deadly pathogen enterohemorrhagic *E. coli* (EHEC) that activates virulence genes in response to interkingdom cross–signaling (*45*). Finally, the third and the second largest group, the cytoplasmic–sensing HKs, has either a membrane–anchored or a cytosolic sensory domain that responds to diffusible or internal stimuli, such as light, O_2 or H_2 (Fig. 1.6A) (*43*). For example, the light sensing EL346 from *E. litoralis* HTCC2594 (see Chapter 4) and the nitrogen regulatory kinase NtrB from Escherichia coli (*32*) are both

members of cytoplasmic HKs. Since the class II HKs mentioned above interact with another sensor protein intermolecularly, it can be considered either as a cytoplasmic sensing HK since the kinase itself is in the cytoplasmic or as a periplasmic sensing HK since its sensory counterpart is located in the periplasm.

Second, HKs are simply separated into two classes, class I and class II, based on the spacing between the H box and the CA domain (Fig. 1.6B) (*46*). In the class I HKs, which predominate among prokaryotes (*47*), the H box is located in the so called *d*imerization and *h*istidine–containing *p*hosphotransfer (DHp) domain that is immediately adjacent in sequence to the CA domain (Fig. 1.6B) (*32*). In the class II HKs, which is represented by the chemotaxis regulatory kinase CheA (*46*), the H box is located in the so called P1 domain, which is a N–terminal domain that is separated from the CA domain in sequence (Fig. 1.6B). In addition, there are substantial differences between the two classes. First, although both classes contain a dimerization domain and a CA domain, the class II HKs do not contains a sensory domain. The typical domain architecture of Class I HKs start with a N–terminal sensory domain that is followed by the DHp and CA domains (i.e. sensor–DHp–CA). The Class II HKs contain a P1, P2 domain followed by the dimerization domain, CA domain, and a P5 regulatory domain (i.e. P1–P2–dimerization–CA–P5) where the P5 domain is the one that is associated with the membrane–localized sensor module (Fig. 1.6B) (*46*). The second difference lies in the recognition site for the response regulators. In class I

HKs, the DHp domain, contains specific residues that would recognize specific RRs. In class II HKs, this site is located in the P2 domain. For the remainder of this dissertation, I will mostly focus on the discussion of the class I HKs.

The third way to classify HKs is based on the sequence identity in each conserved homology boxes: H, N, D, F, and G boxes. Sequence alignments of 384 HKs performed by Grebe and Stock have revealed 11 distinct subgroups in the histidine kinase superfamily (*33*). These homology boxes define subfamilies based on the flanking sequences around the invariant H, N, D, F, and G residues in each homology boxes. These flanking sequences are generally most conserved in the largest of the HKs subfamily, subfamily 1 (HPK_1), and are diverse in other subfamilies when using HPK_1 as the model for comparison. In some cases, the sequences may be so distorted that they can only be identified through detailed sequence alignments (*33*). EL346 (see Chapter 4) is an example of such an HK that I identified as the member of subfamily 11 (HPK_{11}) by manual alignment of these homology boxes.

1.2.1.2 Histidine kinase functions

The C–terminal kinase domain is responsible for ATP binding, the autophosphorylation of the histidine with ATP, recognition of its cognate RR, and the phosphotransfer from the kinase to an invariant aspartate in the receiver domain of the RR. In addition to autokinase and phosphotransfer functions, some

HKs (e.g. the *Escherichia coli* osmosensor EnvZ (*48*)) also exhibit autophosphatase activity toward their cognate response regulators. The N–terminal sensory domain has no catalytic functions itself but is responsible for detecting environmental stimuli. The diversity of the sensory domains used with HKs allows them to respond to a wide range of stimuli, including changes in osmolarity, nutrients, light, cellular redox state, quorum signals, antibiotics, and more (*49*). Additionally, the input stimuli can regulate either the kinase or phosphatase activity of HKs (*32*).

1.2.1.3 Histidine kinases structures

Most HKs are known to form homodimers in solution via their DHp domain (*50*). The DHp domain consists of two α–helical hairpin, and the dimeric unit is held together by forming a four helical bundle, as first shown in the solution structure of EnvZ dimerization domain (*44*). The H–box, located in the DHp domain, contains the site of histidine phosphorylation. This histidine is located at the solvent–exposed face of the first α–helix of the DHp domain (*51*). An example of the kinase domain of HK853 (PDB ID: 3DGE (*52*)) containing both DHp and CA domains is shown in Fig. 1.8.

Structures of CA domain core feature an unconventional Bergerat ATP–binding fold: an α/β sandwich consisting of a four–stranded mixed β sheet and three α helices (*53*). Since the Bergerat fold is shared among other members of

the GHL (gyrase, Hsp90, MutL) ATPase superclass of proteins, together they emerge and form a GHKL ATPase/kinase superfamily (*50, 53*). Structural analysis of this superfamily reveals the important role of each homology box. The conserved asparagine residue of the N box coordinates a bound Mg^{2+} ion, which connects the ATP phosphates to the protein via solvent–mediated hydrogen bonds (*50*). The conserved aspartate residue of the D–box interacts directly with ATP and forms a hydrogen bond with the N6 amino position of the adenine ring (*50*). The conserved G–box interacts with the adenine nucleotide amine and plays a critical role in phosphotransfer (*50, 54, 55*).

Although several sensory domains have been solved individually from different types of HKs (*56-58*), the lack of structural information of full–length histidine kinases containing both sensory and catalytic domains has prevented us from understanding the molecular mechanism of how the sensory domain interact and regulate the kinase domain. Since many structures of the HK domain containing both the DHp and CA domain have been solved, a central goal in the HK field is to obtain the full–length structure. The membrane–associated nature of these HKs has made it difficult for studying full–length structures, and structures of soluble cytosolic HKs are yet to be reported. Nevertheless, a recent structural study has revealed structures of the *Bacillus subtilis* DesK HK domain in various signaling states, providing us some molecular insight of how signal may be transmitted from the sensory domain to the kinase domain (*59*).

1.2.2 Second component: Response regulators (RRs)

HKs activate RRs by phosphorylation. The phosphorylated RR (RR~P) generates the output response by interacting with target genes or other protein targets, which often regulate corresponding cellular responses.

The CheY–like phosphoacceptor domain has been served as the prototype for the receiver (REC) domain (*60*). The REC domain (~120 a.a.) is the minimum functional unit of RRs, and many RRs contain only the REC domain alone (*61*). It is a highly conserved phospho–aspartyl receiver domain that interacts with the kinase domain of the cognate HK. More often, RRs contain various output domains that commonly act as transcription factors (*62, 63*) or as enzymes that possess methylesterase (*64*) and phosphodiesterase (*65*) enzymatic activities. The cognate RR of HKs can be easily identified if its DNA sequence is located in adjacent to the HK sequence in the operon. In other cases, the RRs can be identified through phosphotransfer profiling, and the details of this method will be further discussed in Chapter 4 (*66, 67*).

1.2.2.1 Diversity of response regulators

Previous sequence analysis performed by Grebe and Stock with 298 REC domain sequences with known cognate HKs has revealed 8 different REC domain subfamilies (*33*). Although this classification does not take into account of the

diversity of the output domain in the RRs, this study revealed a significant correlation between HK and RR subfamilies, suggesting that different subclasses of His–Asp phosphorelay systems have evolved independently of one another (*33*).

Classification of RRs based on their output domains provided a more complete picture in terms of predicting their potential functions (Fig. 1.9). In 2006, Galperin grouped 4,610 RRs encoded in complete genomes of 200 bacterial and archaeal species and classified them using their domain architecture (*39*). There are basically five types of RRs: i) RRs containing only the REC domain (the standalone REC), ii) RRs with DNA–binding output domains, iii) RRs with RNA–binding output domains, iv) RRs that exhibit enzymatic activity, and v) RRs that interact with other proteins or ligands (*65*). Nearly 66% of the RRs function as transcriptional regulators, including members of the OmpR, NarL, NtrC, LytTR, Fis, and AraC families (Fig. 1.9) (*65*). The second largest group is the standalone REC, which were found to comprise about 14% of RRs (Fig. 1.9) (*65*). The remaining 20% of RRs represented by coupling to a wild range of domains: RNA–binding domains (ANTAR and CsrA), enzymatic domains (i.e. GGDEF, EAL, HisK, PP2C,), protein– or ligand–binding domains (i.e. PAS, GAF, and HPt), and also domains with no known functions (Fig. 1.9) (*65*). The diversity of domain combinations in the response regulators allows bacterial to regulate transcription factors, enzyme activity, and protein–protein interactions in

response to various environmental challenges and increase their chance for survival.

1.2.2.2 The response regulator structure and function

Structures of the REC domain were solved prior to the structure of the HK domain because the REC domain is more soluble and it forms a stable, compact structural unit consists of 5 α helices and 5 central parallel β strands forming an α/β/α sandwich Fig. 1.10) (*68, 69*). The active site is located on the top of the β sheet, which harbors the phosphate–receiving aspartate. RRs can also function in three ways. Response regulators are activated upon phosphorylation and many studies have shown that REC domains undergo distinct conformational changes between phosphorylated and unphosphorylated states (*70-72*). The stability of the phosphorylated state appears to be tailored to their function. For instance, the phosphorylated RRs can last from seconds (e.g. CheY) to hours (e.g. SpoF) (*73, 74*). Many of these short–lived phosphorylated RRs are due to the fact that these RRs exhibit autophosphatase activity (*73, 74*). A mutation of an asparagine residue that surrounds and associates the conserved Asp to a Lys residue has been found to attenuate the rate of dephosphorylation (*74*), and this residue may be conserved in RRs exhibit the autophosphatase activity.

Although the structure of the histidine kinase including its sensory domain has yet to be determined, lower resolution (3.8 Å from crystallography and 4.2 Å

from SAXS) structural studies of histidine kinases complexed with their response regulator have advanced our knowledge on how the kinase domain (DHp–CA) interacts with the response regulators (*51, 52, 59*). Most recently, the high resolution (3.8 Å) crystal structure of the HK853cp–RR468 complex from *T. maritime* revealed molecular details of how the REC domain binds to the HK domain (Fig. 1.11) (*52*). Specifically, the structure shows the RR makes intimate contacts to both the DHp domain and the CA domain, which provides additional evidence that the DHp domain dictate the substrate specificity of the HKs as suggested by Laub (*36*) and Skerker (*42*).

Figure 1.6 Classification of histidine kinases. In A, the kinases are classified based on the localization of their sensory domain (adapted from Mascher *et. al*. Microbio. & Mol. Bio. Rev., 2006 (*43*)), and B. based on the location of the H–box. The sensory domains that are responsible for receiving the external stimuli are labeled as S and colored in light blue, green, orange, and yellow in A.

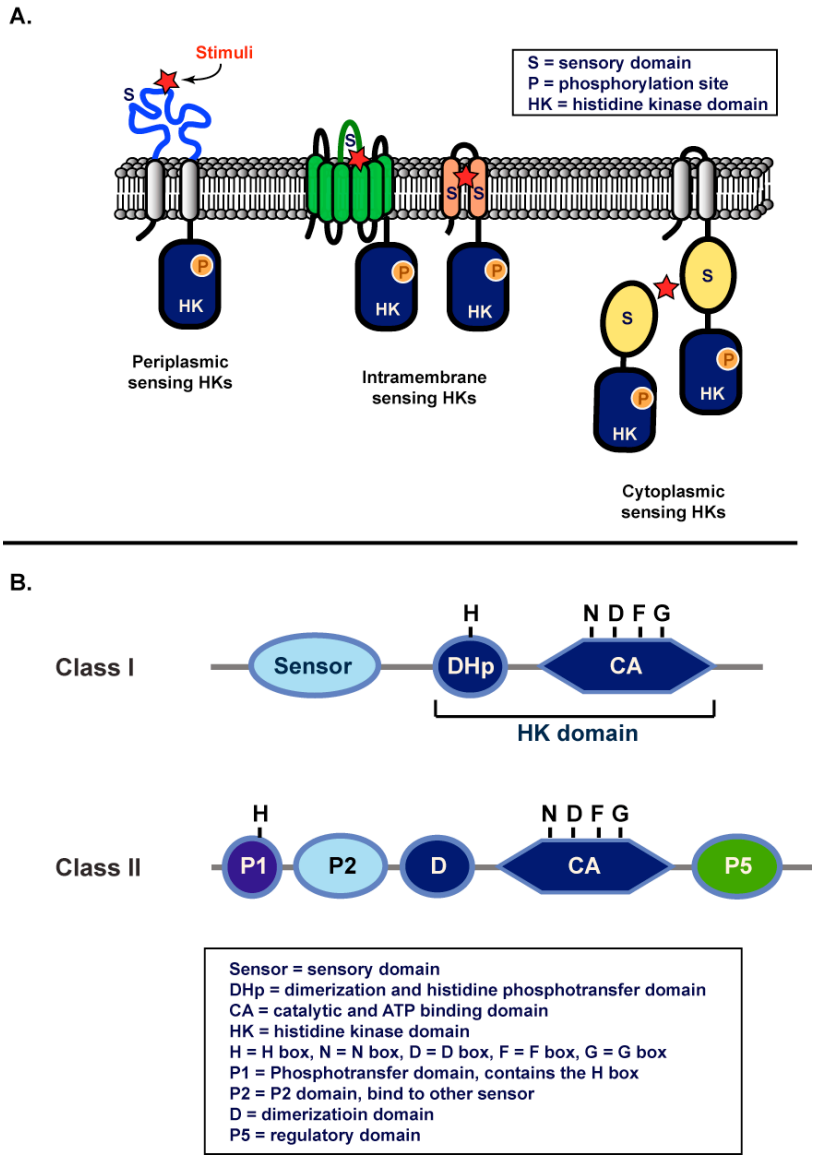


Figure 1.7 Individual subdomain structures of the HK domain: A. the DHp domain and B. the CA domain, taken from HK853 (PDB: 3DGE) (*52*). The N– and C–termini are labeled. As shown in A., the DHp domain forms a homodimer via the coil–coiled fashion. The conserved histidine residue (H260) is shown as a stick figure in magenta, which is located in the first helix from the DHp domain and is exposed to solvent. The conserved residues from the N, D, F, and G boxes are shown in stick figure (N380 in red, D441 in yellow, F425 in green, G441 and G442 in magenta). The ADP ligand is shown in sphere (carbon in green, nitrogen in blue, oxygen in red, and phosphor in orange). The ATP lid motif is shown in light blue.

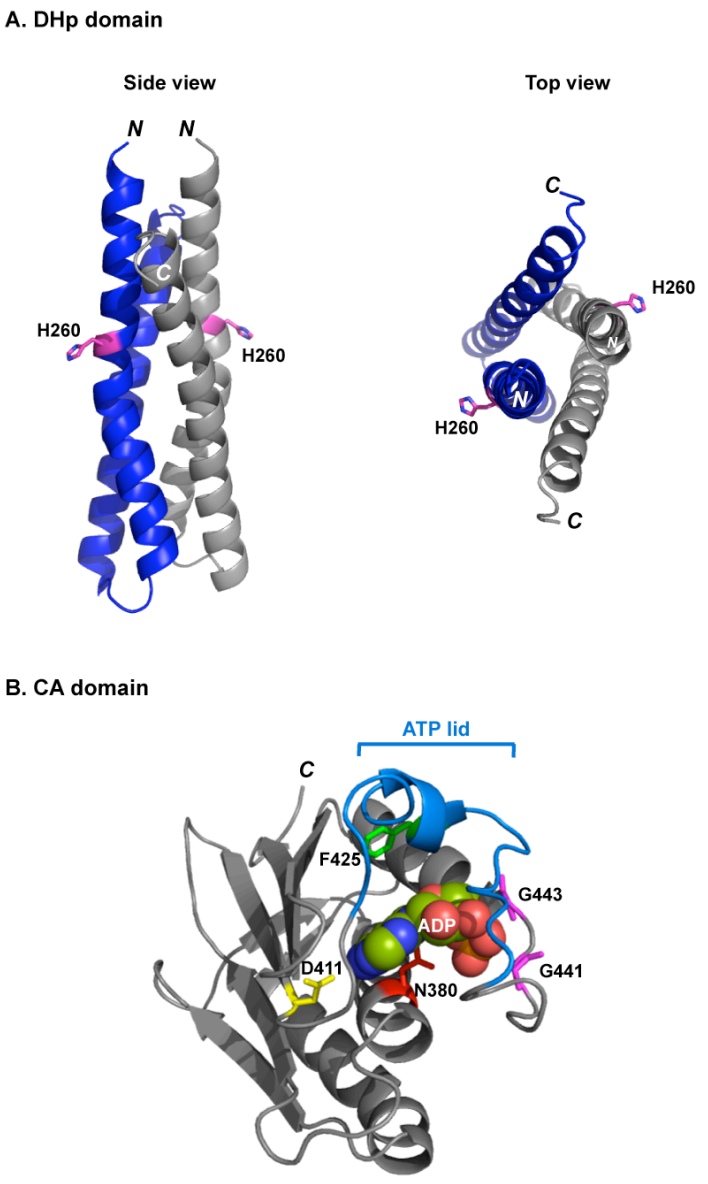


Figure 1.8 An example of the whole HK domain represented here using part of the crystal structure from the HK domain complex with its own response regulator (HK853–RR468) (PDB: 3DGE) (*52*). The response regulators from the original structure are removed for simplicity. The relative orientation of the CA domain to the HK domain is shown.

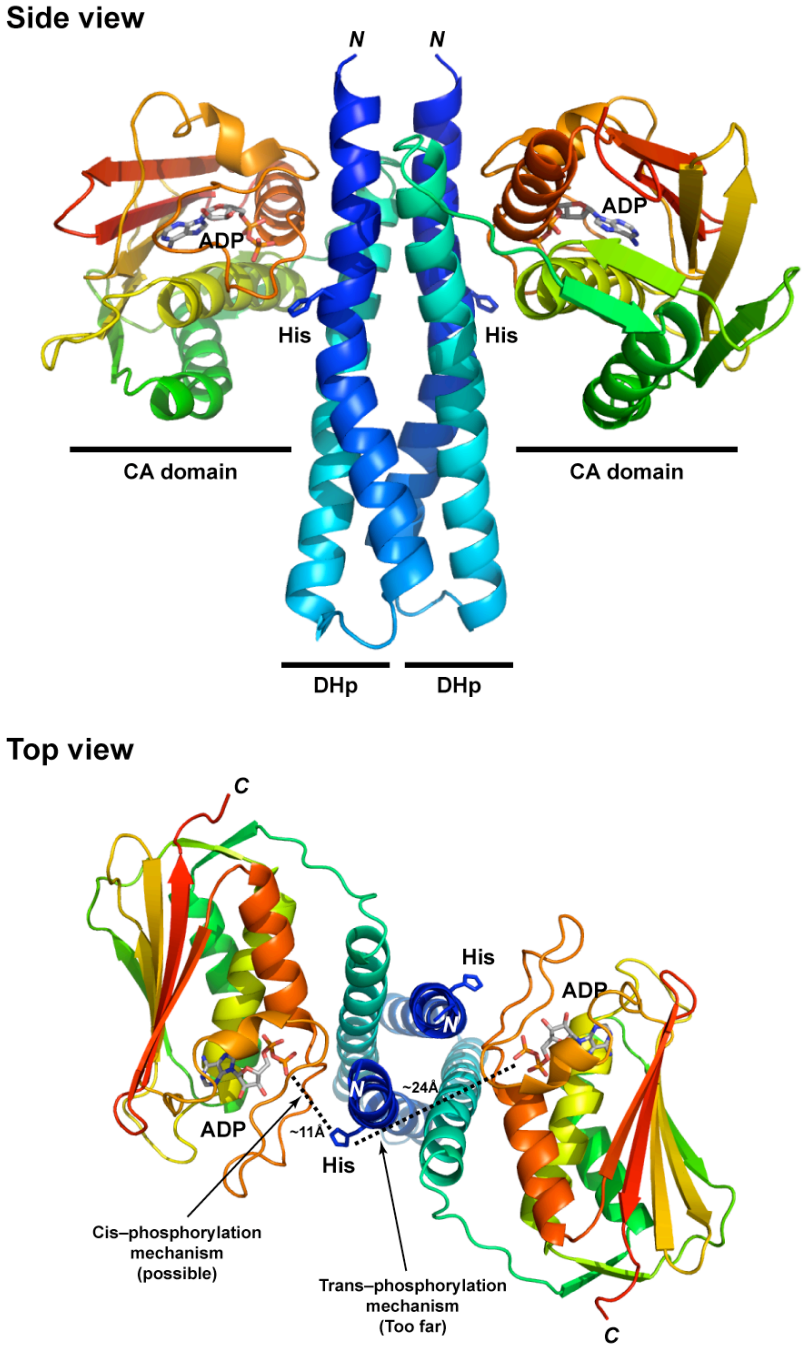


Figure 1.9 Different types of response regulators classified by their domain architectures. Figure adapted from Galperin *et. al*. Journal of bacteriology, 2006 (*39*).

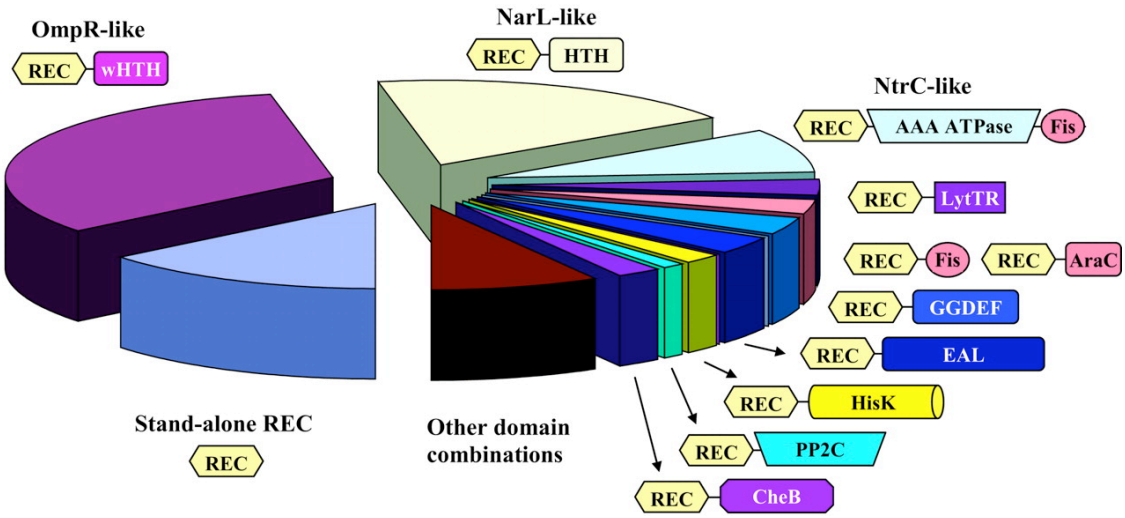


Figure 1.10 An example of the response regulator structure, shown here using the crystal structure of RR468 (PDB: 3DGF) (*52*). Alpha helices are colored in cyan and labeled as α1 to α5. Beta strands are colored in magenta and labeled as β1 to β5. The conserved phosphate receiving aspartate residue is shown in stick figure and labeled as Asp. This structure mimics the phosphorylated form of the RR with the presence of BeF_3^- in solution.

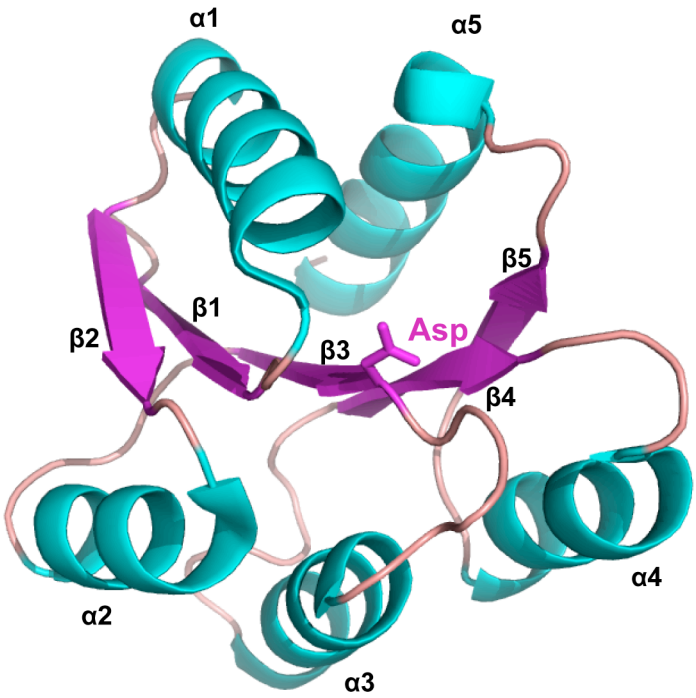
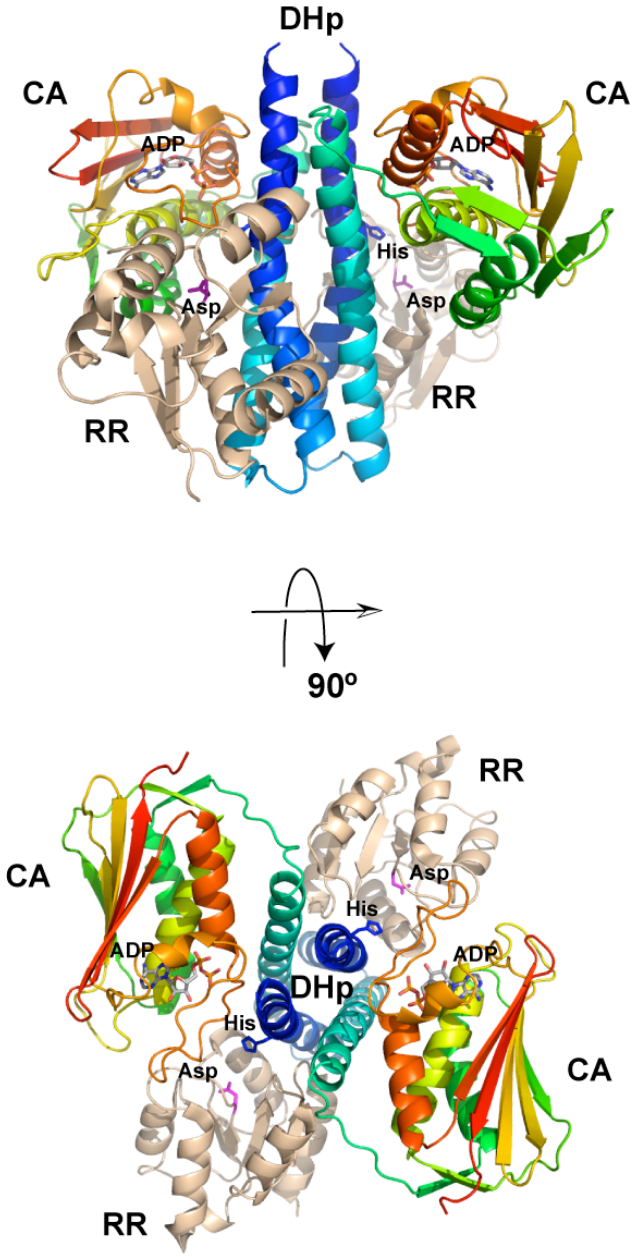


Figure 1.11 An example of the response regulator complex with the histidine kinase domain, shown here using the crystal structure from the HK853–RR468 complex (PDB: 3DGE) (*52*). The histidine kinase domain is shown the same way as in Figure 1.8. The response regulators are shown in wheat and its conserved phosphate receiving aspartate (Asp) residue is shown in stick figure in magenta. The DHp domain not only bears the phosphoryl histidine residue, but also contains specific recognition site for the RR. Side and top views are both shown.



1.3 Light signaling in two–component regulatory systems

The presence of LOV domains in bacteria genomes suggests the ability of bacteria to sense and utilize blue light may regulate many biological functions (*75*). In many cases, these LOV domains are coupled to histidine kinases and are expected to regulate the kinase activity by light. The function of these putative photoreceptors remains mostly unclear with a few exceptions. The LOV–histidine kinase (LovK) from the bacterium, *Caulobacter crescentus*, was found to enhance cell to cell attachment when the cells were exposed to light (*1*). On the other hand, the LOV–histidine kinase from a pathogenic bacteria, *Brucella abortus*, was found to use light to regulate the bacteria virulence by enhancing their ability to survive and multiple in the hostile intracellular environment of host macrophages (*76*). Besides aquatic oligotrophs and pathogenic bacteria, genes encoding LOV–containing histidine kinases have also been identified in other species. These proteins include EL346 and EL368 from *Erythrobacter litoralis* (a marine bacteria) and PS–LOV–HK from *Pseudomonas syringae*, all of which have been identified to increase their autophosphorylation activity when exposed to light (*76*).

Artificially, the oxygen sensing PAS domain from *Bradyrhizobium japonicum* histidine kinase, FixL, has been replaced by the blue light sensing LOV domain from *Bacillus subtilis* YtvA and successfully turn FixL to sense light instead of oxygen (*77*). This example suggests the potential of making LOV

domains into a biological light switch and its application in the therapeutic treatments.

1.4 Importance of TCSs: histidine kinases as new antimicrobial drug target

Bacterial resistance to antibiotics is an increasing concern worldwide. This major threat to human health is driving scientists to search for alternative mechanism and new antimicrobial drugs to treat bacterial infections. Advances in molecular microbiology and genomics have led the researchers to prokaryotic two–component systems (TCSs), which have a potential for therapeutic treatments and/or prevention of numerous bacterial diseases in both plants and animals.

Recently a membrane histidine kinase (QseC) in EHEC was discovered to act as a bacterial adrenergic receptor that activates virulence genes in response to host stress hormones such as epinephrine and norepinephrine (*45*). In fact, many bacterial pathogens rely on this conserved membrane histidine sensor kinase to respond to host adrenergic signaling molecules and bacterial signals in order to promote the expression of virulence factors (*78*). The QseC signaling pathway makes an attractive drug target system for many reasons. First, they are wildly spread in many animal pathogens but are absent in human. Second, inhibition of QseC may not exert a strong pressure for the development of resistance, since

inhibition of this interkingdom signaling pathway does not directly affect bacterial growth (*78*).

In general, there are two main approaches in the development of TCS inhibitors: the inhibitors that inhibit histidine kinase functions or the inhibitors that disrupt the response regulator functions. Examples of histidine kinase inhibitors include nucleotides binding domain inhibitors that are design to competing for the ATP binding site, and small peptide inhibitors or aromatic compounds that would compete for the ATP binding site in the HK domain or the ligand binding site in the sensory domain (*79*). For instance, a lead compound, LED209, has been identified to inhibit the binding of signals (i.e. stress hormones) to QseC, which prevent its autophosphorylation and consequently inhibiting QseC–mediated virulence gene expression without being toxic to the host or inhibiting the bacterial growth (*78*).

Chapter 2 Signal transduction mechanism in the *As*LOV2 domain

(Published as: A conserved glutamine plays a central role in LOV domain signal transmission and duration;

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

Biochemistry, **2008**, 47: 13842–13849.

[‡]: A.I.N. and W.–H.K. contributed equally to this manuscript)

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.

2.1 Introduction

Protein signaling cascades are central to organism 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 (*80*), circadian rhythm–dependent gene transcription (*81*), and phototropism and

chloroplast organization in plants (*82*). A specific subset of PAS proteins, the LOV (light–oxygen–voltage) domains (*7*), is capable of sensing blue light as an environmental signal and converting it into a biochemical signal in a wide variety of proteins.

LOV domains contain a series of highly conserved residues surrounding an internally bound flavin mononucleotide (FMN) or flavin adenine dinucleotide (FAD) chromophore (Fig. 2.1a, b) 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. 2.1c) (*2, 83*). The stability of this photoadduct is variable among LOV domains and ranges from a few seconds to many hours before it spontaneously relaxes back to the noncovalent dark state (*84, 85*).

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 (*11*). 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 (*20*). 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*. (*25*) 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) (*86*). Recent crystallographic data on *As*LOV2 containing the Jα helix also support a role for Jα helix in signal transduction (*87*). 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 (*88, 89*). In the dark state, crystallography shows that the side chain amide of a conserved glutamine residue (Q1029) in the *A. capillus–veneris* phytochrome3 LOV2 domain donates

a hydrogen bond to the O4 atom of FMN (Fig. 2.1c) (*6*). 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 (*88, 89*). 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 (*90*). 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 Q1029L mutation in full–length phototropin demonstrated attenuated autophosphorylation activity in the light versus wildtype protein (*91*). In addition, studies of the corresponding glutamine residue in the fungal receptor, VIVID, also show a leucine mutant diminishes light–induced activity in vivo (*26*).

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 lock the domain into a pseudo–dark state (Q513L) or pseudo–lit state (Q513N) compared to wildtype *As*LOV2. These data underscore the importance of hydrogen bond networks between FMN and the protein b–sheet in tuning properties of the chromophore and communicating light–induced structural changes throughout the domain.

2.2 Materials and methods

2.2.1 Cloning, Expression, and Purification of AsLOV2

Plasmid DNA encoding the *As*LOV2 domain plus the Jα helix (residues 404–560 (*25*)) 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 37ºC to an A_{600} of 0.6–0.8 and then induced with IPTG (0.12 g/L). After 16 hr induction at 20ºC, 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_6–Gβ1 tagged (*25*) 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 4ºC and stopped using a Ni^{+2}–NTA column to remove the His_6–Gβ1 and His_6–TEV protease. Post–cleavage, the resulting proteins contain only GEF (N–terminal) and G (C–terminal) residues as cloning artifacts.

2.2.2 Protein:Flavin Stoichiometry Calculation

UV−visible absorbance spectra (from 250 to 550 nm) were recorded for all three freshly purified proteins following buffer exchange into 50 mM sodium phosphate and 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 the wild type (2.60) as a reference for 1:1 protein:FMN stoichiometry (*30*), 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.

2.2.3 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 250–500 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).

2.2.4 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.

2.2.5 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. CD data were collected using a wavelength range from 195 to 260 nm at 21ºC with 1.5 nm bandwidth and 3 s averaging time. Data for the Q513N mutant was collected with a 1 s averaging time due to the faster rate of dark state recovery. Final data were generated from an average of 3 repeats.

2.2.6 Nuclear Magnetic Resonance spectroscopy

Proteins were concentrated to 1 mM in pH 6.0 buffer containing 50 mM sodium phosphate 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 25ºC, using nmrPipe (*92*) for data processing and NMRview (*93*) 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 (*25*).

2.2.7 Sequence Alignment

A multiple sequence alignment of LOV domains was generated using CLUSTAL W (*94*) and sequences were displayed using ESPript.cgi Version 3.06 (*95*).

2.2.8 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 (*30*). 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 ºC 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 ($A_{446,nonacidic}/A_{446,TCA}$) 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 1**.**

2.3 Results

2.3.1 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. 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. 2.2). 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 (*90, 91*), indicating a change in the electronic

environment surrounding the FMN. This contrasts with the Q513N absorbance spectra that 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 1). Comparison of extinction coefficients calculated for the two mutants further supports a perturbation of the electronic state in the Q513L mutant and relatively little change in the Q513N mutant. Wildtype and Q513N proteins have similar extinction coefficients of ~8000 M^{-1} cm^{-1} and ~6900 M^{-1} cm^{-1} at their 406 nm isosbestic point, respectively (Table 1), while the extinction coefficient of Q513L at its longest wavelength isosbestic point, 403 nm, is significantly lower at ~3000 M^{-1} cm^{-1}. Overall, these spectroscopic data indicate that the electronic environment of the FMN exhibits a more dramatic 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.

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 1). In contrast, the Q513N mutation has a shorter recovery time constant (37 s, Table 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.

2.3.2 Structural effects of Q513 mutations as determined by 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 (*25, 96*). Circular dichroism reflects total 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. 2.3), which demonstrate a clear decrease in molar residue ellipticity in wildtype *As*LOV2 following illumination with white light (*97*). The Q513L mutant shows a less pronounced decrease in ellipticity in this helical region upon illumination, reflecting less change in secondary structure between the dark state and light–induced conformations. Both states have intermediate ellipticity values compared to the two wildtype states, suggesting that Q513L retains a relatively high degree of structure. Similarly to Q513L, Q513N displays little difference between the dark and lit states (Fig. 2.3), but with much lower molar residue ellipticity than either state of the wildtype *As*LOV2. These data suggest that it

contains less secondary structure overall, consistent with solution NMR results (*vide infra*).

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 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 (*84, 97*).

2.3.3 Structural effects of Q513 mutations by 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 (*25*). This is reflected in the light–induced acceleration in the appearance of a lower molecular weight species in SDS–PAGE (Fig. 2.4). 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.

2.3.4 Structural effects of Q513 mutations characterized using 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. 2.6a) (*25*). 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.

Specific analysis of two tryptophan side chain indole (H_e1–N_e1) 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 (*25*). In the dark state of wildtype *As*LOV2, these indole crosspeaks are clearly separated (Fig. 2.5a). Upon illumination, they collapse towards a central position that is near the average location for protein tryptophans in general (Fig. 2.5b) (*98*), 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. 2.5a). After light induction, there is minimal chemical shift perturbation of the Q513L crosspeaks (Fig. 2.5b), indicating that a significant majority of the Q513L population still remains in a pseudo–dark 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 more centralized as seen in the wildtype lit state (Fig. 2.5a). Additionally, the crosspeaks do not collapse toward each other

to the same extent as wildtype upon illumination (Fig. 2.5b), 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. 2.5). Further, signs of peak doubling can be observed in several of these spectra, suggestive of slow (t ~ 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.

An analysis of the full $^{15}N/^{1}H$ HSQC spectra of wildtype, Q513L and Q513N (Fig. 2.6) supports the assignment of Q513L as a pseudo–dark state structure and Q513N as pseudo–lit state. The dark state spectrum of Q513L is quite similar to wildtype, consistent with limited proteolysis and CD results. 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. 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. 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.

2.4 Discussion

While the connection between light–induced covalent adduct formation and protein conformational changes in LOV domains is well established (*25, 26, 84, 86, 87, 90, 99*), 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(*26, 90*). While FTIR studies show that illumination breaks a hydrogen bond between this residue and the FMN O4 position upon adduct formation (*90*), 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(*26, 88, 89, 99*), while other structures argue against it (*87*). 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.

2.4.1 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 phy3 LOV2 (*90, 100*) 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.

2.4.2 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, including the presence of basic compounds (*e.g.* imidazole) (*101*), alkaline pH (*102, 103*) and ionic strength (*102, 103*); 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 $\Delta G^{\ddagger} \sim 14.5$ kcal mol^{-1} based on the temperature dependence of dark state recovery (*84*). 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 (*104*), which is experimentally supported by a light–dependent increase in ^{2}H exchange rates (*84*). 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 (*105*).

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* phy3 LOV2 slowed dark state recovery significantly (7–fold, (*100*)), as does a F1010L mutation at an adjacent position on the neighboring Hβ strand (10–fold, (*106*)). 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 Q513 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.

2.4.3 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 (*89*). 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 (*87*) 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.

2.4.4 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 (*91*). 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 (*26*). 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. 2.8). 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 (*107*), while the other two have been described as effectively irreversible (*81*). 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(*83, 105, 106, 108, 109*). 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 (*106*). 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 (*96*).

Table 1. Comparisons of the kinetic time constants (τ) and the extinction 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 rate equation to obtain the time constant (τ). The dark recovery time constant at A_{446} is measured by UV–visible spectroscopy while the dark recovery time constant at θ_{222} is measured via CD spectroscopy. Extinction coefficients were measured at the longest wavelength isosbestic point for each protein as noted in the table.

Construct	$\tau_{\text{dark recovery}}[A_{446}]$ (s)	$\tau_{\text{dark recovery}}[\theta_{222}]$ (s)	ε_{iso} (M^{-1} cm^{-1})	Isosbestic points (nm)
WT	68.3	72.4	7982	327, 388, and 406
Q513N	37.3	40.6	6921	330, 388, and 406
Q513L	1080	>1000	3038	330, 380, and 403

Figure 2.1 Proposed signal transduction pathway in the *As*LOV2 domain (*90*). Front (a) and side view (b) of *As*LOV2 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.

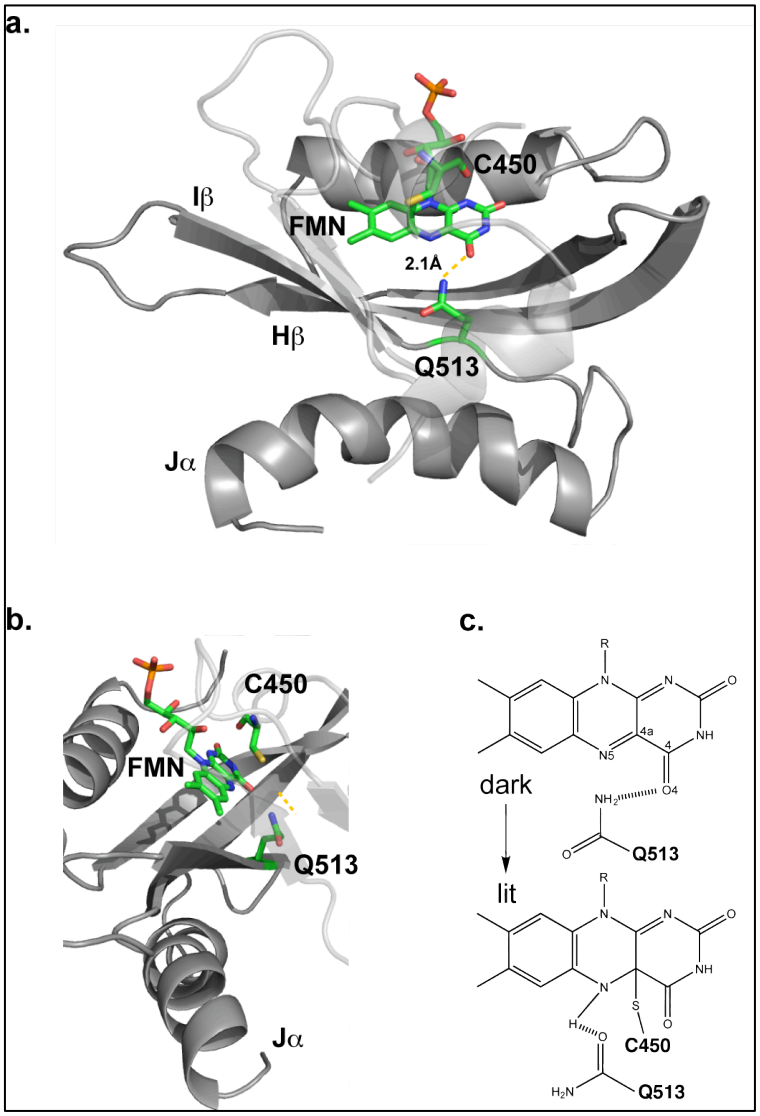


Figure 2.2 UV-visible absorbance profiles of *As*LOV2 wildtype protein and Q513 mutants. Black solid traces represent the dark state spectra and red dashed traces represent the lit state spectra. The mutants all display the same characteristic dark state absorbance profile for typical LOV domains, with 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 *As*LOV2 and Q513N; 403 nm for Q513L).

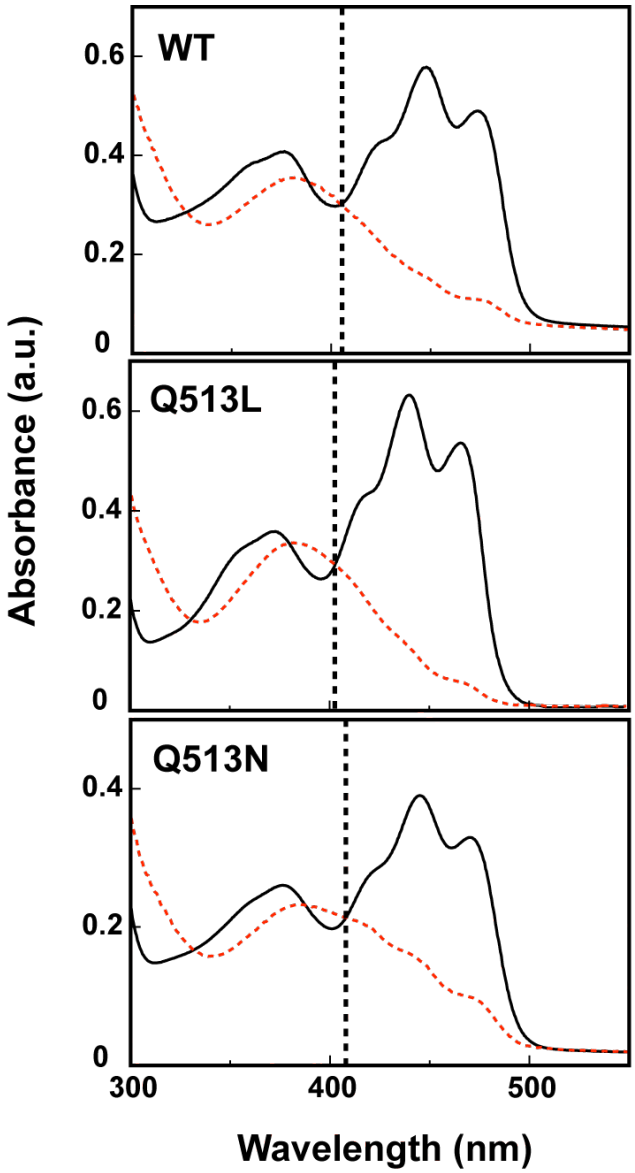


Figure 2.3 Structural effects of Q513 mutations in the dark and lit states monitored by circular dichroism. The room-temperature far UV region (190 nm to 250 nm) CD spectra of the wildtype dark (black) and lit (grey) states, Q513L dark (orange) and lit (yellow) states, and the Q513N dark (dark blue) 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.

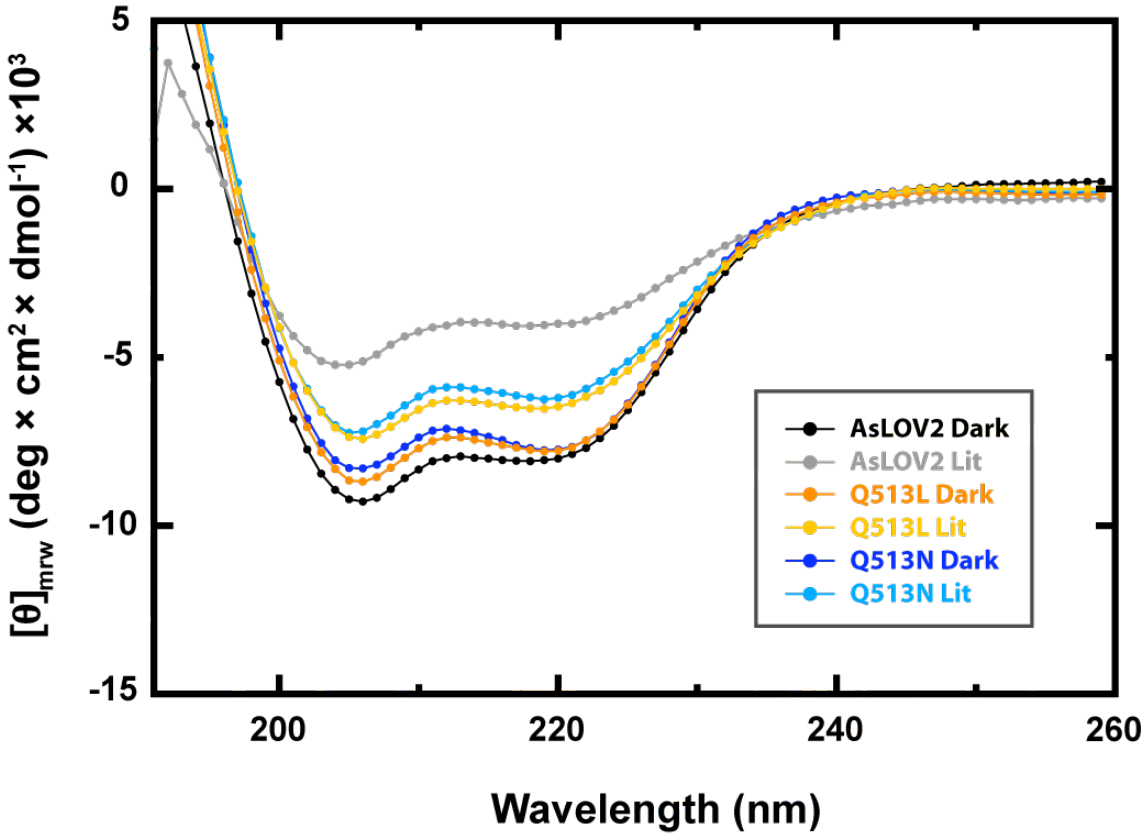


Figure 2.4 Structural effects of Q513 mutations in the dark and lit states monitored by limited proteolysis. Limited proteolysis by chymotrypsin digestion was performed at room temperature for the wildtype protein and Q513 mutants in both the dark (left panel) and lit states (right panel). Time points 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.

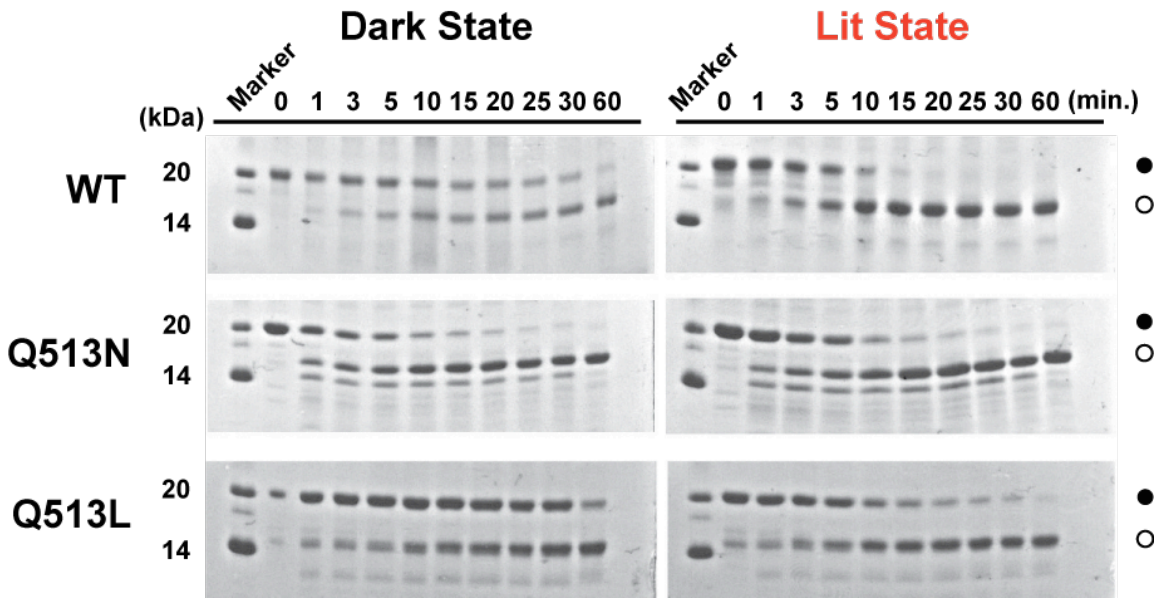


Figure 2.5 $^{15}N/^{1}H$ HSQC spectra of the tryptophan H_e1-N_e1 region. The overlaid spectra of the dark (a) and lit (b) states of the wildtype *As*LOV2 (black), Q513L (blue), and Q513N (red) are shown. The tryptophan indole assignments are indicated in both the dark and lit state panels with subscript D or L respectively. In the lit state, there is a significant shift in this region in wildtype protein (b).

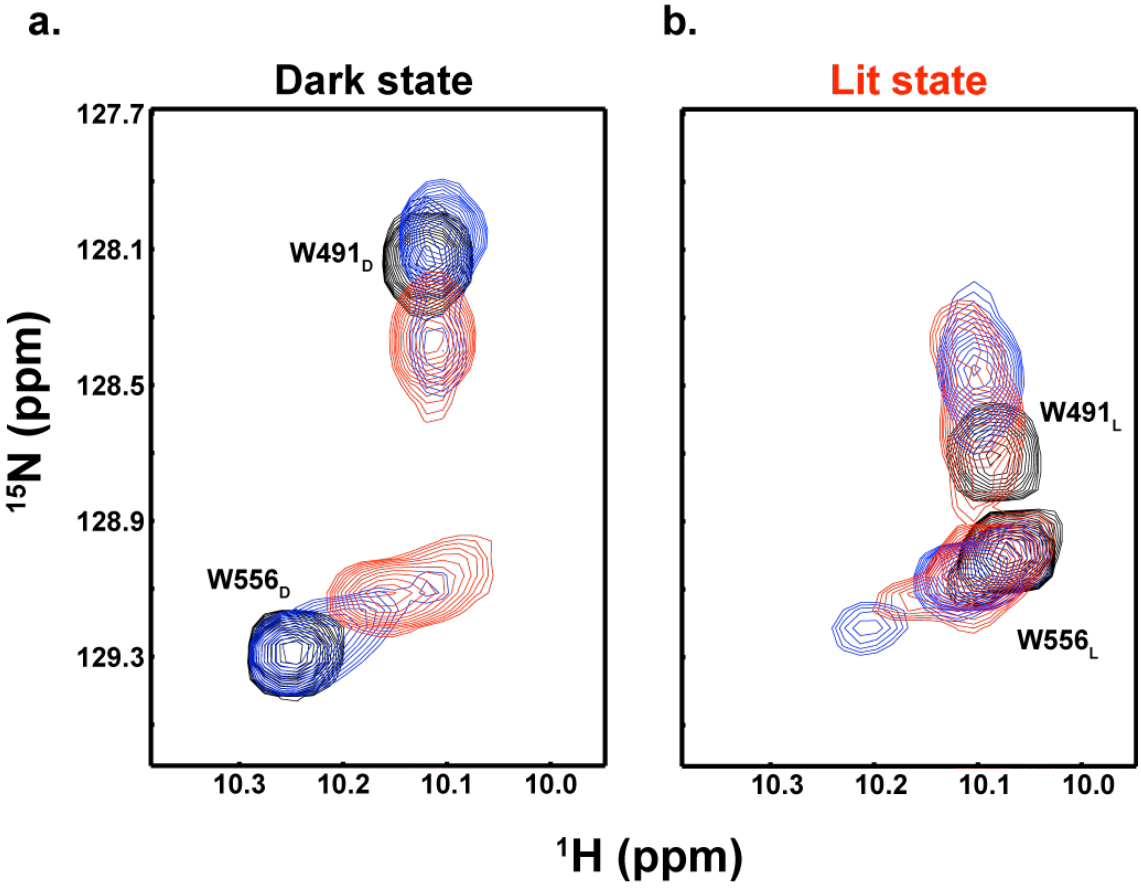


Figure 2.6 Overlay of dark (black) and lit (red) state 15N/1H HSQC correlation spectra for (a) wildtype *As*LOV2 domain, (b) Q513L, and (c) Q513N. Spectra were recorded at 25°C at 500 MHz. Black spectra represent the dark state and red spectra represent the lit state. See materials and methods for more details. Mutations lead to chemical shift changes as well as differential broadening across each spectrum.

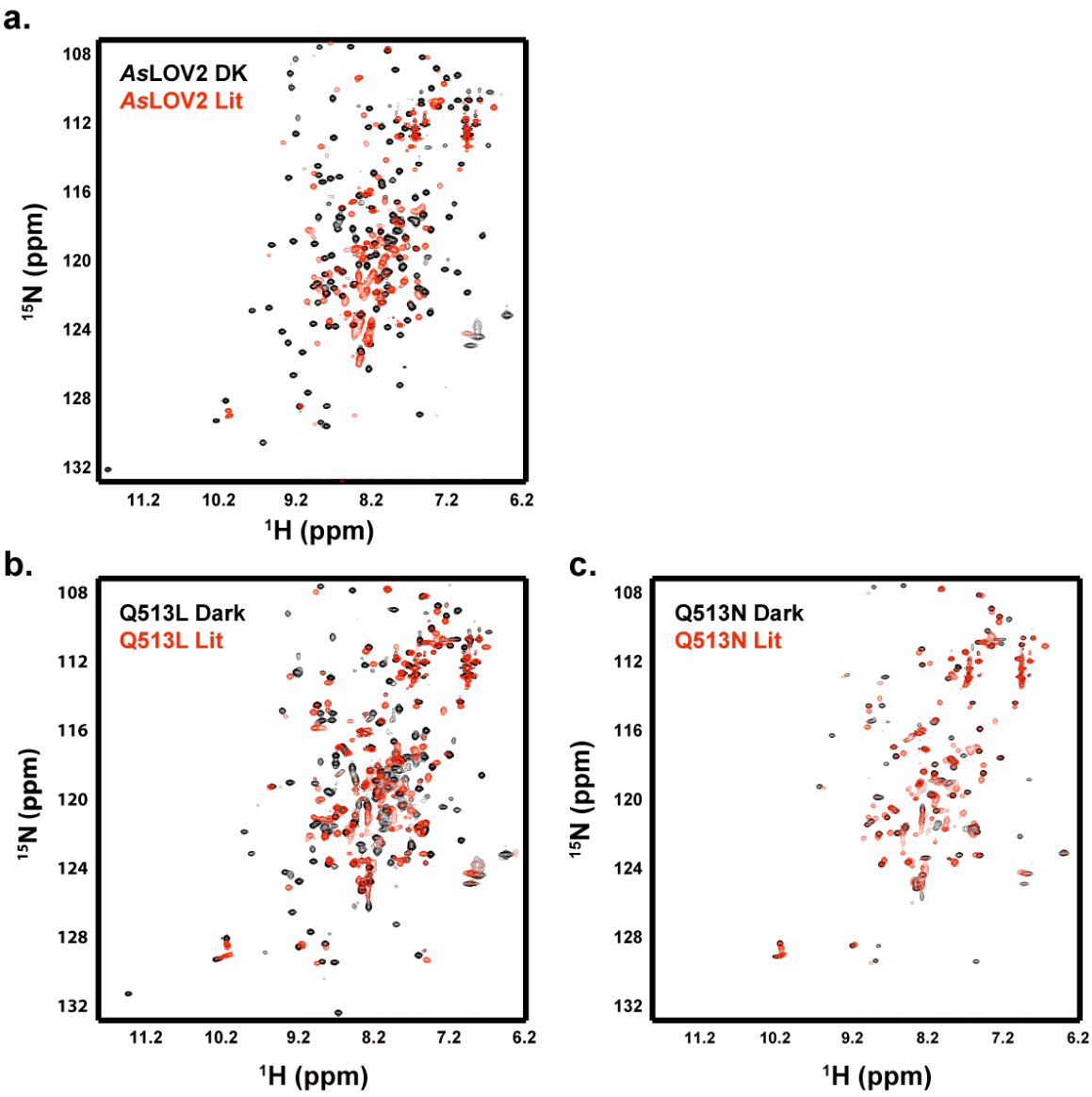


Figure 2.7 UV-visible absorbance spectra of wildtype (*As*LOV2), Q513L and Q513N mutants. UV-visible absorbance spectra shown here were recorded from 250nm to 550nm for each protein at concentrations between 50 μM and 70 μM in buffer containing 50 mM sodium phosphate (pH 6.0) and 100 mM NaCl. Calculated 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).

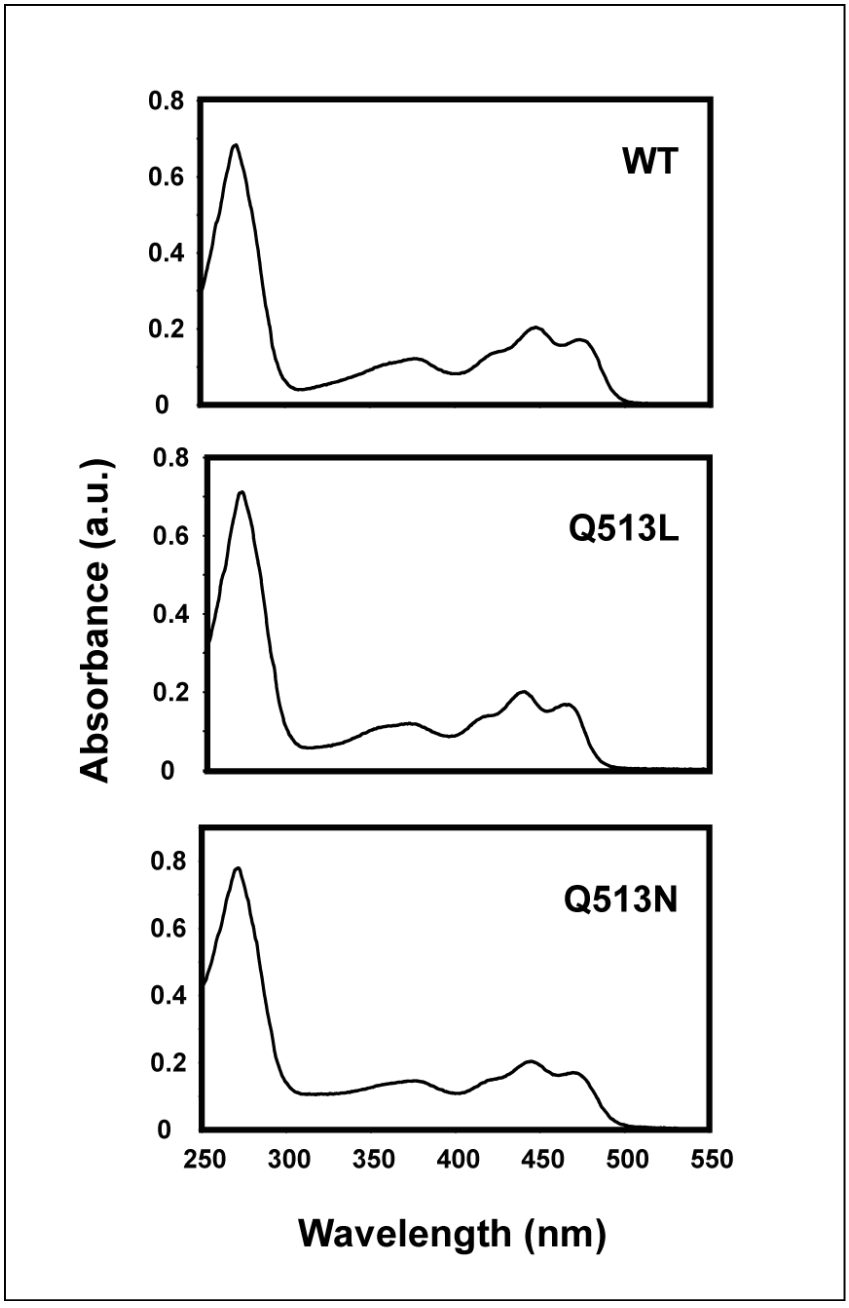
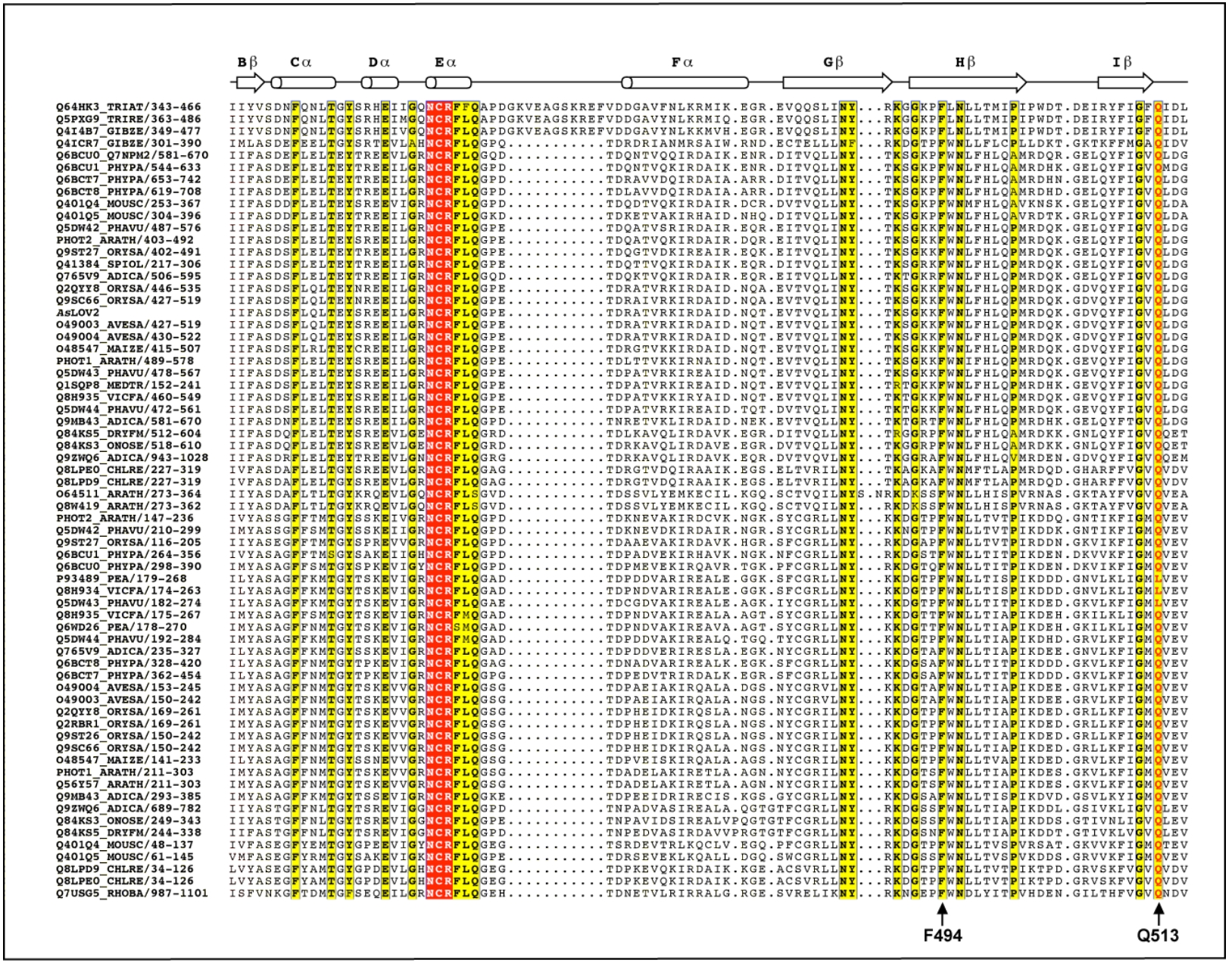
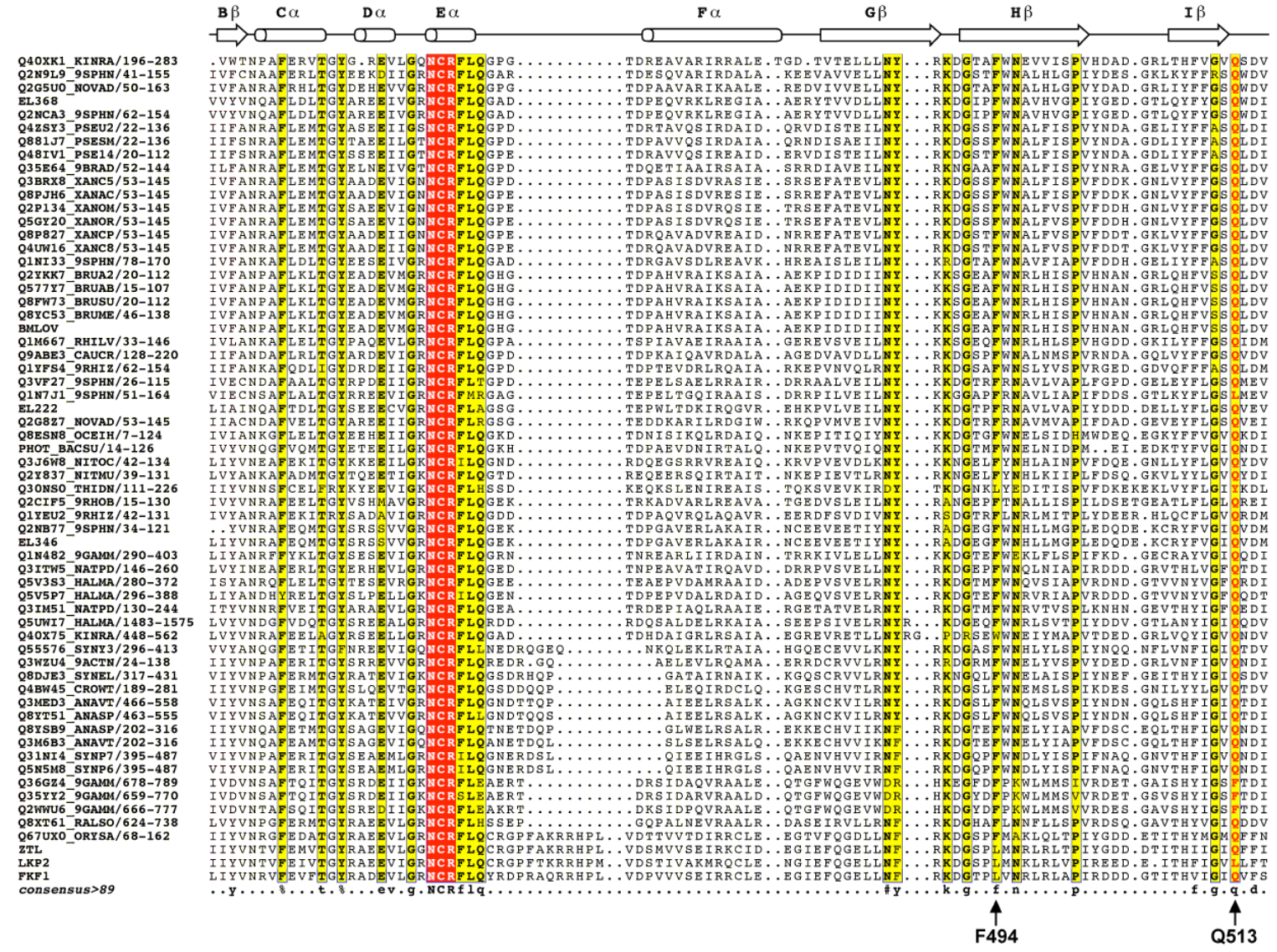


Figure 2.8 Sequence alignments of known LOV domains. Sequence alignments of 128 known LOV domains are shown with the predicted secondary structure indicated on the top of the sequences. Residues Q513 and F494 are indicated on the bottom. Red boxes with white characters indicate strict identity. Red characters or bold black characters indicate similarity within a group, while yellow boxes indicate similarity across groups (see next two pages).


Bβ
Cα
Dα
Eα
Fα
Gβ
Hβ
Iβ
Q64HK3_TRIAT/343-466 IIYVSDNFQNLTGYSRHEIIGQNCRFFQAPDGKVEAGSKREFVDDGAVFNLKRMIK.EGR.EVQQSLINY...RKGGKPFLNLLTMIPIPWDT.DEIRYFIGFQIDL
Q5PXG9_TRIRE/363-486 IIYVSDNFQNLTGYSRHEIMGQNCRFLQAPDGKVEAGSKREFVDDGAVYNLKRMIQ.EGR.EVQQSLINY...RKGGKPFLNLLTMIPIPWDT.DEIRYFIGFQIDL
Q4I4B7_GIBZE/349-477 IIYVSDNFQNLTGYSRHEIVGQNCRFLQAPDGKVEAGSKREFVDDGAVYNLKKMVH.EGR.EVQQSLINY...RKGGKPFLNLLTMIPIPWDT.DEIRYFIGFQIDL
Q4ICR7_GIBZE/301-390 IMLASDEFEELTGYSRTEVLAHNCRFLQGPQ..........TDRDRIANMRSAIW.RND.ECTELLLNF...RKDGTPFWNLLFLCPLLDKT.GKTKFFMGAQIDV
Q6BCU0_Q7NPM2/581-670 IIFASDEFLELTEYTREELIGRNCRFLQGQD..........TDQNTVQQIRDAIK.ENR.DITVQLLNY...TKSGKPFWNLFHLQAMRDQR.GELQYFIGVQLDG
Q6BCU1_PHYPA/544-633 IIFASDEFLELTEYTREELIGRNCRFLQGPD..........TDQNTVQKIRDAIK.ENR.DITVQLLNY...TKSGKPFWNLFHLQAMRDHK.GELQYFIGVQMDG
Q6BCT7_PHYPA/653-742 IIFASDEFLELTEYTREELIGRNCRFLQGPD..........TDRAVVDQIRDAIA.ARR.DITVQLLNY...TKSGKPFWNLFHLQAMRDHN.GELQYFIGVQLDG
Q6BCT8_PHYPA/619-708 IIFASDEFLELTEYTREELIGRNCRFLQGPD..........TDLAVVDQIRDAIA.ARR.DITVQLLNY...TKSGKPFWNLFHLQAMRDHD.GELQYFIGVQLDG
Q401Q4_MOUSC/253-367 IIFASDDFLELTEYSREEVIGRNCRFLQGPD..........TDQDTVQKIRDAIR.DCR.DVTVQLLNY...TKSGKPFWNMFHLQAVKNSK.GELQYFIGVQLDA
Q401Q5_MOUSC/304-396 IIFASDDFLELTEYTREEIIGRNCRFLQGKD..........TDKETVAKIRHAID.NHQ.DITVQLLNY...TKSGKPFWNLFHLQAVRDTK.GRLQYFIGVQLDA
Q5DW42_PHAVU/487-576 IIFASDSFLELTEYTREEILGRNCRFLQGPE..........TDQATVSRIRDAIR.EQR.EITVQLINY...TKSGKKFWNLFHLQPMRDQK.GELQYFIGVQLDG
PHOT2_ARATH/403-492 IIFASDSFLELTEYSREEILGRNCRFLQGPE..........TDQATVQKIRDAIR.DQR.EITVQLINY...TKSGKKFWNLFHLQPMRDQK.GELQYFIGVQLDG
Q9ST27_ORYSA/402-491 IIFASDSFLELTEYTREEILGRNCRFLQGPE..........TDQGTVDKIREAIR.EQK.EITVQLINY...TKSGKKFWNLFHLQPMRDQK.GELQYFIGVQLDG
Q41384_SPIOL/217-306 IIFASDSFLELTEYTREEILGRNCRFLQGPE..........TDQTTVQKIRDAIK.EQR.DITVQLINY...TKSGRKFWNLFHLQPMRDQK.GELQYFIGVQLDG
Q765V9_ADICA/506-595 IIFASDSFLELTEYTREEIIGRNCRFLQGQD..........TDQKTVQKIRDAIR.EQR.EITVQLLNY...TKTGKRFWNLFHLQPMRDQK.GELQYFIGVQLDG
Q2QYY8_ORYSA/446-535 IIFASDSFLQLTEYNREEILGRNCRFLQGPE..........TDRATVRKIRDAID.NQA.EVTVQLINY...TKSGKKFWNLFHLQPMRDQK.GDVQYFIGVQLDG
Q9SC66_ORYSA/427-519 IIFASDSFLQLTEYNREEILGRNCRFLQGPE..........TDRAIVRKIRDAID.NQA.EVTVQLINY...TKSGKKFWNLFHLQPMRDQK.GDVQYFIGVQLDG
AsLOV2 IIFASDSFLQLTEYSREEILGRNCRFLQGPE..........TDRATVRKIRDAID.NQT.EVTVQLINY...TKSGKKFWNLFHLQPMRDQK.GDVQYFIGVQLDG
O49003_AVESA/427-519 IIFASDSFLQLTEYSREEILGRNCRFLQGPE..........TDRATVRKIRDAID.NQT.EVTVQLINY...TKSGKKFWNLFHLQPMRDQK.GDVQYFIGVQLDG
O49004_AVESA/430-522 IIFASDSFLQLTEYSREEILGRNCRFLQGPE..........TDRATVRKIRDAID.NQT.EVTVQLINY...TKSGKKFWNLFHLQPMRDQK.GDVQYFIGVQLDG
O48547_MAIZE/415-507 IIFASDSFLRLTEYCREEILGRNCRFLQGPE..........TDRGTVKKIRDAID.NQT.EVTVQLINY...TKSGKKFWNLFHLQPMRDQK.GDVQYFIGVQLDG
PHOT1_ARATH/489-578 IIFASDSFLELTEYSREEILGRNCRFLQGPE..........TDLTTVKKIRNAID.NQT.EVTVQLINY...TKSGKKFWNIFHLQPMRDQK.GEVQYFIGVQLDG
Q5DW43_PHAVU/478-567 IIFASDSFLELTEYSREEILGRNCRFLQGPE..........TDPATVRKIREAID.NQT.EVTVQLINY...TKSGKKFWNLFHLQPMRDQK.GEVQYFIGVQLDG
Q1SQP8_MEDTR/152-241 IIFASDSFLELTEYSREEILGRNCRFLQGPE..........TDPATVRKIREAID.NQT.EVTVQLINY...TRTGKKFWNLFHLQPMRDHK.GEVQYFIGVQLDG
Q8H935_VICFA/460-549 IIFASDSFLELTEYSREEILGRNCRFLQGPE..........TDPATVKKIRYAID.NQT.EVTVQLINY...TKTGKKFWNLFHLQPMRDQK.GEVQYFIGVQLDG
Q5DW44_PHAVU/472-561 IIFASDSFLELTEYSREEILGRNCRFLQGPE..........TDPATVRKIREAID.TQT.DVTVQLINY...TKTGKKFWNLFHLQPMRDQK.GEVQYFIGVQLDG
Q9MB43_ADICA/581-670 IIFASDNFLELTEYSREEILGRNCRFLQGPD..........TNRETVKLIRDAID.NEK.EVTVQLLNY...TKTGRTFWNLFHLQPMRDHK.GELQYFTGVQLDG
Q84KS5_DRYFM/512-604 IIFASDQFLELTEYSREEVLGENCRFLQGRD..........TDLKAVQLIRDAVK.EGR.DITVQLLNY...TRGGRPFWNLFHLQAMRDKK.GNLQYFIGVQQET
Q84KS3_ONOSE/518-610 IIFASDQFLELTEYSREEVLGENCRFLQGRD..........TDRKAVQLIRDAVE.EGR.DVTVQLLNY...TKGGRPFWNLFHLQAMRDKK.GNLQYFIGVQQET
Q9ZWQ6_ADICA/943-1028 IIFASDRFLELTEYTREEVLGNNCRFLQGRG..........TDRKAVQLIRDAVK.EQR.DVTVQVLNY...TKGGRAFWNLFHLQVMRDEN.GDVQYFIGVQQEM
Q8LPE0_CHLRE/227-319 IVFASDAFLELTGYSREEVLGRNCRFLQGAG..........TDRGTVDQIRAAIK.EGS.ELTVRILNY...TKAGKAFWNMFTLAPMRDQD.GHARFFVGVQVDV
Q8LPD9_CHLRE/227-319 IVFASDAFLELTGYSREEVLGRNCRFLQGAG..........TDRGTVDQIRAAIK.EGS.ELTVRILNY...TKAGKAFWNMFTLAPMRDQD.GHARFFVGVQVDV
O64511_ARATH/273-364 IIYASDAFLTLTGYKRQEVLGQNCRFLSGVD..........TDSSVLYEMKECIL.KGQ.SCTVQILNYS.NRKDKSSFWNLLHISPVRNAS.GKTAYFVGVQVEA
Q8W419_ARATH/273-362 IIYASDAFLTLTGYKRQEVLGQNCRFLSGVD..........TDSSVLYEMKECIL.KGQ.SCTVQILNY...RKDKSSFWNLLHISPVRNAS.GKTAYFVGVQVEA
PHOT2_ARATH/147-236 IVYASSGFFTMTGYSSKEIVGRNCRFLQGPD..........TDKNEVAKIRDCVK.NGK.SYCGRLLNY...KKDGTPFWNLLTVTPIKDDQ.GNTIKFIGMQVEV
Q5DW42_PHAVU/210-299 IMYASSGFFSMTGYSSKEIIGRNCRFLQGPD..........TDKNEVDKIRDAIR.NGR.SYCGRLLNY...KKNGTPFWNLLTVTPIKDDK.GNTIKFIGMQVEV
Q9ST27_ORYSA/116-205 IIYASEGFFTMTGYSPREVVGRNCRFLQGPD..........TDAAEVAKIRDAVK.HGR.SFCGRLLNY...RKDGAPFWNLLTVTPIRDDN.GKVIKFIGMQVEV
Q6BCU1_PHYPA/264-356 IVYASAGFFTMSGYSAKEIIGHNCRFLQGPD..........TDPADVEKIRHAVK.NGK.NFCGRLLNY...RKDGSTFWNLLTITPIKDEN.DKVVKFIGMQVEV
Q6BCU0_PHYPA/298-390 IMYASAGFFSMTGYSPKEVIGYNCRFLQGPD..........TDPMEVEKIRQAVR.TGK.PFCGRLLNY...RKDGTQFWNLLTITPIKDEN.DKVIKFIGMQVEV
P93489_PEA/179-268 ILYASAGFFKMTGYTSKEVIGRNCRFLQGAD..........TDPDDVARIREALE.GGK.SFCGRLLNY...KKDGTPFWNLLTISPIKDDD.GNVLKLIGMLVEV
Q8H934_VICFA/174-263 ILYASAGFFKMTGYTSKEVIGRNCRFLQGAD..........TDPNDVARIREALE.GGK.SFCGRLLNY...KKDGTPFWNLLTISPIKDDD.GNVLKLIGMLVEV
Q5DW43_PHAVU/182-274 ILYASAGFFKMTGYTSKEVIGRNCRFLQGAE..........TDCGDVAKIREALE.AGK.IYCGRLLNY...KKDGTPFWNLLTITPIKDED.GKVLKFIGMLVEV
Q8H935_VICFA/175-267 IMYASAGFFSMTGYTSKEVIGRNCRFMQGAD..........TDPNDVAKIREALA.AGT.SYCGRLLNY...KKDGTTFWNLLTIAPIKDEH.GKILKLIGMQVEV
Q6WD26_PEA/178-270 IMYASAGFFNMTGYTSKEVIGRNCRSMQGAD..........TDPNDVAKIREAVA.AGT.SYCGRLLNY...KKDGTTFWNLLTIAPIKDEH.GKILKLIGMQVEV
Q5DW44_PHAVU/192-284 IMYASAGFFKMTGYTSKEVIGRNCRFMQGAD..........TDPDDVAKIREALQ.TGQ.TYCGRLLNY...KKDGTPFWNLLTIAPIKDHD.GRVLKFIGMQVEV
Q765V9_ADICA/235-327 ILYASAGFFKMTGYSSKEVIGRNCRFLQGAD..........TDPDDVERIRESLA.EGK.NYCGRLLNY...KKDGTAFWNLLTIAPIKDEE.GNVLKFIGMQVEV
Q6BCT8_PHYPA/328-420 ILYASAGFFNMTGYTPKEVIGRNCRFLQGAG..........TDNADVARIREALK.EGK.SFCGRLLNY...KKDGSAFWNLLTITPIKDDD.GKVLKFIGMQVEV
Q6BCT7_PHYPA/362-454 ILYASAGFFNMTGYTPKEVIGRNCRFLQGPG..........TDPEDVTRIRDALK.EGR.SFCGRLLNY...KKDGSAFWNLLTITPIKDDD.GKVLKFIGMQVEV
O49004_AVESA/153-245 IMYASAGFFNMTGYTSKEVVGRNCRFLQGSG..........TDPAEIAKIRQALA.DGS.NYCGRVLNY...KKDGTAFWNLLTIAPIKDEE.GRVLKFIGMQVEV
O49003_AVESA/150-242 IMYASAGFFNMTGYTSKEVVGRNCRFLQGSG..........TDPAEIAKIRQALA.NGS.NYCGRVLNY...KKDGTAFWNLLTIAPIKDEE.GRVLKFIGMQVEV
Q2QYY8_ORYSA/169-261 IMYASAGFFNMTGYTSKEVVGRNCRFLQGSG..........TDPHEIDKIRQSLA.NGS.NYCGRILNY...KKDGTPFWNLLTIAPIKDED.GRLLKFIGMQVEV
Q2RBR1_ORYSA/169-261 IMYASAGFFNMTGYTSKEVVGRNCRFLQGSG..........TDPHEIDKIRQSLA.NGS.NYCGRILNY...KKDGTPFWNLLTIAPIKDED.GRLLKFIGMQVEV
Q9ST26_ORYSA/150-242 IMYASAGFFNMTGYTSKEVVGRNCRFLQGSG..........TDPHEIDKIRQSLA.NGS.NYCGRILNY...KKDGTPFWNLLTIAPIKDED.GRLLKFIGMQVEV
Q9SC66_ORYSA/150-242 IMYASAGFFNMTGYTSKEVVGRNCRFLQGSG..........TDPHEIDKIRQALA.NGS.NYCGRILNY...KKDGTPFWNLLTIAPIKDED.GRLLKFIGMQVEV
O48547_MAIZE/141-233 ILYASAGFFNMTGYSSNEVVGRNCRFLQGSG..........TDPVEISKIRQALA.AGS.NYCGRILNY...KKDGTPFWNLLTVAPIKDED.GRVLKFIGMQVEV
PHOT1_ARATH/211-303 IMYASAGFFNMTGYTSKEVVGRNCRFLQGSG..........TDADELAKIRETLA.AGN.NYCGRILNY...KKDGTSFWNLLTIAPIKDES.GKVLKFIGMQVEV
Q56Y57_ARATH/211-303 IMYASAGFFNMTGYTSKEVVGRNCRFLQGSG..........TDADELAKIRETLA.AGN.NYCGRILNY...KKDGTSFWNLLTIAPIKDES.GKVLKFIGMQVEV
Q9MB43_ADICA/293-385 IMYASAGFFKMTGYTSSEVIGRNCRFLQGKE..........TDPEEIDRIRECIS.KGS.GYCGRLLNY...KKDGSAFWNLLTISPIKDVD.GSVLKYIGMQVEV
Q9ZWQ6_ADICA/689-782 IIYASTGFFNLTGYTSREVIGGNCRFLQGPD..........TNPADVASIREALA.QGTGTFCGRLLNY...RKDGSSFWNLLTIAPIKDDL.GSIVKLIGVQLEV
Q84KS3_ONOSE/249-343 IIYASTGFFNLTGYTSREVIGGNCRFLQGPE..........TNPAVIDSIREALVPQGTGTFCGRLLNY...RKDGSSFWNLLTIAPIKDDS.GTIVNLIGVQLEV
Q84KS5_DRYFM/244-338 IIFASTGFFNLTGYTSREVIGGNCRFLQGPD..........TNPEDVASIRDAVVPRGTGTFCGRLLNY...RKDGSNFWNLLTIAPIKDDT.GTIVKLVGVQLEV
Q401Q4_MOUSC/48-137 IVFASEGFYEMTGYGPEEVIGYNCRFLQGEG..........TSRDEVTRLKQCLV.EGQ.PFCGRLLNY...RKDGTPFWNLLTVSPVRSAT.GKVVKFIGMQTEV
Q401Q5_MOUSC/61-145 VMFASEGFYRMTGYSAKEVIGKNCRFLQGPE..........TDRSEVEKLKQALL.DGQ.SWCGRLLNY...RKDGSSFWNLLTVSPVKDDS.GRVVKFIGMQVEV
Q8LPD9_CHLRE/34-126 LVYASEGFYAMTGYGPDEVLGHNCRFLQGEG..........TDPKEVQKIRDAIK.KGE.ACSVRLLNY...RKDGTPFWNLLTVTPIKTPD.GRVSKFVGVQVDV
Q8LPE0_CHLRE/34-126 LVYASEGFYAMTGYGPDEVLGHNCRFLQGEG..........TDPKEVQKIRDAIK.KGE.ACSVRLLNY...RKDGTPFWNLLTVTPIKTPD.GRVSKFVGVQVDV
Q7USG5_RHOBA/987-1101 ISFVNKGFTDMTGFSEQEILGRNCRFLQGEH..........TDNETVLRIRRALG.RGE.SVRELIKNY...RKNGEPFWNDLYITPVHDEN.GILTHFVGVQNDV
F494
Q513

Sequence alignment Cont.



Chapter 3 Roles of the LOV2 domain in regulating the Serine/threonine kinase activity

3.1 Introduction

It has been shown that LOV2 domains of plant phototropins are primarily responsible for activating the serine/threonine kinase activity of these proteins and subsequently promoting phototropism and other blue light–regulated responses (*20*). Previously, Harper *et. al.* found that light causes an unfolding event in the Jα helix motif and results in the dissociation of Jα from the β sheet surface of the LOV domain core (*25*). However, the exact mechanism of how light activated LOV2 domain activates the kinase domain is unknown. Based on general principles of protein regulation, two major models have been proposed (Fig. 3.1) (*86*). The first model for the kinase activation would be for the LOV2 domain to act as a dark state inhibitor, reminiscent of the type I transforming growth factor β receptor kinase (TβR–1) (*110*). In this case, the Jα somehow interacts with the kinase domain and keeps it in the inactive conformation in the dark state. Upon light illumination, the unfolding of Jα leads to the release of the kinase inhibition, activating the kinase. The second model suggests that the LOV2 domain acts as a lit state activator. In this case, only the light activated LOV2 domain will interact with the kinase domain and trigger the kinase activity, either through the interaction with the exposed β sheet surface of the LOV core or the unfolded Jα.

To deduce the mechanistic role of the disrupted LOV2–Jα plays in activating serine/threonine kinase activity, *in cis* and *in trans* LOV2 kinase constructs will be needed for testing these proposed models. In this chapter, I summarize the constructs I have generated.

3.2 Materials and methods

3.2.1 Cloning, expression, and purification

DNA encoding different fragment of LOV2Kinase and kinase constructs from *A. sativa* phototropin1 and *A. thaliana* phototropin1 and phototropin2 were subcloned into the expression vectors pHis–parallel1, pHis–Gb1–parallel1, GST, and/or MBP vectors (see Fig. 3.2 for details). All constructs were overexpressed in *E. coli* BL21 (DE3) cells (Novagen) cells in LB media. Cultures were grown at 37°C, induced at an A_{600} of 0.7–1.0 with 0.12 g/L isopropyl b–D–thiogalactoside (IPTG), and allowed to grow in the dark at 20°C for an additional 16 h.

For solubility test (see Fig. 3.3 and 3.4), I used the following freeze/thaw protocol: I prepared the desired buffer with 0.2% Triton X100 (for 5 mL buffer, add 10 μL 100% Triton, vortex to dissolve). I resuspend 50 mL culture pellet in 2 mL prepared buffer, add 5 μL of lysozyme (25 mg/mL), and let it sit in R.T. for 10 minutes. I then froze the cell using liquid nitrogen and thaw it in room temp and repeated this step twice. Next, I add 2 μL of Dnase (1 mg/ml) and let it sit in R.T. for additional 30 to 50 minutes. The supernatant is separated from the pellet

by centrifugation at 4ºC for 1 h and proteins from the supernatant and pellet fractions were verified by 20% homogenous PhastGel gels (Amersham Biosciences).

For purifying MBP–tagged proteins, cells were harvested and resuspended into 50 mM sodium phosphate (pH 7.0), 500 mM NaCl buffer and 1 mM EDTA, lysed by extrusion, and clarified by centrifugation. The supernatant was filtered (0.45 μm) and loaded into an affinity column packed with 7 mL of Amylose resin (New England BioLabs). Proteins were eluted with a 100% 50 mM sodium phosphate (pH 7.0), 500 mM NaCl, 1 mM EDTA, and 50 mM Maltose. The desired protein fractions were buffer exchange into 50 mM sodium phosphate (pH 7.0), 100 mM NaCl, and 1 mM EDTA buffer, and pass through the MonoQ column as an additional purification step. The results for MBP–Asphot1KD(592–923) and MBP–Asphot1L2K(387–923) were shown in Fig. 3.5.

Alternatively, for MBP–His6–tagged proteins, they were first purified using the Ni^{2+} column packed with 10 mL of High Performance Ni Sepharose resin (GE Healthcare). Protein was eluted with a 20 C.V. gradient set from 15–500 mM imidazole using buffer containing 50 mM sodium phosphate (pH 7.8), 500 mM NaCl, and 500 mM imidazole, and were further purified by an affinity column packed with 7 mL of Amylose resin (New England BioLabs) and MonoQ column. The MBP and His6 tag were then cleaved by addition of TEV protease and separated by passing through another Ni^{2+} column. Size–exclusion

chromatography (SEC200) was used as the final step of purification to access protein oligomeric state using 50 mM sodium phosphate (pH 7.5) and 100 mM NaCl buffer. For LOV2–containing proteins, concentrations were determined using the absorption coefficient (ε) (ε_{446} = 1,1800 M^{-1} cm^{-1}) for flavin–containing proteins. For Kinase domain only protein, ε_{280} = 56,840 M^{-1} cm^{-1}, which is predicted from the ProtParam tool from ExPASy website were used (*111*).

3.3 Summary and perspective on phototropin project

From most of the constructs I tested, I did not get any soluble proteins that can be detected by the coomassie blue stained SDS–PAGE gel. During the time when I was trying to generate soluble kinase alone and LOV2Kinase constructs from various source, a paper from Matsuoka *et. al.* came and showed the phototropin kinase activity using a GST–tagged LOV2, GST–tagged Kinase, and GST–tagged LOV2Kinase proteins from *At*Phot2 (*112*). Although I tried to regenerate the same exact constructs as indicated from the paper (see fig. 3.2 for details), I was not able to generate sufficient amount of protein to do any biophysical or biochemical characterization (no visible protein band in the gel can be detected by eyes, data not shown). Therefore, I was not able to repeat their result and verify them in other system like *As*Phot1 or *At*Phot1. Since then I have focused my attention to generate various LOV2Kinase and kinase constructs using only *As*Phot1 with His_6, MBP or MBP–His_6 affinity tags (see Fig. 3.2).

Although among all constructs tested, I was able to obtain some soluble MBP– or MBP–His_6 tagged LOV2Kinase and kinase, they either degraded (in the kinase case) or was confirmed to be in the form of soluble aggregate by gel filtration (in the LOV2Kinase case) and, therefore, could not be used for future biochemical and biophysical characterization, or structural determination.

Very recently, functional studies from Kaiserli *et. al*., have also demonstrated that the role of the LOV2 domain as a dark state repressor of the phototropin1 activity both *in vitro* and in the plant Arabidopsis thaliana (*113*). The role of the LOV1 domain, however, has been consistently shown not to act as the dark–state inhibitor, although it has been shown to be involved in the chloroplast accumulation at high light intensities (*112, 113*).

Since no structure of the LOV domain coupled to its downstream effector domain has been unpublished, the question of how the LOV2 domain interacts with its kinase domain remains. Since the full–length protein have been shown to be successfully expressed in Sf9 (*Spodoptera frugiperda*) insect cells (*91*), it is likely that the LOV2Kinae and the kinase domain can be expressed using the insect cells instead.

Figure 3.1 Models for *As*Phot1 kinase regulation. First model suggests a dark state autoinhibition mechanism of the LOV2 domain. The second model suggests that the LOV2 domain acts as a lit state activator. Models adapted from Harper *et. al.* (*86*).

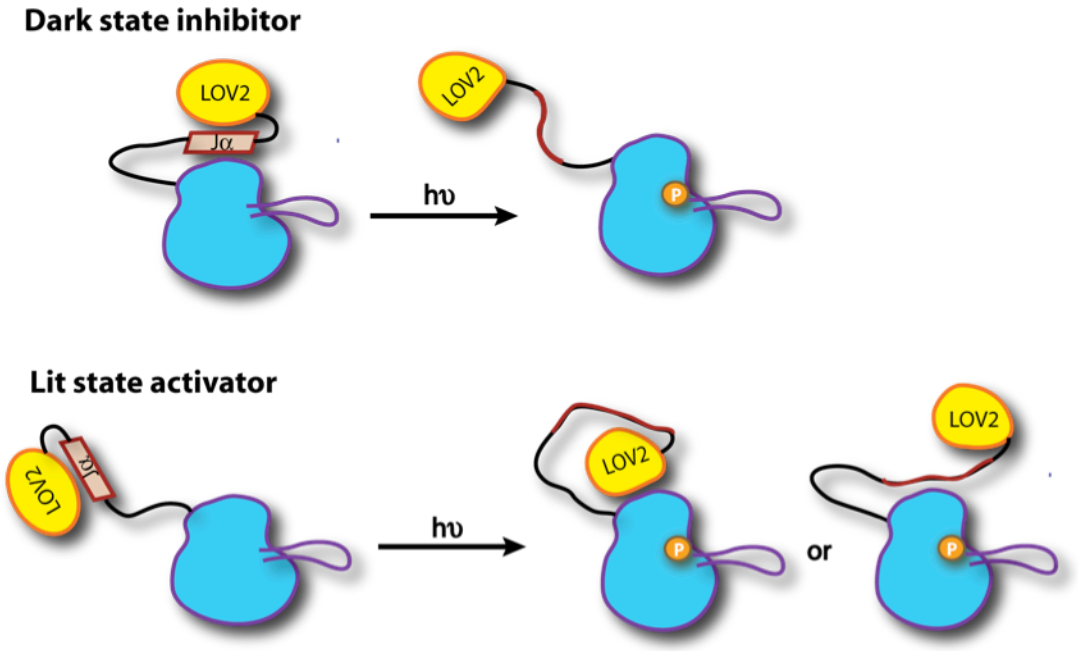
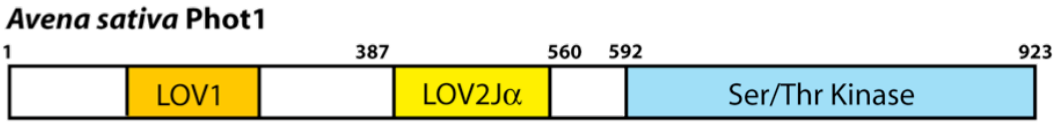


Figure 3.2 Total constructs generated for *in trans* and *in cis* LOV2–Kinase domains. Here *As*Phot1 is used to represent the general phototropin domain architecture. Three different sources of phototropins were used: *Avena sativa* phototropin1 (*As*Phot1) and *Arabidopsis thaliana* phototropin1 and 2 (*At*Phot1 and *At*Phot2). The N– and C–terminal amino acid numbers in each kinase and LOV2kinase constructs are listed in the figure. These constructs were cloned into different expression tags (labeled with boxes) to see which condition yields soluble proteins. Proteins that were insoluble (in the pellet) are indicated in black numbers. The constructs that show some soluble protein in the supernatant fraction were indicated in blue numbers.



MBP	***387-923***	His_6Gβ1	387-923	MBP-His_6	***387-923***
	400-923		400-923		***518-923***
	404-923		404-923		***592-923***
	518-923		518-923		
	556-923		556-923	GST	387-923
	567-923		567-923		592-923
	581-923		581-923		
	592-923		592-923		

***At*Phot1**

His_6	661-996	His_6Gβ1	658-996	GST	661-996
	449-996		661-996		449-996
			449-996		

***At*Phot2**

His_6	575-915	His_6Gβ1	575-915	GST	575-915
	363-915		363-915		363-915

Figure 3.3 Sample solubility test from various MBP tagged constructs in *As*Phot1. A: MBP–JαKinase (518–923), B: MBP–Kinase (556-923), C: MBP–Kinase (567-923), D: MBP–Kinase (581–923), and E: MBP–LOV2kinase (400–923). Various buffer conditions were tested for each constructs to see which condition yield the most soluble proteins. Pi indicates 50 mM phosphate buffer was used, and Tris indicates 50 mM Tris buffer was used. 100, 300, and 500 indicate the concentration of NaCl in mM added in each buffer. Here only the sup fractions from the lysate were shown in the gel. While most fractions were insoluble, MBP–Kinase (518–923) and MBP–Kinase (581–923) seemed to be soluble in some conditions shown in A and D.

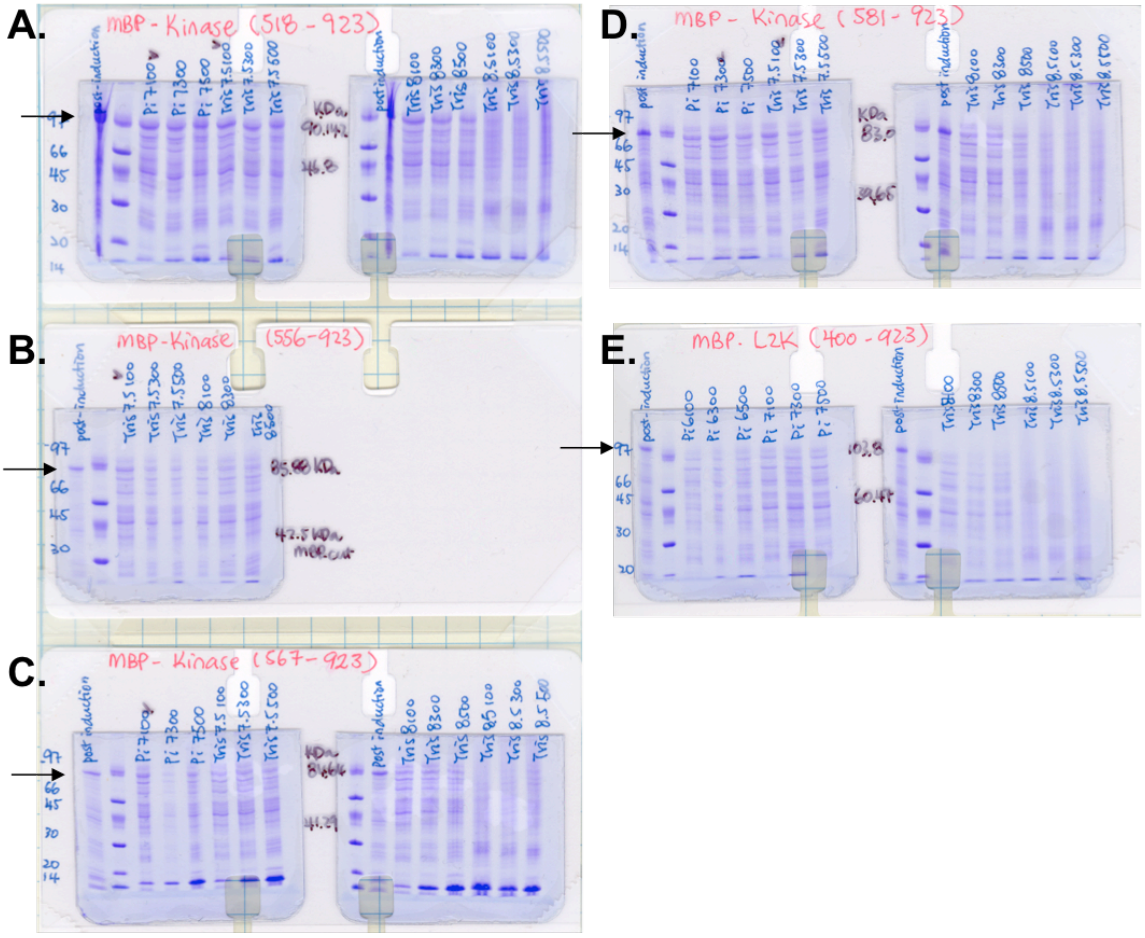


Figure 3.4 Solubility test for MBP tagged AsPhot1 kinase domain (Asphot1 KD, residue 592 to 923) and LOV2Kinase domain (Asphot1L2K, residue 387 to 923). Various buffer conditions were tested for both constructs to see which condition yield the most soluble proteins. A: pH6.0 buffer containing 50 mM sodium phosphate with 100, 300, and 500 mM NaCl. B: pH7.6 buffer containing 50 mM Tris with 100, 300, or 500 mM NaCl. C: pH8.1 buffer containing 50 mM sodium phosphate with 100, 300, and 500 mM NaCl. D: pH7.0 buffer containing 50 mM sodium phosphate with 100, 300, and 500 mM NaCl. Pellet and sup fractions were shown in the gel. While there were no dramatic differences between different conditions for L2K, the best condition for KD is shown in C (in red box).

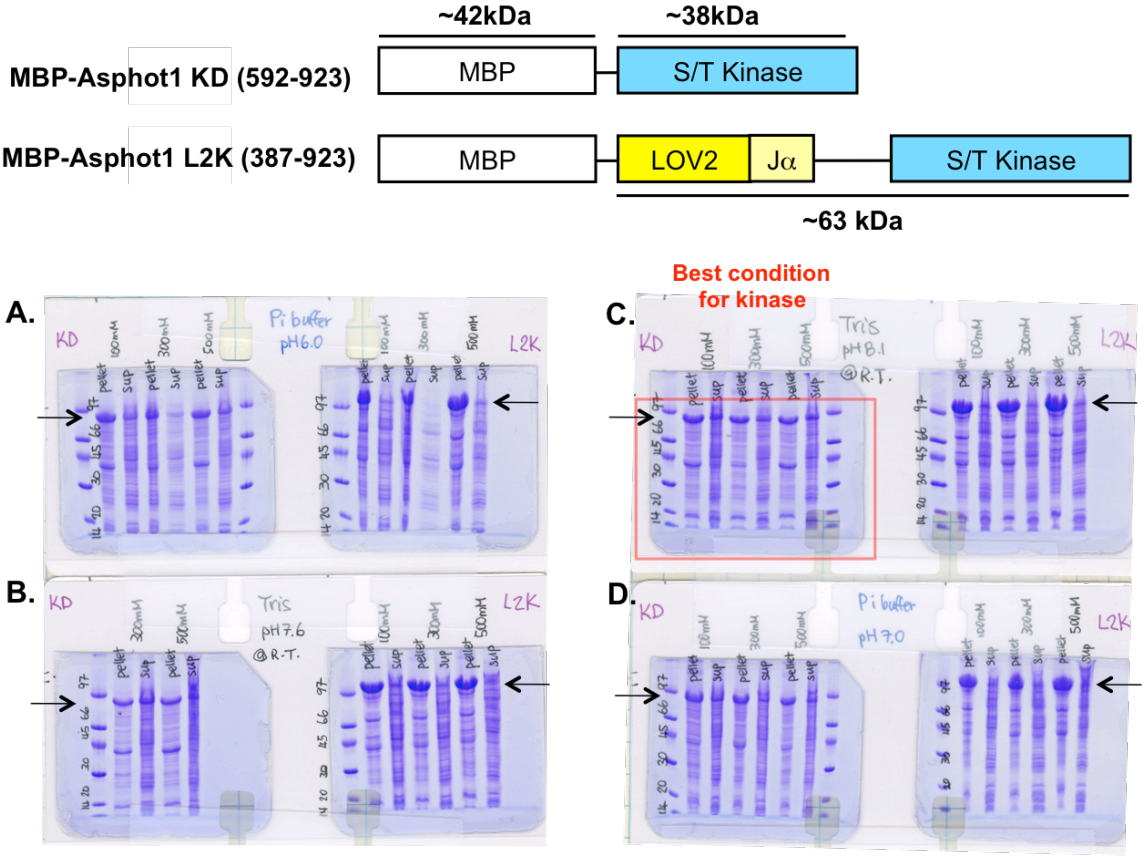


Figure 3.5 Initial characterization of LOV2–Kinase domains. Domain architectures of MBP tagged AsPhot1 kinase domain (Asphot1 KD) and LOV2Kinase domain (Asphot1L2K) were shown. The N– and C–terminal amino acid numbers in the kinase and LOV2kinase constructs are listed and the expected molecular weight were indicated. Proteins were purified by amylose affinity column followed by MonoQ column. At the end, the protein is in pH7.5 buffer containing 50 mM Tris, 200 mM NaCl, and 1 mM EDTA. Protein purity was verified by Phast gel. The UV-visible spectra of Asphot1L2K protein are collected at 10ºC to slow down the dark state recovery. The dark state spectrum is shown in black (with three maximum between 410–500 nm, indicating the protein is bound to FMN), and lit state spectrum is shown in red. The recovery of the protein from the lit state to the dark state is monitored at every 50s as shown in purple, orange, green, and blue spectra.

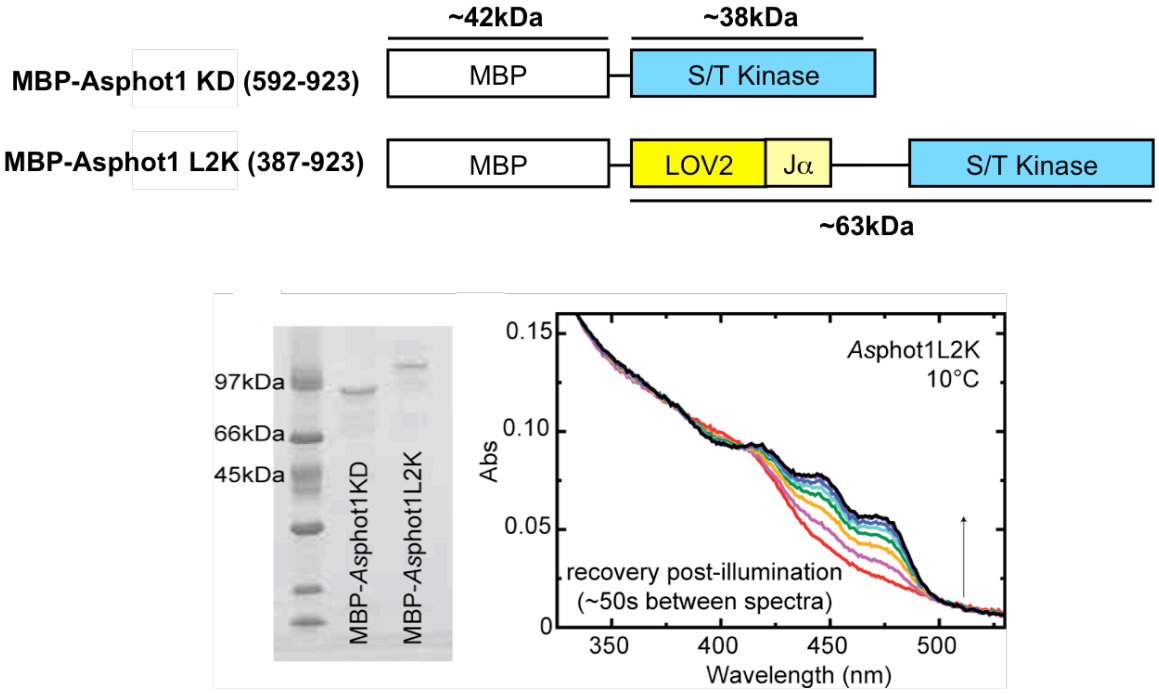


Figure 3.5 Summary of *As*Phot1 constructs tested. The complete list of the proteins that were cloned is shown in page 77.

MBP

518-923 : Protein starts to degrade during 1st affinity purification. Completely gone after 2nd step purification (MonoQ)

MBP-His_6

387-923 : protein degrades during purification, but was able to isolated partial truncated fragments. Mass Spec. result shows 2 fragments at M.W. of 64.2kDa and 63.6kDa (the original uncleaved is 106.7kDa, tag included)

518-923 : protein degrades during purification, at the end all got chew up.

592-923 : Protein also degraded. But was still able to get full length protein. However, gel filtration determine they are all soluble aggregate.

Chapter 4 Biochemical and biophysical characterization of EL346

Abstract

Light governs and regulates a wide variety of biological functions in plants. Recently, different photoreceptor domains such as BLUF, LOV, PYP, phytochrome, and sensory rhodopsin have been identified in bacteria. Here I present a detailed biophysical and biochemical characterization on a photosensory histidine kinase named EL346, a 346 residue bacterial photoreceptor from *Erythrobacter litoralis* HTCC2594 that contains a LOV (Light–Oxygen–Voltage) domain as its sensory domain. Nucleotides such as ATP or AMPPNP have two important effects in EL346: first, they significantly enhanced EL346 structural stability as shown by NMR and limited proteolysis studies. Second, the addition of either ATP or AMPPNP accelerates the decay of the photogenerated protein–FMN bound within the LOV domains by nearly two–fold. These data demonstrate that elements external to the LOV core domain affect the photocycle. Another molecule, TCEP, was found to have an even more profound effect than imidazole in facilitating dark–state recovery rate in the EL346 LOV domain. Blue light regulates EL346 autokinase activity by modulating the affinity for ATP, and this light signaling effect is directly propagated into two of the twenty–three *Erythrobacter litoralis* response regulators. Finally, our data show that

EL346 is a monomeric histidine kinase that undergoes self–phosphorylation, and the LOV domain acts as a weak inhibitor for the autokinase activity.

4.1 Introduction

Two–component signaling systems (TCS) allow bacteria to sense and adapt to changes in their environment to increase their chances for survival (*114*). Classically, these systems are composed of two components. The first component, a histidine kinase, detects specific environmental stimuli and undergoes autophosphorylation at one of its own histidine residues. This phosphate group is later transferred to an aspartic acid residue on the second response regulator (RR) component. These proteins are typically composed of a conserved N–terminal receiver domain and a diverse C–terminal effector domain that is involved in transcriptional regulations (reviewed in (*39, 114*)).

Of particular interest is the mechanism by which the sensory domains convert changes in environmental stimuli into altered activity of the kinases. Most histidine kinases are dimeric, in which the autophosphorylation occurs in a *trans*–phosphorylation manner between the two subunits. A typical histidine kinase contain two functionally and structurally distinct parts: a variable N–terminal input sensory domain and a conserved C–terminal histidine kinase (HK) domain that in turn is composed of two subdomains, the dimerization and histidine phosphotransfer (DHp) domain and the catalytic and ATP–binding (CA)

domain. While the sensory domains differ based on the type of signals they can detect, the output histidine kinase domain contains several conserved motifs that are located near the ATP binding pocket (*74*). Additionally, the CA domain shares a common structural motif with the ATPase family (*50*).

As an important environmental cue that regulates HK, light is involved in many biological functions in living organisms particularly in plants, which can sense and response to changes in light using a variety of photoreceptors (*115*) (*116*). Photosensitive proteins provide a number of favorable characteristics for biophysical and biochemical studies of conformational changes and signaling, making them well–suited for studies of TCS. Here, we focused on a light sensing histidine kinase (EL346) from the marine bacteria *Erythrobacter litoralis HTCC2594*. A previous study has shown that EL346 contains a N–terminal light sensing LOV (Light–Oxygen–Voltage) domain and a C–terminal histidine kinase domain, which can undergo autophosphorylation in a light regulated manner (*76*). LOV domains belong to a subset of the family of PAS (Per–ARNT–Sim) sensory domains, a group of small (~120 aa) protein domain senses a wide array of environmental stimuli such as light, redox potential, oxygen, and small ligands and transmitting this information downstream through altered protein–protein interactions (*8, 23*). The LOV domain utilizes a flavin cofactor (FMN in EL346) to sense blue light, and can exist as a single domain protein or coupled to a wide variety of domains to regulate many different type of biological functions (*7*).

Upon light stimulation, photochemistry converts a non–covalent LOV:FMN complex into a covalent one by forming a reversible covalent linkage between a conserved Cys in the LOV domain and the C4a from the FMN isoalloxazine ring (*2, 6, 83*). This process induces conformational changes that propagated from the flavin–bound LOV core to external elements that are N– or C–terminal to the LOV core, resulting in the activation of effector domains or downstream signaling partners (*25, 26*). Previous studies in the *A. sativa* Phot1 LOV2 (*As*LOV2) domain have shown that light induced conformational change in the LOV domain by unfolding a C–terminal α helix (Jα), which docks onto the β sheet surface in the dark state of *As*LOV2 (*25*). Subsequent mutagenesis studies showed that light signal is propagated from the core of the LOV domain through the β sheet to the Jα helix in the full–length protein and that these are essential to phot1 regulation (*118*). Currently, the mechanism of how light–induced conformational changes are propagated from the LOV domains to their effectors or downstream targets is not well defined.

A common problem in studying full–length histidine kinase structures is that the majority of histidine kinases are membrane–bound or membrane–associated, making it challenging to obtain soluble and homogenous full–length proteins that are most amenable to biophysical study. Several soluble heme–based sensor histidine kinases that contain both the sensory and histidine kinase have been well characterized, including DosT, DevS, and FixL (*119-123*).

However, structural information is only available for the sensory domains of these proteins (*10, 24, 29, 124-128*). On the other hand, there are many high resolution structures of the CA domains without their sensory counterparts (*46, 51, 129-133*). Structural modeling and small–angle X–ray scattering (SAXS) are the main tools used to elucidate multidomain interaction of these histidine kinases and to provide insights into how the sensory and kinase domains could interact. In addition, limited structural information is available on how activated kinases interact with response regulator as an intact two–component complex (*134, 135*). Overall, the molecular detail of how sensory domains regulate kinase activity remains unclear.

Taking advantage of the stability and solubility of the EL346 histidine kinase, we have performed a detailed biophysical and biochemical study of the full–length EL346 in an effort to understand how the LOV domain regulates the histidine kinase activity via light. Here, we describe experiments that demonstrate the importance of nucleotides in ordering and stabilizing EL346 structure and show light induces conformational changes in EL346, making it more susceptible to proteolytic cleavages in the light state. Our kinetic data suggest a close interdomain interaction between the LOV and the HK domains, showing the first evidence that an accessory domain can affect the LOV domain photocycle. Chromatography and light scattering results confirm EL346 is a monomer in both the dark and light states, revealing a novel *cis*–

autophosphorylation mechanism in histidine kinases. Our functional data show that light plays an important role in regulating EL346 autophosphorylation by changing K_m for ATP under limiting ATP conditions, an effect that is propagated through phosphotransfer to two response regulators we identified. Finally, the HK domain alone autophosphorylates better than the full–length protein and is capable of transferring the phosphate group to either response regulators. The proposed roles of the LOV domain are discussed in the end.

4.2 Materials and methods

4.2.1 Cloning

DNA encoding the full–length *Erythrobacter litoralis* EL346 histidine kinase (*76*) and fragments of the EL346 (residue 1–134 and 121–346; designated EL134 and HK) were subcloned into the pHis–Gb1–Parallel1 (*25*) expression vector between the StuI and XhoI restriction enzyme sites. There is a N–terminal TEV protease cleavage site designed to remove the affinity tag, and 6 additional N–terminal residues (GEFKEL) are left as cloning artifact after the His–Gb1 tag is cleaved. A H142Q mutation of EL346 was generated from wildtype DNA using QuickChange II XL site–Directed Mutagenesis Kit (Stratagene) followed by the manufacturer's instructions. Response regulators ELI07655 (GeneID: 3871024, designated RR1) and ELI10215 (GeneID: 3869106, designated RR5) were amplified by PCR from *Erythrobacter litoralis* strain HTCC2594 using

genomic DNA as a template and pairs of primers including 5' NcoI and 3' KpnI restriction endonuclease sites (5'–TTTTTTCCATGGGTATGCCCAAAGTTCTC GTGCTCGAGAC–3' and 5'–TTTTTTGGTACCTCAGGCCGGTTCCTTGTCC CC–3' for ELI07655 (RR1), and 5'–AAAATTCCATGGGTATGTCTGCTTCAC AAAAAATCGCTGCC–3' and 5'–AAAAAAGGTACCTCAAACGGCCGTAG GGCTGTT–3' for ELI10215 (RR5)). These purified PCR products were subcloned into pHis–parallel1 expression vector (*136*). All constructs were verified by DNA sequencing.

4.2.2 Protein expression and purification

All constructs were overexpressed in *E. coli* BL21 (DE3) cells (Novagen) in either LB or M9 minimal media containing 1 g/L $^{15}NH_4Cl$ for U–^{15}N samples and supplemented with 3 g/L $^{13}C_6$–glucose for U–$^{15}N/^{13}C$ labeled samples. Cultures were grown at 37°C, induced at an A_{600} of 0.7–1.0 with 0.12 g/L isopropyl b–D–thiogalactoside (IPTG), and allowed to grow for an additional 16 h in the dark at 20°C. Cells were harvested, resuspended into 50 mM sodium phosphate (pH 8.0) and 100 mM NaCl buffer, lysed by extrusion, and clarified by centrifugation at 10,000 g for 45 min. The supernatant was filtered (0.45 μm) and loaded into an affinity column packed with 7 mL of High Performance Ni Sepharose resin (GE Healthcare). Protein was eluted with a 20 C.V. gradient set from 15–500 mM imidazole using buffer containing 50 mM sodium phosphate

(pH 8.0), 100 mM NaCl, and 500 mM imidazole. The desired protein fractions were buffer exchange into 50 mM sodium phosphate (pH 8.0) and 17 mM NaCl buffer, and the fusion tag was cleaved by increasing 1 mg His_6–TEV protease per 15 mg of protein for 2 h at room temperature. The tagless protein was separated from its affinity tag and His_6–TEV protease by the Ni^{2+} column, and was further purified by size–exclusion chromatography with 50 mM Tris (pH 8.2) and 100 mM NaCl buffer. The final product was concentrated and stored at –20°C with 50% (v/v) glycerol. For NMR samples, however, we stored the protein at 4°C without glycerol. Sample purity was accessed by SDS–PAGE analyses using 20% homogenous PhastGel gels (Amersham Biosciences) and concentrations are determined from the predicted absorption coefficient (ε) (ε_{446}= 1,1800 M^{-1} cm^{-1} for flavin–containing proteins, ε_{280}= 1,615 M^{-1} cm^{-1} for RR1, and ε_{280}= 14,440 M^{-1} cm^{-1} for RR5) using the ProtParam tool from ExPASy website (*111*). All these steps were performed under dim red light condition for all flavin–containing proteins, and purifications of EL346 in particularly were finished in one day to minimize oxidation/aging of this protein. We observed that EL346 lost its activity as it aged; therefore, any kinase assays involving EL346 that was more than 2 days old were avoided.

4.2.3 NMR Spectroscopy

Proteins were prepared in buffer containing 50 mM Tris (pH 7.5), 100 mM NaCl, and 10% (v/v) D_2O unless mentioned otherwise. Additional factors such as $MgCl_2$, DTT, ATP or AMPPNP were added later depending on the experiments. All NMR experiments were performed at 25°C on a Cryoprobe–equipped Varian Inova 800 MHz spectrometer. Data were processed by nmrPipe (*92*) and analyzed with NMRview (*93*). The photoexcited state of EL346 for the NMR experiment was achieved by blue light illumination through a quartz fiber optic (10 m long, 0.6 mm in diameter) (*25*). Blue light was generated from a Coherent Sapphire laser running in a single wavelength mode at 488 nm with power levels set at 50 mW. The output end of the optic was inserted into a coaxial insert tube to prevent direct contact with the protein sample, and this coaxial insert tube was then placed inside of a 5 mm NMR tube and merged onto the top of the protein sample. $^{15}N/^{1}H$ TROSY and constant time $^{13}C/^{1}H$ HSQC spectra of light state of EL346 were recorded by preceding each transient in the NMR experiment with a 50 ms laser pulse during the 1.06 s recycle delay.

4.2.4 UV–visible Absorbance Spectroscopy and Photocycle Kinetics

UV–visible absorbance profiles were collected using a Cary50 (Varian) spectrophotometer. The dark and light states of all flavin–containing proteins were verified from the 250–550 nm absorption spectra. Dark state samples were

collected in the dim red light, and light state samples were generated via a photographic flash illumination. The dark state recovery profile of EL346 was monitored immediately after illumination by recording scans from 250–550 nm every 15 min for a total of 2.5 h.

To obtain time constants for the dark state recovery of flavin–protein adducts, we monitored changes in the absorbance at 446 nm after illumination. A_{446} was recorded every 30 s with an integration time of 0.1125 s for a period of time that was at least 5 times greater than τ. Time constants were determined by fitting measurements to monoexponential decays (GraphPad Software, San Diego, CA). All proteins tested were prepared in the pH 8.0 buffer containing 50 mM Tris, 100 mM NaCl, and 10% glycerol. Additional compounds such as nucleotides (ATP or AMPPNP), $MnCl_2$, and tris (2–carboxyethyl) phosphine (TCEP, Thermo Scientific) were added later to the protein solution prior to the kinetic measurements.

4.2.5 Limited Proteolysis

Limited proteolysis experiments were conducted on EL346 (11 mg) in buffer containing 50 mM Tris (pH 8.0) and 100 mM NaCl at room temperature. For nucleotide–containing samples, 5 mM AMPPNP was used. Trypsin was added to a final concentration of 11 μg/ml in a total reaction volume of 300 μL. Dark state experiments were done under dim red light while light state

experiments were performed under constant irradiation via a mercury arc lamp supplied with 200 W power and a blue glass filter (ThermoOriel instruments). 10 μL aliquots of reaction mixture were removed at various timepoints over the next 2 h and reactions were stopped by addition of 10 μL 2× SDS gel–loading buffer (100 mM Tris (pH 6.8), 0.2% Bromophenol Blue, 4% (v/v) SDS, 40 mM DTT and 20% (w/w) glycerol). SDS samples were heated at 90°C for 5 min prior to loading (5 μL) to a 15% homogenous Bis–Tris SDS–PAGE gel. Electrospray ionization mass spectrometry (ESI–MS) analysis were performed on selected proteolytic fragments treated with 1% TFA solution instead of 2× SDS gel–loading buffer, and the resulting fragment masses to the full–length protein sequence were analyzed using ProteinProspector v 5.2.2 proteomics tools (UCSF).

4.2.6 Kinase Assays

Autophosphorylation reactions were carried out at room temperature in 100 μL of reaction buffer (50 mM Tris (pH 8.2), 100 mM NaCl, 5 mM $MnCl_2$, and 10% (v/v) glycerol) containing 10 μM EL346 and 10 or 50 μCi (γ–^{32}P) ATP (6,000 Ci/mmol; PerkinElmer) mixed with unlabeled ATP (Sigma) to give a final ATP concentration of 1.8 or 500 μM, respectively. Reactions were initiated by addition of hot/cold ATP mixture and stopped by adding 3.3 μL of 4× SDS gel–loading buffer (500 mM Tris (pH 6.8), 200 mM NaCl, 40 mM EDTA, 0.2%

Bromophenol Blue, 10% (v/v) β–Mercaptoethanol, 4% (w/w) SDS, and 20% (v/v) glycerol) to 10 μL aliquot of the reaction mixture at desired time points. Phosphotransfer reactions with RR1 and RR5 (25 μM each) were carried out in the same manner as the autophosphorylation. Experiments that did not require time point collection were scaled down to 10 or 20 μL sample volumes instead of 100 μL.

Dark state experiments were performed under dim red light, while light state experiments were performed under constant illumination via a 200 W mercury arc lamp supplied and a blue glass filter (ThermoOriel instruments). Samples were kept on ice before loading onto SDS–PAGE gels. Unincorporated cold/hot ATP was separated using 15% SDS–PAGE gels running under constant voltage at 220 V for 40 to 50 min until the dye front reached to the end of the gel. The lower portions of these gels containing the dye front with unincorporated ^{32}P–ATP were removed, dried under vacuum at 80°C for 45 min and exposed to Fujifilm phosphorimager screens for 0.5 to 3 h at room temperature. Bands containing ^{32}P–phosphorylated proteins were visualized using a Fujifilm FLA–5100 phosphorimager and the intensity of these bands were quantified using the ImageQuant software included from the phosphorimager. To visualize protein bands, the gels were treated with Coomassie Brilliant Blue stain instead without drying.

4.3 Results

4.3.1 EL346 is a blue light sensing histidine kinase that belongs to HPK_{11} subfamily

EL346 is a cytosolic histidine kinase that contains a N–terminal light sensory LOV domain and a C–terminal HK domain that is in turn composed of DHp and CA subdomains (Fig. 4.1a). The domain architecture shows that EL346 belongs to class I histidine kinase. Sequence alignment (Fig. 4.2) revealed that EL346 belongs to the subfamily 11 (HPK_{11}) based on Grebe and Stock's classification (*33*).

It has been previously shown that EL346 displays the signature LOV domain UV–visible absorbance profile and undergoes light–regulated changes in autophosphorylation activity(*76*). Single LOV and HK domain constructs were also made from EL346, allowing us to study the features of the individual domains. The domain boundaries between the LOV and HK domains in EL346 are unclear due to lack of structural information. We therefore relied on secondary structure predictions via Jpred (*137*) and sequence/structural comparisons using the LOV domain structure from *As*LOV2 (*25*) and kinase domain structure from HK853 (*51*). Jpred prediction suggested that there was approximately a ten–residue linker between the core of LOV domain and the beginning of the HK domain, suggesting that the α helix from the DHp domain might serve analogously to the *As*LOV2–like Jα helix and could potentially

interact with the LOV domain in EL346. After screening different constructs for expression and solubility, we obtained a soluble LOV–domain only construct (EL134, residue 1 to 134) and the histidine kinase domain construct (HK, residue 121 to 346) (Fig. 4.1).

4.3.2 Photocycle of EL346: Effects of nucleotide and other small molecules

Like other LOV–containing proteins, the dark state UV–visible absorption profile of EL346 has been previously demonstrated to possess three absorbance maxima between 400–500 nm (*76*). These three maxima diminished upon illumination, and three isosbestic points (at 336, 393 and 410 nm) were observed in the dark and light state spectral overlay (Fig. 4.3a). Our kinetic data (Fig. 4.3c and Table 2) showed that the dark recovery of the full–length EL346 alone is quite slow (τ = 54.6 ± 1.2 min) as measured at pH 8.2. Surprisingly, addition of nucleotide, ATP or AMPPNP (a derivative of ATP that contains a non–hydrolysable βP–γN bond), supplemented with 5 mM $MnCl_2$ as the divalent cation slightly accelerated recovery (τ = 41.5 ± 1.0 min for ATP and τ = 32.2 ± 0.4 min for AMPPNP, Table 2). Adding $MnCl_2$ to EL346 without nucleotide did not change $t_{1/2}$ (data not shown), confirming that the accelerating effect of rates is solely due to the addition of nucleotide.

Giving that nucleotide binding is mediated by the CA domain of the histidine kinase, we assumed that nucleotide–dependent changes in the LOV

domain photocycle would result from an interdomain interaction between the sensory and catalytic domains. To test this hypothesis, we performed the same kinetic experiments on the LOV–only protein, EL134. We observed a similar dark state recovery in EL134 (τ = 70.8 ± 1.1 min, Table 2) in the same buffer condition without nucleotide, and addition of nucleotide such as AMPPNP did not have any effect on the photocycle (data not shown).

Of a practical note, our initial characterization revealed that TCEP, a stable and powerful reducing agent used to eliminate disulfide linkage of proteins, significantly accelerated the recovery with increasing TCEP concentration (τ = 4 min at 45 mM of TCEP, Fig. 4.4a). Kinetic experiments performed using various concentrations of TCEP with fixed concentration of EL346 showed that TCEP bound to EL346 specifically with an estimated K_d of 26.2 ± 4.5 mM (Fig. 4.4a). This contrasts with imidazole, which is commonly used to accelerate LOV domain dark state recovery kinetics to facilitate biophysical studies (*101*). Similar experiments done with various imidazole concentrations showed a much weaker effect on the photocycle time constant, apparently is non–specific in the range of concentration we measured (Fig. 4.4b). To see whether the kinase domain is involved in the changes in photocycle by TCEP, we repeated the same kinetic experiment using EL134. We found that the dark state recovery in this LOV domain alone protein was accelerated in a very similar manner as the full–length protein (τ = 6.2 min at 35 mM TCEP, data not shown). Therefore, we

concluded that TCEP affects the EL346 photocycle by binding specifically to the EL346 LOV domain.

4.3.3 EL346 is monomeric both in the dark and light states

Trans–phosphorylation is a well–accepted autophosphorylation mechanisms among histidine kinases since the histidine kinase domains studied to date undergo dimerization using the DHp domain. Under this assumption, we expected EL346 to dimerize upon kinase activation via its DHp domain. EL346 slow photocycle allows for characterization of oligomeric state using gel filtration. Analytical size exclusion chromatography (SEC) showed that dark state EL346 is monomeric at 100 μM concentration (Fig. 4.5a). A small dimer peak was observed in the elusion profile but it was eliminated by addition of reducing agents using DTT or TCEP. Repeating SEC with addition of 5 mM ATP in the sample and the buffer were analogous to experiments without ATP (Fig. 4.5a). To confirm the molecular weight of the monomeric species we observed in all these conditions, we coupled the size exclusion chromatography to multiangle light scattering (MALS) (Fig. 4.5b). MALS analysis showed that the weight–average molar mass (Mw) for EL346 in the ATP free condition is 39.1 kDa in the dark state, and 37.9 kDa in the light state, and it is 39.3 kDa for the dark state and 36.4 kDa for the light state when ATP is present. In all cases, the molecular weights we obtained agreed well with the theoretical monomeric

molecular weight of 38.6 kDa, supporting the monomeric results we obtained from the size exclusion chromatography.

4.3.4 Nucleotide binding confers additional structural stability and improves light–induced structural changes of EL346

Initial dark and light state $^{15}N/^{1}H$ TROSY spectra of nucleotide–free EL346 acquired at 800 MHz were reminiscent of a poorly–behaved protein (Fig. 4.6a). Notably, these spectra demonstrated significant peak overlap between 7.8 and 8.8 ppm (^{1}H) with widely variable linewidths, consistent with protein of marginal stability and possible aggregation. Few peaks were observed outside the crowded regions (Fig. 4.6a), presumably due to protein dynamics or size (EL346 = 38.6 kDa). Although we could observe few changes in peak location or linewidth between the dark and the light state spectra (Fig. 4.6a), it was obvious that more detailed information could not be obtained without improvement in the spectral quality.

We observed spectral improvement by including 5 mM ATP into our EL346 samples without changing buffer, salt, or pH conditions (Fig. 4.6b). The resulting spectra have significantly more peaks of uniform linewidth as mentioned above. Importantly, these spectra are still missing about 30% of peaks, presumably due to the larger size of EL346 or slightly basic pH buffer (pH 7.5). Notably, we can still observe light–induced structural changes between the dark

(black spectrum in Fig. 4.6b) and the light state spectra (red spectrum in Fig. 4.6b). This improvement by nucleotide addition can also be observed in the CT–$^{13}C/^{1}H$ HSQC spectra in the methyl region (Fig. 4.7a, b). Addition of AMPPNP also improves the overall numbers, distribution, and intensity of the peaks, as similarly observed with ATP. Notably, light induced peak changes occurred in the 8 methionine residues in EL346 (represented in the grey boxed region in Fig. 4.7a, b) are more pronounced in the AMPPNP spectra versus the AMPPNP–free spectra.

Limited proteolysis provides another way to evaluate the structural changes between the dark and light states in EL346. Light clearly caused some changes in the EL346 structure that made it more susceptible to trypsin cleavage. We observed a two–fold increase in the light state cleavage rate on the first four proteolytic fragments in nucleotide–free samples (white and black circles, Fig. 4.8a, b). At this condition, full–length EL346 completely disappeared within 20 min in the dark state (Fig. 4.8a) and 10 min in the light state (Fig. 4.8b). There were no obvious differences in the cleavage rate observed on the subsequent proteolytic fragments (triangles) between the dark and light state (Fig. 4.8a, b). Light also induced structural changes that facilitate the cleavage of light state EL346 in the present of 5 mM AMPPNP, however, with a difference of having a more sustained intermediate (black circle and triangle, and grey triangle) than in the nucleotide–free sample. In this case, the first three proteolytic fragments

(white circles) generated were digested in a very similar time manner in the dark or in the light states (both disappeared around 15 min, Fig. 4.8c, d). The main light–induced differences lied on the subsequent proteolytic fragments (black circle and triangle, and grey triangle), which again showed a two–fold accelerated cleavage between the dark and the light state but in the present of AMPPNP. Notably, these fragments appeared to be significantly stabilized for a longer period of time compared to the nucleotide free condition. These newly stabilized fragments most likely arised from the kinase domain, since binding of AMPPNP would help ordering and stabilizing the kinase domain as we observed from our NMR data. These results showed light induced conformational changes that led to a less stable lit state structures with or without AMPPNP. As a result, the light state is cleaved two–fold faster than the dark state with or without AMPPNP. On the other hand, nucleotide plays an essential role for stabilizing EL346 structure, most likely by stabilizing the kinase domain. In all condition, the dark and light species is digested into a single proteolytic fragment after 6 h, with the molecular weight around 15 kDa. Analyzing the ESI–MS results showed this final fragment is the LOV domain.

4.3.5 Autophosphorylation and phosphotransfer of EL346 is light regulated in an ATP dependent manner

To complement these analyses of light–induced structural changes in EL346, we optimized an EL346 autophosphorylation assay under varying pH and divalent cation concentration and species. Our studies demonstrated EL346 is most efficient under basic (pH 8.0 to pH 9) conditions in the presence of 5 mM $MnCl_2$, yielding a maximum ^{32}P incorporation of 15–20%. Notably. EL346 was 2–fold less efficient in the presence of 5 mM $MgCl_2$. Unexpectedly, we initially found that EL346 was equally effective at autophosphorylation under dark and light conditions (with nearly identical initial rate) under 500 μM ATP condition (Fig. 4.9a). To verify that the small light/dark difference was not an artifact of using excess of cold ATP (500 μM, 50–fold more concentrated than EL346 and expected to be greater than Km for ATP) we performed analogous experiments using cold ATP concentration from 500 to 1.8 μM (Fig. 4.9b). At limiting ATP concentrations (1.8 μM), the light state initial rate of autophosphorylation is three fold faster than the dark state initial rate. The difference between the dark and light state autophosphorylation diminished as we increased the ATP concentration (data not shown). Together, these results consistent with the effect originating from photo–induced alterations in the K_m.

To see if these light–induced differences at 1.8 μM are relevant to subsequent steps in EL346 function, we performed phosphotransfer assays. To identify a cognate response regulator for EL346, we cloned and tested the 23 predicted response regulators in *Erythrobacter litoralis (60)* by performing

phosphotransfer profiling (*67*). Our initial screen found two response regulators, RR1 and RR5, which were phosphorylated by EL346. Both proteins are soluble, stable and expressed with good yield (> 20 mg/mL) in *E. coli* (Fig. 4.1b). Phosphotransfer of RR1 and RR5 in the dark and in the light states, when performed with 500 μM ATP, followed the same trend as the EL346 autophosphorylation and did not show any sign of light regulation (data not shown). However, light induced differences are observed in both response regulators when we switch to 1.8 μM ATP condition (Fig. 4.9c, d). In this case, light resulted in a 5.6 fold increase in the initial rate in the light state phosphotransfer for RR1 and a 1.6 fold increase in the initial rate in the light state phosphotransfer for RR5. Interestingly, RR1 exhibited additional phosphatase activity in the light state that is not observed in RR5. These data show that light indeed has an effect on EL346 autophosphorylation and this effect is transmitted directly to the subsequence phosphotransfer of the response regulators. To complement our previously mentioned biophysical studies where we showed EL346 is a monomer at all states, we performed an autophosphorylation assay over a wide range of EL346 concentration. The result demonstrated a linear dependence of signal intensity on protein concentration over the range of 0.5–50 μM (Fig. 4.10). Together, these data indicated that EL346 did not undergo a monomer–dimer transition like other histidine kinases and remained monomeric even when the kinase is active. Together, our results indicated that EL346 must

undergo *cis*–autophosphorylation. The fact that EL346 undergoes cis–autophosphorylation may explain the phenomenon why EL346 is not totally inactive in the dark state.

4.3.6 The Role of the LOV domain in EL346 activity

The results from the kinase assays shown that EL346 is active even in the dark state. Mutagenesis result confirms that residue H142 is phosphorylated (Fig. 4.11a). To see whether the LOV domain is required for the EL346 autokinase activity, we performed autophosphorylation and phosphotransfer assays on the HK domain alone (residues 121 to 346). The kinase assay shows that HK alone is capable of undergoing both autophosphorylation (Fig. 4.11a) and phosphotransfer (for both RR1 and RR5, Fig. 4.11b), indicating that the LOV domain is not required for these two functions. Notably, autophosphorylation was slightly more efficient in HK alone compared to the full–length protein when equal concentrations of these two proteins are used (signal intensity is about 25% less in the full–length) (Fig. 4.11b). These finding, combined with the biophysical studies from NMR, limited proteolysis, gel filtration, and MALS showed that although the LOV domain is interacting with the HK domain, its interaction with the HK domain is not an absolute requirement for EL346 kinase activity. Rather, the LOV domain acts as a weak inhibitor of the kinase activity, in addition to attenuating the K_m between the dark and light state.

4.4 Discussion

To understand the molecular basis of how light signal is detected and transmitted inside the bacteria cells to regulate essential biological functions, besides knowing the structure, one must understand the physical and biochemical properties of bacterial photoreceptors. We are in particularly interested in EL346 because of its blue light sensing property and the potential it has to reveal how the sensory domain regulates the HK domain. Here we report a detailed characterization of this LOV–containing histidine kinase. With these biophysical and biochemical evidence, we show that light plays a regulatory role to affect the signaling transduction in the EL346 initiated two–component system in *E. litoralis*.

4.4.1 Variations in the LOV domain photorecovery: intrinsic vs. small molecules

Depending on the organisms and/or functions that the LOV domains are involved, LOV domains have been shown to have very different photocycle kinetics, ranging from seconds (i.e. *As*LOV2, τ = 68.3 s (*118*)) to hours (i.e. *Bs*YtvA_LOV, τ = 65 min (*138*) and VIVID, τ = 5 h (*139*) or days (i.e. FKF1, τ = 89.5 h (*107*)), and in some cases it is considered effectively irreversible (i.e. ZTL and LKP2 (*81*) and BM_LOV (*76*)). In the isolated EL346 LOV domain, the dark state recovery rate is comparable to *Bs*YtvA LOV domain (τ = 70.8 min

for EL134 and τ = 65 min in *Bs*YtvA_LOV (*138*)). Sequence alignment between these two LOV domains showed that they are 43% identical.

Interestingly, full–length EL346 photocycle (τ = 54.6 min in the nucleotide free condition) is a bit faster than the LOV domain alone (τ = 70.8 min), a similar observation that has also been reported in *Bs*YtvA, where τ for the full–length and isolated LOV domains are 43 and 65 min, respectively (*138*). A possible explanation of the slower ground state return rate in the isolated LOV domains could be a result of increased protein stability when accessory domains or additional factors are present. In fact, when full–length EL346 is structurally stabilized by addition of nucleotides as shown by NMR and limited proteolysis studies, the photocycle becomes faster for almost two fold when compared to the nucleotide free condition, in these cases τ is 32.2 min when AMPPNP is added and τ is 41.5 min when ATP is added.

As expected, additional factor such as high concentration of imidazole can accelerate the ground state recovery rate in EL346 and EL134 as reported in other flavin proteins such as LOV domain and BLUF domain. The above results demonstrated the first time that there are factors (i.e. protein domains other than imidazole) outside of the LOV domain that can have an accelerated effect on the dark state recovery rate. Additionally, a reducing agent TCEP was found to accelerate the ground state recovery rate in both EL134 and EL346 even more significantly and efficiently than imidazole. Unlike imidazole, which has been

previously shown to facilitate recovery rate by acting like a base catalyst in the LOV domain core (*140*), the mechanism of TCEP acceleration of dark state recovery is unclear. More tests need to be done to other LOV domains to ascertain whether the acceleration of the dark state recovery by TCEP is general.

4.4.2 Diversity of the histidine kinase domains

Grebe and Stock classified histidine protein kinases (HPKs) into 11 subfamilies based on the conserved sequence in the homology boxes located in the DHp and CA domains, with the majority of HPKs belong to HPK_1 and HPK_4 (*33*). Although the structures of these domains remained similar, the sequence identity of these conserved boxes varies among different subfamilies. Amino sequence reveals that EL346 belongs to HPK_{11}(*33*). The H–box bears the phospho–accepting histidine (H142 in EL346) that is located in the first α helix in the DHp domain like all other HPKs, but with very low similarity to all other subfamily of HPK in the distribution of charged and hydrophobic residues around this histidine, particularly not having the conserved H–box proline. EL346 also contains the X–, N–, D–, G–boxes but no F–box, a homology box that is not definitely conserved in HPK_{11} based on the sequence alignment. Some of the bacterial strains have most of their HPKs belong to a single subfamily (i.e. most HPKs in *Methanobacterium thermoautotrophicum* belong to HPK_{11}) but it does

not seem to be the case in *Erythrobacter litoralis* HTCC2594 after close examination.

4.4.3 Factors driven conformational changes

Conformational changes are a common mechanism for protein–protein or protein–ligand interactions. There are two important factors that drive conformational changes inside EL346: nucleotide and light. ATP not only is required for the kinase activity but also necessary for ordering and stabilizing histidine protein kinases, primarily in the CA domain where the ATP is bound. This is demonstrated by NMR and limited proteolysis studies for EL346, where ATP significantly improved the spectral quality and became more resistance to proteolysis. This importance has also been demonstrated previously by a NMR structure of EnvZ catalytic domain, in which the structure was solved in the present of the ATP analogue, AMPPNP (*129*). Overall, increased stability will facilitate future histidine kinase structural studies. Light is another factor that causes conformational changes, however, in the opposite direction than ATP. Limited proteolysis has demonstrated that EL346 is susceptible to proteolysis in the light state, suggesting a less rigid or stable structure in the light than in the dark states. NMR data supports the idea that EL346 is less order or stable and more dynamic, as evidently by decreasing in peak intensity due to linewidth broadening. However, light induce structural changes are reversible, suggesting

the changes is non–destructive. Ultimately, light causes conformational changes that activate the proteins. In our case, the changes may lead to the release of kinase inhibition or it may lead to a higher affinity binding for the ATP, which achieved a higher level of autophosphorylation.

4.4.4 Bifunctional EL346 histidine kinase and its light dependent two–component signaling cascade

Light induced structural changes in EL346 is linked to the light induced functional changes, as we have observed with EL346 where light affects the kinase autophosphorylation by changing K_m for ATP and this effect is propagated to the phosphotransfer of response regulators. The decrease in the phosphorylation signal observed in phosphotransfer of EL346 to RR1 led us to suspect that EL346 possesses autophosphatase activity, since HPKs are multifunctional enzymes that can have autophosphorylation, phosphotransferase and phosphatase activities (reviewed in (*32*)). However, it is unlikely that EL346 would selectively dephosphorylate only one of its response regulators, RR1, and not RR5 (Fig. 5c and 5d). On the other hand, some response regulators (i.e. CheY and Spo0F) have intrinsic autophosphatase activity (*73, 74*). Indeed, sequence alignment has revealed that RR1 contains an asparagine residue (N53, equivalent to N56 in CheY), analogous to one that has been identified to play a critical role in determining autophosphatase activity in other response regulators (*74*).

Substitution of non–Asn residue to Asn (Lys56 to Asn) has resulted in further increases in autophosphatase activity in Spo0F (*74*). However, substitution of Asn53 to Lys in RR1 abolished the autophosphatase activity in another HPK also found in *Erythrobacter litoralis* (unpublished data), suggesting that the diminishing phosphorylation signal observed in phosphotransfer reaction to RR1 is most likely due to autophosphatase activity by RR1, and EL346 does not contribute in autophosphatase activity and is a bifunctional histidine kinase.

4.4.5 Trans– or Cis–autophosphorylation: dimer vs. monomer

Biochemical assay confirms that EL346 is a bifunctional histidine kinase. However, it does not tell us exclusively how autophosphorylation is achieved. Biophysical characterization using gel filtration chromatography and MALS showed that EL346 is a monomer both in the dark (ground) and the light (signaling) states, indicating EL346 undergoes *cis*–autophosphorylation. This hypothesis is supported by our functional assays showing EL346 did not undergo monomer–dimer transition upon light stimulation, as seen in the case of the O_2–sensitive FixL histidine kinase from *Rhizobium meliloti* (*Rm*FixL)(*122*). This result contrasts with the vast majority of other histidine kinases studied to date, which have been shown to dimerize in solution and undergo trans–phosphorylation (*74, 141*). An exception came from EnvZ mutant, which was engineered to contain an additional DHp domain N–terminal to the DHp–CA

domain to form an DHp–DHp–CA mutant and showed that this protein was monomeric and underwent *cis*–phosphorylation (*142*). However, this mutant kinase only phosphorylates the histidine in the first DHp domain but not the second DHp domain that is adjacent to the CA domain. Notably, most phosphorylation studies are done with the histidine kinase domains (CA or DHp–CA domains) alone without their sensory domains. Some full–length histidine kinases are monomeric in the ground state and dimeric in the signaling state (i.e. *Rm*FixL (*122*)). Others, particularly members of LOV–containing histidine kinase from *Caulobacter crescentus* (LovK) (*1*) and EL368 from *Erythrobacter litoralis* (unpublished data), are dimeric in both dark and light states.

To examine why EL346 does not dimerize, we took a closer look at the sequence in the DHp domain. This domain has been shown to dimerize by forming coiled–coil interaction between the two helices from each histidine kinase monomer. Using the LearnCoil program (*143*), we found that there is no predicted coiled–coil–like motif anywhere in EL346, which may be a possible reason why we do not observe dimerization in this protein. To see if this is just an isolated case in EL346, we used this program to analyze all members of HPK_{11} full–length proteins that are listed from Grebe *et. al. (33)*. Of the 20 proteins that we analyzed, we found that only 40% of the HPK_{11} analyzed has a clear predicted coiled–coil–like motif in their DHp domain, another 40% (including EL346) does not have any coiled–coil–like motif predicted in the DHp domain, and the other

20% shows a predicted coiled–coil–like motif but with very low probability and very short coiled–coil sequence (results not shown). These numbers could vary since it is likely that more HPK_{11} members will be identified in the future (EL346 was not included in the HPK_{11} listed in Grebe *et. al.*, and possibly other HPK members (HPK_{1-10}) that we did not examine could also have monomeric HPKs. Nevertheless, this result suggests that there are potentially more monomeric histidine kinases, and that *cis*–autophosphorylation is another activation mechanism for histidine kinases like EL346.

4.4.6 Molecular Models

Our biochemical data showed that the autophosphorylation signal is 25% weaker when the LOV domain is intact, indicating that the LOV domain acts as a weak inhibitor for kinase autophosphorylation. Conformational changes upon photoactivation in EL346 subsequently could lead to the relief of kinase inhibition. On the other hand, we have also shown that photoactivation of EL346 result in difference affinity for ATP, which potentially coupled with the light–driven conformational changes. Together, we propose that the LOV domain plays a duel role in regulating EL346 kinase activity. First, it acts a photoregulator by inhibiting the autokinase activity in the dark state. Second, light–induced conformational changes that originated from the LOV domain subsequently lead to enhanced activities by increasing affinity for ATP in the light state at limiting

ATP conditions. Future structural information of the full–length EL346 in the dark or lit states should elucidate the molecular details for the inhibition and the cis–autophosphorylation mechanism, and future *in vivo* studies in the organism of *Erythrobacter litoralis* should reveal the physiological function of EL346.

4.5 Acknowledgements

We thank Dr. Tong–Seung Tseng (Carnegie Institution of Washington, Stanford) for providing us with initial guideline. We are grateful to Dr. Carlos Amezcua (Baxter Healthcare) for countless NMR technical support and maintenance, and Drs. Gaya Amarasinghe and Michael Rosen for assistance and sharing of the equipment/space. We thank Dr. Xiaolan Yao and people from the Gardner lab for discussions. Many thanks go to Jason Tuckerman and Dr. Marie–Alda Gilles–Gonzalez for the encouragement and guidance throughout the kinase assay optimizations. This work was supported by grants from the US National Institutes of Health (R01–GM81875) and the Robert A. Welch Foundation (I–1424). The *E. litoralis* genome sequence data was provided by Stephen Giovannoni's laboratory (Oregon State University) and the J. Craig Venter Institute with grant support from the Gordon and Betty Moore Foundation Microbial Genome Sequencing Project.

Table 2. Dark state recovery time constants and rates for EL346 and EL134.

Sample[a]	τ [A_{446}] (min)[b]
EL346	54.6 ± 1.2
EL346 + AMPPNP	32.2 ± 0.4
EL346 + ATP	41.5 ± 1.0
EL134	70.8 ± 1.1

a: Dark state return of all samples were monitored at 446 nm at room temperature via UV–vis spectrophotometer in the buffer containing 50 mM Tris (pH 8.2), 100 mM NaCl, 5 mM $MnCl_2$ (only in the nucleotide included samples), and 10% glycerol. Protein concentration: ~50 μM EL346, 5 mM nucleotide (either AMPPNP or ATP), and ~25 μM EL134.

b: Rates (τ) were determined by fitting measurements (Fig. 2c) to a single exponential decay function via Prism (GraphPad Software, San Diego, CA). All data represented here are the average of three separate measurements.

Figure 4.1 Overviews of EL346 and RR constructs used in this work. a. Domain architecture of full length EL346, truncated EL346 constructs, and two EL346 response regulators. EL346 is the full–length protein, which is composed of one LOV domain (in yellow) followed by the histidine kinase domain that contains two subdomains named the DHp domain (in black) and the CA domain (in red). RR1 contains only the receiver domain (in orange) while RR5 contains a C-terminal receiver domain (in orange) and a N-terminal output domain called σ70 (in blue). b. Protein expression and purity accessed by homogenous 20% SDS–PAGE gel. All proteins shown here are free of fusion tags. LMW indicates low molecular weight marker.

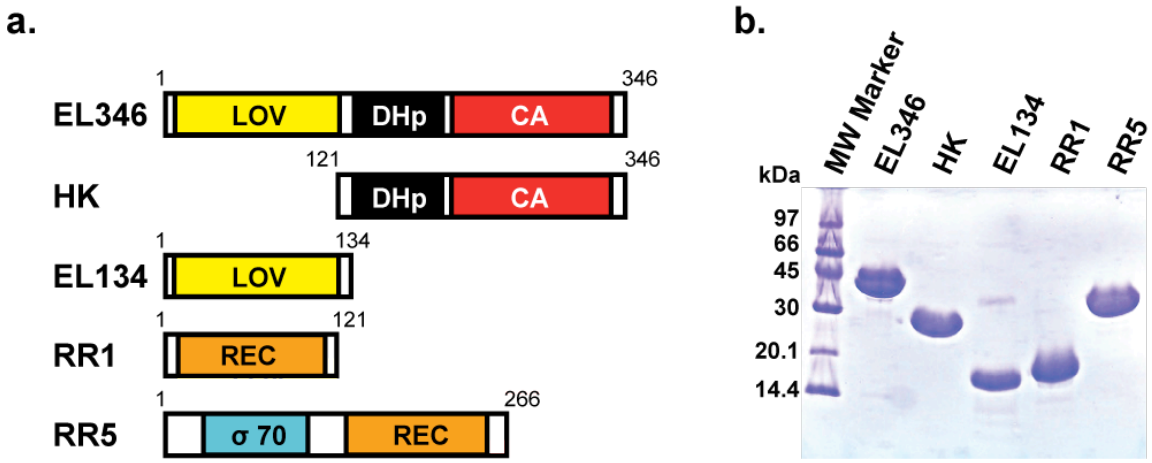


Figure 4.2 Sequence alignment of the HK domain for EL346 and other members of the HPK_{11} class of histidine kinase. Sequence alignment of HK domain alone between EL346 and 16 other known members of HPK_{11} histidine protein kinases is shown in the next page (*33*). Red shaded characters indicate identical residues, while red characters indicate conserved (>70%) residues. Beginning and ending points of each sequence within its own native sequences are indicated after each protein names. MTH902_mth_261–462: residue 261 to 462 from *M. thermoautotrophicum* MTH902; MTH1124_mth_176–348: residue 176 to 348 from *M. thermoautotrophicum* MTH1124*;* MTH901_mth_144–352: residue 144 to 352 from *M. thermoautotrophicum* MTH901*;* MTH446_mth_377–583: residue 377 to 583 from *M. thermoautotrophicum* MTH446*;* MTH174_mth_583–785: residue 583 to 785 from *M. thermoautotrophicum* MTH174*;* MTH468_mth_356–554: residue 356 to 554 from *M. thermoautotrophicum* MTH468*;* MTH360_mth_492–700: residue 492 to 700 from *M. thermoautotrophicum* MTH360*;* MTH459_mth_293–495: residue 293 to 495 from *M. thermoautotrophicum* MTH459*;* MTH292_mth_362–564: residue 362 to 564 from *M. thermoautotrophicum* MTH292*;* MTH356_mth_368–567: residue 368 to 567 from *M. thermoautotrophicum* MTH356*;* MTH619_mth_545–749: residue 545 to 749 from *M. thermoautotrophicum* MTH619*;* MTH985_mth_156–365: residue 156 to 365 from *M. thermoautotrophicum* MTH985*;* MTH823_mth_500–677: residue 500 to 677 from *M. thermoautotrophicum* MTH823*;* MTH123_mth_151–356: residue 151 to 356 from *M. thermoautotrophicum* MTH123*;* G2072665_mtu_289–501; residue 289 to 501 from *M. tuberculosis* G2072665; EL346_eli_128–346: residue 128 to 346 from *E. lightoralis* EL346; ExsG_sme_706–907: residue 706 to 907 from *S. meliloti* ExsG. Predicted secondary structures shown on top of the sequence are based on Jpred and structural/sequence comparison of HK853 crystal structure (RCSB accession: 2C2A) and the order of α helices and β strands are listed the same as in HK853 (*51*). Black helices represent the DHp domain while red helices and β strands represent the CA domain. The ATP-binding Bergerat fold that is shared among GHKL superfamily is started from α4 and ended at βG (*50*). This analysis suggested that residue H142 is the conserved phosphoaccepting histidine in EL346, which we subsequently validated (see Fig. 4.11a).

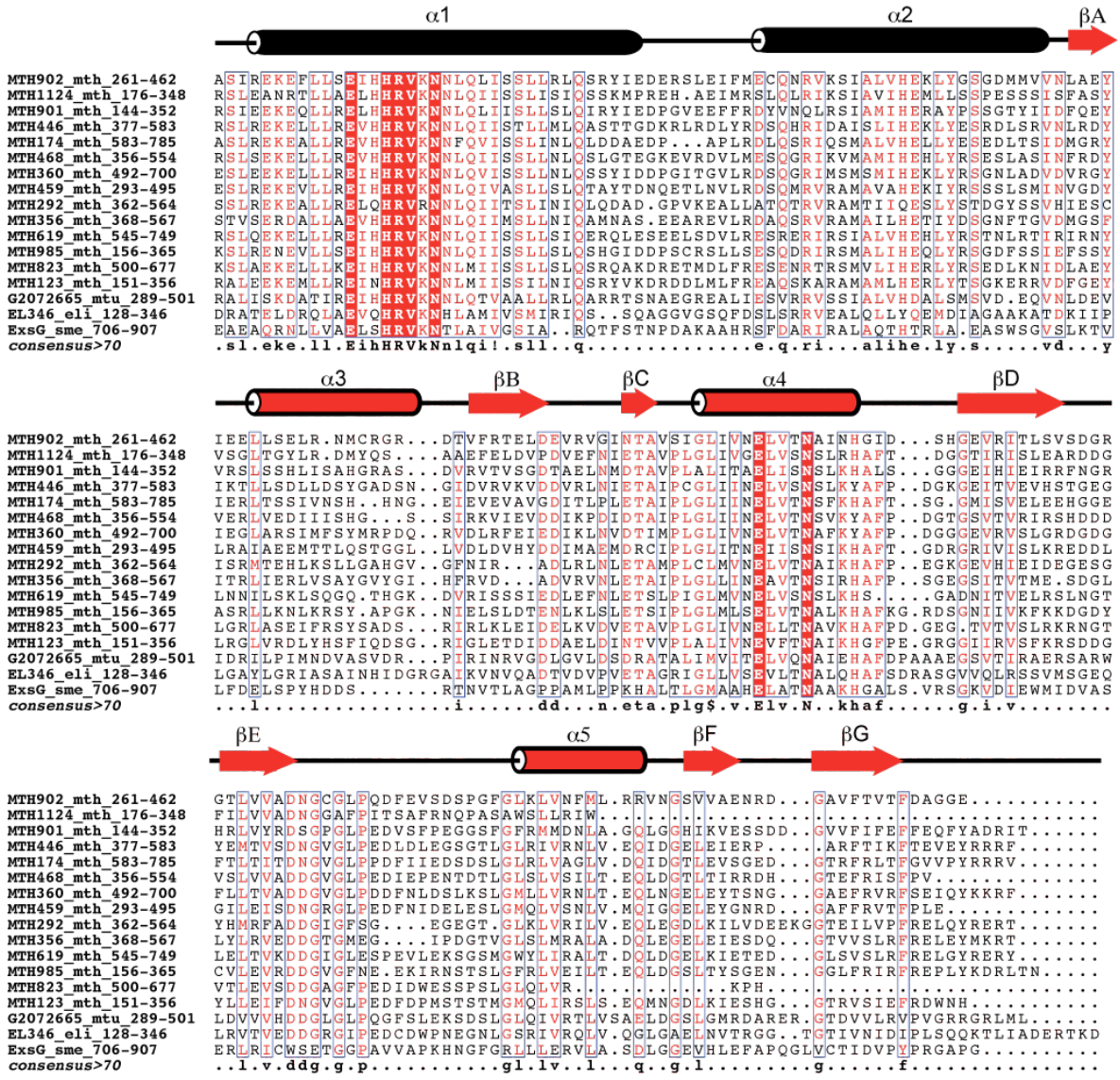

α1
α2
βA
α3
βB
βC
α4
βD
βE
α5
βF
βG
MTH902_mth_261-462
MTH1124_mth_176-348
MTH901_mth_144-352
MTH446_mth_377-583
MTH174_mth_583-785
MTH468_mth_356-554
MTH360_mth_492-700
MTH459_mth_293-495
MTH292_mth_362-564
MTH356_mth_368-567
MTH619_mth_545-749
MTH985_mth_156-365
MTH823_mth_500-677
MTH123_mth_151-356
G2072665_mtu_289-501
EL346_eli_128-346
ExsG_sme_706-907
consensus>70

Figure 4.3 Photocycle and dark state recovery kinetics of EL346. **a.** UV-visible absorbance spectra for EL346 in the dark (black) and light states (red) are shown here between 290–520 nm. Three isosbestic points at 336, 393 and 410 nm are observed on the overlaid dark and the light state spectra. The recovery of the light state to the dark state was monitored via scans with 15 min interval and shown here by traces between the dark (black) and the light (red) state traces. **b.** Demonstration of the interaction between FMN and the LOV domain in the dark (top panel) and in the light (bottom panel) states using *As*LOV2 X-ray structures (dark state PDB ID: 2V1A and light state PDB ID: 2V1B) as our model (*87*). In the LOV domain, FMN is non-covalently associated with the LOV core in the dark state. In its photoactivated state, FMN forms a covalent linkage with a conserved Cys residue (C450 in the *As*LOV2) from the LOV domain and this covalent adduct is reversible in most LOV domains. **c.** Comparisons of dark state recovery kinetics of EL346 alone (black), EL346 with AMPPNP (blue), EL346 with ATP (purple), and EL134 LOV domain alone (red). Each dark state recovery kinetic trace was monitored at 446 nm at room temperature for a period of time that was at least five times longer than τ for each case. The plots shown here are means of three averaged data points, and these data were fitted into single exponential decay function (shown in red) using Prism (GraphPad Software, San Diego, CA) and the resulting rates are listed in Table 2.

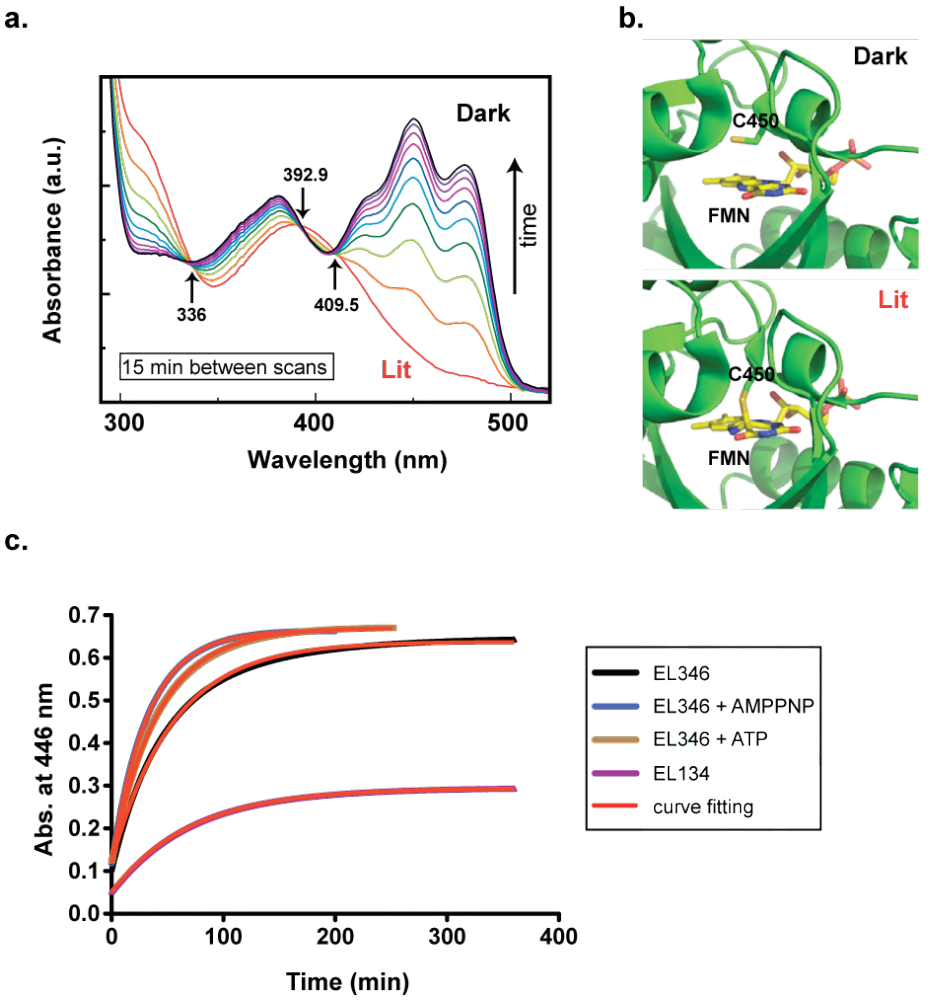


Figure 4.4 TCEP and Imidazole affect the dark state recovery rate of EL346. **a**. The kinetic of the dark state recovery at different TCEP concentrations is plotted against TCEP concentrations from 0 to 200 mM. Each dark state recovery kinetic is monitored at 446 nm at room temperature. Data were fitted using single site specific binding via Prism (GraphPad Software, San Diego, CA). **b**. The kinetic of the dark state recovery at different imidazole concentrations is plotted against imidazole concentrations from 0 to 400 mM. Again, each dark state recovery kinetic is monitored at 446 nm at room temperature. Data were fitted using single site total binding via Prism (GraphPad Software, San Diego, CA).

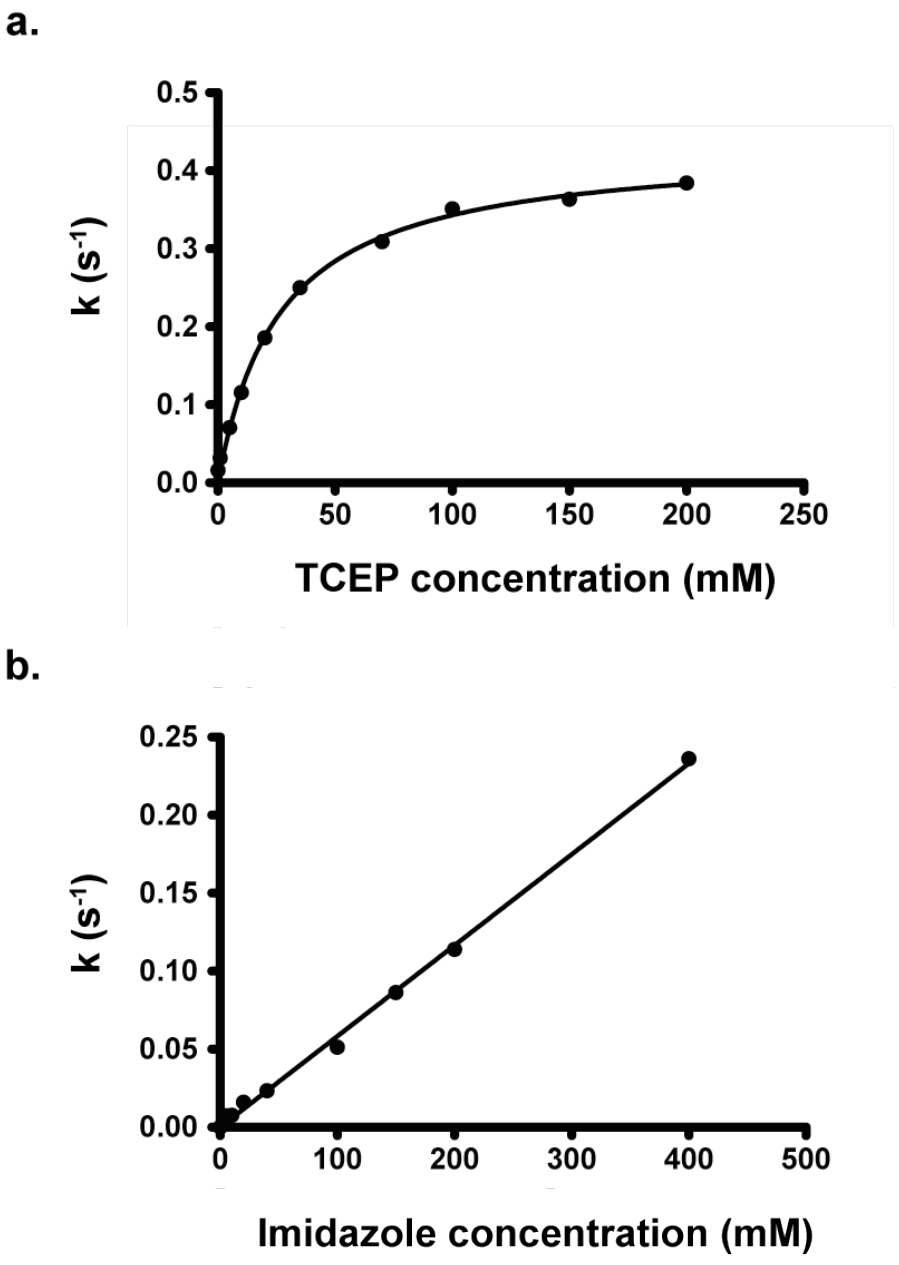


Figure 4.5 EL346 oligomeric states accessed by gel filtration and MALS. a. Gel filtration elution profile of EL346 in the dark and light states with and without ATP. The A_{280} signal intensity was plotted against the elution volume. EL346 eluted around 10.5 mL (Ve) in the dark and light states either with or without ATP. Calculation using this Ve yields a molecular weight that represents monomeric EL346. **b.** For the MALS results, the differential refractive index (dRI) on the left y–axis and the measured molecular weight (MW) on the right y–axis, both plotted against the elution volume. A small peak eluted around 9.5 mL is a dimeric fraction that is caused by disulfide linkages and can be break down to monomer by reducing agents. The majority of protein is eluted around 10.8 mL and is determined to be monomeric. In both **a** and **b** figures, black represents dark state EL346 in the dark state without ATP, red represents light state EL346 without ATP, grey represents dark state EL346 with 5 mM ATP, and orange represents light state EL346 with 5 mM ATP. The slightly higher elution volume in the MALS data was due to the inclusion of additional tubing in the MALLS system. Both methods show that EL346 is monomeric in both the dark and light states, either with or without the presence of ATP.

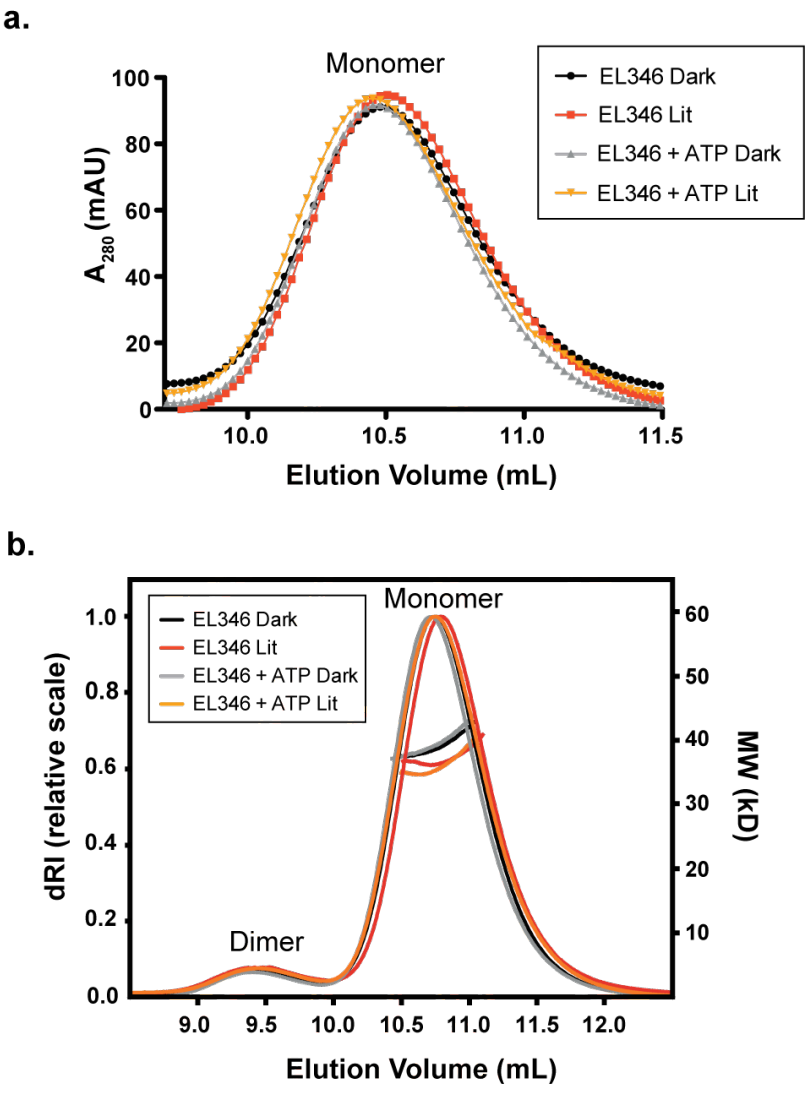


Figure 4.6 Comparisons of $^{15}N/^{1}H$ TROSY spectra of EL346 with and without ATP in the dark and light states. Overlay of dark (black) and light (red) state $^{15}N/^{1}H$ TROSY spectra of EL346 a. without ATP, and b. with 5 mM ATP.

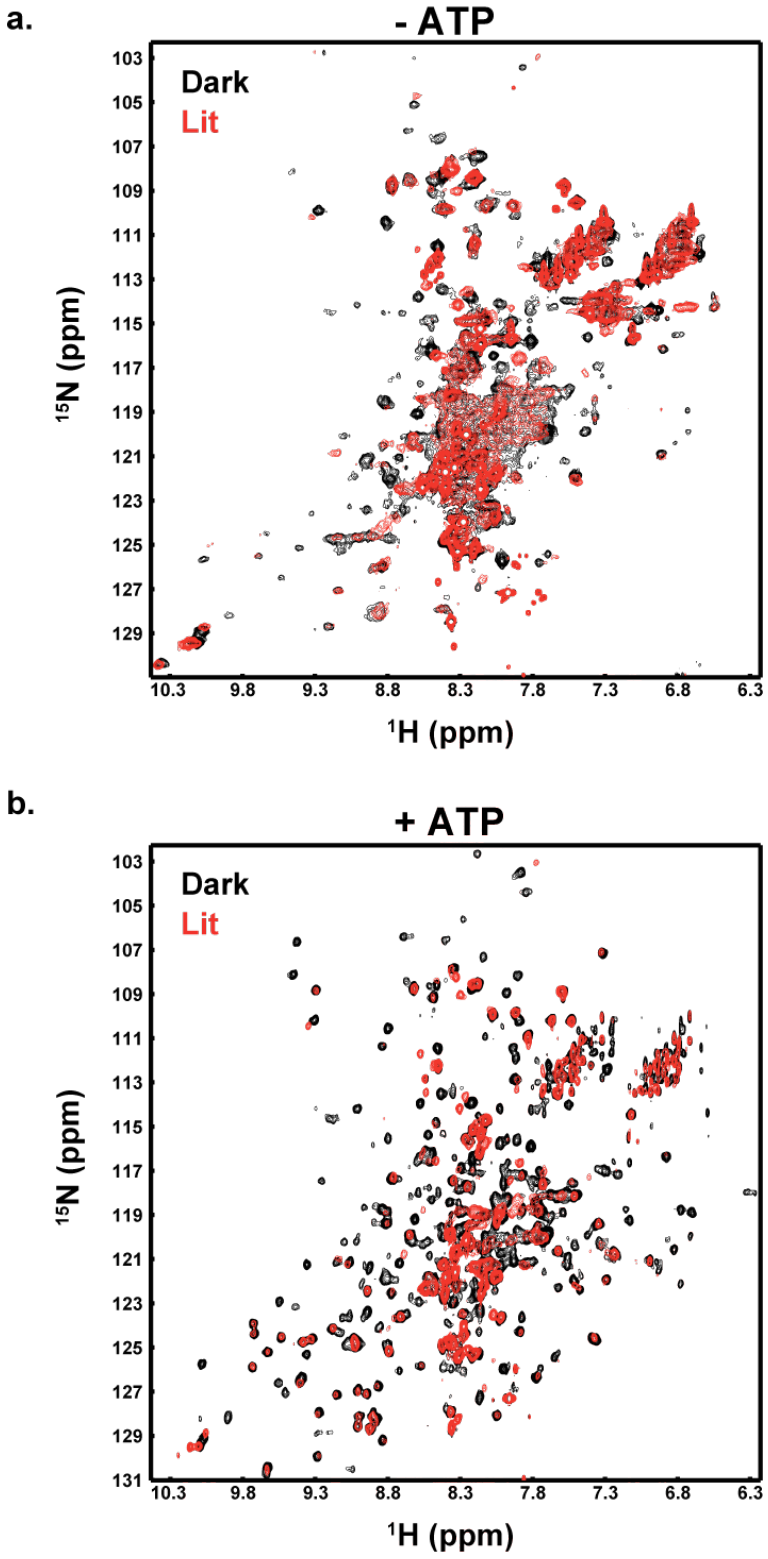


Figure 4.7 Comparisons of CT $^{13}C/^{1}H$ HSQC dark and light states spectra of the methyl region for EL346. Dark and light states spectral overlays at the methyl region for EL346 are shown here **a.** without AMPPNP, and **b**. with AMPPNP. Red and black peaks indicate the positive peaks while the purple and blue peaks indicate the negative peaks. Enlarged regions (show as a large grey box on the right side of each spectrum) indicate the methyl region where the methionine residues (there are a total of 8 Met in EL346) localized, and they are shown as negative peaks in blue (dark state) and magenta (light state). Brown dashed areas and three arrows in the large grey box in **b.** emphasize the nucleotide–enhanced changes between the dark and light states.

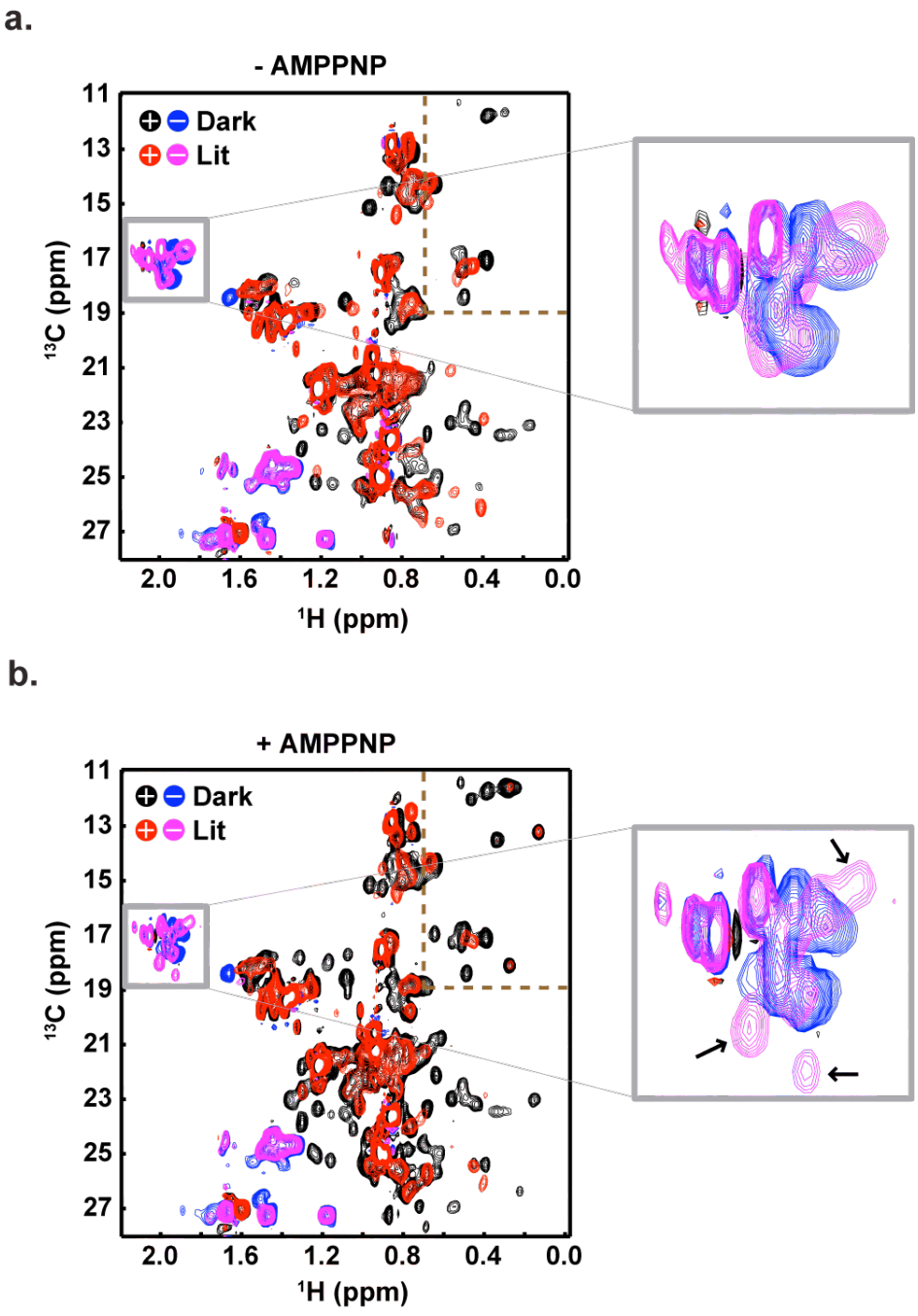


Figure 4.8 Comparisons of trypsin digestion of EL346 with and without AMPPNP in the dark and light states. Various limited proteolysis experiments performed by trypsin on EL346 (1:100 w/w ratio, respectively) are shown **a.** in the dark state without AMPPNP, **b.** in the light state without AMPPNP, **c.** in the dark state with 5 mM AMPPNP, and **d.** in the light state with 5 mM AMPPNP. Major proteolytic fragments are labeled with white and black circles, and white, grey, and black triangles. The stars shown on the right side of the gels came from the dimerized final proteolytic fragment that formed due to insufficient reducing agent. The final proteolytic fragment (white triangle) was not further cleaved and remained stable for more than 6 h in all conditions, and ESI-MS analysis on this final proteolytic fragment (cleaved overnight) for both the dark and light state in EL346 sample supplemented with ATP indicating that it is the LOV domain from residue G4 to R134 (data not shown).

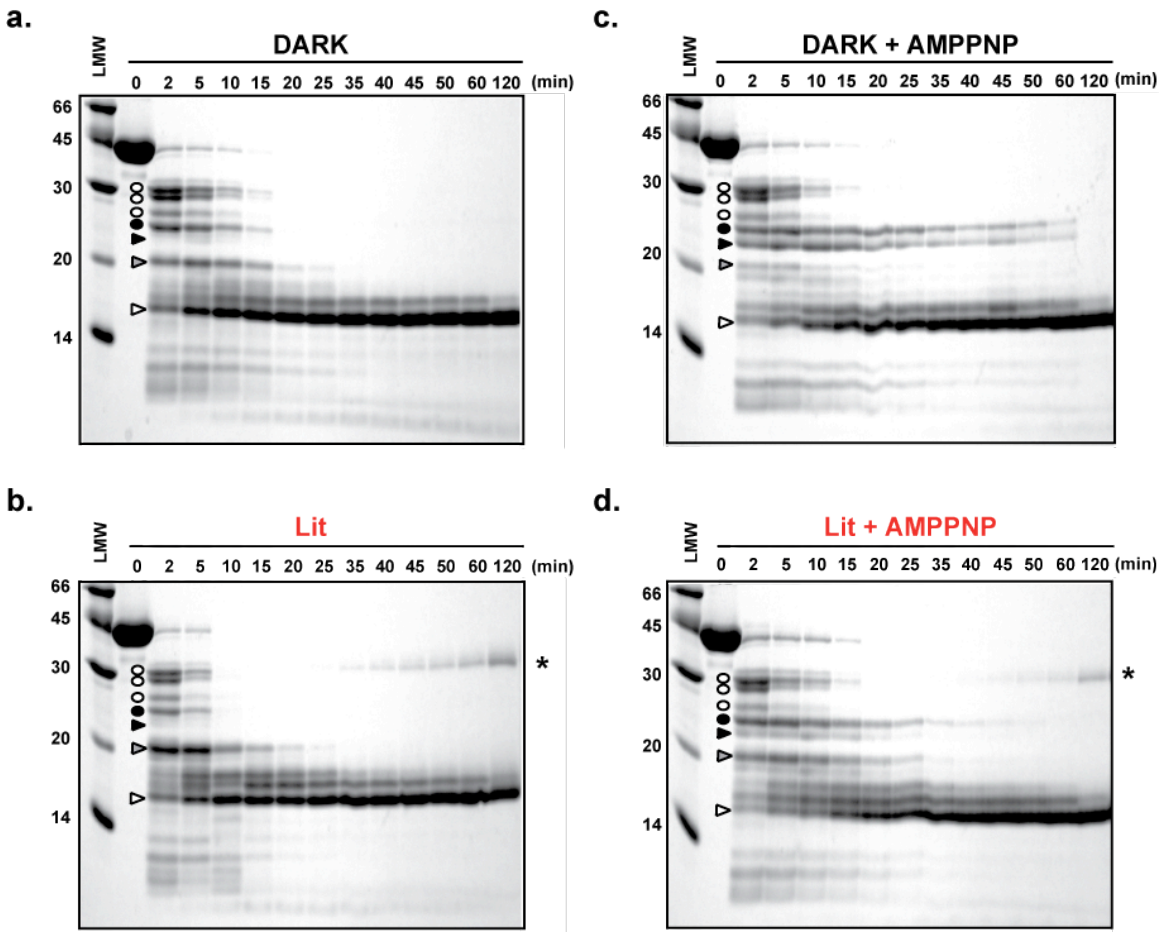


Figure 4.9 Light-dependent autophosphorylation and phosphotransfer of EL346. Dark (black circle) vs. light (open triangle) autophosphorylation of EL346 using **a.** 500 μM cold ATP and **b.** 1.8 μM cold ATP. Dark (black circle) and light (open triangle) phosphotransfer of EL346 to **c.** RR1 and **d.** RR5 under 1.8 μM cold ATP condition. In each figures, ^{32}P phosphorylation gels are shown on the top and the plot of the phosphorylation signal versus time are shown on the bottom.

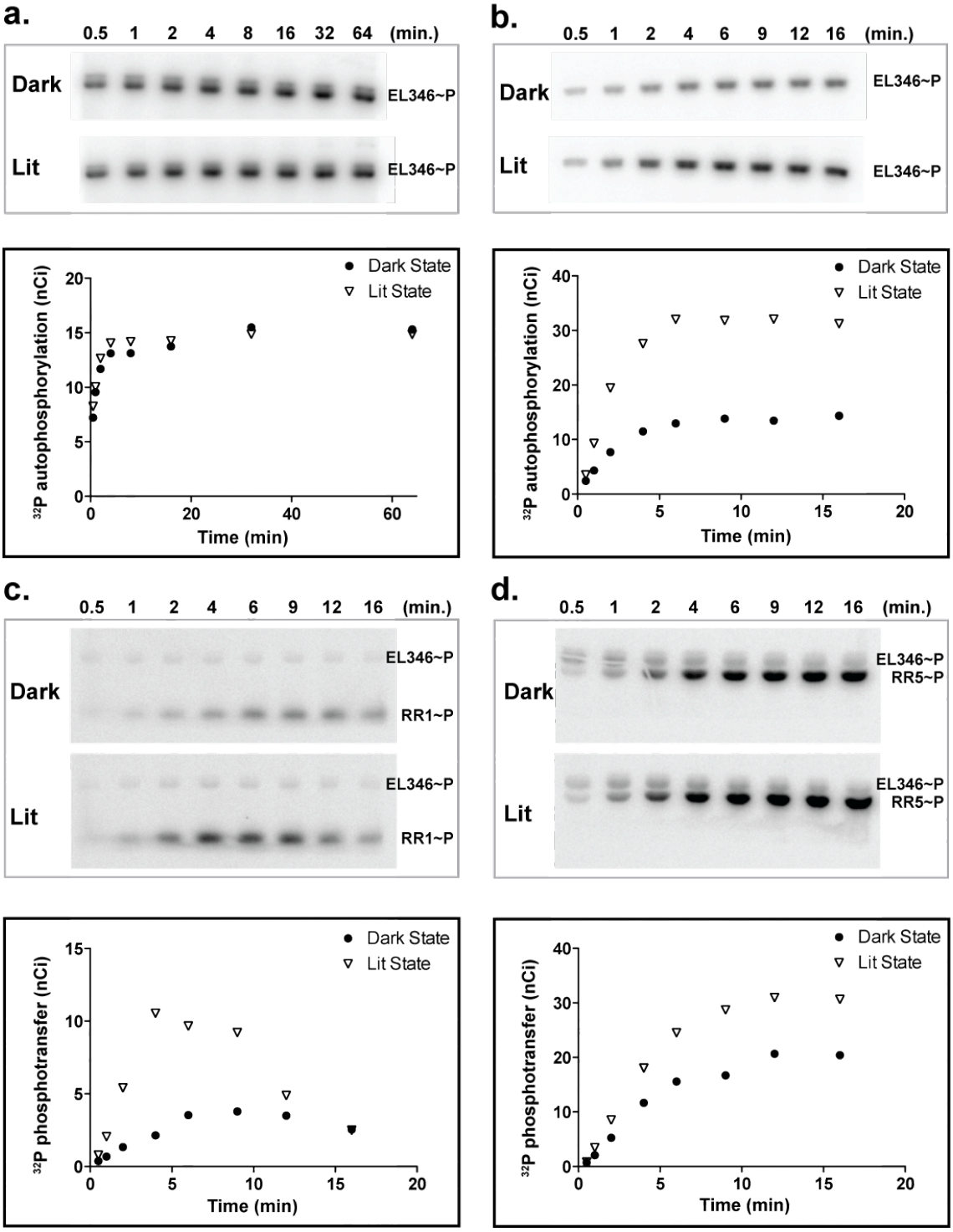


Figure 4.10 Oligomeric state of EL346. Autophosphorylation of EL346 is shown here with various concentrations from 0.5 to 50 μM. The resulting signal intensity is plotted against EL346 concentrations. The linear dependency of signal intensity (see also insert) to the concentration suggesting the autophosphorylation is independent of EL346 concentration and thus dimerization is not required for EL346 kinase activity.

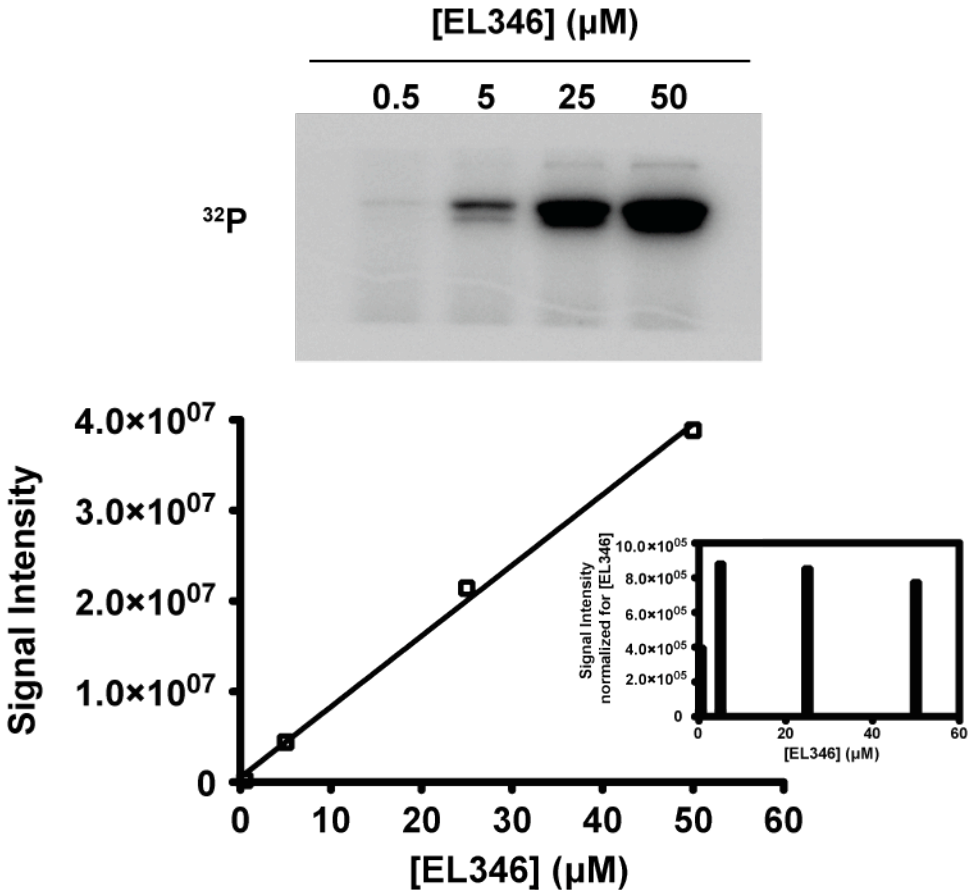


Figure 4.11 The HK domain alone is capable of undergoing autophosphorylation and phosphotransfer. a. ^{32}P phosphorylation and Coomassie Brilliant Blue stained gels are both shown for autophosphorylation of WT, H142Q, and HK, and **b.** phosphotransfers of HK to RR1 and RR5. LMW indicates the low molecular weight marker.

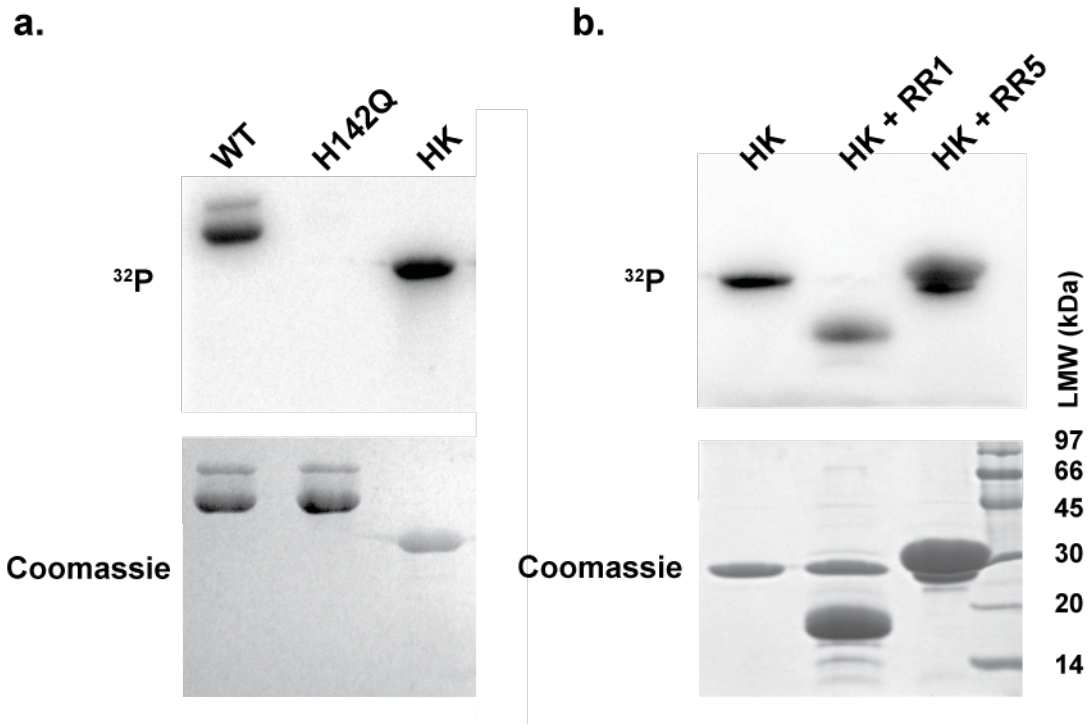
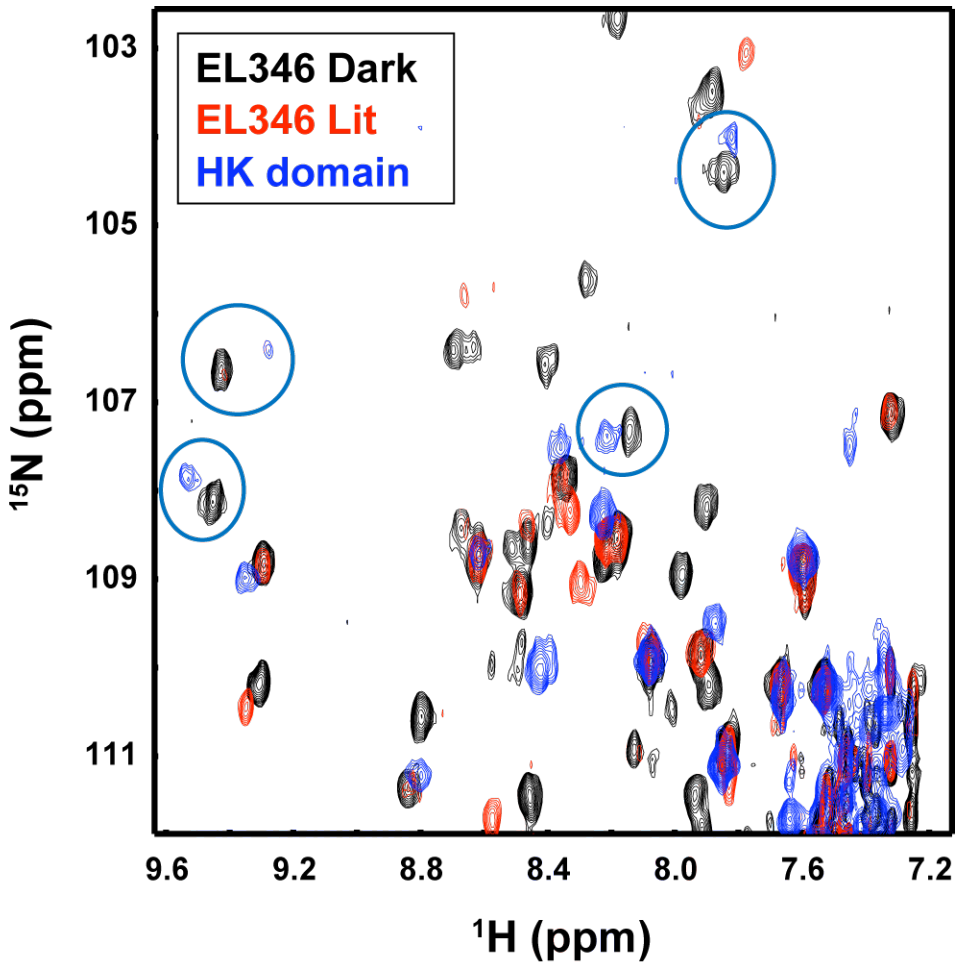


Figure 4.12 Overlay HSQC spectra of HK domain alone with full–length EL346 dark and light states.



Chapter 5 Summary and perspectives of studies on EL346

It has been over two decades since TCSs were first described; however, the regulation mechanism of these proteins remains largely undefined. The molecular mechanism of how the activities of the HKs are regulated by their own sensory domain remains unclear, and structures of the full–length HKs are still missing. Central questions concerning LOV domain signaling mechanisms in general are also remained unclear. Solving the structure of the LOV domain coupled to its downstream effector domain will shed light into the molecular insight of how LOV domains regulate their effector domains. Light regulated bacterial TCSs protein, EL346, serves as an ideal model to study how the LOV and HK domains interact, and how autokinase activity is regulated by light. Determining the structures of the full–length EL346 in various signaling states (i.e. dark or lit, ATP bound or ADP bound, and phosphorylated or non–phosphorylated) would shed new light on how these systems are regulated.

I have initiated the work in characterizing the dark state EL346 structure using both NMR and X–ray crystallography. Structure determination using NMR was only partially successful due to poor protein stability during the lengthy acquisition times required for 3D datasets. As a result, I was only able to assign a portion of the peaks in thee spectra due to poor data quality. On the other hand, I obtained crystals of EL346, and had one crystal diffracted to less than 3 Å (Fig. A.14, 15). However, this condition (D10, Fig. A.14) could not be

reproduced by hand and most of the other EL346 crystals only diffracted to 8Å. The use of an additive screen may improve crystal quality and yield better diffraction pattern but are yet to be tested. Additionally, EL346 mutants such as dark state mimic C55A, lit state mimic C55M, or kinase inactive mutants H142A or H142Q may be used for crystallization and to solve the structure.

Although the work done here has provided extensive *in vitro* biophysical and biochemical characterization of EL346, we do not yet know what is the physiological role of EL346 in *E. litoralis*. Much of this awaits results of ongoing studies in the lab of the response regulators actuvated by EL346. Some hits to this including RR1 and RR5, which I discussed in Chapter 4. A recent study from Francez–Charlot A *et al*. provided us clues on one of the EL346 cognate response regulator, RR5, suggesting it could be involved in the general stress response in the bacteria (*144*). This study showed that although these PhyR–type of response regulators (RR5 in our case) contain a sigma–70–like DNA binding domain, they do not actually bind DNA (*144*). Instead, upon phosphorylation, this PhyR type protein disrupts the interaction between the sigma factor (σ) and its anti–sigma factor (anti σ) by binding to the anti σ, allowing σ to associate with RNA polymerase to transcribe stress genes (*144*). Testing this hypothesis will identify the EL346/RR5 signaling pathway and the consensus DNA binding sequence can be confirm by performing DNA gel shift assays. Lastly, one could also knock out the gene of EL346 or its response regulators, RR1 and RR5, to see how they

affect the bacteria growth or behavior, once we master the growth and handling of *E. litoralis*.

Appendix I

A. Additional characterization of EL346

A1. Domain architecture of the full–length EL346 and generation of various constructs and mutations in EL346.

Secondary structural predictions are based on Jpred (*137*) and sequence/structural comparison of known HK domain from HK853 (*51*) and *As*LOV2 domain (Fig A.1, and A.2) (*25*). EL346 exhibits characteristic homologous boxes belong to the member of HPK_{11} (*33*), and the conserved boxes in the kinase domain were mapped as indicated in green (H box), grey (X box), cyan (N box), purple (D box), and brown (G boxes) (Fig. A.2). Unlike other HK subfamilies, HPK_{11} does not contain an F box.

Various constructs were made to acquire stable isolated LOV, DHp, and CA domains as well as the LOV–DHp and DHp–CA domain constructs (Fig. A.3). These proteins were purified in the general buffer of 50 mM Tris or sodium phosphate with 100 mM NaCl in pH7.0 to 8.0. At the end, both the LOV domain and the kinase domain are stable in pH 7.5 50mM Tris, 100mM NaCl, and/or 5 mM $MgCl_2$ (for kinase only) buffer, while the full–length protein is usually purified in pH8.0 buffer instead. The results of the construct expression, their solution properties (oligomeric state), and kinetics were summarized in Fig A.3. Constructs for the LOV domain (residue 1 to 134) and the DHp–CA domain

(residue 121 to 346) were consequently used in the studies to identify the role of the LOV domain as described (see Chapter 4.3.6).

Additionally, I generated several mutations in the LOV domain (C55A, C55S, C55M) of the full–length EL346 construct to study the role of the LOV domain, and potentially to be used for crystallography. The lit state mimic C55S mutation did not generate a stable soluble full–length protein (data not shown); however, another lit state mimic (C55M) and the dark state mimic (C55A) can be expressed. The UV–visible absorbance profile of C55M shows that it can be excited to the lit state (Fig. A.4). Once it is photoexcited to the lit state, the protein is irreversibly converted to a red absorbing species by the new absorption maxima between 600 to 700 nm (Fig. A.4a). A similar phenomenon has been reported in Phot1–LOV1 domain from *C. reinhardtii* C57M mutation (*108*). Compared to the wild type protein, the $^{15}N/^{1}H$ HSQC TROSY spectrum of lit state C55M closely resembles the wild type lit state spectrum (Fig. A.5). However, some cross peaks in the WT lit state spectra were missing in the C55M spectra, suggesting that the overall conformation between the C55M mutant and WT lit state may be similar, but specific residues may experience different electronic environments. On the other hand, the dark state mimic mutant, C55A, has an absorbance profile identical to the wild type as expected, and it does not convert to the lit state specie upon illumination (Fig. A.4b). Again, the $^{15}N/^{1}H$ HSQC TROSY spectrum of C55A in the dark state (green) agreed fairly well to the wild

type dark state (black) spectrum (Fig. A.6). While minor differences were between the dark (green) and lit (magenta) state C55A spectra, their overall correspondence was quite good (Fig A.6).

I also generated several mutations in the DHp domain (H142A and H142Q) of the full–length EL346 to determine if H142 is really the phosphate–receiving residue. $^{15}N/^{1}H$ HSQC TROSY spectra of both mutants were also collected to verify the protein folding, and spectral quality were compared to the wild type protein. Overall they are similar to the WT protein (Fig A.7 and A.8). Autophosphorylation of these mutants were also tested and compared to the wild type protein. Both of these predicted kinase dead mutation, where the conserved histidine residue was mutated to Ala or Glu, lost their phosphorylation activity (data not shown), providing reasonable evidence that H142 is the phosphate–receiving histidine in the EL346 protein histidine kinase as discussed in Chapter 4. However, the C55A mutant has a similar level of autophosphorylation as the wild type protein (Fig A.9). Notable, these assays were performed at 500 μM of cold ATP, at which level the phosphorylation between the dark and lit state is likely similar. The fact that the dark state is a bit more active than the lit state might be due to errors during handling.

A.2 Optimization for EL346 kinase assay

The original kinase activity for EL346 reported in Swartz *et. al*. waw very low (*76*) and is not ideal for further characterization for both the autokinase and phosphotransfer activities that is proposed to be mediated by light. I, therefore, went through a series of optimizations to obtain higher activity from EL346. pH optimization results showed that the enzyme activity is dramatically improved by increasing the pH from 7.5 (as reported in (*76*)) to 8.0 or higher (Fig. A.10). Additionally, EL346 has a bit slight preference for Mn^{2+} as its divalent source than Mg^{2+} (Fig. A.11). EL346 also loses its activity dramatically when stored in the –80 ºC when compared to the –20 ºC, complicating long term storage of this protein under these conditions (Fig. A.12). In fact, EL346 loses most of its activity two days after being purified. Therefore, fresh proteins were prepared for all kinase assays by purifying EL346 on the first day and performing kinase assays on the following day.

A.3 EL346 phosphotransfer profiling

Phosphotransfer profiling (*36*) of EL346 were performed in pH 8.0 buffer containing 50 mM Tris, 100 mM NaCl, 2 mM TCEP, 0.2 mM $MgCl_2$, and 5mM $MnCl_2$. A total of 21 out of the total of 23 known RRs in the *Erythrobacter litoralis* HTCC 2594 were tested against EL346 for 30 second and 10 min time points. The result from the 30–second phosphotransfer experiment indicated that RR1 and RR5 are both the cognate response regulators for EL346, with RR1 to be

a bit more potent than RR5 (Fig. A.13). Additionally, RR7 can also be phosphorylated by EL346 (as seen in the 10 min result) but with a much lower phosphorylation signal, suggesting it may not be the cognate RR for EL346 (Fig. A.13).

A.4 Initial results for EL346 crystallization

EL346 was purified in one day following the same purification protocol as described in Chap 4.2.2 with the exception that EL346 was stored in the dark at 4 ºC without any glycerol. The trays were set up the following day under ambient light condition using Phoenix robot and kept at 20 ºC in the dark afterward. Protein samples (10 mg/ml, ~260 μM) were buffer exchanged and prepared into two different buffer conditions: i) pH 8.0 10 mM Tris, 100 mM NaCl, 5 mM $MgCl_2$, 2 mM TCEP, and 5 mM AMPPNP, and ii) pH 8.0 10 mM Tris, 100 mM NaCl, 5 mM $MgCl_2$, 2 mM TCEP, 5 mM AMPPNP, and 50 mM L–Arg/L–Glu. Out of 960 conditions, I got three hits from the Hampton index screen: D10, F6 and F7. D10 contains 20% PEG MME 5000 and 0.1 M BIS–Tris buffer in pH 6.5, F6 contains 20% PEG 3350, 0.2 M ammonium sulfate (AMS), and 0.1 M BIS–Tris buffer in pH 5.5, and F7 contains the same condition as F6 but at pH 6.5 instead. Single and/or multiple crystals were observed. Diana and I have collected the diffraction data for the D10 crystal here at UTSW and it diffracted to an estimation of 3 Å or less (Fig. A.14). However, this condition was not

reproducible by hand and the original crystal that diffract to ~3 Å was used in the seeding procedure.

Optimizations of these three crystallization conditions were carried by hand. Variations of buffer conditions were done by altering the pH (from 5.0 to 7.0), the percentage of PEG3350 (from 16 to 30%), the percentage of PEGMME 5000 (from 16 to 24 %), inclusion and/or deletion of 0.2 M AMS, and testing different incubation temperature (4, 16, and 20 ºC). Trays are set both under ambient light or dim red light in room temperature. Surprisingly, no crystals were formed from the trays that were set under dim red light. Additionally, I was only able to reproduce crystal in condition similar to F6 and F7 but not D10, which originally gave the best diffraction data (3Å from D10 crystal compared to 5 to 8Å from the F6/F7 conditions). The diffraction data collected for D10 crystal is shown in Fig A.15. The conditions that can reproduce crystals by hand were listed as the following with the two most reproducible conditions labeled in bold:

1) 25% PEG3350, 0.2 M AMS, and 0.1 M BIS–Tris at pH 5.5, 6.0, and 6.5.
2) 26%, **28%**, and 30% **PEG3350, 0.2 M AMS, and 0.1 M BIS–Tris at pH 6.5.**
3) 26% and 28% PEG3350, 0.2 M AMS, 0.1M BIS–Tris, and 1% glycerol at pH 6.5.

4) 26% PEG3350, 0.2 M AMS, 0.1 M Bis–Tris, 3% glycerol or 3% ethylene glycol at pH6.5.
5) **24% PEG3350, 0.2 M AMS, 0.1 M Bis–Tris at pH 5.5.**

General problems in these reproducible conditions are that although these crystals are initially observed as small single crystal after a week and a half of incubation in the 20 ºC, it tends to grow as multiple crystals instead of forming bigger singular crystals as the incubation time increase. Additional optimization (i.e. additive screen or changing constructs) will be needed to acquire better EL346 crystals.

Figure A.1 Sequence alignment of various LOV domains: *As*LOV2 domain from *Avena sativa* phototropin1, the LOV domains from EL346, EL368, and EL222 from marine bacteria *Erythrobacter litoralis*, and BMLOV domain from human/animal pathogen *Brucella melitensis (2)*. The red letters represent the invariant residues while the blue letters represent the highly conserved residues. The secondary structure (in green and light green) shown here are based on the *As*LOV2 structure (PDB: 2V1A) (*87*).

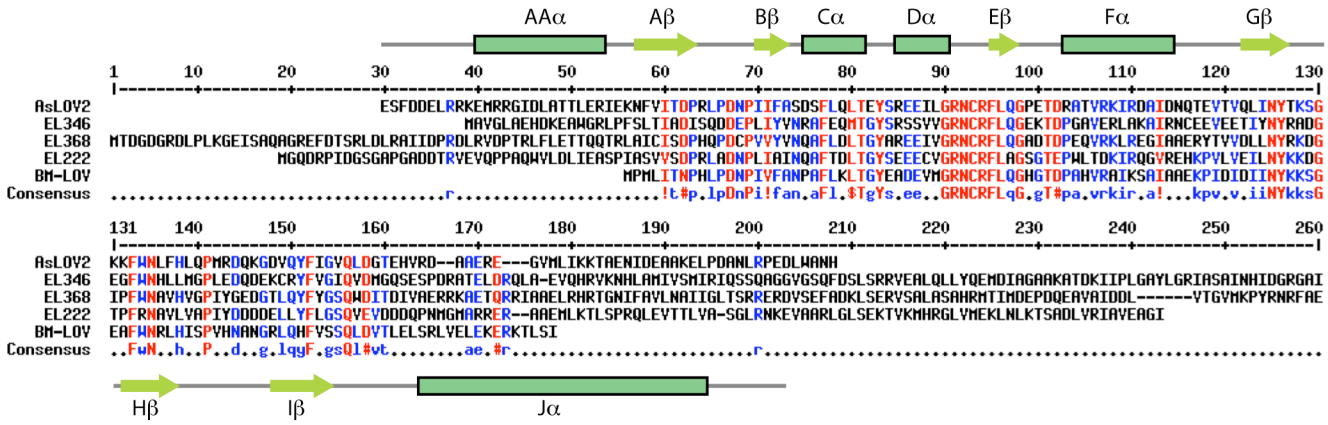


Figure A.2 Domain architecture and secondary structural prediction of EL346. Structural predictions are based on Jpred and sequence/structural comparison of known HK domain from HK853 (*51*). Based on Grebe's classification (*33*), EL346 belongs to member of HPK11 and the conserved boxes in the kinase domain were mapped as indicated in green (H box), grey (X box), cyan (N box), purple (D box), and brown (G boxes).

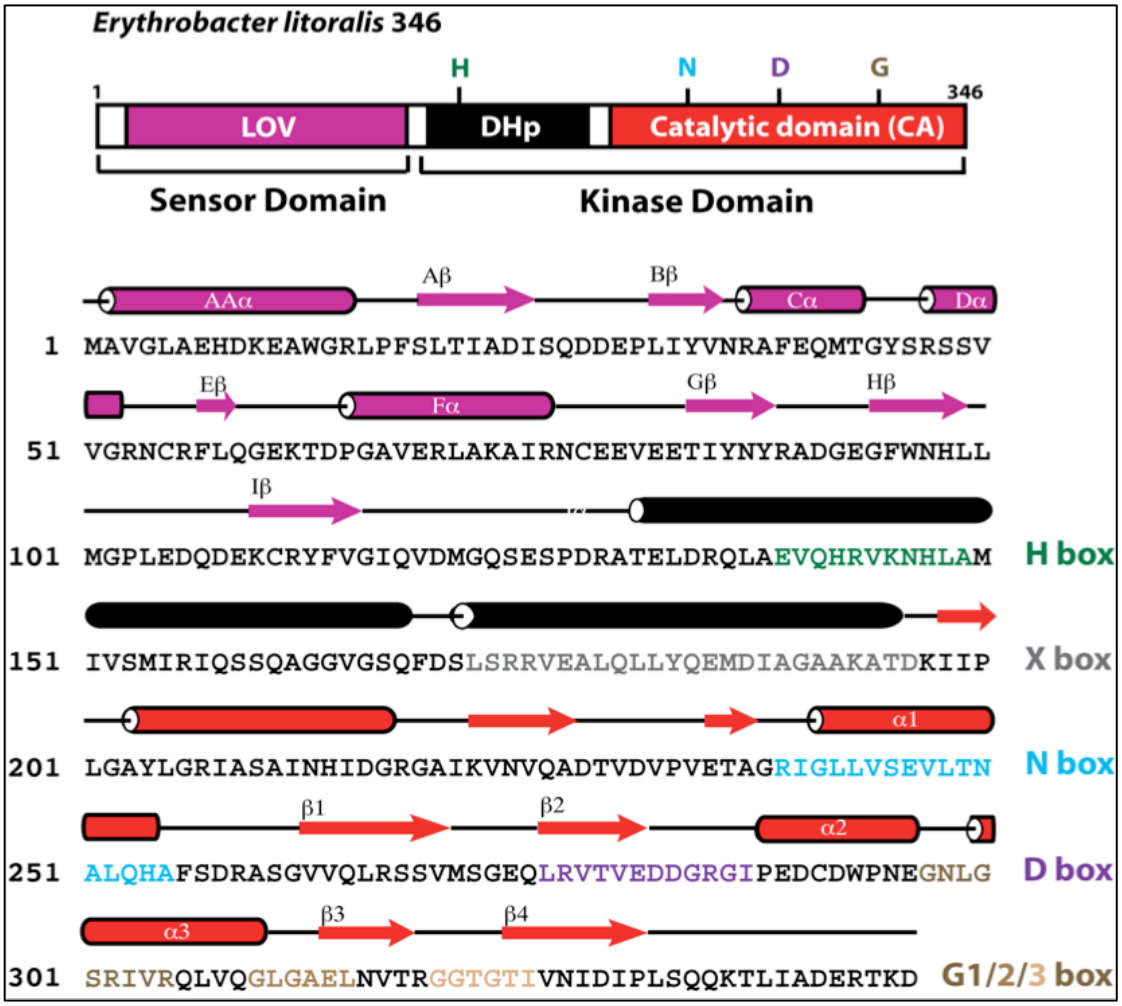
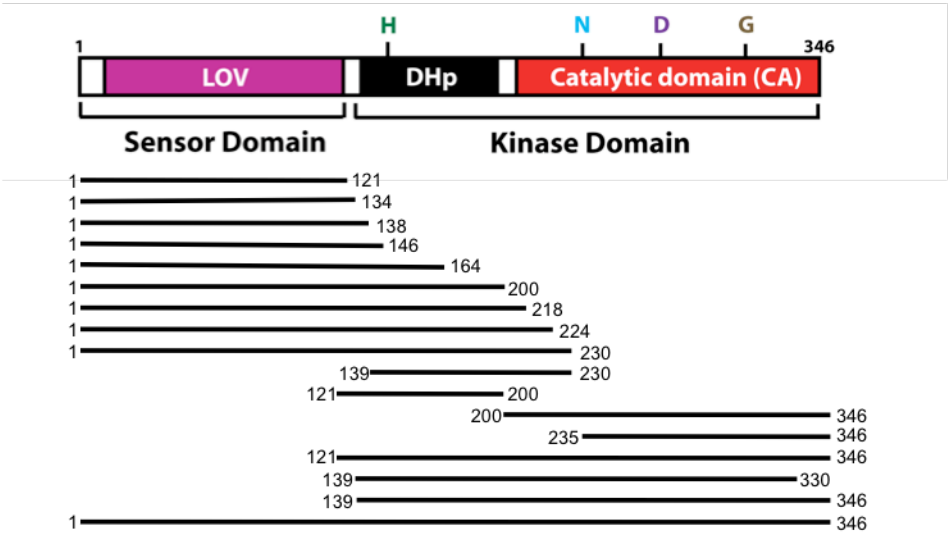
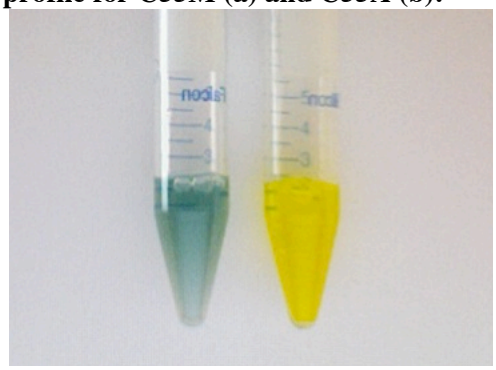


Figure A.3 Various lengths of constructs tested for EL346 isolated domains generated to final soluble and stable constructs of the LOV, DHp, CA, LOV–DHp, and DHp–CA domains. Successful candidates were further tested for their solution properties in terms of the oligomeric state and dark state recovery kinetics (τ). Preferred purification conditions for these proteins see A1 in the Appendix I.

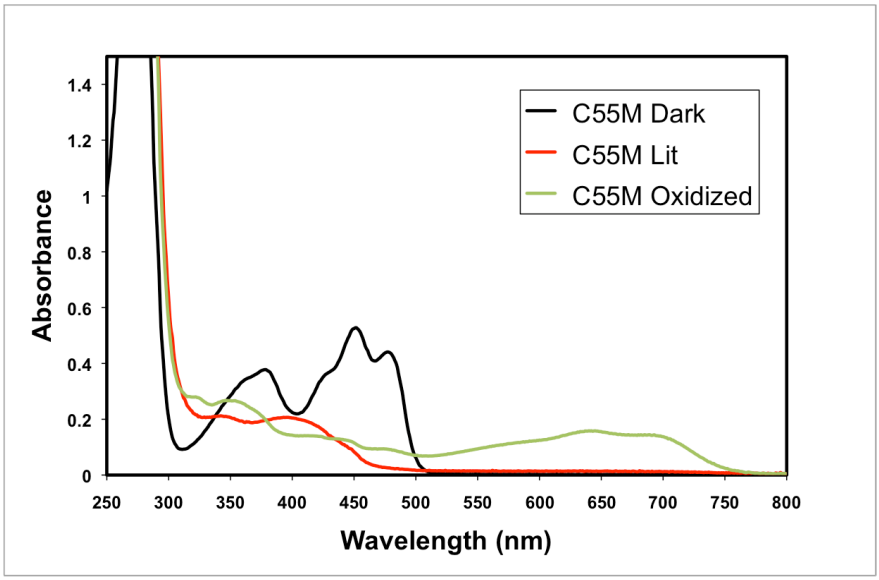


Constructs	Solubility	SEC75 profile	τ (min.)
1-121	Yes but not yellow	--	--
1-134	Yes	Dimeric	~87
1-138	Yes	Dimeric	--
1-146	Yes	Dimeric	31.9
1-164	Yes but unstable	Monomeric	71.4
1-200	Yes but not yellow	--	--
1-218	Barely & unstable	--	--
1-224	No	--	--
1-230	Barely & unstable	--	--
139-230	No	--	--
121-200	Yes but low yield	--	--
200-346	Some but unstable	--	--
235-346	Very low yield	Dimeric	--
121-346	Yes	Dimeric	--
139-330	No	--	--
139-346	Barely	Monomeric	--
1-346	Very soluble	Monomeric	56.8
1-346 with His_6 tag	Very soluble	Monomeric	66.7 (w/ tag on)

Figure A.4 A picture of EL346 lit state mimic C55M mutant (green, in its oxidized state) vs. dark state mimic C55A mutant (yellow, in its lit state) in solution, and the UV–visible absorption profile for C55M (a) and C55A (b).



a.



b.

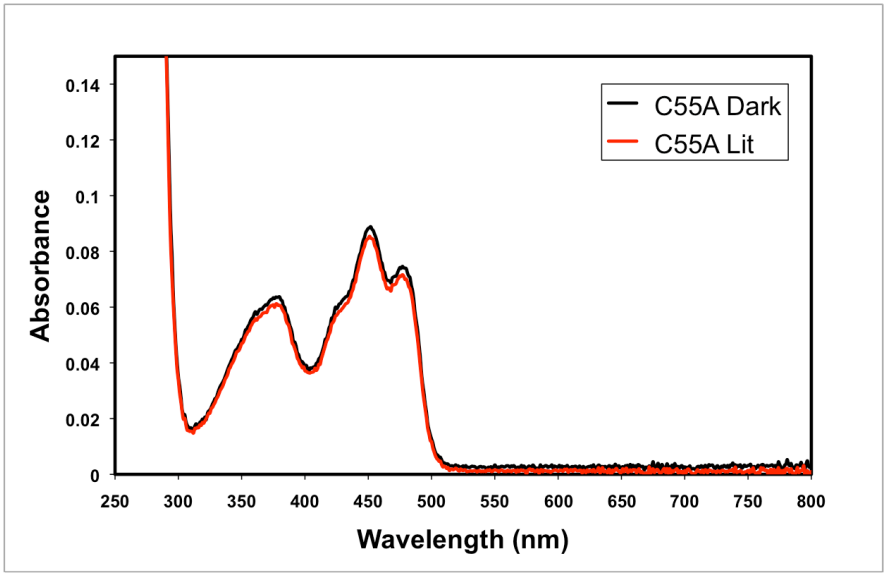


Figure A.5 Overlay $^{15}N/^{1}H$ TROSY HSQC spectra of the EL346 lit state C55M (green spectra) with the wild type EL346 in the dark (black spectra) and lit (red spectra) states. Data were collected at 800MHz using 250μM of WT EL346 and 265 μM of C55M in pH 7.5 buffer containing 50 mM Tris, 100 mM NaCl, 10 mM $MgCl_2$, 5 mM AMPPNP, 5m M DTT, 10% D_2O, and 0.05% NaN_3.

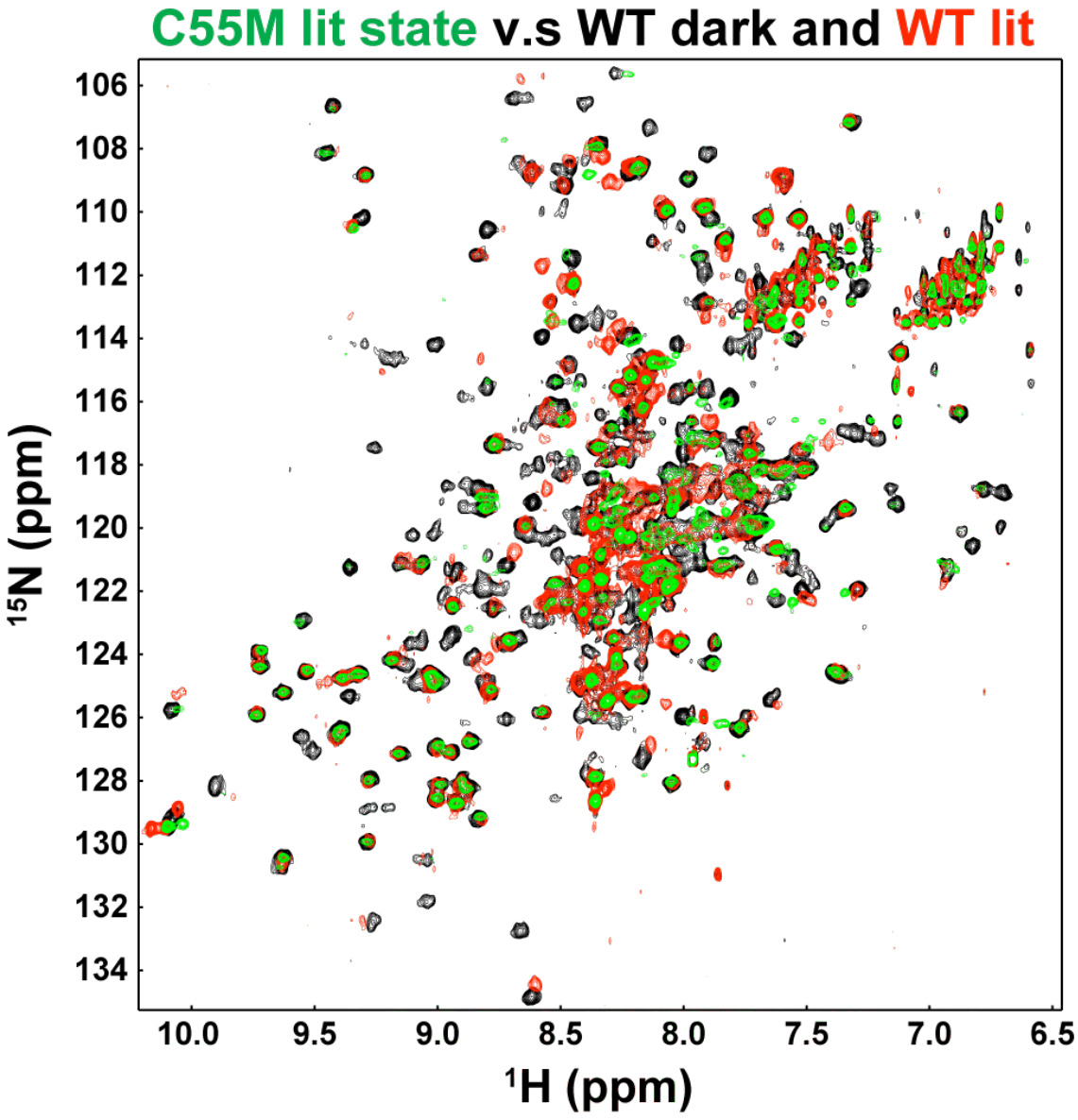


Figure A.6 Overlay $^{15}N/^{1}H$ TROSY HSQC spectra of the EL346 C55A in the dark state (red spectra) and in the lit state (blue spectra) with the wild type EL346 in the dark state (black spectra). The spectrum of the C55A dark state and lit state are almost identical, as shown here the blue (lit state) peaks overlay so well with the red (dark state) peaks that they cover the red peaks almost completely in some cases. There is minor variation of some peaks location between C55A and the WT, but overall they are very similar. Data were collected at 800MHz. 250μM of WT EL346 and 540 μM of C55A samples were prepared in pH 7.5 buffer containing 50 mM Tris, 100 mM NaCl, 10 mM $MgCl_2$, 5 mM AMPPNP, 5 mM DTT, 10% D_2O, and 0.05% NaN_3.

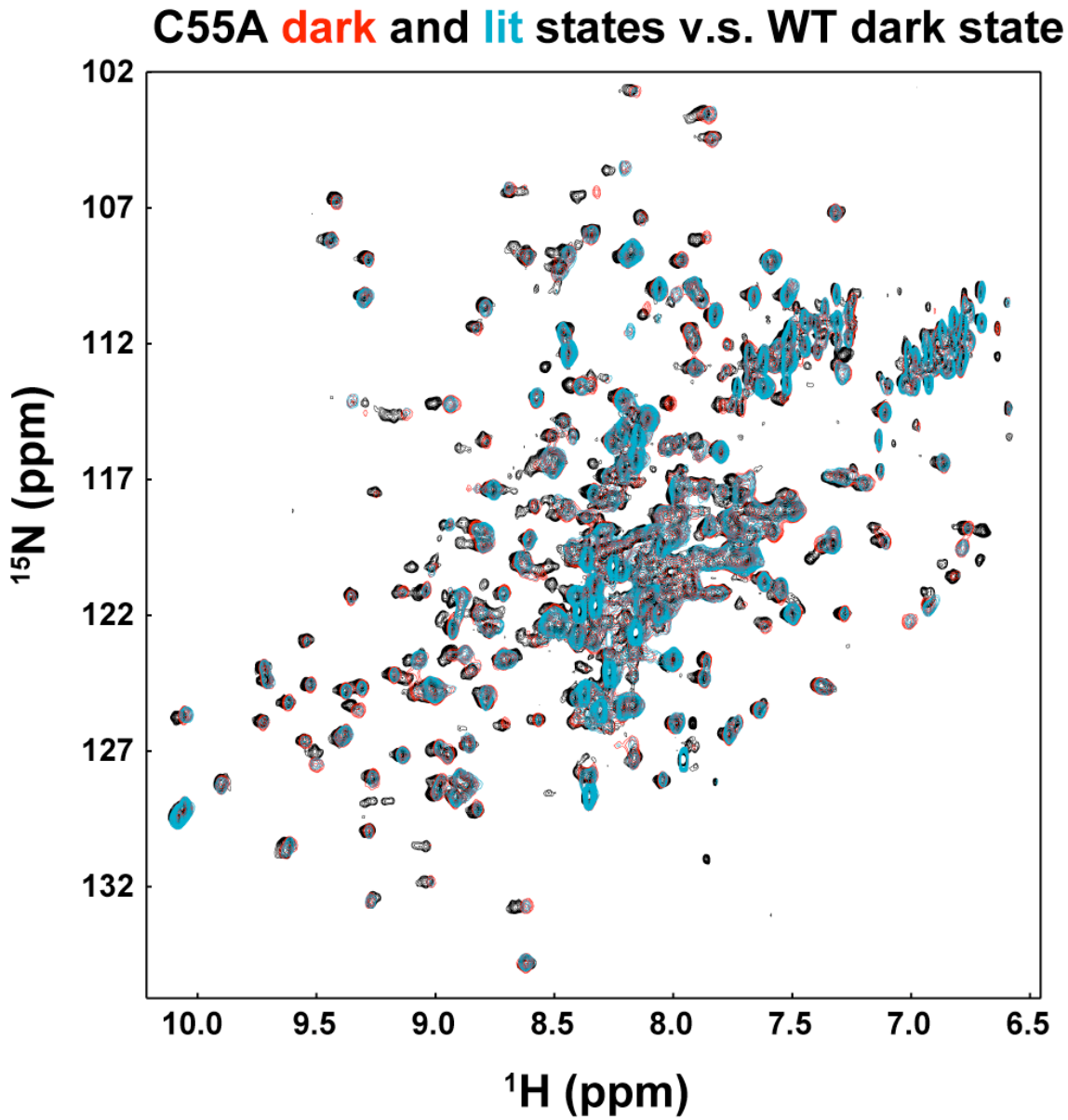


Figure A.7 Overlay $^{15}N/^{1}H$ TROSY HSQC spectra of the EL346 H142A in the dark and with the wild type EL346 in the dark state. Data were collected at 800MHz. 250μM of WT EL346 and 545 μM of H142A samples were prepared in pH 7.5 buffer containing 50 mM Tris, 100 mM NaCl, 10 mM $MgCl_2$, 5 mM AMPPNP, 5 mM DTT, 10% D_2O, and 0.05% NaN_3.

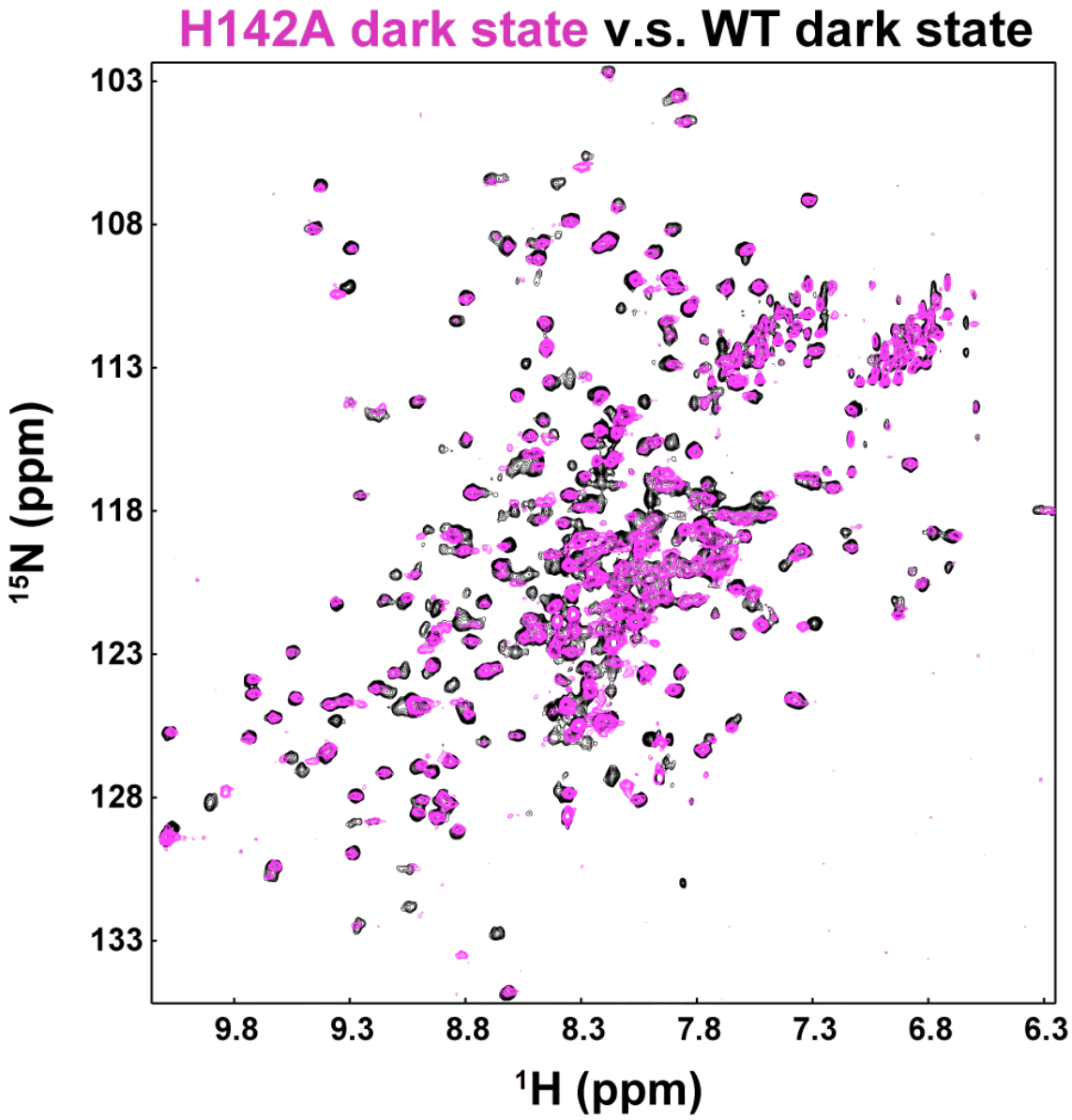


Figure A.8 Overlay $^{15}N/^{1}H$ TROSY HSQC spectra of the EL346 H142Q in the dark and with the wild type EL346 in the dark state. Data were collected at 800MHz. 250μM of WT EL346 and 356 μM of H142A samples were prepared in pH 7.5 buffer containing 50 mM Tris, 100 mM NaCl, 10 mM $MgCl_2$, 5 mM AMPPNP, 5 mM DTT, 10% D_2O, and 0.05% NaN_3. The area crowded with many inhomogeneous cross peaks (located at the lower region in between the brown dashed lines) is an indication of protein degradation. The SDS–PAGE gel of the post–NMR sample showed that H142Q is indeed partially depredated (data not shown). The protein degradation was later found to result from the dirty filter in the FPLC system and was not due to an inherent instability of H142Q.

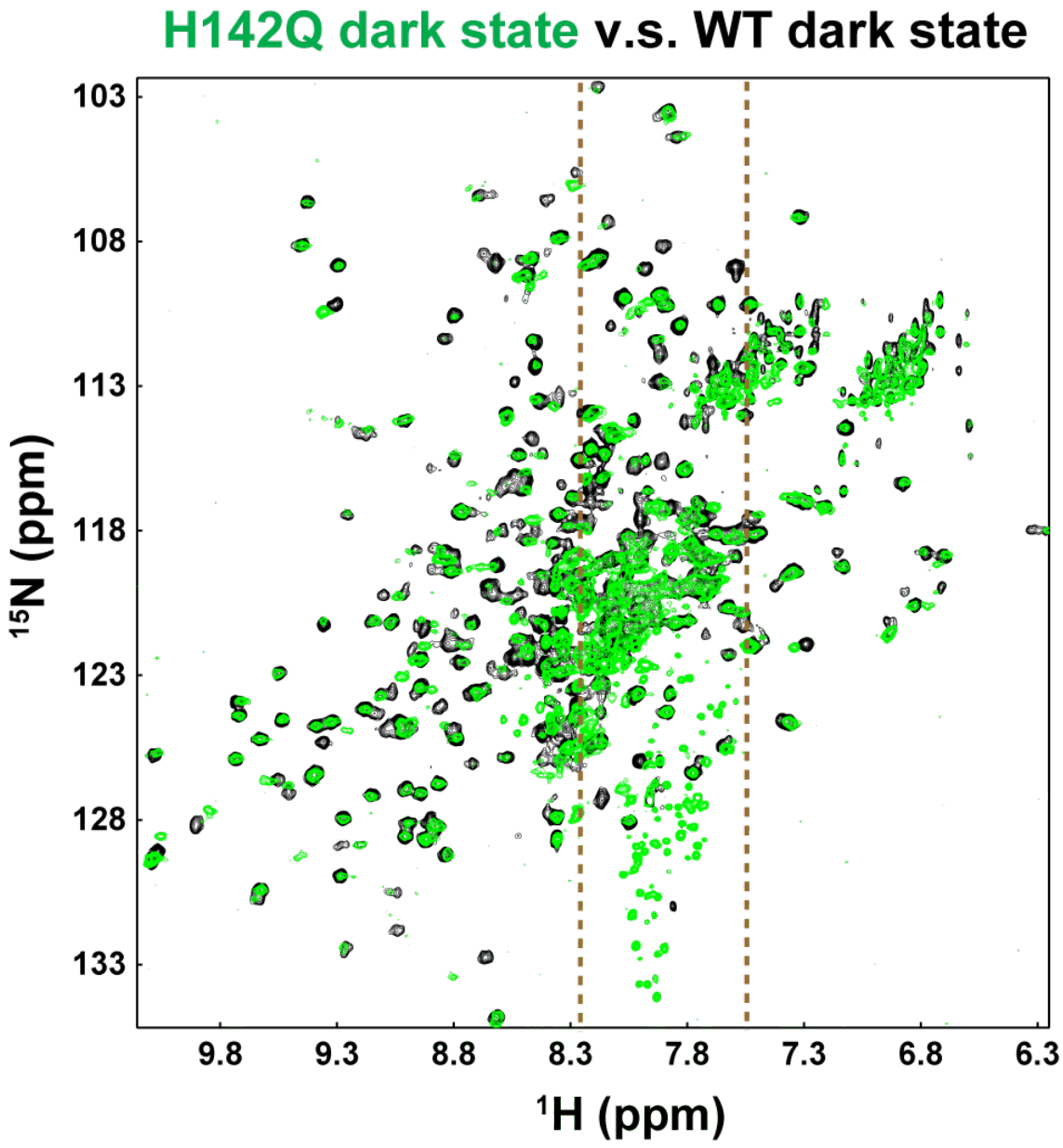


Figure A.9 Autophosphorylation comparison of the WT EL346 with the dark state mutant C55A. Reaction mixtures: 10 μM of either EL346 WT or C55A mutant in pH8.2 buffer containing 50mM Tris, 100mM NaCl, 5 mM DTT, and 5 mM $MnCl_2$. 5 μCi of hot ATP and 500 μM of cold ATP were used to initiate the phosphorylation and reactions were stopped at 16 min by addition of 4X SDS stopping buffer.

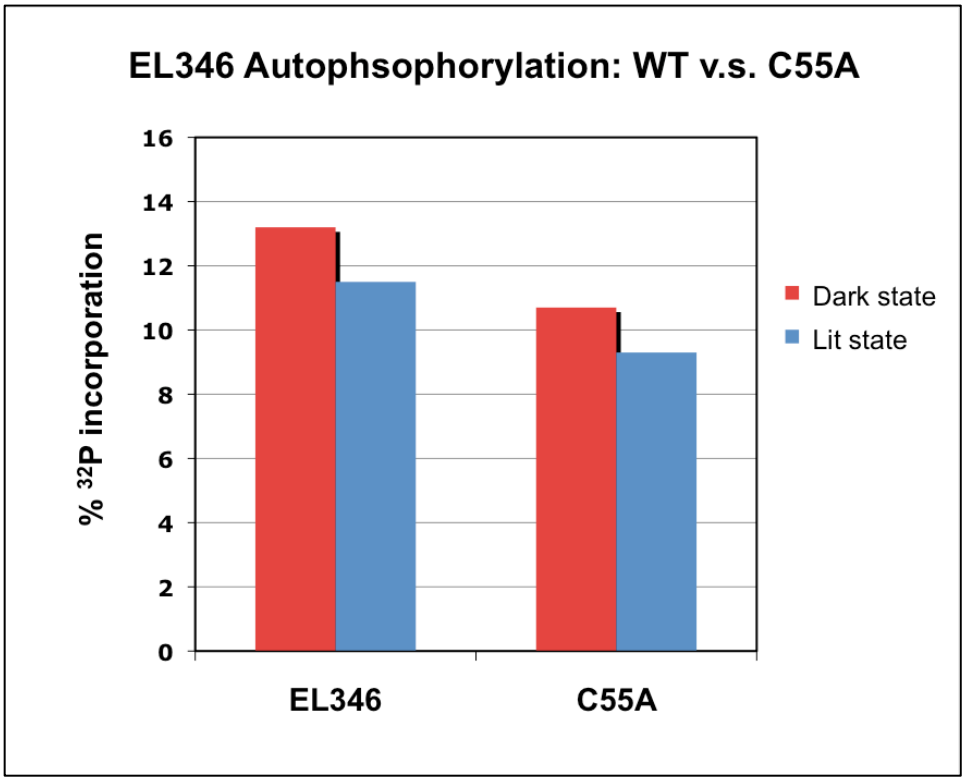


Figure A.10 pH optimization for EL346 kinase assay. Samples (10μM protein with 1μCi hot ATP) were incubated in various pH of buffers containing 50 mM Tris, 5 mM $MgCl_2$, and 100 mM NaCl for 4 min. at room temperature and stopped by adding 2X SDS loading buffer. These were then fractionated with SDS–polyacrylamide gel electrophoresis, and the gels were exposed to imaging plate for 10 h.

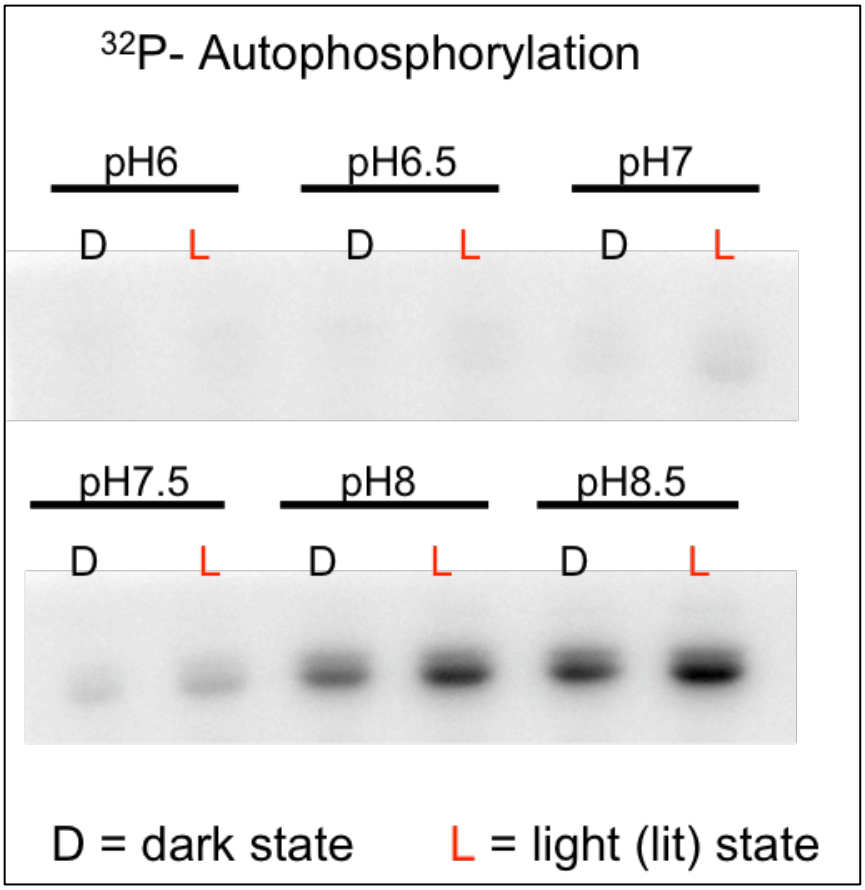


Figure A.11 Divalent preference for EL346 kinase assay. Dark and lit states kinase assays were performed using 10 μM of proteins in pH 8.2 buffer with 50 mM Tris, 100 mM NaCl, 10 mM $MgCl_2$, and 10% glycerol. 5 μCi of hot ATP were used per 10 μL of reaction mixture with final cold ATP concentration of 500 μM. Divalent preference was tested using 5 mM of $MgCl_2$, $MnCl_2$, or $CaCl_2$. Total reaction proceeded for 8 min before the addition of 2X SDS stop buffer.

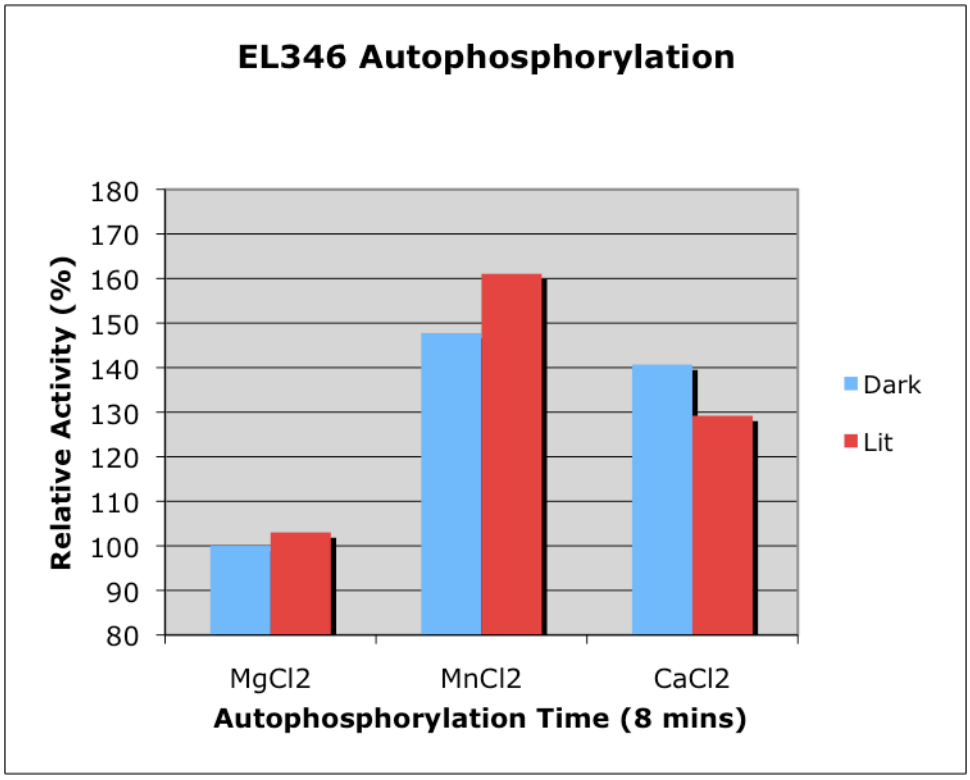


Figure A.12 Temperature optimization for EL346 stability. Tagged and untagged proteins were stored in two different conditions, in -20 ºC (with 50% glycerol) and in -80 ºC (fast freeze with liquid nitrogen without glycerol) overnight prior to the autokinase assay. Kinase assays were performed using 10 μM of proteins in pH 8.5 buffer with 50 mM Tris, 100 mM NaCl, 10 mM $MgCl_2$, and 10% glycerol. 5 μCi of hot ATP were used per 10 μL of reaction mixture with final cold ATP concentration of 300 μM. Total reaction proceeded for 8 min before the addition of 2X SDS stop buffer. Proteins stored at the 4 ºC were also tested and the activity in 4 ºC was weaker than proteins stored at -20 ºC but a bit better than protein stored at -80 ºC (data not shown).

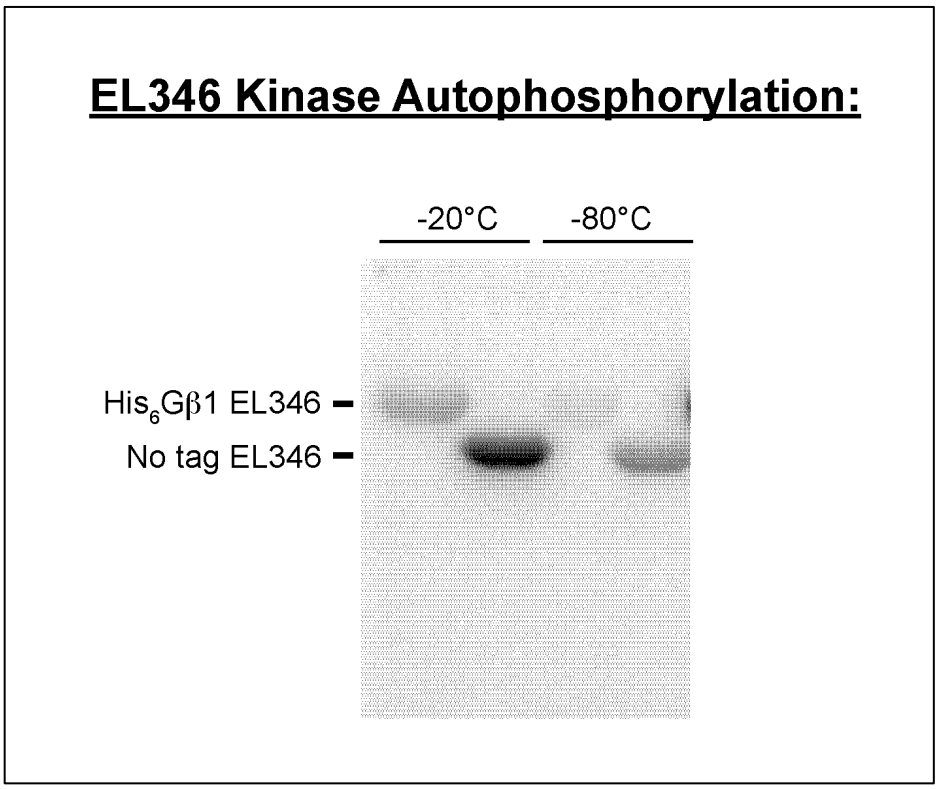


Figure A.13 Phosphotransfer profiling of EL346 tested against 21 out of the total of 23 known RRs in the *Erythrobacter litoralis*. Individual response regulator was randomly assigned to a number as the following: RR1: ELI07655, RR2: ELI04852, RR3: ELI11255, RR4: ELI14085, RR5: ELI10215, RR6: ELI07690, RR7: ELI09195, RR8: ELI02135, RR9: ELI02515, RR10: ELI03645, RR11: ELI05400, RR12: ELI11920, RR13: ELI14230, RR14: ELI09055, RR15: ELI09265, RR16: ELI02325, RR17: ELI05325, RR18: ELI06265, RR19: ELI06275, RR20: ELI12660, RR21: ELI12895, RR22: ELI11250, and RR23: ELI01510. The autophosphorylation and phosphotransfer were conducted in the light state. Autophosphorylation of EL346 was conducted for 10 minutes prior to the addition of equal molar ratio of various RRs. The phosphotransfer reaction was performed at two different lengths of time (30 second and 10 minuets) for comparison. The result from the 30–second phosphotransfer experiment showed that RR1 and RR5 are the cognate response regulators for EL346 and RR1 is a bit more potent than RR5. Additionally, RR7 can also be phosphorylated by EL346 (as seen in the 10 min. result) but with a lesser degree of phosphotransfer. Fernando Corrêa generated profiling results shown here.

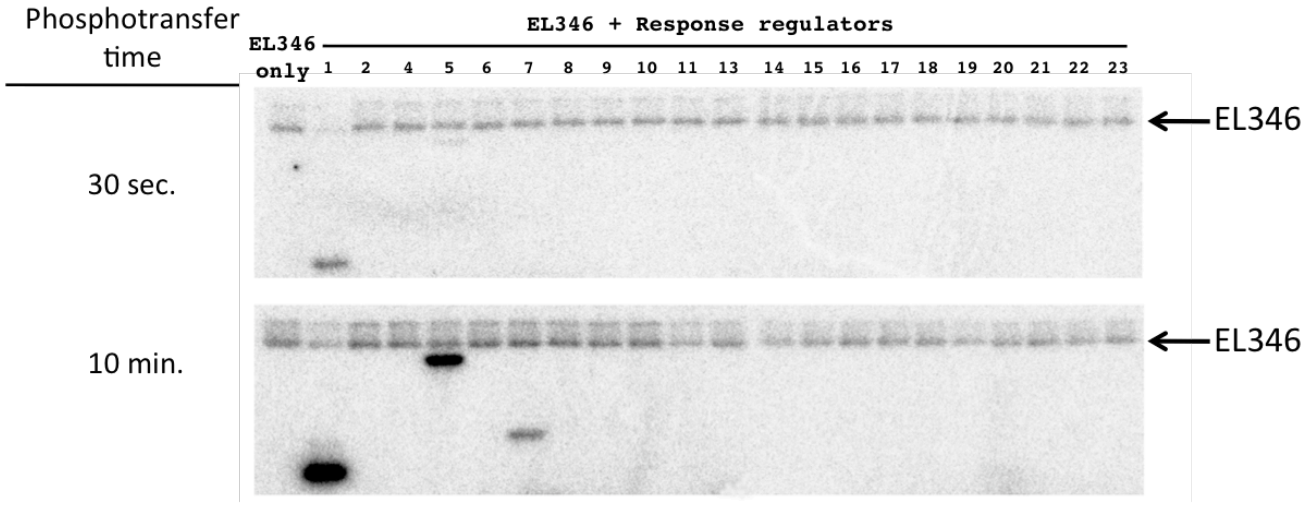


Figure A.14 Initial hit of EL346 crystallization. EL346 was purified in one day and the trays were set up the following day under ambient light condition using Phoenix robot. Proteins (10 mg/ml, ~260 μM) were prepared in pH 8.0 10 mM Tris, 100 mM NaCl, 5 mM $MgCl_2$, 2 mM TCEP, and 5 mM AMPPNP two different conditions: i) with additional 50 mM L–Arg/L–Glu and ii) without. Sample drop size is 300 nl of protein plus 300 nl of reservoir buffer. I did a total of 5 screens: the classic suite, the PACT suite, and the protein complex suite, the JCSG plus suite, and the Hampton index screen, each screen contains a total of 96 conditions. Four hits were obtained from the Hampton index screen in the following three wells: D10, F6 and F7 (buffer conditions see main text). Both F6 and F7 contain many single and multiple hexagonal plate shape crystals, while D10 contains a single log rod shape crystal coupled to other plate shape crystals.

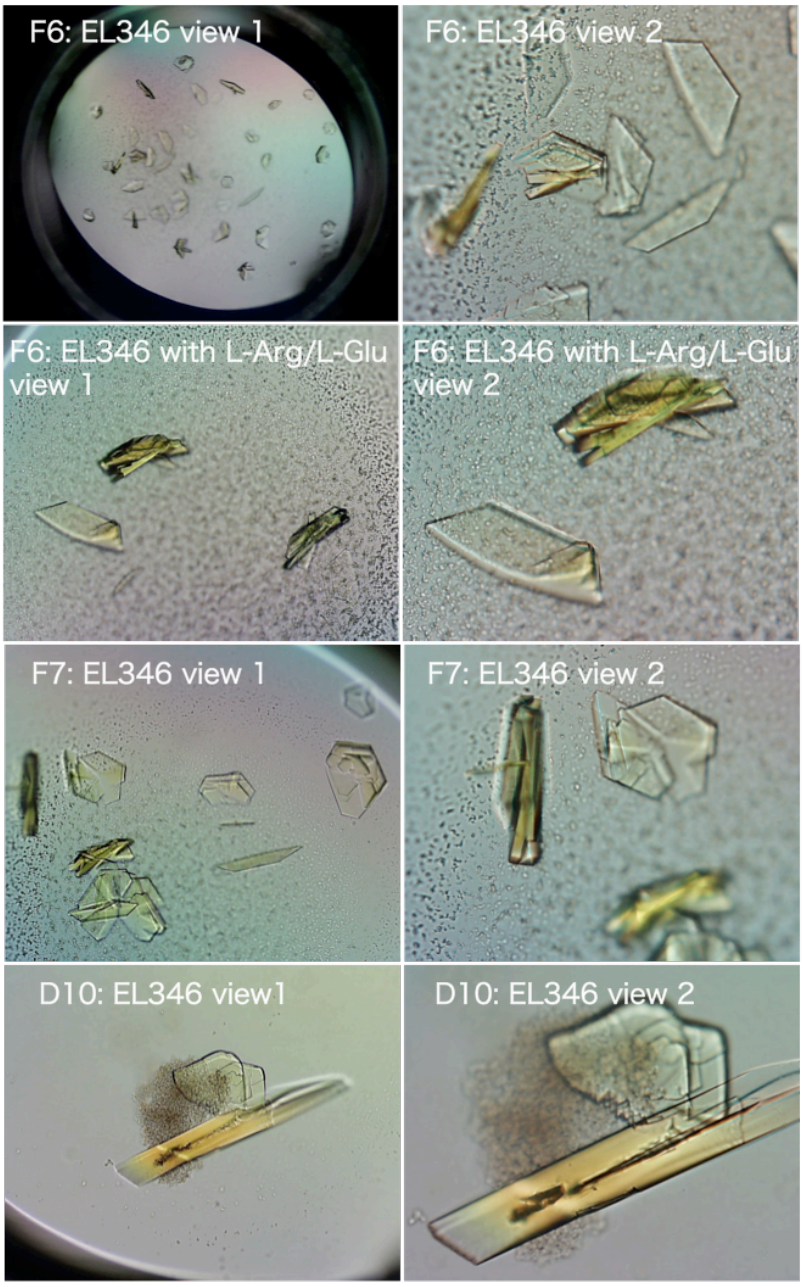
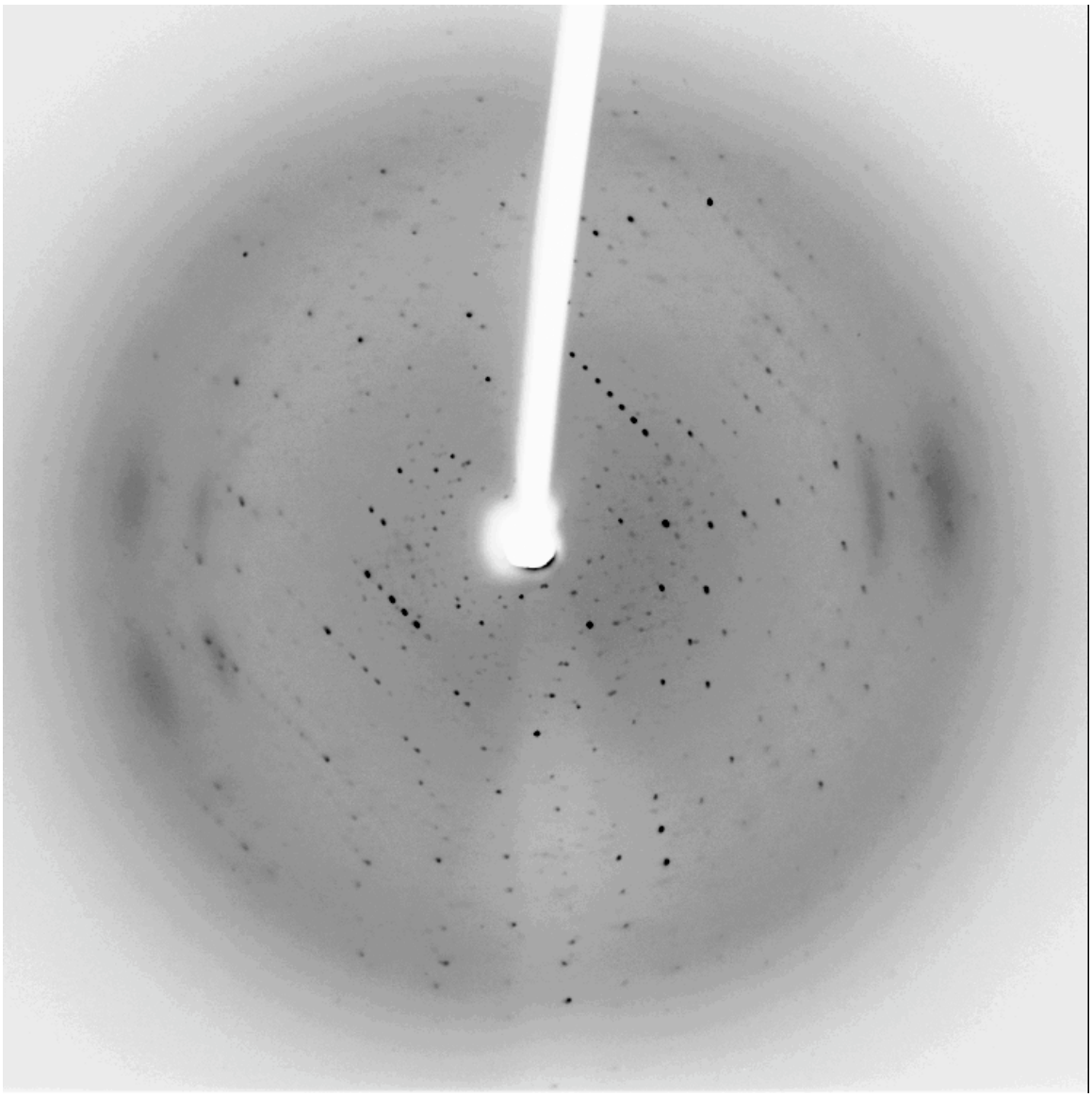


Figure A.15 Diffraction pattern of the EL346 (well: D10) crystal. It diffracts to less than 3 Å. Data collected by Diana Tomchick.



References

1. Purcell, E. B., Siegal-Gaskins, D., Rawling, D. C., Fiebig, A., and Crosson, S. (2007) A photosensory two-component system regulates bacterial cell attachment, *Proc Natl Acad Sci U S A 104*, 18241-18246.
2. 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.
3. Huala, E., Oeller, P. W., Liscum, E., Han, I. S., Larsen, E., and Briggs, W. R. (1997) Arabidopsis NPH1: a protein kinase with a putative redox-sensing domain, *Science 278*, 2120-2123.
4. Purschwitz, J., Muller, S., Kastner, C., and Fischer, R. (2006) Seeing the rainbow: light sensing in fungi, *Curr Opin Microbiol 9*, 566-571.
5. van der Horst, M. A., and Hellingwerf, K. J. (2004) Photoreceptor proteins, "star actors of modern times": a review of the functional dynamics in the structure of representative members of six different photoreceptor families, *Acc Chem Res 37*, 13-20.
6. 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.
7. Crosson, S., Rajagopal, S., and Moffat, K. (2003) The LOV domain family: photoresponsive signaling modules coupled to diverse output domains, *Biochemistry 42*, 2-10.
8. 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.
9. Pellequer, J. L., Wager-Smith, K. A., Kay, S. A., and Getzoff, E. D. (1998) Photoactive yellow protein: a structural prototype for the three-dimensional fold of the PAS domain superfamily, *Proc Natl Acad Sci U S A 95*, 5884-5890.
10. Miyatake, H., Mukai, M., Park, S. Y., Adachi, S., Tamura, K., Nakamura, H., Nakamura, K., Tsuchiya, T., Iizuka, T., and Shiro, Y. (2000) Sensory mechanism of oxygen sensor FixL from Rhizobium meliloti: crystallographic, mutagenesis and resonance Raman spectroscopic studies, *J Mol Biol 301*, 415-431.
11. 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.

12. 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.
13. Morais Cabral, J. H., Lee, A., Cohen, S. L., Chait, B. T., Li, M., and Mackinnon, R. (1998) Crystal structure and functional analysis of the HERG potassium channel N terminus: a eukaryotic PAS domain, *Cell 95*, 649-655.
14. Moglich, A., Ayers, R. A., and Moffat, K. (2009) Structure and signaling mechanism of Per-ARNT-Sim domains, *Structure 17*, 1282-1294.
15. Finn, R. D., Mistry, J., Schuster-Bockler, B., Griffiths-Jones, S., Hollich, V., Lassmann, T., Moxon, S., Marshall, M., Khanna, A., Durbin, R., Eddy, S. R., Sonnhammer, E. L., and Bateman, A. (2006) Pfam: clans, web tools and services, *Nucleic Acids Res 34*, D247-251.
16. Hefti, M. H., Francoijs, K. J., de Vries, S. C., Dixon, R., and Vervoort, J. (2004) The PAS fold. A redefinition of the PAS domain based upon structural prediction, *Eur J Biochem 271*, 1198-1208.
17. Bateman, A., Birney, E., Cerruti, L., Durbin, R., Etwiller, L., Eddy, S. R., Griffiths-Jones, S., Howe, K. L., Marshall, M., and Sonnhammer, E. L. (2002) The Pfam protein families database, *Nucleic Acids Res 30*, 276-280.
18. 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.
19. Miller, S. M., Massey, V., Ballou, D., Williams, C. H., Jr., Distefano, M. D., Moore, M. J., and Walsh, C. T. (1990) Use of a site-directed triple mutant to trap intermediates: demonstration that the flavin C(4a)-thiol adduct and reduced flavin are kinetically competent intermediates in mercuric ion reductase, *Biochemistry 29*, 2831-2841.
20. 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.
21. 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.
22. 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.

23. Gu, Y. Z., Hogenesch, J. B., and Bradfield, C. A. (2000) The PAS superfamily: sensors of environmental and developmental signals, *Annu Rev Pharmacol Toxicol 40*, 519-561.
24. Key, J., Srajer, V., Pahl, R., and Moffat, K. (2007) Time-resolved crystallographic studies of the heme domain of the oxygen sensor FixL: structural dynamics of ligand rebinding and their relation to signal transduction, *Biochemistry 46*, 4706-4715.
25. Harper, S. M., Neil, L. C., and Gardner, K. H. (2003) Structural basis of a phototropin light switch, *Science 301*, 1541-1544.
26. 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.
27. Zoltowski, B. D., and Crane, B. R. (2008) Light activation of the LOV protein vivid generates a rapidly exchanging dimer, *Biochemistry 47*, 7012-7019.
28. 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.
29. Gong, W., Hao, B., and Chan, M. K. (2000) New mechanistic insights from structural studies of the oxygen-sensing domain of Bradyrhizobium japonicum FixL, *Biochemistry 39*, 3955-3962.
30. 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.
31. Ronson, C. W., Nixon, B. T., and Ausubel, F. M. (1987) Conserved domains in bacterial regulatory proteins that respond to environmental stimuli, *Cell 49*, 579-581.
32. Stock, A. M., Robinson, V. L., and Goudreau, P. N. (2000) Two-component signal transduction, *Annu Rev Biochem 69*, 183-215.
33. Grebe, T. W., and Stock, J. B. (1999) The histidine protein kinase superfamily, *Adv Microb Physiol 41*, 139-227.
34. West, A. H., and Stock, A. M. (2001) Histidine kinases and response regulator proteins in two-component signaling systems, *Trends Biochem Sci 26*, 369-376.
35. Parkinson, J. S., and Kofoid, E. C. (1992) Communication modules in bacterial signaling proteins, *Annu Rev Genet 26*, 71-112.
36. Laub, M. T., and Goulian, M. (2007) Specificity in two-component signal transduction pathways, *Annu Rev Genet 41*, 121-145.

37. 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.
38. Ulrich, L. E., and Zhulin, I. B. (2007) MiST: a microbial signal transduction database, *Nucleic Acids Res 35*, D386-390.
39. Galperin, M. Y. (2006) Structural classification of bacterial response regulators: diversity of output domains and domain combinations, *J Bacteriol 188*, 4169-4182.
40. Alex, L. A., and Simon, M. I. (1994) Protein histidine kinases and signal transduction in prokaryotes and eukaryotes, *Trends Genet 10*, 133-138.
41. Swanson, R. V., Alex, L. A., and Simon, M. I. (1994) Histidine and aspartate phosphorylation: two-component systems and the limits of homology, *Trends Biochem Sci 19*, 485-490.
42. Skerker, J. M., Perchuk, B. S., Siryaporn, A., Lubin, E. A., Ashenberg, O., Goulian, M., and Laub, M. T. (2008) Rewiring the specificity of two-component signal transduction systems, *Cell 133*, 1043-1054.
43. Mascher, T., Helmann, J. D., and Unden, G. (2006) Stimulus perception in bacterial signal-transducing histidine kinases, *Microbiol Mol Biol Rev 70*, 910-938.
44. Tomomori, C., Tanaka, T., Dutta, R., Park, H., Saha, S. K., Zhu, Y., Ishima, R., Liu, D., Tong, K. I., Kurokawa, H., Qian, H., Inouye, M., and Ikura, M. (1999) Solution structure of the homodimeric core domain of Escherichia coli histidine kinase EnvZ, *Nat Struct Biol 6*, 729-734.
45. Clarke, M. B., Hughes, D. T., Zhu, C., Boedeker, E. C., and Sperandio, V. (2006) The QseC sensor kinase: a bacterial adrenergic receptor, *Proc Natl Acad Sci U S A 103*, 10420-10425.
46. Bilwes, A. M., Alex, L. A., Crane, B. R., and Simon, M. I. (1999) Structure of CheA, a signal-transducing histidine kinase, *Cell 96*, 131-141.
47. Dutta, R., Qin, L., and Inouye, M. (1999) Histidine kinases: diversity of domain organization, *Mol Microbiol 34*, 633-640.
48. Zhu, Y., Qin, L., Yoshida, T., and Inouye, M. (2000) Phosphatase activity of histidine kinase EnvZ without kinase catalytic domain, *Proc Natl Acad Sci U S A 97*, 7808-7813.
49. 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.
50. Dutta, R., and Inouye, M. (2000) GHKL, an emergent ATPase/kinase superfamily, *Trends Biochem Sci 25*, 24-28.

51. Marina, A., Waldburger, C. D., and Hendrickson, W. A. (2005) Structure of the entire cytoplasmic portion of a sensor histidine-kinase protein, *EMBO J 24*, 4247-4259.
52. Casino, P., Rubio, V., and Marina, A. (2009) Structural insight into partner specificity and phosphoryl transfer in two-component signal transduction, *Cell 139*, 325-336.
53. Bergerat, A., de Massy, B., Gadelle, D., Varoutas, P. C., Nicolas, A., and Forterre, P. (1997) An atypical topoisomerase II from Archaea with implications for meiotic recombination, *Nature 386*, 414-417.
54. Yang, Y., and Inouye, M. (1993) Requirement of both kinase and phosphatase activities of an Escherichia coli receptor (Taz1) for ligand-dependent signal transduction, *J Mol Biol 231*, 335-342.
55. Stewart, R. C., VanBruggen, R., Ellefson, D. D., and Wolfe, A. J. (1998) TNP-ATP and TNP-ADP as probes of the nucleotide binding site of CheA, the histidine protein kinase in the chemotaxis signal transduction pathway of Escherichia coli, *Biochemistry 37*, 12269-12279.
56. Pappalardo, L., Janausch, I. G., Vijayan, V., Zientz, E., Junker, J., Peti, W., Zweckstetter, M., Unden, G., and Griesinger, C. (2003) The NMR structure of the sensory domain of the membranous two-component fumarate sensor (histidine protein kinase) DcuS of Escherichia coli, *J Biol Chem 278*, 39185-39188.
57. Reinelt, S., Hofmann, E., Gerharz, T., Bott, M., and Madden, D. R. (2003) The structure of the periplasmic ligand-binding domain of the sensor kinase CitA reveals the first extracellular PAS domain, *J Biol Chem 278*, 39189-39196.
58. Cheung, J., Bingman, C. A., Reyngold, M., Hendrickson, W. A., and Waldburger, C. D. (2008) Crystal structure of a functional dimer of the PhoQ sensor domain, *J Biol Chem 283*, 13762-13770.
59. Albanesi, D., Martin, M., Trajtenberg, F., Mansilla, M. C., Haouz, A., Alzari, P. M., de Mendoza, D., and Buschiazzo, A. (2009) Structural plasticity and catalysis regulation of a thermosensor histidine kinase, *Proc Natl Acad Sci U S A 106*, 16185-16190.
60. Galperin, M. Y. (2005) A census of membrane-bound and intracellular signal transduction proteins in bacteria: bacterial IQ, extroverts and introverts, *BMC Microbiol 5*, 35.
61. Volz, K. (1993) Structural conservation in the CheY superfamily, *Biochemistry 32*, 11741-11753.

62. Martinez-Hackert, E., and Stock, A. M. (1997) The DNA-binding domain of OmpR: crystal structures of a winged helix transcription factor, *Structure 5*, 109-124.
63. 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.
64. Djordjevic, S., Goudreau, P. N., Xu, Q., Stock, A. M., and West, A. H. (1998) Structural basis for methylesterase CheB regulation by a phosphorylation-activated domain, *Proc Natl Acad Sci U S A 95*, 1381-1386.
65. Galperin, M. Y., Nikolskaya, A. N., and Koonin, E. V. (2001) Novel domains of the prokaryotic two-component signal transduction systems, *FEMS Microbiol Lett 203*, 11-21.
66. Skerker, J. M., Prasol, M. S., Perchuk, B. S., Biondi, E. G., and Laub, M. T. (2005) Two-component signal transduction pathways regulating growth and cell cycle progression in a bacterium: a system-level analysis, *PLoS Biol 3*, e334.
67. Laub, M. T., Biondi, E. G., and Skerker, J. M. (2007) Phosphotransfer profiling: systematic mapping of two-component signal transduction pathways and phosphorelays, *Methods Enzymol 423*, 531-548.
68. Stock, A. M., Martinez-Hackert, E., Rasmussen, B. F., West, A. H., Stock, J. B., Ringe, D., and Petsko, G. A. (1993) Structure of the Mg(2+)-bound form of CheY and mechanism of phosphoryl transfer in bacterial chemotaxis, *Biochemistry 32*, 13375-13380.
69. Lewis, R. J., Scott, D. J., Brannigan, J. A., Ladds, J. C., Cervin, M. A., Spiegelman, G. B., Hoggett, J. G., Barak, I., and Wilkinson, A. J. (2002) Dimer formation and transcription activation in the sporulation response regulator Spo0A, *J Mol Biol 316*, 235-245.
70. 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.
71. 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.
72. Robinson, V. L., Buckler, D. R., and Stock, A. M. (2000) A tale of two components: a novel kinase and a regulatory switch, *Nat Struct Biol 7*, 626-633.
73. Lukat, G. S., Lee, B. H., Mottonen, J. M., Stock, A. M., and Stock, J. B. (1991) Roles of the highly conserved aspartate and lysine residues in

the response regulator of bacterial chemotaxis, *J Biol Chem 266*, 8348-8354.

74. Varughese, K. I., Zhao, H., Veldore, V. H., and Zapf, J. (2007) Sporulation phosphorelay proteins and their complexes: crystallographic characterization, *Methods Enzymol 422*, 102-122.
75. Kennis, J. T., and Crosson, S. (2007) Microbiology. A bacterial pathogen sees the light, *Science 317*, 1041-1042.
76. 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.
77. Moglich, A., Ayers, R. A., and Moffat, K. (2009) Design and signaling mechanism of light-regulated histidine kinases, *J Mol Biol 385*, 1433-1444.
78. Rasko, D. A., Moreira, C. G., Li de, R., Reading, N. C., Ritchie, J. M., Waldor, M. K., Williams, N., Taussig, R., Wei, S., Roth, M., Hughes, D. T., Huntley, J. F., Fina, M. W., Falck, J. R., and Sperandio, V. (2008) Targeting QseC signaling and virulence for antibiotic development, *Science 321*, 1078-1080.
79. Matsushita, M., and Janda, K. D. (2002) Histidine kinases as targets for new antimicrobial agents, *Bioorg Med Chem 10*, 855-867.
80. Semenza, G. L. (1999) Regulation of mammalian O2 homeostasis by hypoxia-inducible factor 1, *Annu Rev Cell Dev Biol 15*, 551-578.
81. 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.
82. 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.
83. 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.
84. 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.
85. He, Q., and Liu, Y. (2005) Molecular mechanism of light responses in Neurospora: from light-induced transcription to photoadaptation, *Genes Dev 19*, 2888-2899.

86. Harper, S. M., Christie, J. M., and Gardner, K. H. (2004) Disruption of the LOV-Jalpha helix interaction activates phototropin kinase activity, *Biochemistry 43*, 16184-16192.
87. 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.
88. Crosson, S., and Moffat, K. (2002) Photoexcited structure of a plant photoreceptor domain reveals a light-driven molecular switch, *Plant Cell 14*, 1067-1075.
89. 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.
90. 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.
91. 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.
92. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes, *J Biomol NMR 6*, 277-293.
93. Johnson, B. A. (2004) Using NMRView to visualize and analyze the NMR spectra of macromolecules, *Methods Mol Biol 278*, 313-352.
94. 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.
95. Gouet, P., Courcelle, E., Stuart, D. I., and Metoz, F. (1999) ESPript: analysis of multiple sequence alignments in PostScript, *Bioinformatics 15*, 305-308.
96. Yao, X., Rosen, M. K., and Gardner, K. H. (2008) Estimation of the available free energy in a LOV2-J alpha photoswitch, *Nat Chem Biol 4*, 491-497.
97. 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.

98. 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.
99. 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.
100. 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.
101. 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.
102. 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.
103. Guo, H., Kottke, T., Hegemann, P., and Dick, B. (2005) The phot LOV2 domain and its interaction with LOV1, *Biophys J 89*, 402-412.
104. 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.
105. 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.
106. 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.
107. 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.
108. 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.
109. 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.
110. Huse, M., Chen, Y. G., Massague, J., and Kuriyan, J. (1999) Crystal structure of the cytoplasmic domain of the type I TGF beta receptor in complex with FKBP12, *Cell 96*, 425-436.
111. Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M. R., Appel, R. D., and Bairoch, A., (Eds.) (2005) *Protein Identification and Analysis Tools on the ExPASy Server*, Humana Press.
112. 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.
113. Kaiserli, E., Sullivan, S., Jones, M. A., Feeney, K. A., and Christie, J. M. (2009) Domain Swapping to Assess the Mechanistic Basis of Arabidopsis Phototropin 1 Receptor Kinase Activation and Endocytosis by Blue Light, *Plant Cell*.
114. Gao, R., and Stock, A. M. (2009) Biological Insights from Structures of Two-Component Proteins, *Annu Rev Microbiol*.
115. Briggs, W. R., and Huala, E. (1999) Blue-light photoreceptors in higher plants, *Annu Rev Cell Dev Biol 15*, 33-62.
116. Quail, P. H. (2002) Phytochrome photosensory signalling networks, *Nat Rev Mol Cell Biol 3*, 85-93.
117. van der Horst, M. A., Key, J., and Hellingwerf, K. J. (2007) Photosensing in chemotrophic, non-phototrophic bacteria: let there be light sensing too, *Trends Microbiol 15*, 554-562.
118. 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.
119. Sousa, E. H., Tuckerman, J. R., Gonzalez, G., and Gilles-Gonzalez, M. A. (2007) DosT and DevS are oxygen-switched kinases in Mycobacterium tuberculosis, *Protein Sci 16*, 1708-1719.
120. Sousa, E. H., Tuckerman, J. R., Gonzalez, G., and Gilles-Gonzalez, M. A. (2007) A memory of oxygen binding explains the dose response of the heme-based sensor FixL, *Biochemistry 46*, 6249-6257.
121. Dunham, C. M., Dioum, E. M., Tuckerman, J. R., Gonzalez, G., Scott, W. G., and Gilles-Gonzalez, M. A. (2003) A distal arginine in oxygen-sensing heme-PAS domains is essential to ligand binding, signal transduction, and structure, *Biochemistry 42*, 7701-7708.
122. Tuckerman, J. R., Gonzalez, G., Dioum, E. M., and Gilles-Gonzalez, M. A. (2002) Ligand and oxidation-state specific regulation of the heme-

based oxygen sensor FixL from Sinorhizobium meliloti, *Biochemistry 41*, 6170-6177.
123. Tuckerman, J. R., Gonzalez, G., and Gilles-Gonzalez, M. A. (2001) Complexation precedes phosphorylation for two-component regulatory system FixL/FixJ of Sinorhizobium meliloti, *J Mol Biol 308*, 449-455.
124. Hao, B., Isaza, C., Arndt, J., Soltis, M., and Chan, M. K. (2002) Structure-based mechanism of O2 sensing and ligand discrimination by the FixL heme domain of Bradyrhizobium japonicum, *Biochemistry 41*, 12952-12958.
125. Key, J., and Moffat, K. (2005) Crystal structures of deoxy and CO-bound bjFixLH reveal details of ligand recognition and signaling, *Biochemistry 44*, 4627-4635.
126. Ayers, R. A., and Moffat, K. (2008) Changes in quaternary structure in the signaling mechanisms of PAS domains, *Biochemistry 47*, 12078-12086.
127. Podust, L. M., Ioanoviciu, A., and Ortiz de Montellano, P. R. (2008) 2.3 A X-ray structure of the heme-bound GAF domain of sensory histidine kinase DosT of Mycobacterium tuberculosis, *Biochemistry 47*, 12523-12531.
128. Cho, H. Y., Cho, H. J., Kim, Y. M., Oh, J. I., and Kang, B. S. (2009) Structural insight into the heme-based redox sensing by DosS from Mycobacterium tuberculosis, *J Biol Chem 284*, 13057-13067.
129. Tanaka, T., Saha, S. K., Tomomori, C., Ishima, R., Liu, D., Tong, K. I., Park, H., Dutta, R., Qin, L., Swindells, M. B., Yamazaki, T., Ono, A. M., Kainosho, M., Inouye, M., and Ikura, M. (1998) NMR structure of the histidine kinase domain of the E. coli osmosensor EnvZ, *Nature 396*, 88-92.
130. Bilwes, A. M., Quezada, C. M., Croal, L. R., Crane, B. R., and Simon, M. I. (2001) Nucleotide binding by the histidine kinase CheA, *Nat Struct Biol 8*, 353-360.
131. Marina, A., Mott, C., Auyzenberg, A., Hendrickson, W. A., and Waldburger, C. D. (2001) Structural and mutational analysis of the PhoQ histidine kinase catalytic domain. Insight into the reaction mechanism, *J Biol Chem 276*, 41182-41190.
132. Guarnieri, M. T., Zhang, L., Shen, J., and Zhao, R. (2008) The Hsp90 inhibitor radicicol interacts with the ATP-binding pocket of bacterial sensor kinase PhoQ, *J Mol Biol 379*, 82-93.
133. Bick, M. J., Lamour, V., Rajashankar, K. R., Gordiyenko, Y., Robinson, C. V., and Darst, S. A. (2009) How to switch off a histidine kinase: crystal

structure of Geobacillus stearothermophilus KinB with the inhibitor Sda, *J Mol Biol 386*, 163-177.
134. Yamada, S., Akiyama, S., Sugimoto, H., Kumita, H., Ito, K., Fujisawa, T., Nakamura, H., and Shiro, Y. (2006) The signaling pathway in histidine kinase and the response regulator complex revealed by X-ray crystallography and solution scattering, *J Mol Biol 362*, 123-139.
135. Evans, K., Grossmann, J. G., Fordham-Skelton, A. P., and Papiz, M. Z. (2006) Small-angle X-ray scattering reveals the solution structure of a bacteriophytochrome in the catalytically active Pr state, *J Mol Biol 364*, 655-666.
136. 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.
137. Cuff, J. A., Clamp, M. E., Siddiqui, A. S., Finlay, M., and Barton, G. J. (1998) JPred: a consensus secondary structure prediction server, *Bioinformatics 14*, 892-893.
138. Losi, A., Quest, B., and Gartner, W. (2003) Listening to the blue: the time-resolved thermodynamics of the bacterial blue-light receptor YtvA and its isolated LOV domain, *Photochem Photobiol Sci 2*, 759-766.
139. Zoltowski, B. D., Vaccaro, B., and Crane, B. R. (2009) Mechanism-based tuning of a LOV domain photoreceptor, *Nat Chem Biol*.
140. Kennis, J. T., van Stokkum, I. H., Crosson, S., Gauden, M., Moffat, K., and van Grondelle, R. (2004) The LOV2 domain of phototropin: a reversible photochromic switch, *J Am Chem Soc 126*, 4512-4513.
141. Surette, M. G., Levit, M., Liu, Y., Lukat, G., Ninfa, E. G., Ninfa, A., and Stock, J. B. (1996) Dimerization is required for the activity of the protein histidine kinase CheA that mediates signal transduction in bacterial chemotaxis, *J Biol Chem 271*, 939-945.
142. Qin, L., Dutta, R., Kurokawa, H., Ikura, M., and Inouye, M. (2000) A monomeric histidine kinase derived from EnvZ, an Escherichia coli osmosensor, *Mol Microbiol 36*, 24-32.
143. Singh, M., Berger, B., Kim, P. S., Berger, J. M., and Cochran, A. G. (1998) Computational learning reveals coiled coil-like motifs in histidine kinase linker domains, *Proc Natl Acad Sci U S A 95*, 2738-2743.
144. Francez-Charlot, A., Frunzke, J., Reichen, C., Ebneter, J. Z., Gourion, B., and Vorholt, J. A. (2009) Sigma factor mimicry involved in regulation of general stress response, *Proc Natl Acad Sci U S A 106*, 3467-3472.