STRUCTURE–FUNCTION RELATIONSHIPS IN THE PHOTOACTIVE YELLOW PROTEIN FAMILY OF PHOTORECEPTORS

By

RACHANA RAMILA SURESH RATHOD

Bachelor of Science in Botany
University of Mumbai
Mumbai, Maharashtra India
2004

Master of Science in Botany
University of Mumbai
Mumbai, Maharashtra India
2006

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
DOCTOR OF PHILOSOPHY
December, 2014

STRUCTURE–FUNCTION RELATIONSHIPS IN THE PHOTOACTIVE YELLOW PROTEIN FAMILY OF PHOTORECEPTORS

Dissertation Approved:

Dr. Wouter D. Hoff

Dissertation Adviser

Dr. William Picking

Dr. Jeff Hadwiger

Dr. Marianna Patrauchan

Dr. Aihua Xie

Name: RACHANA RAMILA SURESH RATHOD

Date of Degree: DECEMBER, 2014

Title of Study: STRUCTURE–FUNCTION RELATIONSHIPS IN THE PHOTOACTIVE YELLOW PROTEIN FAMILY OF PHOTORECEPTORS

Major Field: MICROBIOLOGYAND MOLECULAR GENETICS

Abstract: The amino acid sequence of a protein determines its three-dimensional structure, which in turn determines its functional properties. An intensively studied but still partially unresolved question is how the structure of a protein relates to its functional properties. Here we use photoactive yellow protein (PYP) as a model system to examine questions on protein structure-function relationships. PYP is a bacterial blue light photoreceptor, a prototype of the diverse PAS domain superfamily, and a model system for functional protein dynamics. The work in this thesis was directed at three aims: (1) developing tools to identify the structural change that triggers intramolecular proton transfer during the PYP photocycle; (2) the functional role of PAS-conserved residue Ile39 in PYP; and (3) determining to what extent the extensively studied structure-function relations in the PYP from *Halorhodospira halophila* apply to the PYP from *Rhodospirillum centenum*. (1) The molecular events that cause directional proton transfer in proteins are largely unknown. We develop tools to allow the testing of the specific hypothesis that the disruption of the Tyr42-*p*CA hydrogen bond during the PYP photocycle causes proton transfer. We developed an effective approach for obtaining Tyr-D_4-labeled PYP that can be used in infrared studies to identify Tyr side chain signals. (2) The PAS domain superfamily is defined by weak but characteristic amino acid sequence conservation, but the functional role of PAS-conserved residues remains poorly understood. We examined PAS-conserved residue Ile39 through biophysical characterization of the I39A PYP mutant. This work revealed that Ile39 is at the core of a set of hydrophobic interactions conserved in PAS domains, is not an essential part in the transmission mechanism of allosteric structural changes during PYP signaling and affects both signaling kinetics and folding cooperativity. (3) We found that structure-function rules for Hhal PYP qualitatively transfer to Rcen PYP, including the role of Glu46 as the electrostatic epicenter for driving conformational changes. The resulting set of Rcen PYP mutants with altered photocycle rate and reduced conformational changes provides a powerful tool for future studies on the photocycle events that are needed for *in vivo* signaling by PYP.

TABLE OF CONTENTS

Chapter Page

LIST OF TABLES

Table Page

LIST OF FIGURES

Figure Page

Figure Page

.

CHAPTER I

Introduction

1.1 Overview of this thesis

This study examines protein structure-function relationships in photoactive yellow protein (PYP), a bacterial photoreceptor that serves as a model system for understanding functional protein dynamics and the PAS domain superfamily. Structure-function relationships in PYP are examined in three expanding steps: (i) Structure-function relationships in a ***single protein*** (the PYP from *Halorhodospira halophila*), with emphasis on developing approaches to understand directional proton transfer within proteins (Chapter III). (ii) Site-directed mutagenesis of the PYP from *Rhodospirillum centenum* to examine the degree to which the knowledge gained from extensive studies of *H. halophila* PYP transfers to other members of the ***entire protein family*** of PYP photoreceptors (Chapter V). (iii) Using PYP to explore structure-function relationships that apply to an ***entire protein superfamily***, in this case the PAS domain superfamily. (Chapter IV).

1.2 Structure function relationships in proteins

A multitude of cellular functions are performed by a variety of protein molecules. Studying protein structure-function relationships is important for understanding these processes. In general, the question is how the biological function of a protein is encoded in its amino acid sequence. Such studies are of primary focus in diverse fields like structural and molecular biology, genetics, biochemistry, and biotechnology. Apart from exploring the basic precepts, structure-function studies are believed to have broader implications in the therapeutic development of diseases associated with protein misfolding and dysfunction. Two different approaches to examine protein structure-function relationships are generally used: i) using the information from primary amino acid sequence and the associated three-dimensional structure to study the functional properties of different point mutants, and ii) to determine the properties of naturally occurring homologs of the protein of interest, often involving bioinformatics approaches. In this thesis we will combine elements of these two approaches and examine various properties of two different PYPs and their respective mutants, particularly their stability, protein-ligand interactions, allosteric switching, and proton transfer. These topics are briefly introduced below.

1.2.1 Protein folding/stability

The 3D structure of a protein acquired by proper folding (acquiring a minimal-energy conformation (*3*)) of the polypeptide chain, determines its size, shape, and function. The stability of the native state of a protein is usually quantified by the difference in free energy between the folded state and the unfolded state ($\Delta_{F\text{-}U}$ or ΔG_U). The ΔG_U for the wild type protein can then be compared with the value for different mutants (*4*). Experimentally the unfolding of proteins is often studied using chemical denaturants such as urea or guanidinium hydrochloride (Gdm-HCl). For proteins that exhibit two-state folding, meaning that no folding intermediates are accumulated during equilibrium denaturant titrations, and that the folded state is directly converted into the fully unfolded state, the

ΔG_U for the protein in the absence of denaturant can be accurately extracted by analysis of the equilibrium denaturant unfolding process as monitored by some spectroscopic or biophysical probe such as absorbance or fluorescence. Both denaturant and thermal induced unfolding have been reported for various proteins such as *E. coli* thioredoxin (*4*) and *H. halophila* photoactive yellow protein (*5*). The unfolding of PYP follows two-state behavior in denaturant induced unfolding experiments and evidence has been reported for a direct relation between protein folding and signaling in this protein (*5*).

The complex 3D structure of a protein is held together by various weak molecular interactions, particularly hydrogen bonds, ionic bonds, and hydrophobic interactions. Hydrogen bonds are important for the structural stability of proteins, and such interactions between different amino acids are critical for defining the 3D fold of the native state and the physiological or enzymatic function of a protein (*6*). Dissociation of even a single hydrogen bond can render the protein unstable, as is observed in the blue light photoreceptor, PYP from *H. halophila*: the Y42F mutation, which perturbs an active site hydrogen bonding network, significantly reduces the stability of the protein (*7*). The activity of proteins can be changed by a number of factors (pH, ionic strength, temperature, mutations, etc.) that affect the molecular interactions holding its 3D structure together. Such environmental stresses or genetic factors can also cause misfolding of a protein. Misfolded states (toxic conformations) are of great medical relevance since they have been associated with several neurodegenerative as well as other diseases (*8*).

While the role of misfolded proteins in neurodegenerative diseases is beyond the scope of this thesis, it indicates the general relevance of the field of protein folding and misfolding. Accumulation of misfolded proteins form protein aggregates and can cause amyloid diseases (*3*). Mutations in proteins can cause conformational defects leading to diseases like Parkinson's, Alzheimer's (*3, 9, 10*), Huntington's disease (*11, 12*), scrapie in sheep (*13*), and mad cow disease (*14*). Many proteins are robust against single point mutations and generally do not lose their function, but many examples

with loss of normal function due to single residue deletion have been reported in genetic diseases like cystic fibrosis. A well-studied example is hemoglobin, where a single mutation in the β-globin protein causes the protein molecules to stick to each other. The altered structure of hemoglobin has a dramatic effect on the red blood cells and causes them to form a sickle shaped structure instead of the normal round shape, resulting in sickle cell anemia (*15*, *16*). Here we study the effect of point mutations on the stability and folding cooperativity of PYP.

1.2.2 Protein-ligand interactions

In the completely folded structure of proteins with catalytic activity we often observe interactions between amino acids that are distantly located in primary sequence. The amino acids that participate in the catalytic reaction form a special site called active site of the protein that binds specific ligands. Many proteins bind to appropriate ligand/ligands to be fully functional. Functional protein molecules often exhibit some kind of binding or interaction with other ions or small or large ligand molecules. An excellent example is that of antibodies produced by the immune system, which bind to very specific target molecules (haptens). There are several other proteins that require interactions with non-protein molecules for their functions. For example, the signaling photoreceptor protein in our eyes, rhodopsin, requires a retinal chromophore for sensing light. Upon illumination, the retinal chromophore embedded in the active site of rhodopsin absorbs a photon and undergoes photoisomerization that triggers a series of conformational changes in rhodopsin associated with signaling (*17*). Another method used by some proteins is the covalent addition or removal of a chemical group like phosphate that can render the protein active or inactive, respectively. In such cases, the addition of the phosphate group often triggers large conformational changes in the protein which are required for signaling, whereas the removal of this group returns the protein to its inactive conformation. Various computer programs use protein databases to predict the potential ligand binding sites of new proteins and this area of research has gained interest through its potential for the

design of novel drugs, especially for cancer therapy (*18*). Here we study protein-ligand interactions in PYP that alter the absorbance maximum and pK_a of the *p*-coumaric acid (*p*CA) chromophore in PYP.

1.2.3 Protein–protein interactions

A number of biological/signaling mechanisms are facilitated by the transient or continuous formation and dissociation of protein-protein interactions (*19-25*). The experimental identification and bioinformatic prediction of protein-protein interactions is of growing interest in functional genomics and is also important in the development of novel drugs and therapeutics (*26*). As discussed above for ligand binding, large and diverse conformational changes are a common occurrence upon protein-protein interactions/binding. In the case of allosteric proteins, the binding of the protein or ligand to one area of the protein can cause conformational changes at a distant area away from the binding site indicating the importance of studying the role of these interactions in regulating the structure and the associated function of the protein. Here we study the effect of point mutation on allosteric switching during the PYP photocycle.

1.2.4 Proton transfer reactions and hydrogen bonding networks in proteins

With the development of sophisticated spectroscopic instruments and pulsed lasers we can perform in-depth structural analyses of the conformational changes during protein function, enzymatic activity and regulation at the atomic level with time resolution as fast as 10^{-15} s. Proton transfer reactions play a key role in a number of biological processes, signaling transduction mechanisms and enzyme catalysis. In many proteins, proton transfer or a coupled motion of electrons and protons is observed. Proton transfer events in proteins often occur between acidic and basic amino acids that act as proton donors or acceptors (*27*). In bulk solvents, the energy barriers for moving charges during proton transfer events are stabilized by proton carriers like water molecules. Proteins are not simple solvents and the interior of a protein or a membrane has restricted access to such charge carriers. Therefore, proton transfer reactions within a protein usually occur along a chain

of charged (ionizable) or polar amino acids and isolated water molecules. These polar species form a hydrogen bonding network that functions as the proton transfer pathway (*27*, *28*).

Proton transfer reactions have been associated with various conformational changes that drive protein function; for example, deprotonation of the phenolic chromophore and subsequent protonation of Glu222 (due to rearrangement of hydrogen bonding network) in Green Fluorescent Protein and upon photoexcitation (*29*), a series of electron and proton transfer reactions causing the release of the quinone Q_B in the oxygen evolving complex in Photosystem II (*28*), proton pump proteins such as bacteriorhodopsin that contribute to the formation of the proton gradient (*30*, *31*), and an intramolecular proton transfer from Glu46 to the chromophore leading to the formation of the presumed signaling state in PYP (*32*, *33*). Here we develop spectroscopic approaches to examine the mechanism of proton transfer in PYP using vibrational spectroscopy.

1.2.5 Protein dynamics and vibrational spectroscopy

Understanding the structural changes occurring in a protein at a molecular or atomic/subatomic level during a biochemical process is a key to understanding the associated biological function. Recent advancements in the macromolecular structure determination methods like X–ray diffraction and NMR spectroscopy have resolved the three-dimensional structure of many molecules. However, these techniques are largely limited to static structures and stable intermediates of a biochemical reaction and cannot be used to monitor conformational changes occurring during the reaction (*34*). Atoms in protein molecules vibrate constantly and changes in the protein structure will alter the vibrational dynamics of the protein. Any changes in vibrational modes can be measured by a number of vibrational spectroscopy techniques which are known to be sensitive to such changes (*34*). A number of traditional and newly developed vibrational spectroscopic techniques are now widely used in the field of structural biology particularly Resonance Raman spectroscopy and Fourier transform infrared spectroscopy (FTIR).

FTIR spectroscopy is highly sensitive to small structural changes, has high time resolution (ns - µs), requires relatively small amounts (10-100 µg) of sample for measurements, and can be applied to small soluble proteins as well as large membrane proteins. The IR spectra can provide a range of information, such as the chemical structure of the vibrating group and its interactions, proton transfer, dissociation of hydrogen bonds, and changes in the secondary structure, bond angles, conformations and properties, redox state, as well as local electric fields (*35*). A number of methods have been described in the literature (summarized in (*35*)) to trigger protein reactions, particularly light, concentration jump (of ions or molecules), temperature or pressure jumps, and electrochemistry (for redox reactions). For the scope of this project, only light induced reactions will be discussed. Light-induced infrared difference spectroscopy has been one of the pioneer techniques used to study light flash induced photoreactions on photoreceptors like bacteriorhodopsin (*36-38*), where the light driven proton pump including photoisomerization and proton transfer steps, have been probed using different types of FTIR techniques. Other examples include cytochrome *c* oxidases and related heme-copper oxidases as reviewed in (*39*) using rapid scan FTIR, the kinetic resolution of the main intermediates in the pump cycle of Ca^{2+}-ATPase after ATP release (*40*), probing proton transfer and chromophore isomerization and protonation in *H. halophila* PYP (*32, 33*), and in the photosynthetic reaction center. Thus, time-resolved FTIR difference spectroscopy can be used as an important tool towards understanding protein dynamics, structure and function.

A range of different organisms utilize visible light as a source of energy or information to perform various enzymatic activities triggered by protein conformational changes. In case of photoreceptors these changes are generally associated with signal transduction mechanisms in the cell. Elucidation of the structure-function relationships in proteins is an important goal towards understanding such signal transduction mechanisms that govern several life processes. For this study, the photoreceptor Photoactive Yellow Protein from *H. halophila* is used as a model system to

investigate structure-function relationships in the PYP which represents the growing family of PYPs that are part of larger group, the PAS domain superfamily.

1.3 PAS domain superfamily

Classical protein families share a common type of substrate and enzymatic mechanism, and typically exhibit a substantial level of sequence similarity, similar functions, and three-dimensional structures. They also contain conserved active site residues. Development of high-end sequencing technologies and complex analysis tools and databases capable of identifying very low levels of amino acid sequence similarities have discovered small conserved domains or motifs within protein sequences (*41, 42*). The duplication or rearrangement (domain shuffling) of these functional domains is an important part of protein evolution (*43*). These conserved domains form diverse groups called protein superfamilies. Unlike protein families, superfamilies have low amino acid sequence similarities, varied active sites, and different functional properties, but share a common three-dimensional fold. This project focuses on the PAS domain superfamily and its structural prototype.

PAS (Per, Arnt, Sim) domains are important signaling modules found in diverse group of proteins from all three kingdoms of life. These domains are capable of monitoring a range of various stimuli like light, redox potential, oxygen and cellular energy by binding various cofactors within their hydrophobic cores, and interact with a range of different signal transduction chains through protein-protein interactions. PAS domains can be cytosolic or periplasmic, unlike several transmembrane proteins (*44*). They have been identified in many signaling proteins (*44, 45*), transcriptional factors, circadian clock proteins (*46, 47*), phytochromes, and several photo and chemoreceptors taxis and tropism (*48*). In eukaryotes, medically relevant PAS domains have been identified; regulators of responses to hypoxia that is reported to cause myocardial and cerebral ischemia (*49*) and tumor hypoxia (*50*), regulators of embryological development of central nervous system (*51*), and the PAS domain in human ERG potassium channel, mutations in which were

reported to cause the long QT syndrome (cardiac arrhythmia) (*52*). Though a number of PAS proteins have a known function, for the majority of these proteins such functional information is not available. With the growing members of this superfamily (currently ~50,000), analysis of the diverse functions based on the limited sequence similarities is challenging; such analyses can be performed more reliably based on their three-dimensional structures. Photoactive yellow protein is believed to be the structural prototype of the PAS domains (*1, 53, 54*), making PYP an important model system for biological signaling by this superfamily.

1.4 Photoactive yellow protein: A structural prototype of the PAS domain superfamily

Photoactive yellow protein (PYP) was discovered around 30 years ago in the extremely halophilic purple photosynthetic bacterium *Halorhodospira halophila (55)*. PYP is small (14 kDa, 125 amino acids), highly thermostable (T_m ~ 85°C) cytosolic blue light photoreceptor (*55, 56*) and it mediates the negative phototaxis (bacterial cells moves away from the light source) in *H. halophila* in response to blue light (*48*). The yellow color of the protein is caused by its unique chromophore (prosthetic group that captures a photon), an anionic cinnamic acid derivative called 4-hydroxycinnamic acid or *p*-coumaric acid (*p*CA) that is covalently linked via a thioester bond with Cys69 (sole Cys) of PYP (*57, 58*). Similar proteins were found in different bacterial species of the Proteobacteria (*2, 59-63*) and they were grouped into the Xanthopsins family, which is all blue light photoreceptors with *p*CA chromophore (*60*). The biochemical and biophysical properties of PYP have been extensively studied (*53, 64*) due to the rich biophysics of its light triggered function and its extraordinary ease of handling, high solubility, thermostability, and excellent overexpression in *E. coli* (*60, 65*).

The amino acid sequence conservation between the PAS domain consensus sequence and PYP was reported for the first time in 1995 (*66*). The high resolution three-dimensional structure has been resolved by X-ray crystallography (*67, 68*) and in solution by NMR (*69, 70*). It consists of an

α/β fold structure with a central six-stranded antiparallel β-sheet as a scaffold and five short α-helices (*67, 71*) (Fig. 1-1A). This 3D fold is shared by all members of the PAS domain superfamily (*54, 70*). Unlike the multidomain nature of most other PAS domain containing proteins, PYP can function by itself as a single domain.

1.5 Structure of PYP, PYP-*p*CA interactions, and protein stability

PYP consists of two parts, the N-terminal region (residues 1-28) and the PAS fold (residues 29-125). In addition, in the literature the structure of PYP is sometimes described as consisting of four parts: i) the N-terminal cap (residues 1-28), ii) PAS core, the first three strands of the β sheet (residues 29-69), iii) the helical connector (residues 70-87), and iv) the last three strands of PYP, the β scaffold (residues 88-125) (*54*) (Fig. 1-1A). The basic structure of PYP comprises of two hydrophobic cores, one on each side of the central β-sheet (*67*). The N-terminal cap forms the small hydrophobic core and the PAS fold that comprises the *p*CA chromophore binding pocket forms the large hydrophobic core (*72*). The crystal structure provides detailed information on the *p*CA chromophore and the side chains interacting with it. In the ground state of PYP, the *p*CA chromophore is ionized (deprotonated) and is in the *trans* conformation (*57, 73*). The side chains of Glu46 and Tyr42 form hydrogen bonds to the phenolic oxygen of the *p*CA thereby stabilizing the buried negative charge on *p*CA (*67*). Thus, the active site of PYP consists of the hydrogen bonding network formed between *p*CA and Glu46 and Tyr42 with an additional hydrogen bond between Thr50 and Tyr42 (*67*) and an additional stabilization that may be provided by the positive charge on Arg52 (*67, 74*) (Fig. 1-1B, C). In addition, a hydrogen bond exists between the C=O group of *p*CA and the NH backbone of Cys69.

The properties of *p*CA (nonpeptide ligand) are tuned in PYP for specific biochemical purposes. The color (absorbance maximum, λ_{max}) is tuned to make the protein sensitive to a specific wavelength regime. In Hhal PYP, the *p*CA in dark shows λ_{max} of 446 nm in the visible region of the

spectrum. In the dark state, the *p*CA in PYP that has one phenolic oxygen and one isomerizable double bond is deprotonated and in the *trans* state (*57, 73, 75*). The thioester linkage of the chromophore in native PYP is more stable than what is observed in aqueous solutions and is cleaved by hydroxylamine, dithiothreitol, performic acid and high pH (*75-77*). The pK_a of *p*CA is tuned to enable the proton transfer switch between the chromophore and Glu46, which is coupled to the global conformational changes in PYP upon the formation of the presumed signaling state, pB during the PYP photocycle.

The stability of native PYP has been studied in the past by probing the folding of PYP in the native state using common denaturants like Gdm-HCl. Denaturants cause unfolding of the protein, exposing its interior to the solvent resulting in protonation of *p*CA that causes a strong blue shift in the λ_{max} of PYP, which can be used to probe the transition of PYP from fully folded to unfolded state (*5, 56, 57, 78*). The nature of this transition has been reported to be two-state (no folding intermediates), described by a midpoint of the equilibrium titration at 2.72 M Gdm-HCl (unfolding free energy is zero), whereas the unfolding free energy in water for native PYP is 37 ± 1 kJ/mol (*5*).

1.6 PYP photocycle and related molecular events

The overall photocycle and kinetics of PYP is similar to that of the sensory rhodopsins from halobacteria (*79*) and also shows some degree of resemblance to the photocycles of bacteriorhodopsin and halorhodopsin (*55*). When illuminated, Hhal PYP interacts with light with the help of its *p*CA chromophore. Absorption of a blue photon by PYP causes *p*CA trans-cis isomerization (*80*) that triggers a complex photocycle (Fig 1-2) (*56, 78*) with several photocycle intermediates. The key intermediates are: pG dark state – initial receptor state of the photocycle, pR state (ps–ns time scale) - short lived red-shifted state formed immediately after chromophore photoisomerization, and pB state (ms time scale) – photoactivated long lived blue shifted state, presumed to be the signaling state in PYP that has a downstream effect on cell motility. The process of pB formation induces the pB' state,

which is formed upon the occurrence of an intramolecular proton transfer from Glu46 to *p*CA (*32*). This event eliminates the charge on the chromophore through protonation and creates a new buried negative charge on Glu46 (*33*) and within 2 ms the pB' state is converted to the pB state. The proton transfer event and the new charge located in the hydrophobic environment creates instability that leads to large conformational changes in the pB state compared to pG state (*5, 33, 81*). At neutral pH, the pB state decays back to the initial pG dark state in approximately 350 ms. Based on several reports it is believed that the pB state is partially unfolded (*5, 81-87*). The pB decay state or the recovery of the pG ground state of PYP involves three important events; the refolding of the protein to its native state, thermal *p*CA reisomerization (cis → trans), and the proton transfer from *p*CA chromophore back to Glu46.

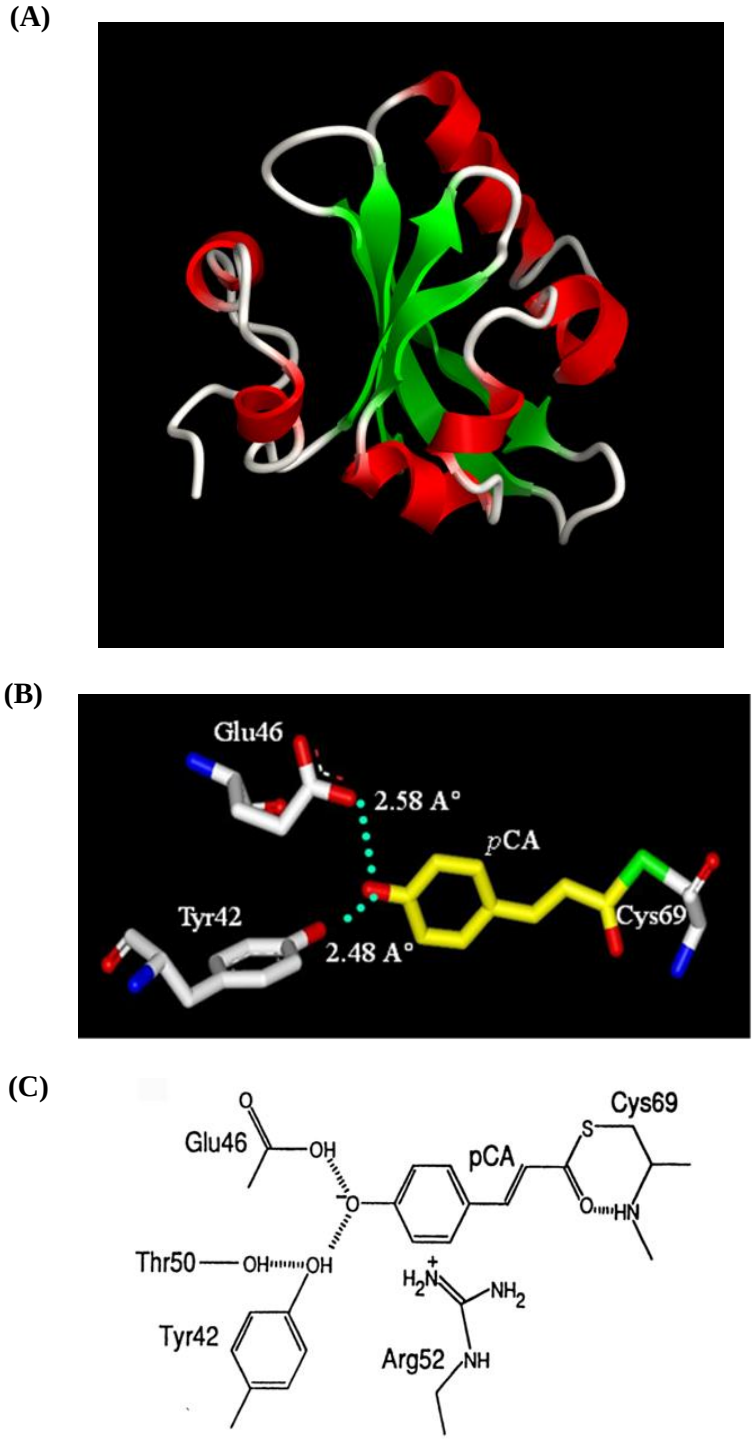


Fig. 1-1 Structure of PYP and its chromophore. (A) Schematic representation of the crystal structure of PYP (α–helices in red, β–sheets in green, and loops in white). (B) Active site of PYP with Glu46 and Tyr42 hydrogen bonded to *p*CA and the respective hydrogen bond distances are shown. (C) Schematic model of the protein-chromophore interactions in PYP.

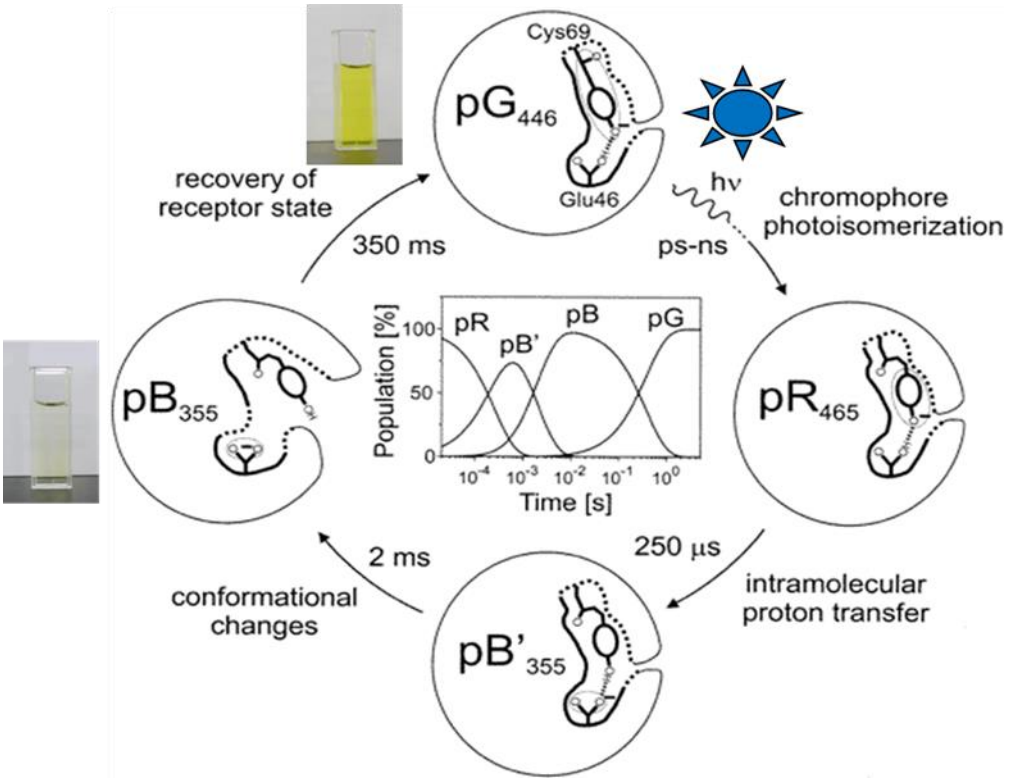


Fig. 1-2 Adapted from (*33*) A model of the PYP photocycle, based on transient UV visible absorbance spectroscopy. The time constants for the various photocycle transitions, two main photocycle intermediates (pR and pB) and the initial pG state of PYP are indicated. The subscripts indicate the absorbance maximum of the three main states.

1.7 PYP family of photoreceptors and the diversity in the biologic function

Studies on various halophilic and phototropic bacteria and bacterial genome/metagenome projects have revealed the presence of photosensory proteins in a large number of bacteria including diverse chemotrophs. This unexpected abundance of photosensory proteins is involved in triggering a range of different photobiological responses, but in many cases their biological function is unknown (*2, 63, 88-90*). These studies have resulted in the strong growth of all six families of photosensory proteins, including photoactive yellow protein. Over the past six years, the number of PYPs or PYP related genes identified has increased from 14 to 140 (*2, 63*) with representatives in three different bacterial Phyla: Proteobacteria, Bacteroides, and Spirochetes. The large majority of these novel PYPs remain uncharacterized (*63*).

An extensive *in vitro* biophysical characterization of different functional properties of PYP has been reported. Three main biophysical properties have been reported for 10 different PYPs: i) spectral tuning of the absorbance maximum (λ_{max}) of pG dark state of PYP (*74*, *91-94*), ii) the p*K*a of *p*CA (*95-98*), and iii) the lifetime of the pB intermediate (presumed signaling state) (Table 1-1).

Organism (w/PYP)	λ_{max} (nm)	pB lifetime (s)	pK_a	pI	Genetic tools	Function reported
Thermochromatium tepidum	**465/358**	**240**	**10.2**	**10.2**	**NO**	**NO**
Idiomarina loihiensis	**447**	**0.23**	**~3.4**	**4.8**	**NO**	**YES**
Rhodothalassium salexigens	**446**	**0.5**	**ND**	**4.4**	**NO**	**NO**
Halorhodospira halophila 1	**446**	**0.5**	**2.8**	**4.8**	**NO**	**YES**
Halochromatium salexigens	**446**	**0.5**	**ND**	**5.3**	**NO**	**NO**
Rhodobacter sphaeroides	**446/360**	**0.002**	**3.8/6.5**	**9.4**	**YES**	**NO**
Halorhodospira halophila 2	**440**	**55**	**ND**	**5.3**	**NO**	**NO**
Rhodobacter capsulatus	**435/375**	**0.001**	**~6**	**9.4**	**YES**	**NO**
Rhodospirillum centenum	**434**	**50**	**ND**	**5.8**	**YES**	**YES**
Salinibacter ruber	**432**	**>60min**	**<3**	**4.0**	**NO**	**NO**

Table 1-1 Adapted from (*2*), Key molecular properties of 10 PYPs reported in literature.

These results revealed a vast difference in key molecular properties within the PYP family; the λ_{max} ranges from 432 nm (*S. ruber*) (*2*) to 465 nm (*T. tepidum*) (*99*), pK_a of the *p*CA ranges from pH 2.8 (*H. halophila*) (*55*, *95*) to 10.2 (*T. tepidum)* (*99*) and pB lifetime varies from 1 ms (*Rb sphaeroides*) (*100*) to >1 h (*S. ruber*) (*2*). *H. halophila* is known to be negatively phototactic with a wavelength dependence matching the absorbance spectrum of PYP (*48*) and this bacterial taxis

response at sub-second scale corresponds well with pB lifetime (~0.5 s) in Hhal PYP. However, this proposed function has never been confirmed *in vivo* due to lack of genetic tools for *H. halophila.* Studies on the PYP from the purple photosynthetic bacterium *Rhodospirillum centenum* are of growing interest since it is the only PYP containing bacteria that is genetically accessible and has a known function which is the blue light regulation of chalcone synthase gene expression upon its light-induced autophosphorylation (*101*). The longer lifetime of pB (50 s) of *R. centenum* PYP correlates with its role in gene expression. It is evident from these data that different PYPs from different bacteria tune these properties, presumably to match their physiological role in the cell.

1.8 Current Research – *Specific aims*

In this study, the blue light photoreceptor PYP from two different bacteria (*H. halophila* and *R. centenum)* that share a highly similar crystal structure but have substantially different amino acid sequences and functional characteristics (*101*) have been used to probe the structure-function dynamics in an expanding circle of proteins: within *H. halophila* PYP (Aim 1), in the PYP family (Aim 2), and in the PAS domain superfamily (Aim 3). This work is aimed at obtaining a deeper understanding of the role of structural changes included in key molecular events in the PYP photocycle.

Aim 1 (Chapter 3) is to develop and apply spectroscopic tools to identify the structural change that causes the proton transfer event in the PYP photocycle that is responsible for driving the signal transduction mechanism of PYP. The structural changes that drive intramolecular proton transfer in proteins is a central open question in protein biochemistry. The PYP system offers an excellent model system to examine this question. The PYP photocycle involves an intramolecular proton transfer which causes subsequent global protein conformational changes (*32, 33*). However, the structural change leading to this proton transfer event remains elusive. We have developed the hypothesis that the disruption of the *p*CA-Tyr42 hydrogen bond leads to proton transfer from active

site residue Glu46 to the *p*CA chromophore. We developed and used high yield isotope editing of PYP with ring-D_4-Tyrosine. The labeling was confirmed and analyzed by LC-MS/MS. The labeled sample was then used to probe the PYP photocycle by assigning signals to tyrosine in the FTIR spectra in a spectral region that is highly sensitive to Tyr-hydrogen bonding (ongoing work in the Xie lab).

Aim 2 (Chapter 5) is to conduct an extensive in vitro characterization of *R. centenum* PYP mutants to examine structure-function studies in the PYP family**.** Almost all of the extensive mutagenesis work on PYP by multiple research groups has been performed using *H. halophila* PYP. We aimed to determine to what degree the structure – function rules derived for *H. halophila* PYP transfer to *R. centenum* PYP. Based on the highly similar crystal structures of these two proteins and the information available for *H. halophila* PYP point mutants we hypothesized that the structure–function relationships derived for *H. halophila* PYP will apply to *R. centenum* PYP. We aimed to design mutants of Rcen PYP that will provide a powerful tool for future research on the *in vivo* function of PYP. To date, no studies have been performed due to the lack of genetic tools in *H. halophila*. Since *R. centenum* is genetically accessible and contains a PYP with a known *in vivo* function, we selected Rcen PYP for this work. Here we aimed to perform the first step towards achieving this goal through studies on the biophysical properties of mutants of *R. centenum* PYP. We aimed to generate two classes of *R. centenum* PYP mutants:

1) Mutants that are expected to block light-induced partial unfolding: E46Q.
2) Mutants that are expected to alter the lifetime of the pB state of the photocycle: P6F, V28F, N43A, M100A.

Aim 3 (Chapter 4) is to determine the functional role of Ile39 in PYP. The importance of this residue is that it is one of the most conserved residues in the PAS domain superfamily. Protein superfamilies are defined by a weak but characteristic pattern of amino acid conservation. Little information is available regarding the function of these superfamily-conserved residues. We investigate the functional role of these residues for the PAS domain superfamily. PYP is a prototype

of this superfamily, and is revealing functional insights into superfamily-conserved residues. We hypothesized that the PAS-conserved residue Ile39 may be involved in the transmission mechanism of allosteric structural changes in the PAS domain superfamily and in the stability of the PAS domain fold. Our approach was to examine the effect of the I39A mutation on the kinetics and structural changes during the photocycle of *H. halophila* PYP using time-resolved visible and rapid-scan FTIR spectroscopy and to probe the stability of I39A PYP using denaturant titrations.

CHAPTER II

Experimental procedures

2.1 Bacterial strains used and site directed mutagenesis

Mutagenesis was performed using Stratagene's Quikchange site-directed mutagenesis kit with primers (Invitrogen) designed (Table 2-1) to introduce the I39A substitution in *H. halophila* PYP and the P6F, V28F, E46Q, and M100A substitutions in *R. centenum* PYP. For I39A Hhal PYP, wt Hhal PYP in pQE-80L was used as a template. The template for the PCR reactions of Rcen PYP mutants was wt Rcen PYP in the cloning vector pBluescript (SK-) (Novagen) modified in the Hoff lab by deleting the *Kpn*I restriction site and replacing it with 2 new restriction sites, *Nco*I and *Nde*I in the multiple cloning site. The final PCR products were digested with *Dpn*I and then transformed into *E. coli* DH5α. In the case of constructing the N43A mutant of Rcen PYP, due to unexplained complications multiple nucleotide repeats were observed in the PCR product when the above method using QuikChange kit was used. Therefore an alternative approach was used to obtain the N43A mutant of Rcen PYP. Two different PCR reactions were carried out with i) N43A FW and wt Rcen PYP RV ii) wt Rcen PYP FW and N43A RV and for the last PCR reaction the PCR products from i) and ii) were used as primers and the final product was ligated in pBluescript (SK-). The DNA from all reactions was purified using a QIAPrep Spin Miniprep kit (QIAGEN) and the mutations in all cloned mutant *pyp* genes were confirmed by DNA sequencing in the OSU DNA core facility.

Each mutant DNA was cloned into the respective expression vectors listed in Table 2-2, which was transformed into *E. coli* BL21 (DE3) cells grown on LB agar plates in the presence of the antibiotics corresponding to the expression vector used (Table 2-2).

2.2 Protein expression and reconstitution

The transformed *E. coli* cells expressing the recombinant *apo*PYP variants were grown overnight (12 – 14 h) in either Luria Broth medium (in most cases) or minimal medium (for isotopic labeling of Tyr-D_4 – see chapter 3) with the appropriate antibiotic (Table 2-2) in a shaker at 37°C. Expression of *apo*PYP in the overgrown *E. coli* cells was induced by the addition of 1mM IPTG (Sigma-Aldrich). After 4 h, the cells were harvested at 4000 rpm for 15 min and the *E. coli* cells were resuspended in Tris-HCl buffer (pH 7.5). The cells were lysed and the protein was liberated by sonication on ice for 30 seconds on and off for ~10 times, and reconstituted with *p*-hydroxycinnamic anhydride to form *holo*PYP according to the procedure described in (*77*). This step was followed by centrifugation of the samples at 20,000 rpm for 20 min after which the pellet with the lysed cell debris was separated and the supernatant containing PYP was stored at 4°C. For wt Hhal PYP, after harvesting the cells they were resuspended in 8M Urea (100 mL/L of *E. coli*) and incubated for 30 min on ice causing lysis of the cells. These cells were centrifuged for 20 min at 15,000 rpm and the supernatant was diluted with an equal volume (100 mL) of Tris-HCl at pH 7.5 followed by the addition 100 µL/1L culture of the *p*-hydroxycinnamic anhydride as described in (*102*). The *holo*PYP was dialyzed overnight against 10 mM Tris-HCl, pH 7.5 at 4°C in dark. The urea extraction procedure was not used for the other PYP samples to avoid the risk of excessive unfolding of the samples.

PYP mutant	Primer sequence
I39A Hhal PYP	Primer design used from ref (*1*)
wt Rcen PYP	FW - 5'-ggcagccatatgccggaccggaccaccga-3' RV - 5'-agctcgaattctcacagcttgcgcacgaa-3'
P6F Rcen PYP	FW - 5'-caccgacgatttcggcttcttcaccgagcagatc-3' RV - 5'-gatctgctcggtgaagaagccgaaatcgtcggtg-3'
V28F Rcen PYP	FW - 5'-gatgccttgccgttcggcgccatcc-3' RV - 5'-ggatggcgccgaacggcaaggcatc-3'
N43A Rcen PYP	FW - 5'-ggtcatccaccgttacgccaggacggaaagccgg-3' RV - 5'-ccggctttccgtcctggcgtaacggtggatgacc-3'
E46Q Rcen PYP	FW – gtcatccaccgttacaacaggacgcaaagccggctcag RV – ctgagccggctttgcgtcctgttgtaacggtggatgac
M100A Rcen PYP	FW - 5'-tcttcgatttccaggcggcgccggtgcggg-3' RV - 5'-cccgcaccggcgccgcctggaaatcgaaga-3'

Table 2-1 List of mutagenic primer pairs used in this study - for the introduction of desired point mutation in both *H. halophila* PYP and *R. centenum* PYP.

PYP mutant	Plasmid	Restriction site	Antibiotic resistance (50µg/mL)
wt Hhal PYP	pQE-80L	*Bam*HI – *Hind*III	Amp
wt Hhal PYP, Tyr-D_4	pET16b	*Nco*I – *Bam*HI	Amp
I39A Hhal PYP	pQE-80L	*Bam*HI – *Hind*III	Amp
N43A Hhal PYP (*103*)	pQE-80A	*Bam*HI – *Hind*III	Amp
wt Rcen PYP	pET28A	*Nde*I – *Eco*RI	Km
P6F Rcen PYP	pET16b	*Nde*I – *Bam*HI	Amp
V28F Rcen PYP	pET16b	*Nde*I – *Bam*HI	Amp
N43A Rcen PYP	pET16b	*Nde*I – *Bam*HI	Amp
E46Q Rcen PYP	pET16b	*Nde*I – *Bam*HI	Amp
M100A Rcen PYP	pET16b	*Nde*I – *Bam*HI	Amp

Table 2-2 List of all expression vectors used in this study. Columns describe the different mutant PYPs they harbor within the two restriction sites and their respective antibiotic resistant marker used for selection on LB media.

2.3 Protein purification

All protein samples were purified by Ni-NTA (MCLAB) affinity chromatography and anion exchange chromatography with a DEAE-Sepharose CL6B column (GE Healthcare). Both columns were pre-equilibrated with 10mM Tris-HCl, pH 7.5. For Rcen PYP the purification method was substantially modified compared to published methods for both Hhal PYP and Rcen PYP.

Hhal PYP: Hhal PYP was purified by Ni-NTA wherein the protein was eluted in fractions with 200 mM imidazole and anion exchange chromatography wherein the protein samples were eluted with 100 mM NaCl. The chromatography steps were repeated until the optical purity index, PI (Abs_{278}/Abs_{446}) of PYP reached approximately 0.42 for wt Hhal PYP and approximately 0.44 for I39AHhal PYP. Size exclusion chromatography was used only for purification of Tyr-D_4 labeled Hhal PYP (methods explained in detail in Chapter 3). The N-terminal 6X His tag was cleaved by treatment with *Enterokinase* (Genscript) at room temperature for 14-16 h to obtain I39A PYP samples without the histidine tag.

Rcen PYP: wt Rcen PYP and its mutants were first run through a Ni-NTA column and eluted with 200 mM imidazole. The samples were then run through a DEAE-Sepharose column. Since the first run through the anion exchange column showed low affinity (of the total amount of protein applied to the column only 20-30% was bound) for Rcen PYP, the flow through fractions were collected and the remaining protein bound to the column was eluted by a gradient of NaCl, 100 – 200 mM. Alternate runs through Ni-NTA column and the anion exchange column were performed using concentrated protein sample. The resulting protein samples were concentrated using an ultrafiltration membrane (Amicon, Centriprep-10). It was observed that as the purity of the PYPs increased, the binding affinity of the protein for the DEAE-Sepharose column became weaker, such that ~90-95% of the PYP could be eluted with Tris-HCl buffer, pH 7.5 followed by 50 -100 mM NaCl to elute any bound fraction. The process was repeated several times and all Rcen PYPs were purified to obtain a PI $\leq$ 0.40. The purity of the samples was confirmed by SDS PAGE.

2.4 SDS PAGE analysis of the purity of Rcen PYP samples

SDS PAGE (BIORAD) was performed using a 4% stacking gel and a 12% resolving gel. A pre-stained protein marker (BIORAD) was used and the samples were prepared as described in

the instructor's manual. Increasing concentrations of the proteins (wt and E46Q) were applied to the gel (1µg, 2µg, 5µg, 10µg, and 50µg). The gel was run as per the instructor's protocol, followed by staining the gel in Coomassie Brilliant blue and destaining in destaining buffer. The gels were visualized using an Alphaimager (Alpha Innotech Corp).

2.5 Optical spectroscopy

UV/Vis absorbance spectra were measured at room temperature using an HP-8453 (Hewlett-Packard) diode array spectrophotometer using a broad range of buffer conditions. The pH adjustment for all buffers used in this study was done using 5N NaOH and HCl. A 150 W halogen quartz visible light source (Cuda) with a broadband blue filter (band-pass filter 59855, Oriel) was used to illuminate the PYP samples in buffer and photocycle kinetics were monitored at pH 7.5. The absorbance maxima (λ_{max}) of the *p*CA chromophore in wt Hhal PYP and wt Rcen PYP were measured in dark and the λ_{max} of the mutants was compared with the respective wt PYP.

2.5.1 Molar extinction coefficient

The relative value of the molar extinction coefficient of Hhal PYP, Rcen PYP and their mutants was determined by measuring the effect of denaturation of the protein samples in 10 mM Tris-HCl, pH 7.5 with 2% SDS on the absorbance spectrum (*96*). The extinction coefficient of I39A Hhal PYP and wt Rcen PYP was derived by comparing the amplitude of the absorbance band at 345 nm of the denatured proteins to that of native and SDS-denatured wt Hhal PYP. Native Hhal PYP has a molar extinction coefficient of 45,500 $M^{-1}cm^{-1}$ at its absorbance maximum of 446 nm (*104*). The molar extinction coefficient for Rcen PYP mutants was calculated using the same approach.

2.5.2 Photocycle kinetics

PYP samples were illuminated with blue light until they were photobleached to their steady state photo-equilibria, after which the actinic light was manually shuttered. The thermal recovery of the pG dark state was measured with the UV/Vis spectrophotometer. The photocycle kinetics at 446 nm for wt Hhal PYP and I39AHhal PYP, at 434 nm for wt Rcen PYP, P6F, V28F, N43A, M100A, and at 451 nm for E46Q mutants of Rcen PYP was measured with an appropriate time resolution depending on the lifetime of the pB state for each PYP sample, allowing sufficient time to observe the entire recovery process of the respective mutants. The kinetics was described as a monoexponential decay for all PYPs measured in this study. The pH dependence of the kinetics of pB decay was measured for all PYPs using a mixed buffer (glycine, succinate, MES, and MOPS, 100 mM each) for a pH range 3.5 to 11 for Hhal PYP and pH 3.4 to 12 for Rcen PYP. The effects of the mutations on the rate of photocycles for both Hhal PYP and Rcen PYP were determined as described in Chapter 4 and 5.

2.5.3 pH titrations

For Hhal PYP: Acid titration for both wt PYP and I39APYP was performed to determine the effect of the mutation on the pK_a of the *p*CA. PYP samples were prepared in the dark in 10 mM Tris-HCl, pH 7.5 and the samples were gradually titrated with increasing amounts of HCl in the same cuvette in a dark room with the minimal intensity of red light required for sample handling to avoid PYP photoexcitation by ambient light. The absorbance at 446 nm and the pH of the sample was measured following each addition of HCl. The acid titration was continued until the protein was completely unfolded (*p*CA was protonated) as detected by the complete disappearance of the native absorbance band near 445 nm. The acid denaturation curves were measured using the HP-8453 spectrophotometer.

For Rcen PYP: The acid titration method used for wt Rcen PYP and its mutants was modified because of protein aggregation and resulting high light scattering when the above method was used. For all Rcen PYP samples, an equal concentration of the protein sample was mixed with an appropriate pH buffer and the absorbance spectrum was measured immediately following mixing of the PYP sample in the buffer with the desired pH value. pH buffers ranging from neutral pH to very low pH were used until the protein was completely acid-denatured. To correct small fluctuations in protein concentration in samples at different pH values due to pipetting precisions, after every measurement the sample was mixed with 2% SDS to normalize the molar concentration of the protein used through the entire titration.

In addition, base titrations were performed for wt Rcen PYP and its E46Q mutant. Both titrations were performed in a continuous fashion in a single cuvette where increasing concentrations of NaOH was added to the protein sample prepared in dark in 10 mM Tris-HCl, pH 7.5. The absorbance spectra were measured each time NaOH was added to the sample, and the titration was continued until complete alkaline denaturation of the protein sample was observed. Base titrations were performed to determine the pK_a of the active site residue involved in the deprotonation process.

2.5.4 Gdm-HCl-denaturant titrations

Protein stability was measured by titration of PYP in 10 mM Tris-HCl, pH 7.5 with chemical denaturants using a stock of 6M and 7M Gdm-HCl for Hhal PYP and Rcen PYP, respectively. All measurements were taken under dark conditions to minimize interference due to photoexcitation of PYP.

For Hhal PYP: Starting with 0 mM Gdm-HCl, increasing concentrations of the denaturant were added to the PYP sample, and the unfolding of the protein was monitored by the loss of visible color recorded with the diode array spectrophotometer. Titration experiments were

ended when PYP was fully unfolded as marked by the complete loss of absorbance of the native pG state at 446 nm. The nature of the transition from the fully folded state to the fully unfolded state was monitored by checking for isosbestic points in the spectra.

For Rcen PYP: As observed for acid titrations, Gdm-HCl titration of wt Rcen PYP and its mutants were found to be complicated by protein aggregation and light scattering when the titration was carried out using a single sample in the same cuvette. Therefore, individual protein samples in 10 mM Tris-HCl, pH 7.5 were prepared in different test tubes with increasing concentrations of Gdm-HCl. The rest of the procedure was performed as described above for the acid titrations of wt Rcen PYP.

2.5.5 Ninety-six well experiments for alkaline stability

To examine the behavior of the pG dark state of wt Rcen PYP and its E46Q mutant in high pH, a mixed buffer (CPSB: CAPS, Piperidine, Sodium dibasic phosphate, and Boric acid, 50 mM each) was used. All sample preparation was done in a dark room and the buffer was mixed with the PYP sample just before the measurement to minimize pG decay through thioester hydrolysis. Both wt Rcen PYP and its E46Q mutant were mixed with a range of pH buffers, from 10 – 14. The samples were mixed in 96-well plates and immediately inserted into a Synergy HT plate reader (Bio-Tek) for single-wavelength measurements at the absorbance maximum of *p*CA in the native states of the respective PYPs with a time resolution of 60 s for 1h in dark. Using these data the rate of thermal pG decay at high pH was calculated.

2.5.6 High pH jump measurements

High pH jump measurements were performed with purified wt Rcen PYP using an HP-8453 diode array spectrophotometer. Two different pH jumps were performed, from pH 10 to 13.86 and from pH 10 to 14. 50 mM CPSB buffer was used for sample preparation. For high pH measurements, highly concentrated PYP sample was mixed with pH 10 buffer and the absorbance

spectra was measured immediately. Similarly, an equal amount of PYP was mixed with pH 13.86 and 14 buffers and absorbance spectra were measured immediately and monitored for 360 s with a time resolution of 12 s. The resulting data were used to determine the of the rate of the pG decay state at different pH values.

2.6 Data Analysis

2.6.1 pK_a fits

The acid/base titration curves were fit with a sigmoidal relationship described by from the Henderson – Hasselbalch equation. The equation was modified to include two sloping baselines and an *n* value to describe deviations from 1 in the steepness of the transition (*105*). The parameters *a* and *b* are the *y*-intercept and slope of each baseline. The subscripts *u* and *l* denote the terms for the upper and lower baselines respectively.

$$\text{Abs} = [\,(a_u + b_u \times \text{pH}) + (a_l + b_l \times \text{pH})10^{n(\text{p}K - \text{pH})}\,]/\,(1 + 10^{n(\text{p}K - \text{pH})}\,] \qquad \text{Equation 1}$$

For fits of data on the low pH titration of wt Rcen PYP, two dummy points at pH -1 and -2 were added to ensure that the fitted curve reached zero percent native species at very low pH. The pK_a value describes the midpoint of the transition, where half of the sample is ionized. The *n* value describes the slope of the pH-induced transition. The sloping baselines describe the pH dependence of the absorbance of native and acid/base denatured PYP.

2.6.2 Denaturation titration fits

Denaturant titration data were fit with a sigmoidal relationship (a two state model) with sloping baselines (*4*). These fits are defined by two parameters; the free energy of unfolding or folding stability $\Delta G_u°$ in water, the denaturant *m* value (steepness of the slope – dependence of

ΔG_U on denaturant concentration). In this description the denaturant midpoint concentration equals $\Delta G°_{U.}$

$$y = [a_u + b_u \times [D] + (a_l + b_l \times [D]) \times \exp(\frac{- \Delta G° + m[D])}{RT}] / [1 + \exp(\frac{- \Delta G° + m[D])}{RT}]$$

Equation 2 (*4, 106*)

2.6.3 Difference spectra and exponential fits of photocycle kinetics

Time-resolved light-induced difference spectra were calculated by subtracting the dark state absorbance spectra from the photobleached spectra of respective PYP samples. The pG dark state recovery traces of all PYP samples at different pH values were fit as a monoexponential rise process to derive the rate constant for the process.

$$y = y_0 + A \exp [– (x – x_0) / \tau] \qquad \text{Equation 3}$$

Protocols used for isotope labeling of Hhal PYP (Tyr-D_4), analysis of labeling using mass spectrometry and rapid scan FTIR spectroscopy are explained in Chapter 3. Methods used to determine conformational changes in PYP using rapid scan FTIR spectroscopy and the bioinformatics analyses of Ile39 side chain interactions in PAS domain superfamily and associated sequence alignments are described in detail in chapter 4.

CHAPTER III

Towards understanding the structural changes that cause proton transfer in PYP

3.1 Introduction to proton transfer in proteins

Proton transfer is one of the most fundamental processes in the functioning of many proteins and plays a central role in bioenergetics and in some signal transduction mechanisms. The extensively studied light-induced proton pump bacteriorhodopsin from the archaeon *Halobacterium salinarum (107)* is a model system for understanding proton pumping mechanisms in proteins. It consists of a retinal chromophore that is covalently attached to Lys216 by a protonated Schiff base. In the ground state of the protein, the Schiff base forms a hydrogen bonding network with a water molecule, Asp85, and Asp212 (Fig. 3-1B). When illuminated, the system absorbs a photon, which triggers a *trans* $\rightarrow$ *cis* isomerization of the retinal chromophore followed by a series of proton transfers steps and protein conformational changes as the protein moves through various spectroscopically distinct intermediate states. This directional proton transfer reaction results in a mechanism for transmembrane proton pumping from the cytoplasmic to the extracellular side of the plasma membrane (*38, 108-112*). The first step in the proton transport mechanism is proton transfer from the protonated retinal Schiff base to Asp85.

Though numerous studies have reported on this process, the specific structural features that cause proton transfer in bacteriorhodopsin are still unclear. A recent report (*113*) addressing this issue provided interesting insights on the energetics of this initial proton transfer coupled to protein conformational changes. Using quantum computational methods, the authors suggest that the increased steric energy contained in the chromophore induced by light is relaxed by increasing the distance between the Schiff base and Asp85 ion pair, causing a charge separation. The energetics of this destabilized ion pair is proposed to be the trigger for proton transfer (*113, 114*). Despite extensive experimental and computational analyses, it remains undetermined what precise structural change/changes drives a proton from the donor to the acceptor even in this highly studied protein system, making the structural basis for proton transfer reaction in proteins a central open problem in the field of protein biochemistry and biophysics.

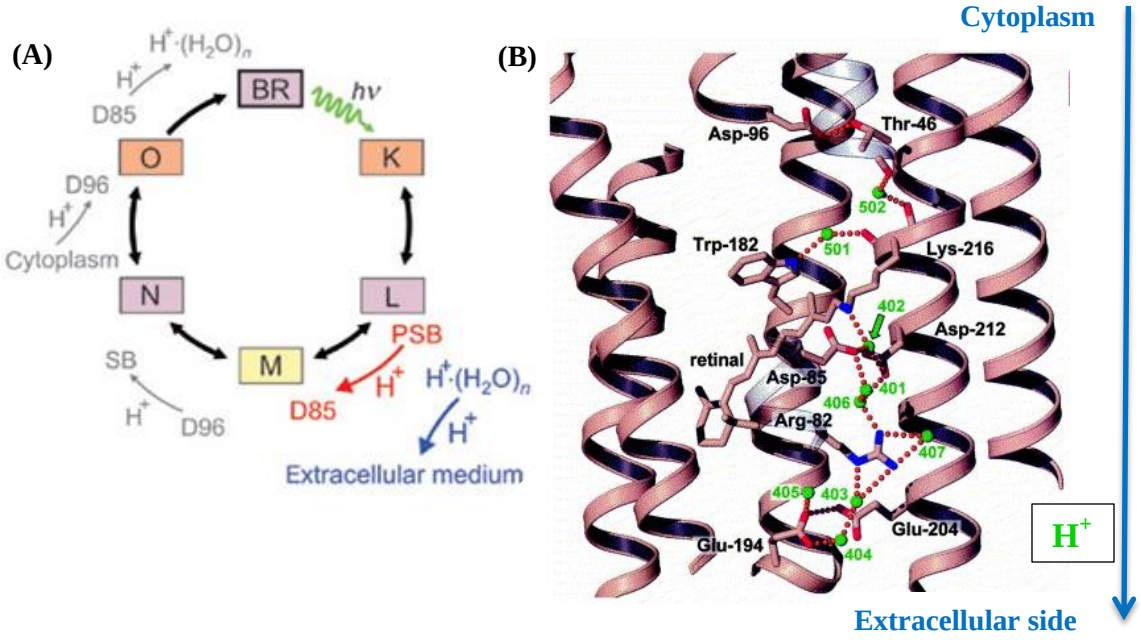


Fig. 3-1 Proton transfer in bacteriorhodopsin (A) Taken from (*115*) Schematic model of the photocycle showing the sequential steps of light proton transport in bacteriorhodopsin; PSB – protonated Schiff base. (B) taken from (*116*). The model of the BR (ground) state, showing the hydrogen bond network in the active site between the water molecule (402), protonated Schiff base, Asp85, and Asp 212 and the detailed display of communication between the cytoplasmic side and the extracellular surface via a string of hydrogen and covalent bonds .

Directional proton transfer reactions are not just restricted to proton pumps but are also commonly observed in enzyme catalyzed reactions where strategically oriented hydrogen bonds are responsible for driving the protons in a specific direction (*117, 118*). In addition, a number a photosensory proteins, including sensory rhodopsins, phytochromes, flavin-based light, oxygen, or voltage (LOV) domains, and blue light using FAD (BLUF) domains exhibit some form of light induced proton transfer during their light-induced functional process that translate light into biological information. We are using the photoactive yellow protein from *Halorhodospira halophila* as a model system to understand the mechanism of proton transfer reaction in proteins.

The basic steps in the PYP photocycle have been discussed in Chapter 1. In the initial dark state of PYP the anionic *p*CA is hydrogen bonded to the neutral side chains of Tyr42 and Glu46. During the light-triggered photocycle of PYP, a direct proton transfer event between two adjacent groups (Glu46 and the *p*CA) occurs, providing a simple and experimentally tractable system to study proton transfer in proteins. The aim of this part of the thesis is to develop tools to experimentally determine the structural change(s) that triggers this intramolecular proton transfer event in the PYP photocycle. In collaboration with the Xie lab at OSU, it was *hypothesized that the disruption of the pCA-Tyr42 hydrogen bond leads to proton transfer from Glu46 to pCA.* To test this hypothesis, the following approach was taken, which uses the concept of vibrational structural markers (*119*) to extract information on side chain hydrogen bonding from FTIR data:

- Microsecond time-resolved FTIR spectroscopy (Xie lab at OSU), which combines sufficient time resolution and sensitivity to Tyr hydrogen bonding to probe the photocycle at stages just before (pR) and just after (pB') proton transfer.
- Tyr-D_4-labeled PYP (mineral medium + Tyr-D_4 + glyphosate) to assign signals in the FTIR spectra to Tyr. Confirmation and analysis of labeling pattern by LC-MS/MS. (*Published: Rathod et al 2012*).

- Expected after starting the photocycle with a laser flash: a single Tyr (Tyr42) will exhibit an FTIR peak shift characteristic of the loss of a hydrogen bond (time constant of 250 μs) (*Ongoing work in the Xie lab, OSU).*

3.2 High yield $TyrD_4$ – labeling of *H. halophila* PYP for infrared structural biology

Abstract

An important bottleneck in the use of infrared spectroscopy as a powerful tool for obtaining detailed information on protein structure is the assignment of vibrational modes to specific amino acid residues. Side-chain specific isotopic labeling is a general approach towards obtaining such assignments. We report a method for high yield isotope editing of the bacterial blue light sensor photoactive yellow protein (PYP) containing ring-D_4-Tyr. PYP was heterologously overproduced in *E. coli* in minimal media containing ring-D_4-Tyr in the presence of glyphosate, which inhibits endogenous biosynthesis of aromatic amino acids (Phe, Trp, and Tyr). Mass spectrometry of the intact protein and of tryptic peptides unambiguously demonstrated highly specific labeling of all five Tyr residues in PYP with 98% incorporation and undetectable isotopic scrambling. FTIR spectroscopy of the protein reveals a characteristic Tyr ring vibrational mode at 1515 cm^{-1} that is shifted to 1436 cm^{-1}, consistent with that from *ab initio* calculations. PYP is a model system for protein structural dynamics and for receptor activation in biological signaling. The results described here open the way to the analysis of PYP using isotope-edited FTIR spectroscopy with side-chain specific labeling.

Introduction

Fourier transform infrared (FTIR) spectroscopy is a promising technique for obtaining highly detailed structural information for protein functional structural dynamics, particularly with respect to hydrogen bonding (*29, 119-121*) and proton transfer (*33, 122-125*). Infrared structural biology is an emerging technology that detects protein structures and structural dynamics using infrared spectroscopy. A variety of time-resolved and temperature resolved infrared spectroscopic techniques have been developed, including rapid-scan FTIR (*40, 88, 126*), step-scan FTIR with microsecond and nanosecond time resolutions, picosecond pump probe infrared spectroscopy (*33, 127-133*), and cryogenic FTIR for cold-trapped intermediate states. A Vibrational Structural Marker (VSM) database library is being developed to translate infrared signals into quantitative structural information (*119*). However, the development of VSM for quantitative structural characterization and applications of VSM for protein structure-function studies have been hampered by the challenges of identification and assignment of key vibrational modes in a protein to a specific side-chain. This important bottleneck limits the application of FTIR spectroscopy as a powerful tool for time-resolved protein structural biology.

Important tools in the assignment of signals to specific side chains are site directed mutagenesis (*36*) and isotope editing (*134-139*). The methods for site-directed mutagenesis are well-established, but are currently limited to the relatively slow process of residue-per-residue attempts to confirm suspected assignments. In the case of isotope editing a strategy for the specific isotopic labeling of residues is needed. One approach is homogeneous ^{13}C or ^{15}N labeling. While this can contribute to band assignment and can help resolve overlap between different signals, it is generally not sufficiently specific to yield unique assignments. Another powerful method is the use of H/D exchange. This can reveal different types of secondary structure (*140*) and changes in solvent exposure during protein function (*83*). However, the large number of exchanging sites in a protein usually precludes the assignment of vibrational signals to

a specific residue. Approaches to isotopically label selected regions of the protein involve the overexpression of proteins in the presence of specific isotopically labeled amino acids (*141*), intein-based segmental labeling (*142-144*), and *in vitro* translation systems (*145, 146*). A final approach is the complete chemical synthesis of the protein under study (*147*). While this is highly useful for the study of small peptides (*148, 149*), it is feasible but technically challenging for entire proteins.

Intein-based techniques, *in vitro* translation, and total chemical synthesis of proteins are powerful but technically demanding approaches. The growth of *E. coli* in defined media in the presence of isotopically labeled amino acids offers an attractive strategy to side-chain specific isotopic labeling. Two possible problems associated with this approach are isotope dilution, in which endogenous synthesis of the labeled amino acid that is added to the growth medium results in a reduced level of isotopic labeling of the side chain under study, and isotope scrambling, where the cellular amino acid metabolism will cause labeling of other side chains. Such problems can be reduced by the use of auxotrophic *E. coli* strains that are unable to synthesize specific amino acids (*138, 141*) or the use of inhibitors that block specific pathways for amino acid biosynthesis. In the case of the biosynthesis of aromatic amino acids the inhibitor glyphosate is available (*150*).

Recent developments in technology have made mass spectrometry a highly attractive tool to determine the degree to which problems in side-chain specific isotopic labeling occur. Significant interest exists in using isotopic labeling to aid in the interpretation of FTIR signals of proteins. The use of mass spectrometry to quantitatively determine the pattern of isotope incorporation has recently been initiated (*138*). Here we use mass spectrometry to show that specific and high-level Tyr labeling is readily achieved by overexpression of the protein under study in *E. coli*, grown in defined media in the presence of glyphosate. These studies were performed using photoactive yellow protein (PYP), a bacterial photoreceptor (*48, 55, 56*) and

model system for functional protein dynamics (*2, 53, 64*). The results indicate that FTIR spectroscopic studies of PYP (*32, 33, 151, 152*) can greatly benefit from side-chain specific isotopic labeling.

Materials and methods

Chemicals

Ring-D_4 L-tyrosine (DLM-451-1 with 98 % isotope enrichment) was purchased from Cambridge Isotope Laboratories. All other amino acids and chemicals were from Sigma-Aldrich. DEAE Sepharose^{TM} Fast flow for Anion Exchange chromatography and Sephadex^{TM} G-50 Superfine for size exclusion chromatography were purchased from GE Healthcare. Ring-D_4 L-tyrosine will be abbreviated as Tyr-D_4 throughout the text below.

Expression strain and plasmid

E. coli strain BL21 (DE3) and the plasmid pET-16b were obtained from New England Biolabs and Novagen respectively and the plasmid was constructed as reported previously (*153*).

Composition of the minimal medium

Growth media for specific Tyr-D_4 isotopic labeling were prepared based on (*154*) with modifications as described in (*141*). First, a 10X stock of M9 salts plus all 20 amino acids (including either Tyr or Tyr-D_4) was prepared as described in Table 3-1. Following the addition of amino acids, the medium was autoclaved for 20 minutes. The remaining ingredients (Table 3-2) were filter sterilized (except $MgSO_4$, which was autoclaved) and were added to the M9 medium containing the amino acids after it had cooled down to room temperature. Finally, 50 mg/mL ampicillin and 1 g of the aromatic amino acid biosynthesis inhibitor glyphosate (dissolved in 2 mL 10M NaOH) were added just before inoculation of the medium with *E. coli*.

10X stock of M9 salts	
128 g	Na_2HPO_4 x $7H_2O$
30 g	KH_2PO_4
5 g	NaCl
10 g	NH_4Cl
Dilute to 1 Liter.	
The following amino acids were added after 100 mL of 10X M9 salts were diluted to 900 mL.	
For 1 Liter culture	
0.5 g	$_L$-alanine
0.4 g	$_L$-arginine
0.4 g	$_L$-asparagine
0.4 g	$_L$-aspartic acid
0.0725 g	$_L$-cystine-HCl x H_2O
0.4 g	$_L$-glutamine
1.413 g	Na-glutamate x 8.3 H_2O
0.55 g	$_L$-glycine
0.143 g	$_L$-histidine-HCl x H_2O
0.23 g	$_L$-isoleucine
0.23 g	$_L$-leucine
0.42 g	$_L$-lysine-HCl
0.25 g	$_L$-methionine
0.13 g	$_L$-phenylalanine
0.1 g	$_L$-proline
0.21 g	$_L$-serine
0.23 g	$_L$-threonine

0.17 g	$_L$-tryptophan
0.23 g	$_L$-valine
0.050g	$_L$-tyrosine or $_L$-tyrosine-ring-D_4

This growth medium is autoclaved.

Table 3-1 Composition of minimal growth medium for Tyr-D_4 labeling of proteins

For 1 Liter medium			
a.	4 mL		1M $MgSO_4$
b.	1 mL		0.01M $FeCl_3$
c.	50 mL		40% glucose
d.	10 mL of solution containing		
	I	2 mg	$CaCl_2$ x $2H_2O$
	Ii	2 mg	$ZnSO_4$ x $7H_2O$
	Iii	2 mg	$MnSO_4$ x H_2O
	Iv	2 mg	H_3BO_4
	V	2 mg	$CuSO_4$
e.	50 mg	Thiamine (dissolved in 5 mL dH_2O)	
f.	1 mg	Biotin (dissolved in 1 mL dH_2O)	
g.	50 mg	Niacin (dissolved in 1-2 mL dH_2O)	
h.	50 mg	Ampicillin (dissolved in 1 mL dH_2O)	
i.	1 g	Glyphosate (dissolved in 2 mL 10M NaOH)- *added just before inoculation.*	

Table 3-2. Additional chemicals to the minimal medium for Tyr-D_4 labeling of proteins

PYP color check of glycerol stocks of pET16b-E. coli BL21 (DE3)

Before using the isotopically labeled medium, a test culture was done using the natural/ unlabeled $_L$-tyrosine to check the viability and protein expression of the transformed *E. coli*

BL21(DE3) cells maintained as glycerol stocks in -20°C. The cells were inoculated in 10 mL M9 medium with ampicillin and were incubated overnight on a shaker at 37°C. After 16 hours, the expression of *apo*-PYP was induced by 0.01 mL 1M IPTG and was reincubated in the shaker. After 4 hours, the cells were harvested at 3750 rpm for 10 min. after which they were resuspended in 1 mL 8M Urea and stirred for 30 min on ice. The lysed cells were spun down at 15000 rpm for 20 min and the supernatant was diluted with equal volume of 10 mM Tris-HCl, pH 7.5 followed by the addition of 100 µL/1L culture of the anhydride of *p*-coumaric acid, which was synthesized as previously described (*77*). The reaction mixture was incubated for 30 min in 4°C and to check proper reconstitution of PYP with its *p*-coumaric acid chromophore, the mixture was spun down at 15,000 rpm for 5 min and the Abs_{446} was recorded before preparing the isotopically labeled medium.

Purification of Tyr-D_4 PYP

Growth medium containing Tyr-D_4 was inoculated with *E. coli* expressing high levels of PYP as determined by the above color check procedure. After reconstitution with the chromophore, the reaction mixture was dialyzed overnight against 10 mM Tris-HCl, pH 7.5 at 4°C in a dark room. The mixture was then applied to a DEAE Sepharose column prewashed and equilibrated with 10 mM Tris-HCl, pH 7.5. After washing the column with the same buffer, Tyr-D_4 PYP was eluted in fractions in the same buffer with 100 mM NaCl and its optical purity index (PI) (Abs_{278}/Abs_{446}) was measured. When needed, a more refined elution was performed by using a gradient of NaCl (50-200 mM), in which PYP eluted around 80-120 mM NaCl. The purification process was repeated and combined with size exclusion chromatography until a PI of 0.43 was obtained. Before applying the Tyr-D_4 PYP sample on size exclusion column, it was pre-equilibrated with 100 mM NaCl. The sample used for application was concentrated using an ultrafiltration membrane (Amicon, Centriprep-10) to $1/100^{th}$ volume of the column. Tyr-D_4 PYP

was eluted in fractions with 100 mM NaCl and the PI for each was determined. Tyr-D_4 PYP with a PI of 0.43 was used for further analysis.

Mass spectrometric analysis of isotopic labeling on PYP

To determine the MW of intact protein, the protein sample was diluted in 60%:40%:10% acetonitrile/water/methanol mixture containing 0.1% formic acid. Samples were ionized by infusion using a New Objective ion source and metal coated infusion tips (Econotips) from New Objective. The positive ion signals were recorded manually using the FT mass analyzer of an LTQ Orbitrap LX mass spectrometer, using a resolution setting of 7,500 to collect a mass spectrum showing the protein's average mass without resolved isotopes and using a resolution setting of 100,000 to resolve individual isotopes. The highly resolved mass spectra were used to estimate charge states, and these estimates were used for manual deconvolution to calculate charge and MW.

To determine the efficiency of protein labeling, the labeled protein was denatured in 8M urea, reduced with TCEP to disrupt disulfide bonds, alkylated with iodoacetamide to prevent reoxidation of Cys residues, and digested overnight with 8 µg/ml trypsin, using 100 mM Tris pH 8.5 to buffer all solutions. Trypsinolytic peptides were purified using C18 affinity tips (OMIX), and analyzed by on a hybrid LTQ-Orbitrap mass spectrometer (Thermo Fisher Scientific) coupled to a New Objectives PV-550 nanoelectrospray ion source and an Eksigent NanoLC-2D chromatography system. Peptides were analyzed by trapping on a 2.5 cm pre-column and analytical separation on a 75 µm ID fused silica column, using a vented column configuration packed in house with 10 cm of Magic C18 AQ and terminating with an integral fused silica emitter pulled in house. Peptides were eluted using a 5-40% ACN/0.1% formic acid gradient performed over 40 min at a flow rate of 300 nL/min. During each one-second full-range FT-MS scan (nominal resolution of 60,000 FWHM, 300 to 2000 m/z), the three most intense ions were

analyzed via MS/MS in the linear ion trap. MS/MS settings used a trigger threshold of 8,000 counts, mono-isotopic precursor selection (MIPS), and rejection of parent ions that had unassigned charge states, were previously identified as contaminants on blank gradient runs, or were previously selected for MS/MS (data dependent acquisition using a dynamic exclusion for 150% of the observed chromatographic peak width). Centroided ion masses were extracted using the extract_msn.exe utility from Bioworks 3.3.1 and were used for database searching with Mascot v2.2.04 (Matrix Science) and X! Tandem v2007.01.01.1 (www.thegpm.org).

Searches of the whole SwissProt database (downloaded 04/10/10 and containing 516,081 authentic sequences and an equal number of reversed decoy sequences) were conducted using the following search parameters: 15 ppm parent ion mass tolerance, 0.8 Da fragment ion tolerance, cleavage with trypsin (allowing one missed cleavage), and the variable modifications pyroglutamate modification of N-terminal Q, oxidization of M, alkylation of C with iodoacetamide, and formylation or acetylation of the proteins' N termini.

Peptide and protein identifications were validated using Scaffold v2.2.00 (Proteome Software) and the Peptide Prophet algorithm (*155*). Probability thresholds were greater than 95% probability for protein identifications, based upon at least 2 peptides identified with 80% certainty, providing an experiment-specific protein false discovery rate of 0.1%. Proteins that contained similar peptides and could not be differentiated based on MS/MS analysis alone were grouped to satisfy the principles of parsimony.

Using the m/z ratios of the peptides thus identified, chromatograms for each PYP-derived peptide ion were manually extracted, smoothed, and peak areas calculated using Xcalibur. Peak areas were compared for the Tyr-D_4 labeled peptides with the corresponding unlabeled peptides to calculate the labeling percentage.

FTIR spectroscopy

Tyr-D_4 PYP and Hh PYP samples for FTIR spectroscopy were prepared by concentrating and then washing with 50 mM phosphate buffer in D_2O, using a Microcon (YM-10, Millipore) centrifuge filter. As is commonly done for setting the pH of protein samples in D_2O for NMR spectroscopy, the pH of the phosphate buffer in D_2O used here for FTIR spectroscopy was set at 7.1 using a pH electrode that was calibrated using pH standard solutions in H_2O, referred to as a pH* 7.1. Each FTIR sample consisted of 2.7 μl of PYP at 8mM protein concentration sandwiched between two 15 mm diameter BaF_2 windows and separated using a 12 μm spacer. A pair of 15 mm diameter BaF_2 windows without any sample was used as reference. The temperature of samples during the IR measurements was maintained at 300 K using a water circulation system (NESLAB, RTE-111). PYP samples were measured at steady state with a Bruker IFS 66v FTIR spectrometer in the range of 4000 – 600 cm^{-1} at 2 cm^{-1} spectral resolution and 40 kHz scanning velocity, using a liquid nitrogen-cooled Mercury Cadmium Telluride (MCT) detector. During the measurement, the water vapor in the optic path was vacuumed out, and the sample chamber was purging with dry nitrogen gas (2.2 liter per minute). The absorption data was averaged over 160 scans. The 2^{nd} derivative of absorbance was calculated using Bruker OPUS software (version 5.5), based on Savitzky-Golay algorithm with 13 smoothing points.

Ab initio vibrational calculations

A 4-propylphenol molecule forming two hydrogen bonds as acceptor and donor with 2 isopropyl alcohol molecules was employed to model the side chain of Tyr residues in protein. The preliminary structural optimization of the input structure was performed usingPM3 method, a semi-empirical method. Density functional theory, an *ab initio* method, was employed for highly accurate structural optimization and vibrational frequency calculation using B3LYP/6-31G(d). All structural and frequency calculations were carried out using Gaussian03 (*156*). A scaling

factor of 0.9668 was used to compensate the overestimated force constants in order to calibrate computational and experimental results.

3.3 Results and discussion

3.3.1. Overexpression of PYP in minimal media

Current overexpression protocols for PYP (*157*) use Luria broth or 2X YT to grow *E. coli* BL21 (DE3) containing the *pyp* gene from *Halorhodospira halophila* in pET or pQE expression vectors. To allow side-chain specific isotopic labeling, growth in defined media is needed. We found that growth in M9 medium supplemented with all 20 amino acids and glyphosate (Table 3-1) yielded approximately 50 mg of PYP per liter culture upon reconstitution and dialysis. Under similar conditions we obtained essentially indistinguishable yields of PYP per liter culture after growth in LB.

While the use of His-tagged PYP for convenient purification by Ni-NTA chromatography has been effectively used, for a range of biophysical measurements it is preferable to use untagged, native PYP. Here we use anion exchange and size exclusion chromatography to obtain untagged Tyr-D_4 PYP of very high purity as indicated by its purity index (ratio of absorbance at 280 nm and 446 nm) of 0.43, which is widely used as the standard for pure *H. halophila* PYP (*157*). Typically 40 to 50% of the PYP observed immediately following sample reconstitution and dialysis was recovered in the final purified protein sample.

To achieve side-chain specific labeling with the Tyr-D_4 derivative, the same conditions were used, but the Tyr in the growth medium was replaced by Tyr-D_4 at 50 mg per liter. The ring-D_4-Tyr used in these experiments was 98% Ring-D_4 labeled. Generally, before performing growth in the presence of isotopically labeled compounds, we performed a check of the quality of the overexpression strain by small scale incubation (color-check procedure; details in Materials

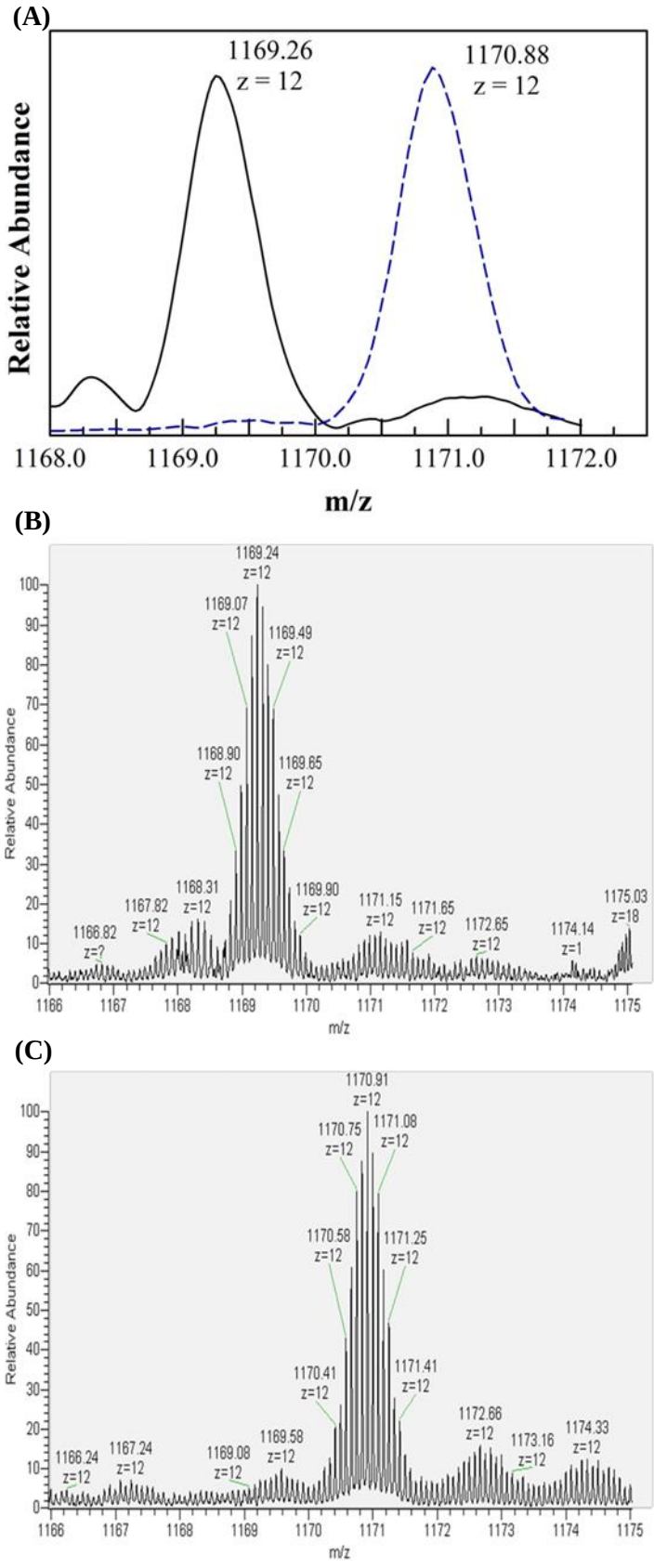
and Methods section). The yield and UV/vis absorption maxima of both labeled and unlabeled PYP reconstituted from *E. coli* grown in either LB or M9 growth medium were comparable.

3.3.2. Mass spectrometric determination of the isotopic labeling pattern.

To determine the labeling efficiency of the five Tyr residues in PYP, and the possibility of isotope scrambling to Phe (or other) side chains, we performed mass spectrometry of full-length unlabeled and Tyr-D_4-labeled PYP. We also analyzed labeling efficiency by measuring the abundance of labeled and unlabeled peptides produced upon trypsin digestion of Tyr-D_4 PYP.

First, we measured the mass spectrum of the unlabeled and Tyr-D_4 labeled samples at lower resolution, in which individual isotope effects were not resolved. The m/z ratios of the five species with +9 to +13 charges yielded experimental molecular weights of 14019.09 ± 0.12 Da for unlabeled PYP. Based on the predicted molecular weight of unlabeled PYP of 14019.12 Da calculated based on its covalent structure, the experimentally observed value is in agreement with the observed mass within 0.04 Da. The experimentally determined mass for Tyr-D_4 PYP was 14038.56 ± 0.06 (Fig. 3-2A). This is 19.44 ± 0.14 heavier than unlabeled PYP. Since the expected increase in mass of the five Tyr-D_4 residues in the labeled PYP is 20.16, this measurement provides an estimate of the labeling percentage (19.44/20.16) of 96.4 ± 0.7 %.

In the isotope-resolved high-resolution spectra (Fig. 3-2B,C) of unlabeled and Tyr-D_4 labeled PYP a series of peaks is detected, which is caused by the natural variation in the number of ^{13}C nitrogen atoms (1.1% natural abundance) in PYP. As a result, series of adjacent peaks that differ by the mass of one neutron are observed in these spectra. The analysis of these data requires that for the unlabeled and Tyr-D_4 labeled samples peaks corresponding to the same level of ^{13}C incorporation are selected. When the central peaks with m/z values of 1169.24 (unlabeled) and 1170.91 (Tyr-D_4) in Fig. 3-2 are selected, this corresponds to an increase in mass of 20.04 Da, very close to the 20.16 Da expected for the increase in mass due to Tyr-D_4 labeling. However,



this selection of corresponding peaks is not fully unambiguous. This complication does not occur in the analysis of the mass spectra measured at lower resolution, as described above.

Fig. 3-2. **Mass spectrometry of Tyr-D_4-labeling of PYP**. (A) Low-resolution mass spectra of intact unlabeled (solid black line) and Tyr-D_4-labeled (dashed blue line) PYP. The observed difference in mass caused by the Ring-D_4 labeling of the five Tyr side chains in PYP is 19.44 Da. Isotope-resolved high-resolution mass spectrum of intact unlabeled (B) and Tyr-D_4-labeled (C) PYP. Comparison of the spectra in B and C reveals a 20.04 Da increase in mass.

More detailed insights into the labeling pattern and labeling efficiency of the Tyr-D_4 PYP sample was obtained through LC-MS/MS analysis of tryptic digests. A total of 12 different trypsinolytic PYP-derived peptides were detected. Together, these peptides covered 91% of all residues in PYP (Fig. 3-3), including all five Tyr residues. All of the residues not detected were present in sequences in which two Lys residues were in close proximity (within 5 residues or fewer); the small size of the resulting peptides can explain their absence in the LC-MS/MS data.

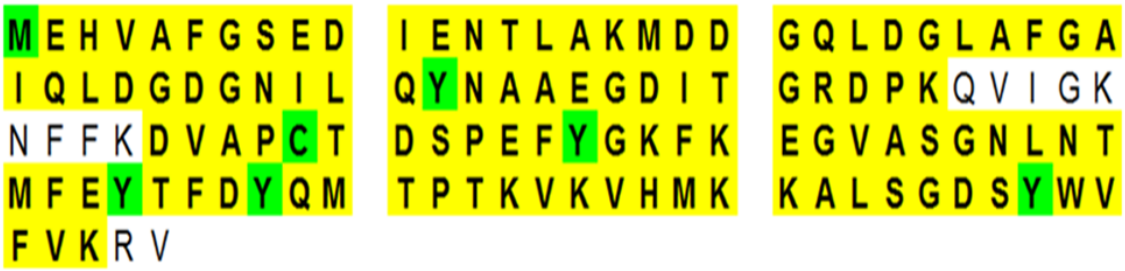


Fig. 3-3 Coverage map of PYP after tryptic digestion. Residues identified in peptides through MS-MS analysis are indicated in color. The modified residues are highlighted in green: M was oxidized, C was carboxyamidomethylated using iodoacetamide, Y was Ring-D_4 labeled. All 5 Tyr side chains of PYP were detected as Ring-D_4 labeled.

Each of the 12 PYP-derived peptides was unambiguously identified by its MS/MS fragmentation pattern (Fig. 3-4), using statistically based proteomics scoring algorithms (see Methods). For Tyr-containing peptides the identification was also inspected manually to confirm that the major ions in the MS/MS spectra represented a contiguous series of b and y fragments derived from the parent peptide. High-quality MS/MS data were obtained, as judged from (i) the high percentage of MS/MS fragments that could be assigned to the trypsinolytic PYP-derived mother peptide, and (ii) the presence of a continuous ladder of fragments (see Fig. 3-4). In these data the majority of the detected were generated through single fragmentation events that cleave the mother peptide into two fragments (b and y ions). Fig. 3-4 provides an example of the mapping of the observed fragments onto the amino acid sequence for the peptide

KALSGDSYWVFVK that contains Tyr118. This analysis unambiguously demonstrated the incorporation of Tyr-D_4 in all five of the Tyr side chains in PYP (Table 3-3).

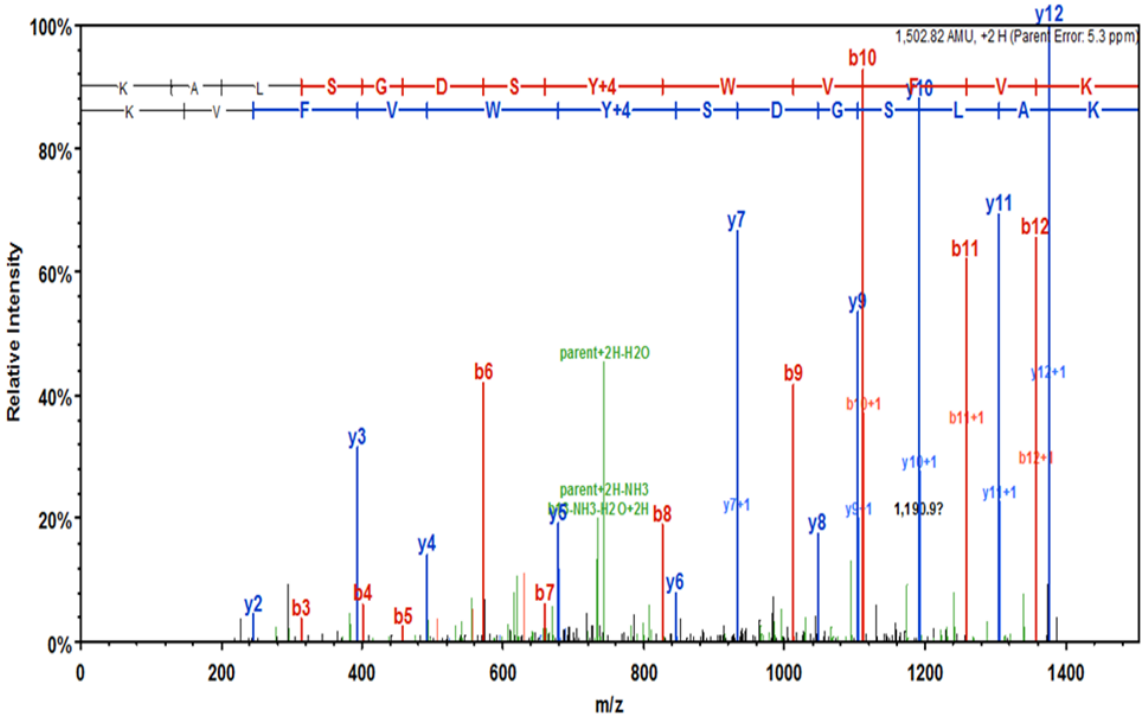


Fig. 3-4 Example of MS/MS identification of Tyr118-ring-D_4 labeled peptide in the tryptic digest of PYP. The fragmentation pattern is annotated to allow identification of the amino acid sequence, with the top row of text representing the peptide sequence in an N- to C-terminal orientation to visualize the b ion fragments containing the N terminus, and the second row of text representing the peptide sequence in the C- to N-terminal orientation to visualize y ion fragments containing the C-terminus. All assigned ions (except when indicated) carry a single charge and are the result of a single fragmentation event at the peptide bond. The vertical bars in the indicated amino acid sequence line up with the m/z of the ion origination from cleavage at that site.

Peptide	Tyr in peptide	D_4Y peptide m/z	Unlabeled peptide m/z	D_4Y peptide peak area	Unlabeled peptides peak area	% unlabel-ed	Charge on the peptide	Mascot Ion Score	X! Tandem Peptide Score
18-52	Tyr42	1209.91	1208.58	1.9×10^5	NM	N/A	3	81.3	6.55
18-55	Tyr42	1323.31	1321.97	2.6×10^5	NM	N/A	3	141.7	6.42
65-78	Tyr76	795.36	793.35	275703418	5720529	2.0	2	79.7	7.37
79-104	Tyr94, Tyr98	1009.82	1007.14	255052	NM	N/A	3	78.6	5.44
81-104	Tyr94, Tyr98	918.1	915.42	182082	NM	N/A	3	33.7	1.72
111-123	Tyr118	752.41	750.40	6055989	146122	2.4	2	96.6	6.96
112-123	Tyr118	688.37	686.35	13806710	327650	2.3	2	82.6	4.89

NM: not measurable. N/A: not applicable

Table 3-3 Summary of LC-MS/MS analysis of Tyr-containing peptides observed in the tryptic digest of PYP obtained after growth in Tyr-D_4 labeling medium.

To further characterize the pattern of incorporation of Tyr-D_4 into PYP, we analyzed two issues that are of general concern for site-specific isotope editing of proteins obtained by *in vivo* expression: isotope dilution resulting in a reduced percentage of labeling of the target sites (in this case Tyr), and isotope scrambling to other side chains. To examine isotope dilution we use the observation that for three different PYP-derived peptides both the unlabeled and the Tyr-D_4 containing species were detected (Table 3-3). While the elution times of the corresponding labeled and unlabeled peptides were essentially identical, they could be distinguished based on their 4 Da increase in mass (Fig. 3-5). Comparison of the areas of chromatographic elution peaks representing each labeled and unlabeled peptide quantitatively revealed very low amounts of unlabeled Tyr-containing peptides in the tryptic digest of Tyr-D_4 labeled PYP. Thus, the high coverage percentage of the amino acid sequence of PYP provides detailed information on the pattern of isotope labeling of PYP. Figure 3-5 depicts the elution profiles for a pair of Tyr-D_4-labeled and unlabeled peptides (for peptide KALSGDSYWVFVK), with the amplitude of the

elution profile of the unlabeled peptide multiplied by a factor 45, illustrating that the unlabeled peptide is present at ~50-fold lower concentration.

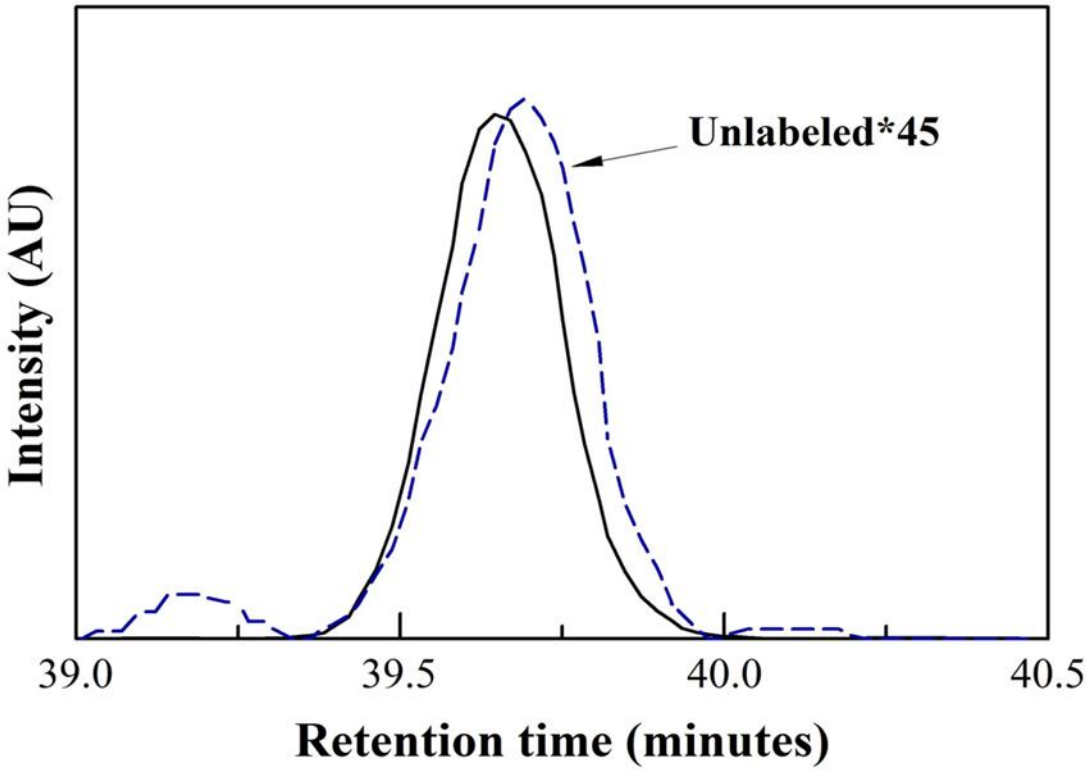


Fig. 3-5 Example of the mass-resolved detection of the elution of a Tyr-D_4 peptide (black line) and the corresponding unlabeled (dashed blue line) peptide. In this figure the intensity of the mass signal of the unlabeled peptide KALSGDSYWVFVK was multiplied by a factor 45, demonstrating a ~98% labeling efficiency at Tyr118 (see Table 3-3 for details).

Quantitative analysis of the peak areas in the elution pattern of the observed unlabeled and Tyr-D_4-labeled peptides yielded values of 2.0%, 2.3%, and 2.4% unlabeled side chains at Tyr76 and Tyr118 (Table 3-3). For the other three Tyr side chains these small contributions to the chromatogram were not detected, presumably due their intensities falling below the noise threshold. Since all five Tyr residues would be expected to be exhibit an identical labeling percentage, and since the major species in the mass spectra of intact Tyr-D_4-labeled PYP exhibited a 20 Da mass increase compared to unlabeled PYP, we conclude that the conditions used here yield the specific isotopic labeling of all Tyr residues in PYP with an efficiency of

97.7%. This corresponds well with the values of 96.4% derived above from mass spectrometry of intact PYP. The Tyr-D_4 that was used to grow the *E. coli* cultures contained a ring-D_4 labeling level of 98%, in excellent agreement with the experimentally detected level of Tyr-D_4 incorporation in PYP. These results therefore demonstrate the quantitative incorporation of the added Tyr derivative into PYP, and thus the absence of isotope dilution.

With respect to isotope scrambling, we considered the possibility of incorporation of Phe-D_4 in PYP (possibly through prephenate as an intermediate). The inclusion of Phe-D_4 in the list of possible derivatives used during the analysis of the LC-MS/MS data did not yield any PYP-derived Phe-D_4 containing peptides. Thus, isotope scrambling to Phe was undetectably low.

3.3.3. FTIR spectroscopy of Tyr-D_4 labeled PYP

Unlabeled and Tyr-D_4-labeled PYP samples were spectroscopically studied using static Fourier transform infrared (FTIR) spectroscopy. To facilitate the identification of shifts in peak positions, second derivative spectra were calculated for the unlabeled and Tyr-D_4-labeled PYP. This revealed a series of characteristic peak shifts (Fig. 3-6). Comparison of these data with calculated vibrational spectra for free Tyr and Tyr-D_4 illustrates that the pattern of peak shifts observed in the protein samples corresponds to that expected for ring-D_4 labeling of the Tyr side chains in PYP. The strong peak at 1515 cm^{-1} has essentially disappeared in ring-D_4 labeled, and has shifted to ~1435 cm^{-1}. Thus, clear signals from Tyr side chains can be detected in FTIR spectra from intact PYP. This opens the way to future time-resolved FTIR difference spectroscopic studies of these signals.

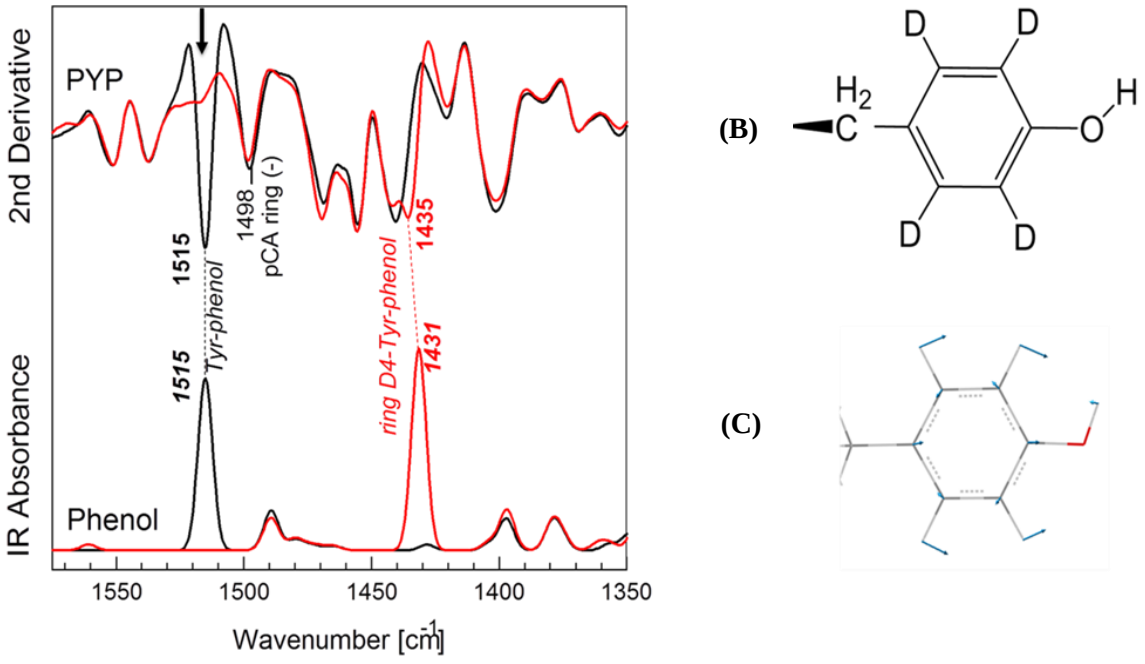


Fig. 3-6 Infrared signals from ring-D_4-Tyr. (A) Top spectra are the second derivative of infrared absorption of Hhal PYP (black) and ring-D_4 isotope edited Hhal PYP (red). The characteristic ring mode at 1515 cm^{-1} is indicated by an arrow. The bottom two are computational data from the phenolic ring resembling the side chain of Tyr (black) and ring-D_4 isotope edited side chain (red) of tyrosine. (B) The side chain structure of ring-D_4-Tyr. (C) The 1515 cm^{-1} vibrational mode of tyrosine with predominately CH bending coupled ring deformation. Upon ring-D_4 labeling, this vibrational frequency is shifted from 1515 cm^{-1} to 1435 cm^{-1} ('A' top), highly consistent with the band shift from ab initio computational data ('A' bottom).

Isotopic labeling is widely used in NMR spectroscopy. An important difference with the application of isotopic labeling to FTIR spectroscopy is that both the labeled and unlabeled variants contribute to the vibrational spectrum. Because isotopic labeling usually causes only fairly small shifts in vibrational frequency, partially isotopic labeling will result in the splitting of the bands under study. This effect therefore results in additional complications of already crowded FTIR spectra. In addition, if the isotopic label is scrambled to other side chains, the vibrational signals of these groups will also be split, hampering the assignment of vibrational modes to specific types of side chains and the unambiguous interpretation of the spectra. Thus, it is greatly preferable to perform vibrational measurements on samples with a very high labeling percentage. The results reported here show that this is entirely feasible for Tyr side chains.

In the case that isotope scrambling is difficult to avoid, such as for the Asp side chain (*138*), detailed knowledge of the level of labeling of the group under study and the pattern of isotope scrambling is of great value for interpreting the FTIR spectra obtained using the samples. Such quantitative information on the pattern of isotopic labeling can be readily obtained using the approach reported here based on LC-MS/MS analysis of tryptic digests of the isotopically labeled protein. The approach reported here for PYP is applicable to studies of Tyr side chains in any protein for which overproduction and purification has been achieved and that is amenable to mass spectrometry. Since glyphosate also inhibits the synthesis of phenylalanine and tryptophan in cells, the same method is expected to work for side-chain specific isotope editing of these two residues in proteins. In addition, in combination with the use of VSM of Tyrosine (*158*), FTIR spectroscopy can provide valuable information on the hydrogen bonding status, protonation state, and redox state of Tyr side chains (*33, 159, 160*). Since structural and/or chemical changes of Tyr side chains frequently are essential events at the catalytic site of proteins (see for example (*159, 161*)), the Tyr labeling and quantification procedure reported here is expected to be of general use.

3.3.4. Preliminary analysis of FTIR signals in terms of changes in Tyr hydrogen bonding (Ongoing work in Dr. Aihua Xie's lab):

The Tyr-D_4 PYP obtained from the above mentioned analysis was used for isotope-edited microsecond step-scan FTIR spectroscopy to identify the Tyr vibrational modes in the vibrational structural marker (VSM) region. The VSM approach (*119*) will allow signals in FTIR difference spectra to be interpreted in terms of specific structural changes in hydrogen bonding and how it can trigger biological proton transfer.

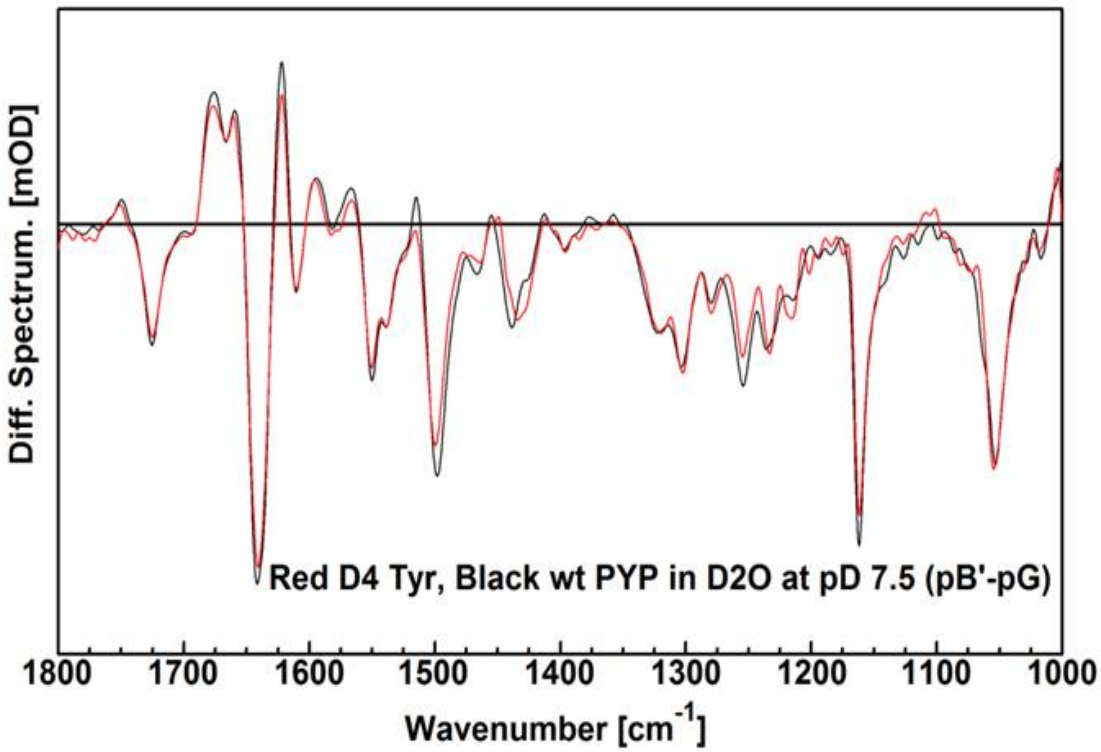


Fig. 3-7 Structural changes in the pB' state detected by time resolved step-scan FTIR difference spectra (pB' – pG) of the unlabeled (black) and D_4Y-PYP (red) (*Data measured by Dr. Lorand Kelemen).*

First side chain isotope edited step-scan FTIR spectroscopy

Microsecond step-scan FTIR spectroscopy was performed using both native and Tyr-D_4 PYP (Kelemen and Xie, unpublished results). To the best of our knowledge this represents the first isotope-edited step-scan FTIR experiment. These data were used to extract proposed pure pR – pG, pB' – pG, and pB – pG difference spectra. In D_2O as well as H_2O, Tyr-D_4 sensitive FTIR peak shifts were observed in these spectra around 1250 cm^{-1}, the spectral region containing VSM signals for Tyr side chain hydrogen bonding (Kang and Xie, unpublished results). The interpretation of these FTIR vibrational signals assigned to Tyr in terms of changes in hydrogen bonding is currently ongoing in the Xie lab.

CHAPTER IV

Understanding the functional role of PAS conserved residue Ile39 in PYP

This chapter is in preparation for submission as a manuscript titled "*Superfamily-conserved residue Ile39 in photoactive yellow protein affects functional kinetics and folding cooperativity*" Rachana Rathod, Jie Ren, Masato Kumauchi, Sandip Kaledhonkar, Aihua Xie, Wouter D. Hoff.

4.1. Abstract

PAS domains form a large and diverse superfamily of signaling proteins that is defined by a weak but characteristic pattern of amino acid sequence conservation. The 9 most strongly PAS-conserved residues are located at structurally inconspicuous and largely unstudied positions. Thus, while evolution indicates that these residues are critical for the PAS domain, their mechanistic role is poorly understood. Two possible functions are in PAS domain folding/stability and in allosteric switching for PAS domain signaling. Here we examine the effects of substituting PAS-conserved residue Ile39 in photoactive yellow protein (PYP), a prototype PAS domain that functions as a photoreceptor based on its *p*-coumaric acid (*p*CA) chromophore. The absorbance spectrum and pK_a of the *p*CA are essentially unchanged in I39A PYP, but the cooperativity of both acid denaturation and Gdm-HCl denaturation are strongly increased. Rapid-scan FTIR difference spectroscopy demonstrates that the allosteric switch in I39A PYP is intact. However, the lifetime of its pB photocycle intermediate is increased by a factor 30.

Bioinformatics analysis of contact maps indicates that Ile39 is part of a network of hydrophobic interactions with Ile31, Val57, and Phe62 that are highly conserved in PAS domains. These results bear a striking resemblance to the role of PAS conserved residue Asn43: in N43A PYP folding cooperativity is perturbed and pB decay kinetics are strongly delayed without disrupting the light-induced allosteric switch. The results imply that PAS-conserved residues Ile39 and Asn43 unexpectedly play a dual role in PYP in determining folding cooperativity and the rate of signaling kinetics.

4.2. Introduction

The Per, Arnt, and Sim (PAS) domain is a large and ubiquitous superfamily of functionally diverse single and multidomain signaling proteins (*44, 54*). PAS domains were initially identified based on their very weak but characteristic pattern of amino acid sequence conservation (*44*). Crystal structures of various PAS domains subsequently confirmed that they share a common 3D structure (a characteristic α/β fold), forming a coherent protein superfamily (*44, 45*). PAS domains can trigger responses to a variety of environmental signals, including light, redox potentials, and ligands (*162*). Structural-function studies on the PAS domains were initiated through work on a select number of proteins, particularly photoactive yellow protein (PYP), involved in negative phototaxis in *Halorhodospira halophila (48),* the N-terminal domain of human ERG potassium channel, involved in regulatory function in cardiac electrical excitability (*163*); the heme binding domain of FixL, a bacterial oxygen-sensor (*164, 165*), and the LOV domain of the phototropin segment of the fern photoreceptor phy3, involved in chloroplast photorelocation (*166*), Over 50,000 different PAS domains have been identified and 43 PAS domains have been found in the human genome as estimated by the SMART database (*41*).

PYP is a small (14-kDa), 125 residues cytosolic blue light receptor that serves as a structural prototype of the PAS domains (*53, 54*). It was the first PAS domain for which the crystal structure was determined (*67*), and consists of two regions: a typical PAS fold (*70*) with a

central antiparallel 6-stranded β-sheet flanked by three α-helices (the hydrophobic core) and an N-terminal region (residues 1-28) containing two α-helices. The PYP family of photoreceptors shares a set of characteristic properties that were identified through studies on the PYP from the halophilic purple bacterium *H. halophila*. PYP exhibits a light-triggered photocycle (*56, 78*) using its *p*-coumaric acid (*p*CA) chromophore (*58*). The phenolic oxygen of *p*CA forms functionally important hydrogen bonds with the side chains of active site residues Tyr42 and Glu46 (*65, 102, 167*). Upon illumination of the initial pG ground state, the *p*CA undergoes a *trans-cis* photoisomerization which triggers a complex photocycle (*56, 78*) that results in the formation of the long-lived photoactivated pB state. During the formation of the pB state, an intramolecular proton transfer occurs from Glu46 to the *p*CA, which triggers global conformational changes in the protein (*5, 81*), that are thought to generate a biological signal. Formation of the pB state involves partial unfolding, including the release of the N-terminal region (*84, 168, 169*), which indicates the allosteric transmission of structural changes from the chromophore to the N-terminal region across the rigid central *β*-sheet. At neutral pH, the pB state decays back to the initial pG state in ~ 350 ms. While the formation of a negative charge on Glu46 upon proton transfer has been identified as a major electrostatic driver for conformational changes during the PYP photocycle (*33*), important questions remain regarding the mechanism for allosteric transmission of structural changes from the chromophore binding pocket to the N-terminal region.

PAS domains are approximately 100 residues in length, and only 9 of these residues are moderately conserved in the PAS domain superfamily (*44*). We will refer to these amino acids as PAS-conserved residues. The molecular role that these PAS-conserved residues play is poorly understood. While classic protein families share conserved catalytically active residues, PAS domains bind diverse cofactors and do not share a conserved active site (*1, 44*). The functional relevance of PAS-conserved residues therefore is not obvious. We are exploring this question using the PYP system. Results from a complete Ala scan of PYP that we previously reported (*1*)

revealed changes in the spectroscopic properties of the N43A and I39A mutants and indicated a functional role for PAS-conserved residues Asn43 and Ile39. We subsequently reported that Asn43 (Fig. 4-1) forms two side chain-backbone hydrogen bonds that are conserved in PAS domains and that strongly affect both pB decay kinetics and folding cooperativity in PYP (*103*). In addition, we reported that a number of PAS-conserved residues are needed to help maintain high heterologous protein overproduction levels in *E. coli* (*1*).

In this study, we examine the molecular role of the PAS conserved residue Ile39 (Fig. 4-1), one of the most conserved residues in the PAS domains, using *H. halophila* PYP. The side chain of Ile39 is part of the main hydrophobic core of PYP (*67*). We use visible and FTIR spectroscopy to examine if the side chain of Ile39 is involved in the transmission mechanism of allosteric structural changes in the PAS domain superfamily, in regulating PYP signaling kinetics, and in the folding of PYP.

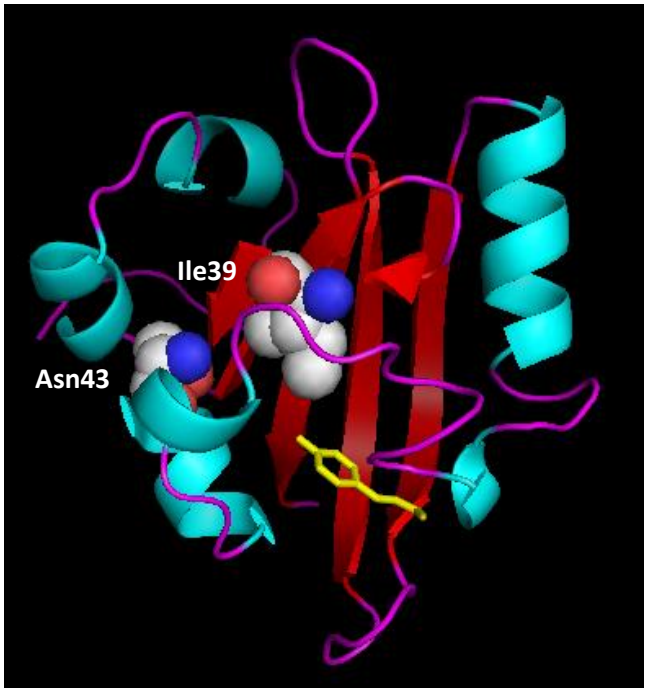


Fig. 4-1 Schematic representation of the crystal structure of PYP. Highlighted in spheres are the PAS conserved residues Ile39 and Asn43. The 6 stranded β-sheet is shown in red, the 5 α-helices are shown in blue, the loops in magenta, and the *p*CA chromophore is shown in yellow.

4.3. Materials and Methods

Site-directed mutagenesis, Protein purification, and UV/visible spectroscopy:

These methods have been described in chapter 2 of the thesis.

FTIR spectroscopy

For FTIR spectroscopy, wt PYP and I39A mutant were used at 8 mM protein concentration in 50 mM phosphate buffer in D_2O at 6.6 as described in ref (*103*). FTIR spectra were obtained on 2.7 µl of PYP sandwiched between two CaF2 windows (15 mm diameter), separated by a 12 µm spacer with a Bruker IFS 66v spectrometer in the spectral range 4000-850 cm^{-1} using a liquid nitrogen-cooled Mercury Cadmium Telluride (MCT) detector.

Bioinformatics analysis

The crystal structure of photoactive yellow protein (1nwzA) was submitted as the query structure on the Dali web server (*170*) to identify proteins with similar three-dimensional structures. The hits included PYP and a range of other PAS domain structures. PAS domains were selected from the resulting list of hits using a published structure-based phylogenetic grouping of PAS domains to cover the diversity of the entire PAS domain superfamily (*171*). The pairwise percentages of sequence identity and similarity among these selected PAS domains were calculated on the SIAS web server according to the structure-based sequence alignment generated from Dali server. A final selection of 15 PAS domains (including Hhal PYP) was made such that these sequences are non-redundant in terms of percentage sequence identity. Side chain contact map of each of these 15 PAS domains was calculated and analyzed using CMview v1.1 with the cutoff value 4.2 Å (*172*). This cutoff value was selected because we were interested in the hydrophobic contacts formed by Ile39 in Hhal PYP and the corresponding residues in the other PAS domains. The resulting set of 15 contact maps was analyzed to identify side chain – side chain contacts that are conserved in most PAS domains. Swiss-PdbViewer v4.10 (*173*) was used

to visualize the crystal structures of the selected PAS domains and to generate figures illustrating these conserved side chain contacts.

4.4. Results

4.4.1. PAS domain sequence alignment based on three dimensional structures

With the analysis of amino acid sequences using hidden Markov models and position-specific scoring matrices, biologically meaningful levels of sequence similarity have been greatly reduced compared to the "twilight zone" (*174-176*) starting near ~25% sequence identity. PAS domains consist of ~100 residues and contain ~9 are moderately conserved (*44*). At such low levels of sequence similarity obtaining reliable sequence alignments involving sequence insertions and deletions is challenging. The availability of high-resolution protein structures can help address this problem. The reliability of the sequence alignment can be greatly improved by superimposing the three-dimensional structures. Figure 4-2A depicts such a 3D-strucure-guided amino acid sequence alignment for a set of 15 diverse PAS domains, including Hhal PYP. These PAS domains were selected to span the entire phylogenetic tree of all PAS domains based on the 3D structures of these proteins reported in (*171*). It highlights the conservation of amino acid residues in the PAS domain superfamily. Of the 9 moderately conserved residues in the PAS domains 6 are observed in Hhal PYP: Asp34, Gly37, Ile39, Asn43, Gly51, and Gly59. Three of these residues are buried (residue 37, 39, and 43) whereas the rest (residue 34, 51, and 59) are solvent exposed. These residues appear to be structurally inconspicuous and are not part of the *p*CA binding pocket.

Based on the amino acid sequence alignment of the 15 PAS domains guided by the crystal structures of these proteins, we calculated all pairwise percentages of sequence identity and similarity for this set of proteins (Fig. 4-2B). Most sequences were found to share 8% sequence identity, with a maximum of ~24% sequence identity, confirming that this is a set of non-redundant (very low sequence identity) PAS domains with respect to amino acid sequence. Interestingly, the percentage pairwise sequence similarity was observed to be much higher, most

frequently 64%. We suggest that this unusual ratio of percentage sequence identity versus similarity reflects that PAS domains have explored the available amino acid sequence space to a relatively high extent.

(A)

```
                  30         40         50         60
                  :          :          :          :
PYP (1nwzA)       AIQLDGDGNILQYNAAEGDITGRDPKQVIGKNFF
FixL (1d06A)      TVVSATDGTIVSFNAAAVRQFGYAEEEVIGQNLR
HisK (2r78A)      IFIXDAEGHYLDVNPAICSAIGYTRDEFLALDWG
HisK (3oloB)      SFCLGDNWQFLYVNDATCRXTEYSREQLLSXNLQ
MmoS (3ewkA)      ILLTDRAGRIIYANPALCRFSGMAEGELLGQSPS
ERG (4lloB)       FVLGNADWPIVYSNDGFCKLSGYHRAEVMQKSSA
LOV (2pr5A)       VVITDPDNPIVYVNQGFVQMTGYETEEILGKNCR
GGDEF (2p7jA)     IRFLDTSGTEKVRI--KYDF----SLIL—RDKSA
Bphy (2oolA)      LFVVSEDLRIASVSANVEDLLRQPPASLLNVPIA
BPhy (2o9cA)      LLTADGSGEVLQMSLNAATFLGQEPTVLRGQTLA
Ykul (2basA)      IYXCDEEGDQLTNVFK-------YQPEYAEKNWS
Per (3rtyA)       AIRHTATGIISHVDSAAVSALGYLPQDLIGRSIM
DctB (3e4oA)      IYLIDRSGVAVAASN------EPTS--FVGNDYA
PhoR (3cwfA)      ASVIDTDGKVLYGSN-------------------
HisK (3licA)      SAIAFKP-LFSPY--VYRS--------DIGADYT
```

(B)

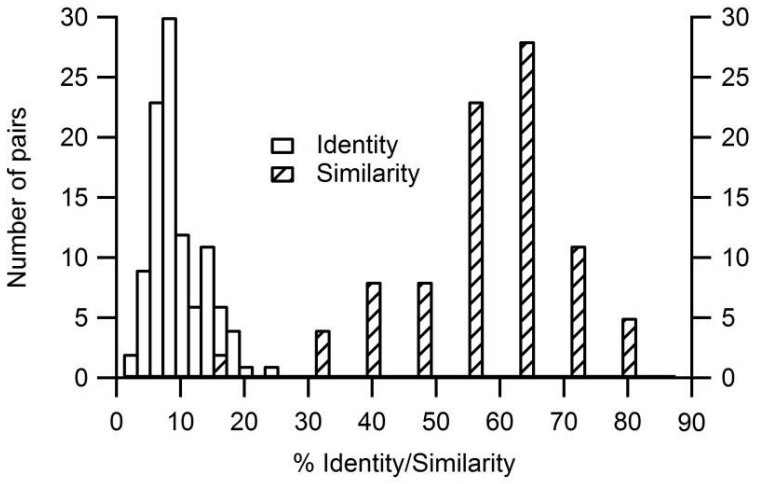


Fig. 4-2 Sequence conservation in the PAS domain superfamily. (A) Amino acid sequence alignment of a set of 15 PAS domains with known crystal structure. Residues 30 to 63 (residue numbering of Hhal PYP) are shown. The depicted alignment is based on the crystal structures of the proteins. Residue Asn43 (*103*) is colored green and residue Ile39 and its interaction partners are colored yellow. (B) Distribution of pairwise percentages in sequence identity and similarity for the same 15 PAS domains.

4.4.2. Interactions of Ile39 that are conserved in the PAS domain superfamily

Based on our finding that Asn43 forms two side chain hydrogen bonds that are highly conserved in the PAS domain superfamily (*103*), we aimed to systematically examine the possibility that Ile39 is also involved in a pattern of side chain interactions that is conserved in PAS domains. To this end, we calculated contact maps for each of the 15 non-redundant PAS domains depicted in Fig. 4-2A. Since the side chain of Ile39 is part of the main hydrophobic core of PYP, we examined interactions between hydrophobic residues, and calculated side chain – side chain contact maps with a cutoff distance of 4.2 Å. We then examined which of the observed side chain – side chain contacts were present in the majority of the PAS domains. Figure 4-3A depicts the section of the contact map containing the conserved contacts for residue 39. This analysis indicates that Ile39 makes conserved contacts with residues 31, 57, and 62 (residue numbering for Hhal PYP).

To analyze the nature of the contacts of these four residues, we examined their relative location and orientation in all 15 PAS domains. The results are depicted in Figure 4-3B for PYP and three other PAS domains from the set of 15 analyzed here. In the case of PYP, the side chain of Ile39 was found to be clamped in between the side chains of Ile31 and Val57, forming a hydrophobic stack of three residues, while the side chain of Phe62 leans against residues 31 and 39 in this stack. A highly similar set of hydrophobic interactions was observed for the other PAS domains. Thus, Ile39 is at the center of a cluster of conserved side chain – side chain interactions in PAS domains.

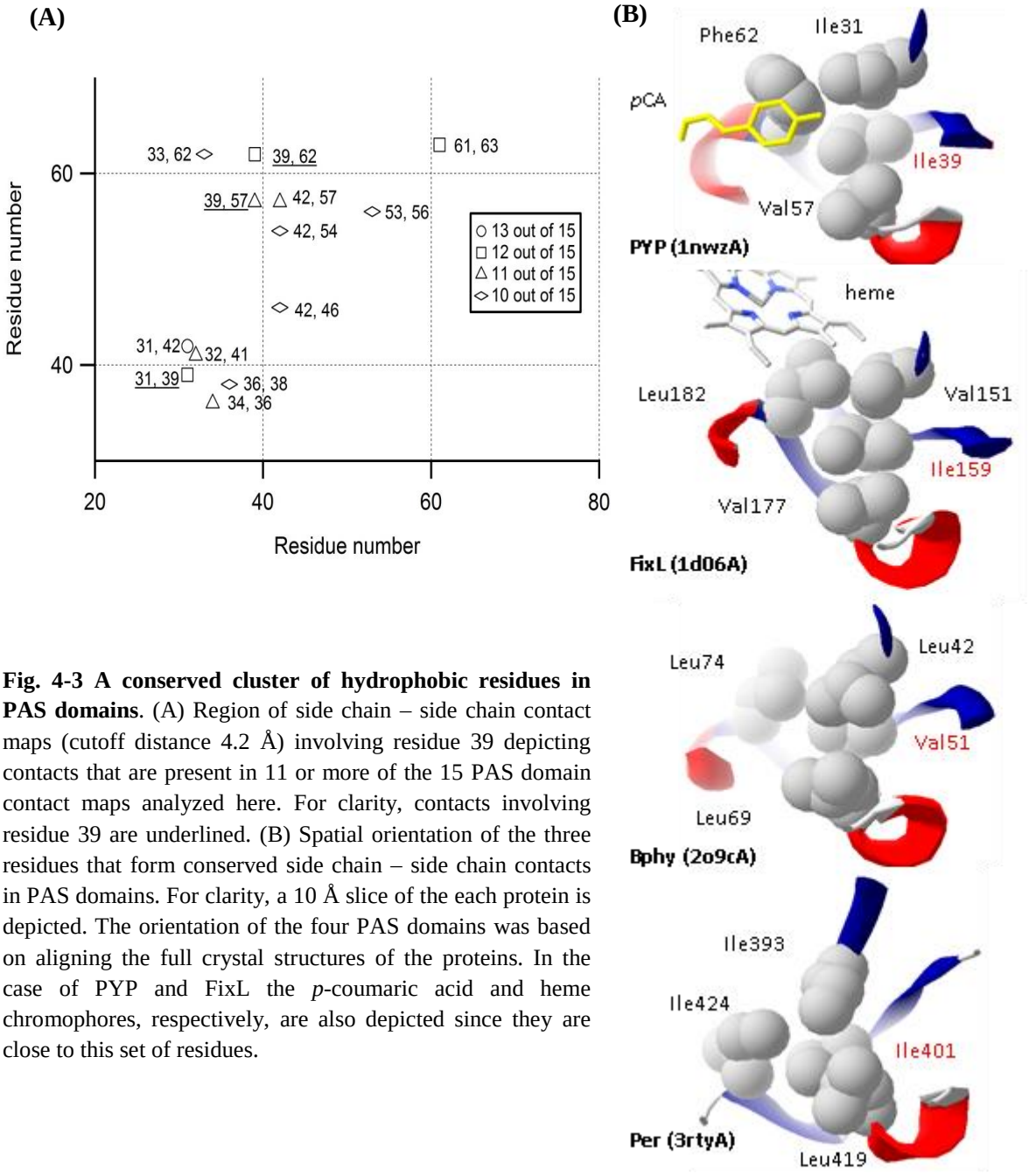


Fig. 4-3 A conserved cluster of hydrophobic residues in PAS domains. (A) Region of side chain – side chain contact maps (cutoff distance 4.2 Å) involving residue 39 depicting contacts that are present in 11 or more of the 15 PAS domain contact maps analyzed here. For clarity, contacts involving residue 39 are underlined. (B) Spatial orientation of the three residues that form conserved side chain – side chain contacts in PAS domains. For clarity, a 10 Å slice of the each protein is depicted. The orientation of the four PAS domains was based on aligning the full crystal structures of the proteins. In the case of PYP and FixL the *p*-coumaric acid and heme chromophores, respectively, are also depicted since they are close to this set of residues.

4.4.3. The I39A mutation affects folding cooperativity in PYP

To determine the role of Ile39 in PYP, we examined the effects of the I39A mutation on the spectroscopic properties and stability of its pG dark state. The effects on PYP of substituting

PAS-conserved residue Asn43 with Ala (*103*) were used for comparison. Both the shape of the absorbance spectrum of the pG state and the absorbance maximum (λ_{max}) of I39A PYP at 446 nm are essentially unchanged compared to wt PYP (Fig. 4-4A), whereas the λ_{max} of N43A PYP is blue-shifted by 5 nm (*103*). The extinction coefficient of the 446 nm absorbance band of I39A PYP was determined to be 38,500 $M^{-1}cm^{-1}$, 15% reduced from its value in wt PYP. The pK_a of the *p*CA chromophore in I39A PYP was determined by acid titration using UV/vis absorbance spectroscopy. Upon acidification of wt PYP the absorbance band of the native pG state at 446 nm is converted to the blue-shifted, acid-denatured species (λ_{max} 350 nm) with a protonated *p*CA chromophore (*55, 95*) (Fig. 4-4B). The same transition was detected for I39A PYP. However, the transition was described by a single pK_a value of 3.3 ± 0.1 with an *n* value of 2.5 ± 0.1 (Fig. 4-4B), whereas under the condition used here in wt PYP this transition exhibits a pK_a value of 3.1 ± 0.1 and an *n* value of 1.6 ± 0.1. Apparently the main effect of the I39A mutation on the acid denaturation of PYP is a substantial increase in the steepness of the transition (Fig 4-4C). For comparison, the pK_a of N43A PYP is 4.6 with an *n* value of 1.2 (*103*).

To examine the stability of I39A PYP we performed denaturant titrations with Gdm-HCl, in which the native pG state is converted into the fully unfolded state that absorbs near 340 nm. The unfolding of wt PYP can be described as a two-state transition (*5*). The well-defined isosbestic point near 384 nm (Fig. 4-5A, B) observed for the unfolding of I39A PYP indicates that the two-state nature of the equilibrium unfolding is maintained in this mutant. This transition is described by two factors; the free energy for unfolding ΔG_U in the absence of denaturant and the denaturant *m* value that describes the steepness of the transition. Unexpectedly, two large effects in the denaturant titration behavior of I39A PYP were found to largely cancel out. In a destabilizing effect, the denaturant midpoint of the unfolding curve of I39A PYP was reduced almost two-fold, from 2.72 M in wt PYP to 1.42 M in I39A PYP. However, the denaturant *m* value of wt PYP of 16.2 ± 1.1 kJ/mol/M^{-1} was increased by ~33% to 21.5 ± 0.7 kJ/mol/M^{-1} in I39A PYP, providing a stabilizing effect. As a result, the ΔG_U of I39A PYP of 30.8 ± 1.0 kJ/mol

is only modestly reduced compared to wt PYP (ΔG_U = 45.1 ± 2.9 kJ/mol). For N43A PYP we previously reported a greatly reduced steepness in the transition (Fig. 4-5C) and a loss of folding cooperativity such that the unfolding curves based on *p*CA absorbance and aromatic fluorescence no longer overlap (*103*).

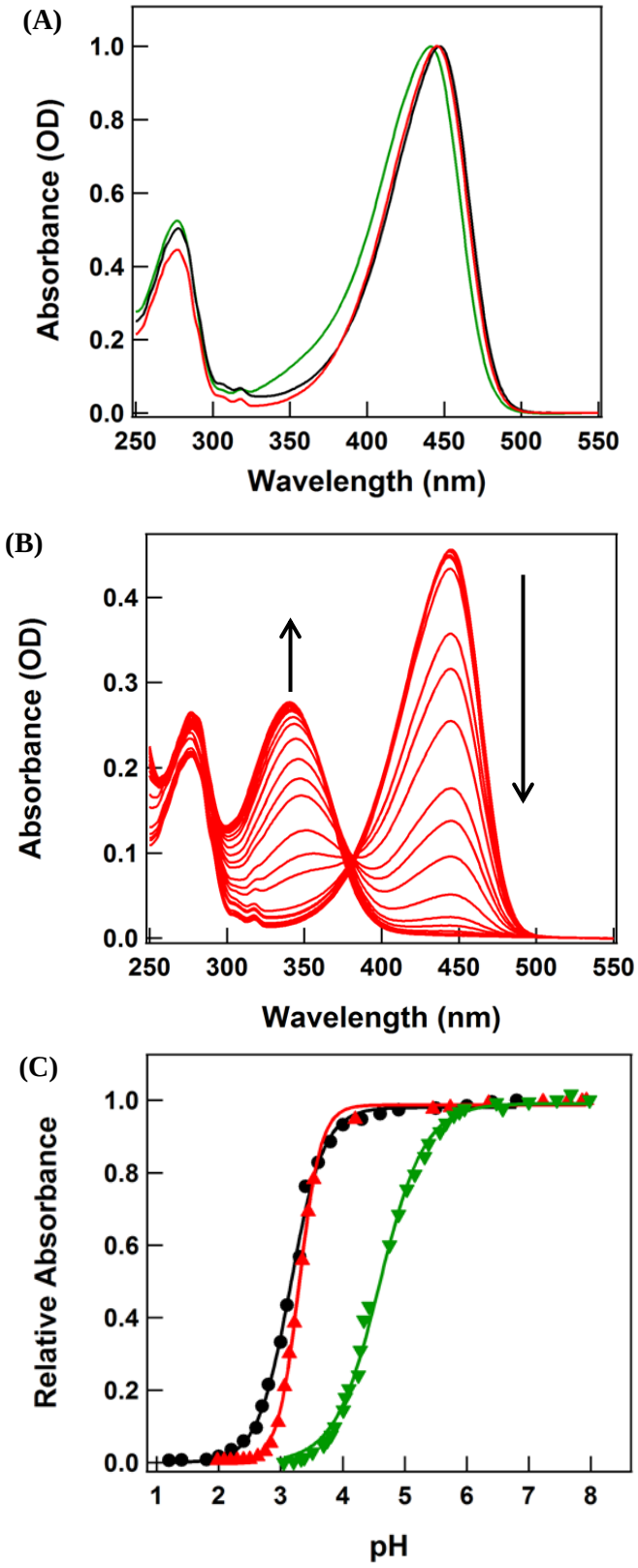


Fig. 4-4 Effect of I39A mutation on the absorbance maximum and pK_a of the *p*CA in PYP. (A) Absorbance spectra of I39A (red), N43A (green), and wt PYP (black) measured in 10 mM Tris-HCl at neutral pH. (B) pH dependent absorbance spectra of I39A PYP for a range of decreasing pH values from pH 8 to pH 2. At low pH the native 446 nm species is converted into the acid denatured 355 nm species. (C) pH titration curves of wt PYP (black), I39A (red), and N43A (green) obtained at 446 nm for wt PYP and I39A and at 441 nm for N43A PYP. The amplitudes of the fits were normalized at their visible absorbance maxima.

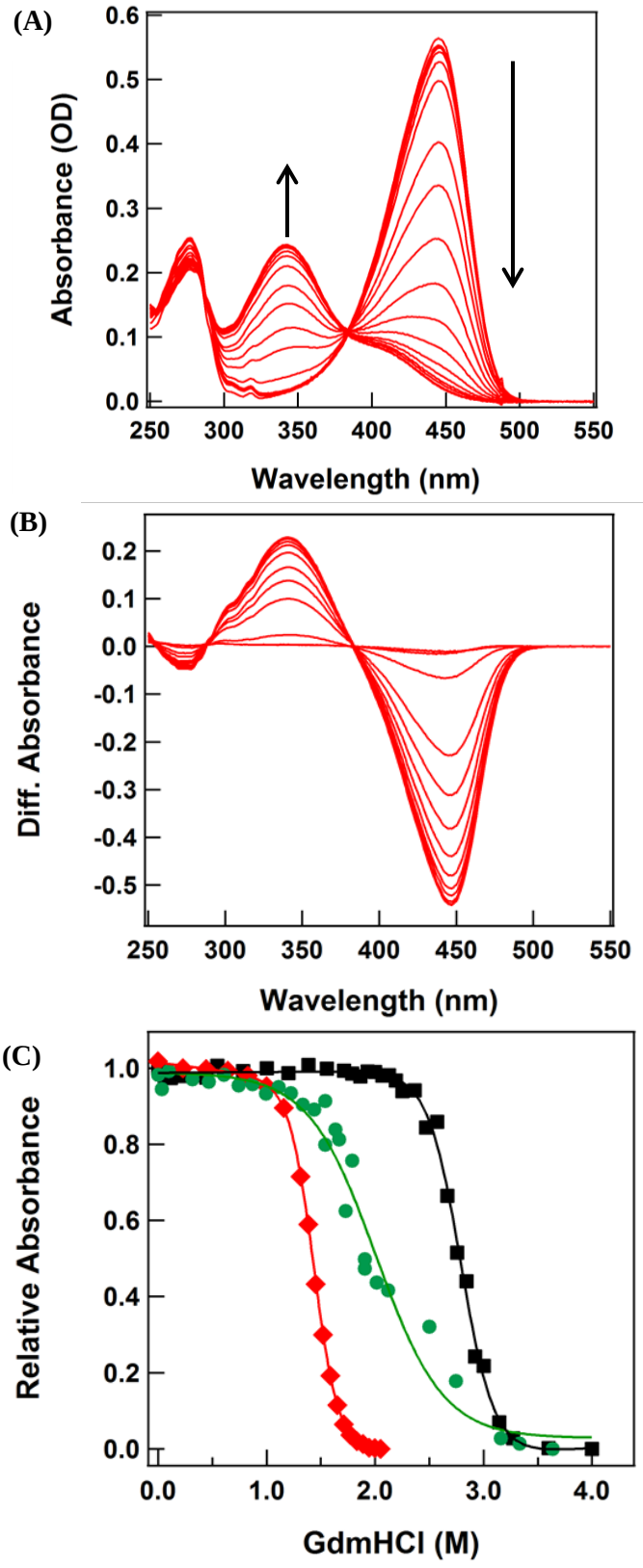


Fig. 4-5 Effect of the I39A mutation on the stability of Hhal PYP. (A) Absorbance spectra as a function of Gdm-HCl concentration for the I39A mutant of PYP. (B) Denaturant-induced difference spectra for I39A PYP. (C) Gdm-HCl denaturant titration curves of wt PYP (black), I39A PYP (red), and N43A PYP (green) mutants obtained at 446 nm for wt PYP and I39A PYP and at 441 nm for N43A PYP. The amplitudes of the fits were normalized at the respective visible absorbance maxima.

4.4.4. The role of Ile39 in tuning pB decay kinetics and allosteric switching upon pB formation

To probe the pB intermediate in I39A we performed UV/vis absorbance difference spectroscopy. The light-dark (pB-pG) difference spectrum of I39A PYP at neutral pH exhibits a negative peak near 446 nm, indicating photoconversion of the pG state, and a positive peak near 345 nm due to the formation of pB state, very similar to wt PYP (Fig 4-6A). We measured the kinetics of thermal pB decay in I39A PYP over the pH range 3.5 to 11 (Fig. 4-6B). At neutral pH pB decay in I39A PYP occurs with a time constant of ~15 seconds, 30-fold slower than wt PYP. Like wt PYP, the pH dependence of pB decay in the I39A mutant can be described reasonably well as bell-shaped, governed by two apparent pK_a values. While for wt PYP the pK_a value are 6.4 and 9.4 (*65*), in I39A PYP these values are 5.4 and 9.6 (measurements on both PYPs without N-terminal His_6 tag). The degree to which pH alters the rate is substantially reduced in I39A PYP. We found that for I39A PYP the presence of an N-terminal His tag measurably altered the pH dependence of pB decay kinetics, and exhibited a sharp feature near pH 6. Since this value is close to the pK_a of the His side chain, this observation suggests that the protonation stage of the His residues in the affinity tag has a modest effect on the rate of pB decay. In summary, at pH 7.0 the I39A mutation slows down the rate of pB decay by a factor ~30. For comparison, the N43A mutation also reduces the rate of pB decay, and in that case the effect is much stronger (a factor 3,400) (*103*). Thus, substitutions at two distinct PAS-conserved residues (Ile39 and Asn43) both slow down pB decay (Fig.4- 6C). Receptor deactivation kinetics is often correlated with *in vivo* signaling mechanisms. In microbial sensory rhodopsins the sensitivity of the *in vivo* signaling response has been reported to be altered by the lifetime of the signaling states (*177*) that matches the timescale of the receptor induced biological response (*178*). A similar response has been shown for two-component regulatory systems (*179, 180*) thus implying that the effect of I39A mutation on Hhal PYP could be related to a biological response.

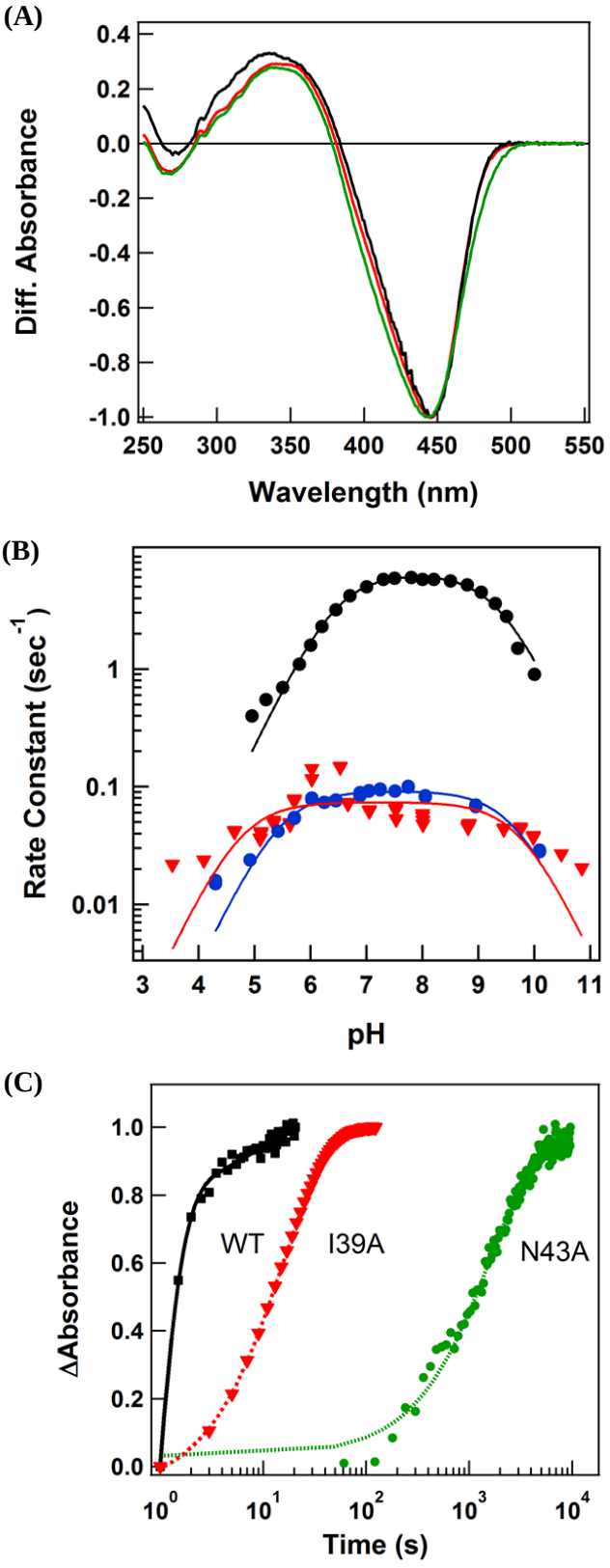


Fig. 4-6 Effect of I39A mutation on the kinetics of pB decay in PYP. (A) The amplitude-normalized pB-pG UV/vis difference spectra of wt PYP (black), I39A PYP (red), and N43A PYP (green). (B) The pH dependence of the kinetics of pB decay of wt PYP (black) (*65*), I39A PYP with the His tag (red), and I39A PYP without the His tag (blue). (C) The amplitude-normalized kinetic traces of the pB to pG transition measured for wt PYP, and the mutants at neutral pH.

To examine the possibility that Ile39 is part of a conserved allosteric switching mechanism in PAS domains, we performed time-resolved rapid-scan FTIR difference spectroscopy to probe the structural changes that occur during the formation of the pB state in wt PYP and its I39A mutant (Fig. 4- 7). The pB–pG FTIR difference spectra extracted from these measurements contain previously assigned characteristic signals for the deprotonation of Glu46 (negative signal at 1726 cm^{-1}) (*32*), protonation of chromophore (negative signal at 1497 cm^{-1} and positive signal at 1515 cm^{-1}) (*33*), and amide I (positive signal at 1624 cm^{-1} and negative signal at 1642 cm^{-1}) (*181, 182*). Since the amplitude of the amide I signals is not affected by the mutation, we conclude that the large conformational changes that occur upon pB formation (*5, 82-84, 86*) proceed essentially unaltered in I39A PYP, indicating that Ile39 is not part of a conserved allosteric switching mechanism. The similarity in the pB-pG FTIR difference spectra indicates that the I39A mutation does not cause substantial structural changes in either the pB or pG states of PYP.

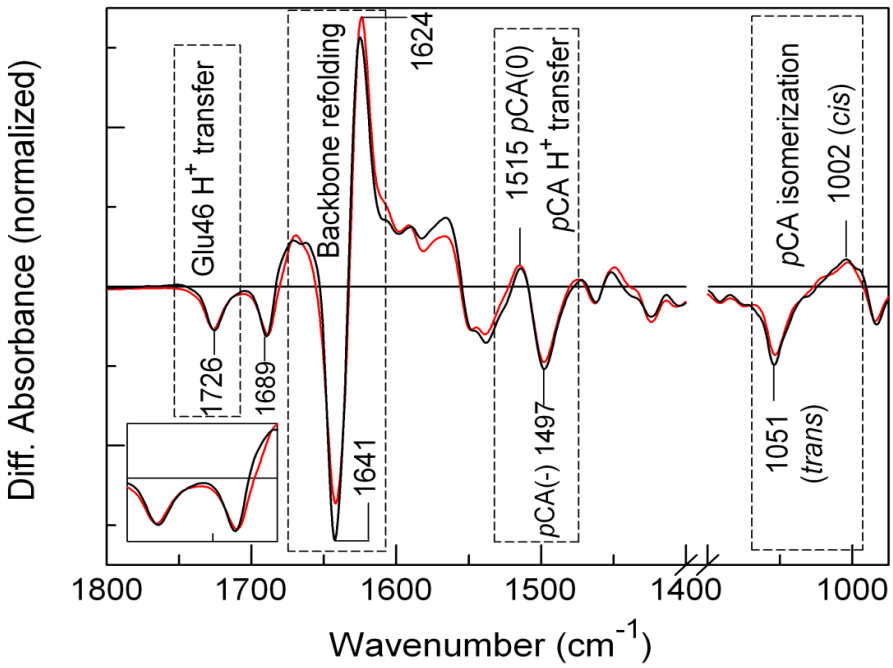


Fig. 4-7 Structural changes upon pB formation in wt PYP and I39A PYP detected by FTIR difference spectroscopy. The pB – pG infrared difference spectra measured at pH 7.5 for wt PYP (black) and I39A PYP (red) derived from flash-induced rapid-scan FTIR difference spectroscopy. The inset depicts an enlargement of the C=O stretching region from ~1750–1675 cm^{-1}.

4.5 Discussion

4.5.1 Molecular mechanism of Ile39 in PYP

The branched-chain hydrophobic residues Ile, Val, and Leu have been studied extensively for their role in contributing to protein stability by being part of tightly packed hydrophobic cores (*183, 184*), but due to their lack of chemical reactivity are generally not viewed as residues that tune functional properties of proteins. The results reported here indicate that Ile39 has a substantial effect on the kinetics of functional transitions in Hhal PYP. The results from our Ala mutagenesis scan of PYP (*1*) also indicated that hydrophobic residues can substantially alter functional properties. Substitution of the six residues in PYP for which Ala substitutions affected three different functional properties of PYP (λ_{max}, pK_a and pB lifetime), two are Ile (residues 31 and 49). While multiple groups independently have reported site-directed mutagenesis studies on PYP guided by its high-resolution crystal structure, Ile31, 39, and 49 were not selected in that work. In this study we provide a detailed analysis of the effect of I39A mutant on different functional properties of PYP.

The results reported here demonstrate that both functional and folding properties are altered in I39A PYP. With respect to function, FTIR difference spectroscopy showed that I39A PYP retains a functional allosteric switch that results in large light-triggered conformational changes. However, the rate of pB decay was reduced by more than an order of magnitude compared to wt PYP, both at neutral pH and over a wide pH range. We propose that this reduction in pB decay rate is related to the partially unfolded nature of the pB photocycle intermediate, such that the hydrophobic interactions involving Ile39 aid in the recovery of the initial pG state from the partially unfolded pB state. The I39A mutation does not have substantial effects on the absorbance spectrum and pK_a of the *p*CA chromophore in the pG state of PYP. These results imply that the I39A mutation does not affect the active site of PYP, but exerts its effects through the energetics of interactions in the hydrophobic core of the PAS domain.

Interestingly, it has been reported that conserved hydrophobic residues are important for signal transduction by the human androgen receptor (*185*), indicating that a functional role for conserved hydrophobic residues in biological signaling may be more widespread.

The I39A mutation also alters the equilibrium folding behavior of Hhal PYP. In the case of denaturant titrations using Gdm-HCl to cause the transition between the native state and the fully unfolded state, the I39A mutation caused two effects with opposing consequences on the stability ΔG_U of PYP. While a destabilizing effect was detected in the ~2-fold reduction in denaturant midpoint concentration, this effect was accompanied by stabilization through a substantial increase in the denaturant *m* value (from 16.2 to 21.5 to kJ/mol/M^{-1}). The final result is that the ΔG_U of I39A PYP was only reduced by 30% (from 45.1 to 30.8 kJ/mol).

The denaturant *m* value has been found to be proportional to the total amount of surface area that is exposed to solvent upon the denaturant-induced unfolding of the protein (*186*). However, the stability, active site properties, and pB – pG FTIR difference spectrum of I39A are quite similar to those of wt PYP, providing strong experimental evidence that the I39A mutation does not significantly change the structure of the pG state of PYP. Thus, it seems quite unlikely that I39A PYP is able to almost bury twice as much surface area as wt PYP. We propose that the large increase in denaturant *m* value in I39A PYP indicates that this mutation affects the amount of residual structure in the "fully unfolded" state of PYP. In this interpretation the I39A mutation substantially reduces the amount of such residual structure, resulting in the observed larger *m* value. Previously we have reported evidence that the isomerization state of the *p*CA exerts a similar effect on the residual structure in fully unfolded PYP (*187*). We previously reported that the N43A mutation also substantially alters the cooperativity of PYP folding (*103*). In that case the cooperativity is perturbed such that N43A PYP no longer exhibits two-state folding.

Unexpectedly, a substantial increase in the steepness of the unfolding process was also observed in the acid denaturation process of I39A PYP. This denaturation process results in a partially unfolded state that retains substantial secondary structure (*105*). In this case the observed

n value, which is usually interpreted as the number of protons involved in the transition, was increased from 1.6 in wt PYP to 2.5 in I39A PYP. The structural interpretation of this observation is not clear, however, it is intriguing that the I39A mutation results in a substantial increase in folding cooperativity in both the complete unfolding process caused by Gdm-HCl and the partial unfolding caused by low pH. In combination, these results indicate that Ala substitution of two PAS-conserved residues, Ile39 and Asn43, substantially affect protein folding cooperativity in PYP.

4.5.2. The functional role(s) of PAS conserved residues

Advances in the bioinformatics analysis of amino acid sequences using hidden Markov models and position-specific scoring matrices has led to the important discovery of protein superfamilies that are characterized by sequence similarities that lie deep within the traditional "twilight zone" (*41, 42*). In the case of PAS domains the determination of high-resolution three-dimensional structures has confirmed the validity of the classification of these sequences by revealing a conserved PAS domain fold (*44, 54*). The functional importance of these PAS-conserved residues remains to be fully determined. Visual analysis of crystal structures to identify functionally important residues often forms the basis of site-directed mutagenesis studies on protein structure-function analysis. PYP has been studied in this regard by a number of independent research groups, however, PAS-conserved residues were not selected for such studies. Instead, the crystal structure of PYP has drawn attention to classic active site residues side chains such as Tyr42 and Glu46. These observations illustrate that the most highly conserved residues in PAS domains are located at seemingly unremarkable positions, and points to an important gap in our current understanding of structure-function relationships.

Here we examine the functional role of Ile39, one of the most highly conserved residues in PAS domains. In the case of classic protein families conserved residues can be part of the active site for catalysis. This pattern of sequence conservation is observed in specific families that

are part of the PAS domain superfamily. In the case of the PYP family of photoreceptors residues that are critical for the active site (particularly Tyr42, Glu46, and Cys69) are highly conserved. However, different members of the PAS domain family bind different cofactors (ranging from *p*-coumaric acid to flavins and heme), and some PAS domains function without cofactors. Therefore, the PAS domain superfamily does not share a conserved active site, and Tyr42, Glu46, and Cys69 are not conserved in PAS domains.

In general, superfamily-conserved residues can be suspected to be involved either in folding/stability or in functional properties. The role of conserved residues in folding or maintaining protein stability is debated (*103, 188, 189*). Studies on PYP have yielded two insights on the role of PAS-conserved residues. First, we previously reported that Ala substitution mutants for all 6 residues in Hhal PYP that follow the pattern of conservation in the PAS domain family are able to fold into a yellow and photoactive protein (*1*), arguing against the necessity of these residues for the folding of the PAS domain fold. However, four of these mutations resulted in a substantial reduction in the degree of heterologous overproduction in *E. coli*, implying an unexpected role for PAS-conserved residues in retaining protein production levels and suggesting that these mutants exhibit an increased tendency for aggregation. Here we report that the I39A substitution substantially increases the cooperativity of protein folding of PYP. Together with our previous results on N43A PYP, these data indicate an unexpected role for PAS-conserved residues in regulating the degree of folding cooperativity. The functional relevance of this finding and the fitness advantage conferred by this effect are not clear at this point.

Regarding possible functional roles of PAS-conserved residues, the finding that PAS domains share a role in signaling suggests a conserved mechanism in signal relay. A possibility is that these residues form a conserved interface for protein-protein interactions for signaling. In that case a patch of conserved residues at the protein surface could occur. No such surface patch is observed for the 9 most conserved residues in the PAS domain superfamily. In addition, PAS domains interact with a range of different signaling partners, including other PAS domains,

histidine kinases, and GGDEF domains. Since these interaction partners have very different structures, a conserved interaction interface appears unlikely. An alternative functional role is that these residues are part of a conserved allosteric switch. The FTIR data reported here for I39A PYP provide strong evidence against this possibility, matching our results on light-induced conformational changes in N43A PYP. However, both of these mutations quite substantially reduce the rate of thermal decay of the pB state of PYP. Therefore, unexpectedly these two PAS-conserved residues are not part of a conserved allosteric switching mechanism but regulate functional dynamics.

4.5.3. The side chains of PAS conserved residues form conserved patterns of interactions

In the case of PAS-conserved residue Asn43 we previously reported that this side chain forms two side chain – backbone hydrogen bonds that are highly conserved in PAS domains (*103*). These two hydrogen bonds were observed for residues including Asn, Ser, Asp, and Thr at this position, indicating that not residue identity, but the pattern of hydrogen bonding interactions is conserved.

Here we report that PAS-conserved residue Ile39 is at the core of a conserved cluster of hydrophobic interactions with three of its immediate neighbors in the PAS domain fold: residues 31, 57, and 62 (residue numbering for Hhal PYP). These residues form a conserved set of interactions in the hydrophobic core of the PAS domain. We inspected a published multiple sequence alignment of 63 different PYPs (*63*) to examine naturally occurring substitutions at position 39 and its three hydrophobic interaction partners. In the PYP family of photoreceptors, position 39 is present as Ile in 54% of the cases; in the remaining 45% it is Val. For residue 31 an Ile was found in 69% of the cases, and the remaining sequences contained Val or Leu; residue 62 was a Phe (83%) or L/I/V (16%) and Tyr (in a single PYP sequence), and finally residue 57 revealed the highest variability with Val (65%) or a hydrophobic residue (Ala or Met in 25% of cases), and Arg or Gln in 10% of the sequences. In the case of the PAS domain superfamily, at each of these four positions the 15 PAS domains examined here contain a range of different

hydrophobic side chains in ~85% of the sequences. Taken together, these results indicate that the hydrophobic nature of these four positions is quite well conserved in both PYPs and the PAS domain superfamily.

The results reported here extend the same conclusion to Ile39 that we derived before for Asn43 that it is not the residue identity but patterns of side chain interactions which causes the patterns of sequence conservation defining the PAS domain superfamily. The data provide insights into the role of these anciently conserved residues. Substitutions at these two superfamily-conserved residues result in unexpected molecular changes: a largely unaltered native structure and allosteric switching, but slower functional kinetics and altered folding cooperativity.

CHAPTER V

Exploring structure–function rules in the PYP from *Rhodospirillum centenum*

5.1. Introduction

The long-term goal of this part of the thesis is to understand and examine the family-wide patterns in the structure-function relationships that regulate different functional and active site properties of PYP. The approach in the Hoff lab towards understanding the structure-function relationships in the PYP family was to examine the pattern of amino acid sequence conservation, identify the most conserved residues or structurally exciting side chains, and utilize site-directed mutagenesis to study the active site and functional properties of the mutants. It is then possible to correlate the tuning of these properties with different amino acid residues and sequence analysis can reveal the extent to which these structure-function rules can be applied to other members of the PYP family. Numerous studies on site-directed *Halorhodospira halophila* PYP mutants (selected based on its crystal structure) and different N-terminal deletion mutants to affect different functional properties of the protein have been reported.

5.1.1 Site directed mutagenesis and PYP function

Our lab has reported an extensive spectroscopic analysis of a complete alanine mutagenesis scan of Hhal PYP (*1*). That work was based on microscale high – throughput biophysical measurements and the dataset reveals the effect of these point mutations on 6 different properties of Hhal PYP *in vitro*: visible absorbance maximum (λ_{max}), pK_a of the *p*CA, *p*CA fluorescence quantum yield, pB lifetime, thermodynamic stability against unfolding (ΔG_U), and protein production level. Interestingly, 124 mutants except C69A (chromophore cannot attach covalently to the protein (*190*)) were functional and exhibited the characteristic PYP photocycle. Though the protein displays strong robustness against point mutation by maintaining the photocycle of PYP, a majority of the substitutions altered different functional properties, revealing that PYP combines robustness and evolvability with respect to point mutations (*1*).

The analysis of the data from this alanine scan was performed using the framework provided by the crystal structure of Hhal PYP (*72*). A number of conserved residues have been identified in the binding pocket of *p*CA and many residues that are important for determining functional properties are present either at the active site (Cys69, Tyr42, Glu46 (*1, 65, 92, 102*)) or in the immediate vicinity like Met100, Thr50, Arg52, and Tyr98 (*1, 2, 68*). These residues can either form a hydrogen bond with *p*CA or have charge-charge and aromatic ring-ring interactions with it. The alanine mutagenesis scan has identified the entire list of residues that tune the functional properties in the active site of PYP and has interestingly revealed four residues (Phe28, Gly77, Gly82, and Tyr94) far (~10Å away) from the active site which when substituted exhibited altered functional properties. A number of residues that are conserved in the PYP family or in the PAS domain superfamily, structurally important as well as inconspicuous residues have been observed to alter (mostly biased towards increasing) the lifetime of the pB state. Structurally, many of these residues were observed to be at the interface between the N-terminus and the central β-sheet of PYP (*1*).

Another important interaction of interest is between the *p*CA and its hydrogen bonding partners; Glu46 and Tyr42. Mutations at these two positions simultaneously affect three properties of PYP; λ_{max}, pK_a, and pB lifetime (*1, 2, 7, 96, 97, 191, 192*). The specific importance of Glu46 and its role in proton transfer and conformational changes in PYP has been discussed in previous chapters. The effect of substitution at residue 46 varies depending on the mutating residue. For example, the lifetime of the photoactivated pB state in the E46Q mutant is significantly reduced (*65*) whereas an opposite effect is observed in E46D mutant of Hhal PYP (*191*). Upon photoactivation, PYP undergoes large global conformational changes (*5, 82-84, 86*) during the formation of the presumed signaling state of PYP (pB), which corresponds to a molecular signal relayed by the protein to achieve a functional goal in the cell. Proton transfer during the PYP photocycle has been reported to be central in the mechanism for initiating this signal relay and the molecular event/events causing functional structural changes in PYP (*33, 193*). After the intramolecular proton transfer from Glu46 to ionized *p*CA during the PYP photocycle, a negative charge is developed on Glu46. Since it is located in the hydrophobic pocket at the active site this buried charge is energetically unstable (*32, 33*). This electrostatic fault causes large conformational changes (*5, 81, 83-85*) in PYP. This conclusion was experimentally confirmed by time-resolved rapid scan FTIR measurements of wt Hhal PYP and E46Q PYP wherein the large amplitude protein quake was significantly reduced, providing evidence for the proposal that the negative charge on Glu46 acts as an electrostatic epicenter driving the protein quake (*33*). Since Glu46 is almost universally conserved in PYP, this mechanism appears relevant to the entire PYP family (*60, 63*).

5.1.2 Active site pK_a shifts in PYP

A key aspect to consider in the biologically important proton transfer from Glu46 to the *p*CA is the pK_a of the groups involved in the process, and the shifts in pK_a during protein function (*194*). Residues with highly perturbed pK_a values are often critical parts of the active site of

proteins (*195*). Functionally important acidic side chains (Asp/Glu) in several proteins especially at the active site exhibit strongly shifted pK_a values. The two highly studied systems - the bacterial photosynthetic reaction center (RC) and the archaebacterial proton pump bacteriorhodopsin (BR) (*194*) – involve pK_a changes during the functional process that are coupled to electron and proton transfer, respectively (*196, 197*). A study on 24 different proteins of known structure and multiple carboxyl pK_a values was reported wherein the distribution of more than 200 aspartates and glutamates was examined (*198*). The mean pK_a values for the aspartate and glutamate carboxyl groups in the proteins were $\leq$ 3.4 (±1.0) and 4.1 (±0.8) respectively. The estimated range of the model compound values were 3.8 – 4.1 for aspartate and 4.1 – 4.6 for glutamate. In that study, the range of pK_a shifts were from < 2 being the lowest to 9.9 being the highest value (*198*). The distribution of solvent exposed residues was much narrower compared to the buried residues. Of the 10 pK_a values higher than 5.5, 9 were identified to be present in the active site of the proteins and are all buried (*198*).

In water, glutamic acid side chains have a pK_a ~4.4 and *p*CA has a pK_a near 9.0 (*91*). In the pG state of Hhal PYP, the pK_a of *p*CA is strongly shifted from its value in water, with a reported value of 2.8 (*55, 95*). It is known that Glu46 (an acid when dissolved in water) is protonated in pG state in spite of being hydrogen bonded to the phenolic oxygen of the ionized *p*CA chromophore, (a base when dissolved in water). It appears reasonable to view this initial state of PYP, in which the protonation states of Glu46 and the *p*CA are inverted from the situation in water, as a state of PYP that is primed for light-induced proton transfer. Therefore, it is important to investigate the pK_a values of these groups and their functional tuning of PYP.

5.1.3 Gap in knowledge about in vivo signaling mechanism in PYP

Here we use the rich dataset of the mutants of Hhal PYP as a reference set to study the structure-function relationship in the entire PYP family and to examine the degree to which the

structure-function rules on *H. halophila* PYP (known from several reports and the alanine mutagenesis scan from the Hoff lab) are widely applicable to other members of the PYP family.

The PYP from *H. halophila* has emerged to be one of the most well studied model systems for understanding the biophysics of functional protein dynamics. The biological function of Hhal PYP has been proposed to be negative phototaxis (*48*), while for PYP in *Idiomarina loihiensis* (a genetically inaccessible organism) a pharmacological approach using locked chromophore analogs has shown that it regulates biofilm formation (*88*). However, the physiological role/function and the *in vivo* signaling mechanisms remain largely unknown for almost all PYPs, and studying these has become a major challenge in field of bacterial photobiology due to lack of genetic tools for most of these organisms.

5.1.4 R. centenum PYP as a model system to expand the structure-function rules in the PYP family and their relation to signaling

To understand the *in vivo* mechanisms and pathways of PYP, a bacterial system is needed that (i) contains PYP; (ii) has a known and measurable *in vivo* output of PYP signaling; and (iii) is genetically accessible. Currently, the only known PYP that fits these criteria is the PYP from the bacterium *Rhodospirillum centenum* (*101*). The PYP from *R. centenum* (Rcen PYP) regulates the expression of chalcone synthase gene (*chs,* a key enzyme that produces photoprotective pigment in plants (*199*)), and is part of a larger chimeric protein (Ppr). Ppr has PYP domain located at the N-terminus followed by bacteriophytochrome domain and a histidine kinase domain (*101*) (Fig. 5-1). The lifetime of the photoexcited pB state of Rcen PYP is 50 s (*101*). The crystal structure of Rcen PYP is available (*200*) and is highly similar to Hhal PYP (Fig. 5-2). The two proteins share a 42% amino acid sequence identity, 62% sequence similarity, and have the same secondary structure except that in Rcen PYP the 6-residue loop containing Met100 residue between –strands 4 and 5 adopts a slightly different conformation (*2*). Also, structures of various

Hhal PYP mutants have been reported, which facilitates the functional interpretation of the amino acid sequence of Rcen PYP. Here we use the extensive analysis of the alanine mutagenesis scan of Hhal PYP to provide a basis for the functional interpretation of different mutants of Rcen PYP studied here.

Fig. 5-1 Schematic representation of PYP-phytochrome related (Ppr) protein in *R. centenum*. The PYP domain consisting of ~135 amino acids is at the N terminus followed by bacteriophytochrome domain of ~500 amino acids and the C-terminal prokaryotic histidine kinase domain with ~250 amino acids.

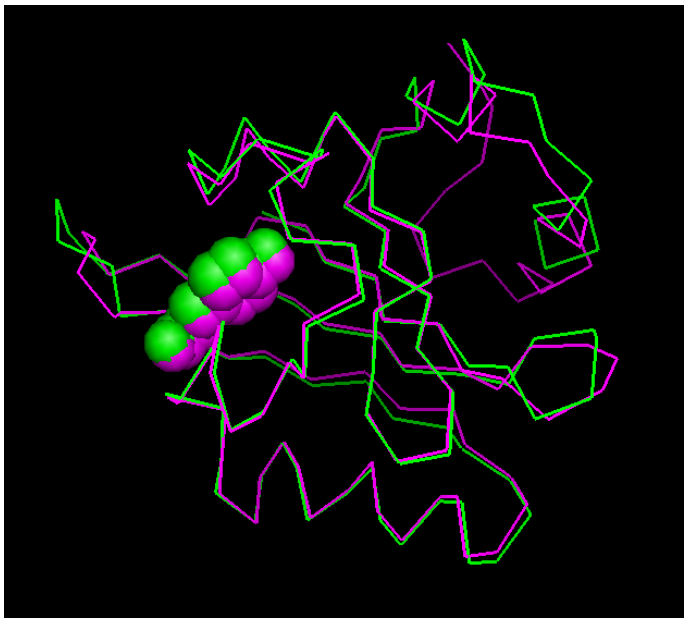


Fig. 5-2 3D structure overlap of Hhal PYP (magenta) and Rcen PYP (green) with the respective chromophores shown as spheres. PDB files 1NWZ for Hhal PYP (*68*) and 1MZU for Rcen PYP (*200*) were used to create the overlap in PyMOL (*201*).

The long-term goal of this work is to understand the *in vivo* signaling mechanism in PYP using the Rcen PYP. The approach used was to connect the extensive knowledge on *in vitro*

biophysics of PYP to *in vivo* signaling. The first step towards achieving this goal, which is also the *main goal of this chapter*, was *to use spectroscopic tools to probe the structural changes in the Rcen PYP photocycle, characterize the biophysical properties of Rcen PYP mutants and determine the degree to which the structure-function rules transfer from Hhal PYP to Rcen PYP*. We selected two classes of Rcen PYP mutants targeted towards understanding specific events in the photocycle of PYP/during signaling in PYP (the residues are numbered as per Hhal PYP amino acid sequence):

1) *Mutants expected to alter the lifetime of the pB state* (presumed signaling state in PYP (*95*)): From the extensive knowledge on Hhal PYP point mutants that alter the pB lifetime to various degrees and based on sequence similarity between Hhal PYP and Rcen PYP we selected the following residues; P6F, V28F, N43A, and M100A (Fig. 5-3).

2) *Mutant/mutants expected to block light-induced partial unfolding*: Several reports suggest that pB photointermediate state is partially unfolded (*5, 81, 83, 85, 86*). As the pB state is considered the signaling state in PYP, our lab has proposed that partial unfolding of the protein is required for signaling in PYP. As discussed earlier, Glu46 is known to drive the light induced protein quake (large conformational changes) in Hhal PYP and the E46Q mutant of Hhal PYP exhibits strongly reduced conformational changes upon the formation of the pB state (*33*) which has also been confirmed by NMR spectroscopy (*193*). Based on this information, we examined the effect of E46Q mutant on light-induced conformational changes in Rcen PYP.

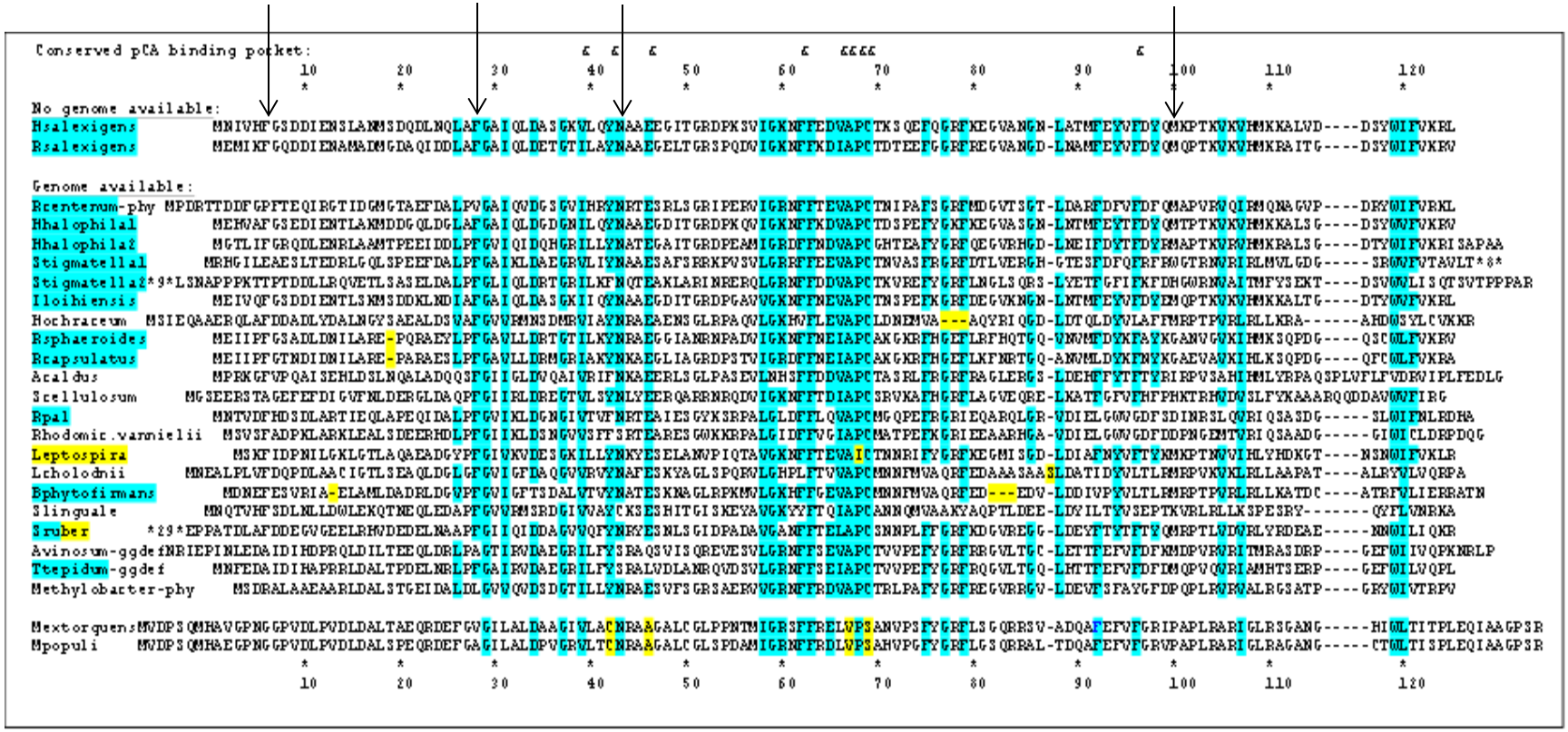


Fig. 5-3 Amino acid sequence conservation in the PYP family (*1, 2*). Alignment of 25 different PYPs shows conserved residues marked in blue. The four residues reported to alter pB lifetime in Hhal PYP and used to alter the photocycle kinetics of Rcen PYP, are indicated by arrows.

5.2. Results

5.2.1. Design of the mutations

Five point mutants (P6F, V28F, N43A, M100A, and E46Q) were designed with the aim at altering the pB lifetime and the degree of light induced conformational changes. The significance of the E46Q mutant has been discussed in the previous section. From previous reports on Hhal PYP two point mutants that strongly reduce the PYP photocycle rate have been identified: N43A and M100A. These residues are obvious candidates for obtaining Rcen PYP mutants with slow pB decay rates. The alanine mutagenesis scan from our lab (*1*) identified several other residues affecting the functional properties; the F6A mutation was observed to significantly reduce the pB decay rate and Phe28 is located far from the active site, and the F28A mutation affects: λ_{max} and pB lifetime. When the amino acid sequences of Hhal PYP and Rcen PYP were compared, Rcen PYP has naturally occurring substitutions at positions 6 and 28, with Phe replaced by Pro and Val respectively. Since the photocycle of Hhal PYP (0.5 s) is faster than Rcen PYP (50 s (*101*)) it was intriguing to check the effect of mutations at these two residues on the pB decay rate in Rcen PYP.

Thus, of the five mutants in Rcen PYP, E46Q is expected to reduce the light induced conformational changes and pB lifetime, N43A and M100A are expected to strongly slow down the rate of pB decay, and P6F and V28F were designed to check if replacing with Phe residues can simulate Hhal PYP, therefore accelerating pB decay.

5.2.2. Purification

Rcen PYP purification was unexpectedly complicated and the method had to be modified as the protein had weak affinity to both anion and cation exchange chromatography columns. During the initial steps of purification partial fraction of the protein was observed to have higher affinity to the anion exchanger and had to be eluted with a gradient 200 – 400 mM of NaCl. As the purity index improved the affinity was reduced almost all of protein could be eluted using 10 mM Tris-HCl buffer. The isoelectric point, calculated based on amino acid sequence of wt Rcen PYP is 5.8 (*2*) and is therefore expected to bind to the anion exchange column. The specific cause of the unpredictable degree of affinity of Rcen PYP to the chromatography columns during multiple runs remains unclear. Possible reasons could be interactions of contaminating proteins with PYP or with the column material. Increasing the pH of the wash buffer (Tris-HCl, pH 10) and running the sample through cation exchangers in dark or dim light conditions did not reduce this issue.

When highly concentrated protein samples were used for runs through weak anion exchange columns, we were successful in obtaining very pure fractions of both wt Rcen PYP and its mutants studied here (probably due to the binding of impurities to the column while Rcen PYP eluted) with optical purity index ~0.38–0.41 (measured by visible absorbance spectroscopy). The purity was confirmed by SDS-PAGE (Fig. 5-4) for wt Rcen PYP and its E46Q mutant. Strong single bands were observed in both PYPs around 16.5 kDa for all five protein concentrations loaded on the gel and almost no other bands of any impurities were seen, confirming the high quality of the protein samples. The protein production level of wt Rcen PYP and the five mutants was similar; the yield of the protein just after reconstitution was estimated to be ~45–50 $OD_{434/451}$/L when the *pyp* gene was overexpressed in *E. coli* and grown in LB medium.

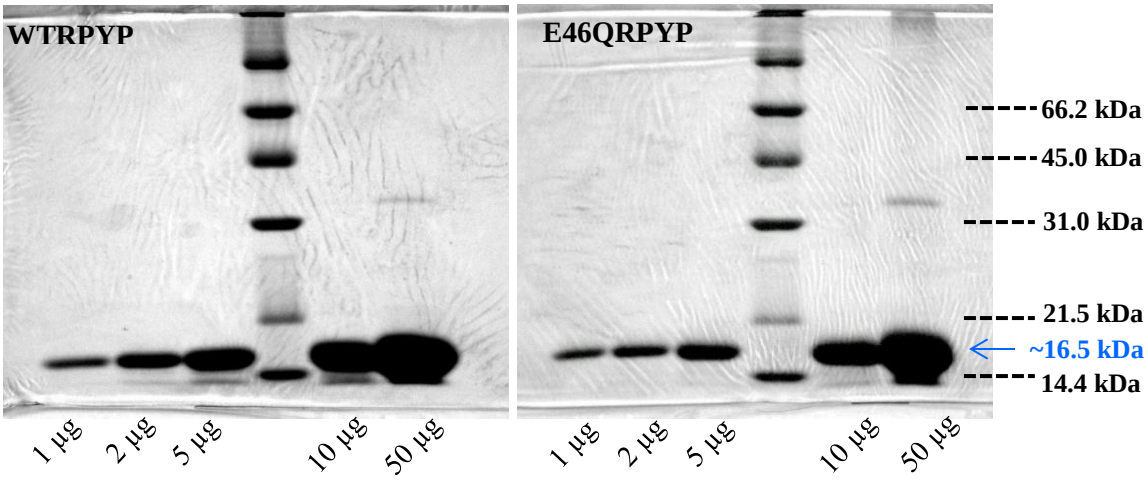


Fig. 5-4 Confirmation of Rcen PYP purification. SDS–PAGE of wt Rcen PYP (PI–0.40) and its E46Q mutant (PI–0.39) loaded at 5 different concentrations as indicated in the figures. The 4th lane shows the molecular weight markers and Rcen PYP with the histidine tag (~16.5 kDa) is indicated by the arrow.

5.2.3. Acid titrations: strongly down-shifted pK$_a$ of the pCA chromophore in Rcen PYP

To study the tuning of the active site properties of the pG dark state of wt Rcen PYP, we measured the absorbance maximum and pK$_a$ of the wt protein and a set of its mutants. The active sites of a wide range of proteins contain functionally important ionizable groups with shifted pK$_a$ values. PYP serves as a model system for functional tuning of the active site residues. The pK$_a$ of the *p*CA in the pG state of PYP is tuned in order to facilitate a proton transfer switch coupled to protein conformational changes. In this study we have examined the active site pK$_a$ shifts in Rcen PYP. We have probed the pK$_a$ of the *p*CA and the ionizable chains in the active site of PYP using acid and alkaline pH titrations respectively. Acid titrations of PYP have been studied by many groups but little data has been reported on alkaline titrations of PYP. Increasing the pH >11 was found to first shift the λ_{max} of Hhal PYP near 400 nm (where ionized *p*CA thioester model compounds absorb), and then spontaneously hydrolyze the thioester bond, shifting the λ_{max} to 340 nm (*76*). Deprotonation of two possible groups, Glu46 and Tyr42, could cause A_{400} formation. In

view of the unusual protonation state of Glu46 in the pG state of PYP we used the E46Q mutant of Rcen PYP to help unravel the high pH transitions.

The UV/vis absorbance spectra of all five mutants of Rcen PYP (Fig. 5-5A) contain an absorbance band with maxima in the range 434–451 nm. The absence of a band near 350 nm in these data indicates that all of these mutants contain a natively folded active site with adequate stability. The absorbance spectrum of wt Rcen PYP has a λ_{max} at 434 nm, identical to the value that was previously reported (*101*). The absorbance spectra of P6F, V28F, N43A, and M100A are essentially identical to wt Rcen PYP with a λ_{max} at 434 nm. However, in the case of the E46Q mutant the value of λ_{max} is significantly red shifted to 451 nm. All five mutants retain a spectral shape of the *p*CA absorbance band that is highly similar to wt Rcen PYP.

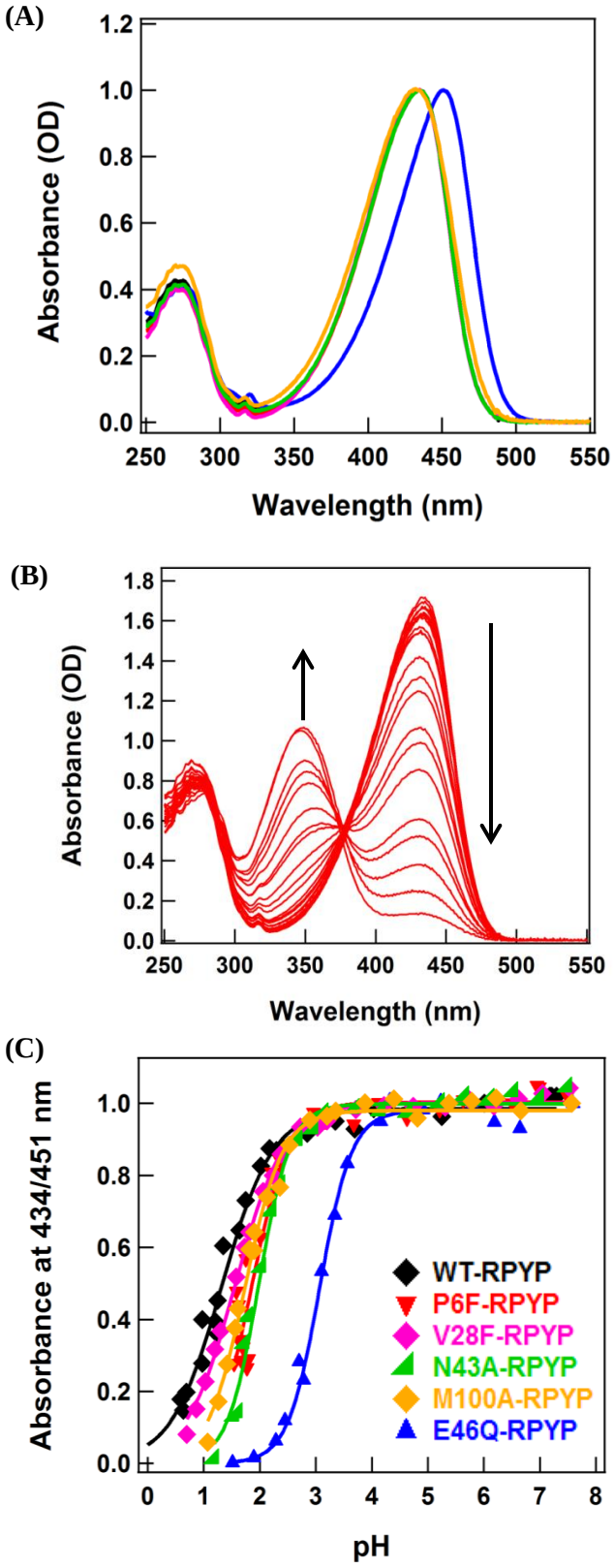


Fig. 5-5 Absorbance and *p*CA pK_a of Rcen PYP. (A) UV/vis absorbance spectra of the pG dark state of wt Rcen PYP and its mutants in 10 mM Tris-HCl buffer, pH 7.5 at RT. (B) Low pH dependence absorbance spectra of purified V28F Rcen PYP as an example acid titration. (C) The pH dependence of the absorbance spectra of wt Rcen PYP and its mutants in the pG dark state. The normalized titration curves were measured at absorbance maxima; 434 nm for wt, P6F, V28F, N43A, and M100A Rcen PYP and at 451 nm for E46Q Rcen PYP derived from the visible absorbance spectra.

w/His tag	Abs. max (nm)	Molar Extinction Coefficient (ε), $M^{-1}cm^{-1}$	pK_a	*n* value
WT	**434**	**36,346**	**1.32 ± 0.02**	**0.95**
P6F	**434**	**37,422**	**1.86 ± 0.04**	**1.30**
V28F	**434**	**35,322**	**1.56 ± 0.01**	**0.98**
N43A	**434**	**37,422**	**1.98 ± 0.02**	**1.54**
M100A	**434**	**33,384**	**1.73 ± 0.02**	**1.29**
E46Q	**451**	**37,422**	**3.07 ± 0.02**	**1.52**

Table 5-1 Active site properties of wt Rcen PYP and its mutants as measured by UV/visible absorbance spectroscopy.

The extinction coefficient of the characteristic *p*CA absorption band for all protein samples was also estimated. For wt Rcen PYP, the ε was estimated to be 36,346 $M^{-1}cm^{-1}$ and the values for all mutants of Rcen PYP studied here had very similar values, except for M100A mutant, which has a slightly reduced value of 33,384 $M^{-1}cm^{-1}$ (Table 5-1). In addition to the calculation of ε of the 434 nm peak, the ε at the 280 nm peak (in water) was calculated based on its amino acid sequence using the ExPASy ProtParam tool to provide an estimate of the expected purity index of Rcen PYP. The extinction coefficients of Hhal PYP were used as the initial reference value. For Hhal PYP, at 280 nm the calculated ε (12,950 $M^{-1}cm^{-1}$) and the experimentally observed ε (19,100 $M^{-1}cm^{-1}$) are somewhat different presumably because of the contribution of the *p*CA chromophore to the 280 nm band. For Rcen PYP the calculated ε at 280 nm is observed to be 8,480 $M^{-1}cm^{-1}$. Therefore, the difference between calculated ε_{280} of Hhal PYP and Rcen PYP (4,470 $M^{-1}cm^{-1}$) was subtracted from the experimental value ε of Hhal PYP (19,100 $M^{-1}cm^{-1}$) to obtain an estimate of ε of Rcen PYP (14,630 $M^{-1}cm^{-1}$). Thus, the expected purity index ($\varepsilon_{280}/\varepsilon_{434}$) of wt Rcen PYP is (14,630/36346) 0.40 which is identical to the value obtained from the purification procedures used in this study.

To examine the pK_a of the *p*CA chromophore in Rcen PYP, we studied the low pH titration behavior of its pG dark state. Acid titrations were performed on the wt and five mutant protein samples and the dependence of the absorbance spectra of these Rcen PYPs on pH were recorded. The continuous titration by the step-wise addition of HCl to a single PYP sample was observed to cause a strong increase in light scatter at low pH values due to protein aggregation. To avoid this complication, the method of titration was modified to measure the absorbance at different pH values with individual PYP samples immediately after the mixing of each sample with acid. These acid titration experiments resulted in the formation of the blue shifted species with λ_{max} near 348 nm (Fig. 5-5B) caused by the protonation of the *p*CA. These spectral changes were observed to be mostly isosbestic in each of the 6 Rcen PYP samples studied. It should be noted that due to a contribution of protein aggregation and the resulting light scattering, the precision of the isosbestic point was somewhat reduced. The dependence on pH of the absorbance of the *p*CA at the λ_{max} of the native species was fit with the Henderson-Hasselbalch equation (Fig. 5-5C). From these fits, the pK_a and *n* values (steepness) for each protein sample were derived (Table 5-1). All transitions could be described by a single pK_a value and *n* values ranging from 0.95 to 1.54 (Table 5-1). The pK_a of wt Rcen PYP was observed to be as low as 1.32. The pK_a of P6F, V28F, N43A, and M100A were only slightly higher than that of wt Rcen PYP, whereas the pK_a of E46Q mutant was upshifted by 1.8 pH units (to 3.1).

5.2.4. Alkaline titrations: strongly up-shifted pK_a of Glu46 in Rcen PYP

We also examined the high pH titration of the pG dark state of wt Rcen PYP and its E46Q mutant (Fig 5-6). Both the proteins were observed to retain a native absorbance band up to pH 14 but underwent thioester hydrolysis at a rate that becomes faster at increasingly high pH values. The extraction of the pK_a value for the alkaline titration should be performed with great care because it is not obvious how to separate the titration of group at high pH from thioester hydrolysis. However, since both wt Rcen PYP and its E46Q mutant retain their native absorbance

band, the two ionizable groups in the *p*CA binding pocket (Glu46 and Tyr42) have a pK_a value of > 13.5. In a gradual steady state alkaline titration it was difficult to separate the titratable acidic side chain from thioester hydrolysis. Therefore, we examined the effect of high pH on stability and the kinetics of pG decay of Rcen PYP in dark and also the transitions of the absorbance maximum of Rcen PYP upon a (manually performed) sudden jump to high pH.

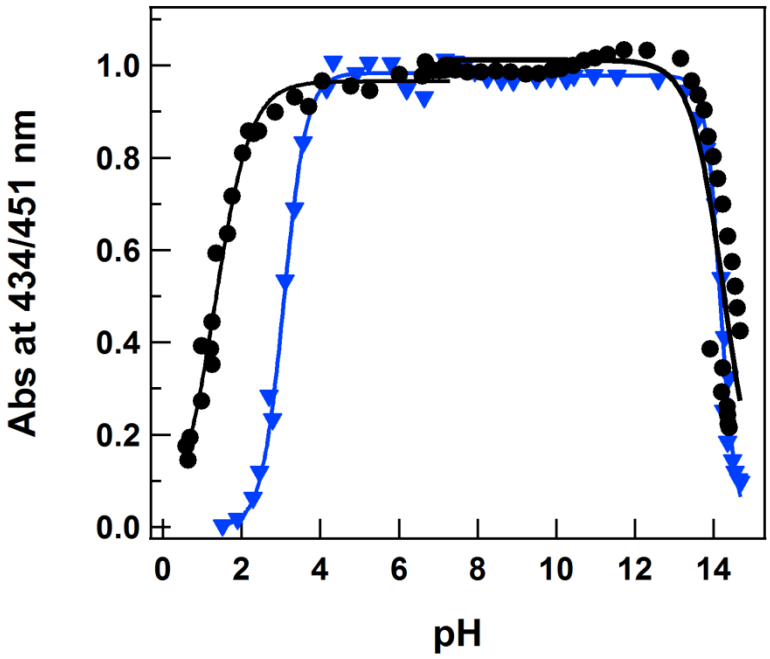


Fig 5-6 The pH dependence of the absorbance spectra of wt Rcen PYP (black) and the E46Q mutant of Rcen PYP (blue) in the pG dark state. The normalized curves for acid and alkaline titrations were measured at absorbance maxima; 434 nm for wt, Rcen PYP and at 451 nm for E46Q Rcen PYP derived from the visible absorbance spectra. Data for pH 8-14 was obtained from gradual titrations.

The kinetics of high pH induced thermal decay of wt Rcen PYP and its E46Q mutant due to thioester hydrolysis were studied for a pH range 10 – 14. Measurements above pH 14 could not be done due to the limitations of the pH meter. In the pH range 10-11.2 both wt Rcen PYP and E46Q were quite stable after 1 h incubation and showed almost no loss of the pG state. Above pH 11.2, the rate of pG decay gradually accelerated and the traces for all pH values were plotted for

wt Rcen PYP at λ_{max} 434 nm (Fig. 5-7A) and for E46Q at λ_{max} 451 nm (Fig. 5-7B). The rate constants for thioester hydrolysis in wt Rcen PYP and E46Q mutant are quite similar (Fig. 5-7C). Importantly, at the highest pH values used here (pH 14), the time constant for thioester hydrolysis is ~75 seconds. Therefore, the first spectra measured upon the manually performed high pH jumps, which were acquired approximately 15-20 seconds after the pH jump, provide a reasonable estimate of the absorbance spectrum of the sample before thioester hydrolysis occurs. In Hhal PYP a shift in λ_{max} from 446 nm to 405 nm occurs above pH 12, which was assigned to the deprotonation of Glu46 (Qureshi and Hoff, unpublished results). We therefore examined the absorbance spectra of the Rcen PYP immediately following the pH jump to examine if the formation of the 405 nm species was detected.

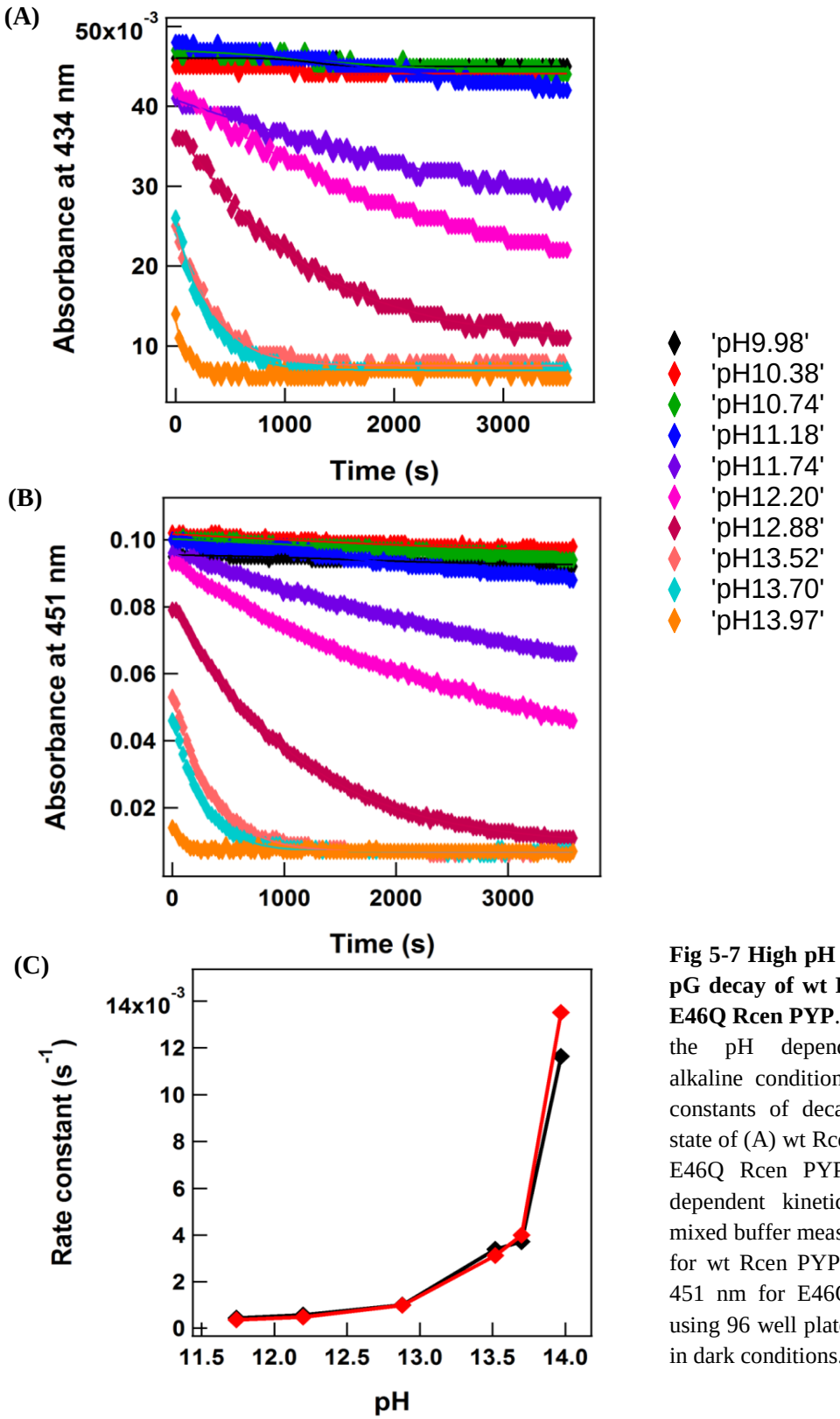


Fig 5-7 High pH dependence of pG decay of wt Rcen PYP and E46Q Rcen PYP. Comparison of the pH dependence (highly alkaline conditions) of the time constants of decay of pG dark state of (A) wt Rcen PYP and (B) E46Q Rcen PYP and (C) pH dependent kinetics in 50 mM mixed buffer measured at 434 nm for wt Rcen PYP (black) and at 451 nm for E46Q mutant (red) using 96 well plate reader for 1 h in dark conditions.

Two different pH jumps were performed for wt Rcen PYP from pH 10 to 13.8 (Fig. 5-8A) and 14.0 (Fig. 5-8C). The rapid mixing pH jump experiments revealed a very small red shift in the absorbance maximum of Rcen PYP from 434 nm to 436 nm. The redshift of Abs max occurs within 15-20s of the pH jump and importantly no other species with a λ_{max} near the 400 nm occurred. Subsequently, an isosbestic transition involving the formation of a species absorbing near 340 nm was detected (Fig. 5-8A, C), which can be attributed to thioester hydrolysis. The kinetics of this transition for pH jumps in shown in Fig. 5-8 B, D and time constants for the pG decay were calculated to be 94 s and 86 s for pH 13.8 and 14 respectively, very similar to the values depicted in Fig. 5-7.

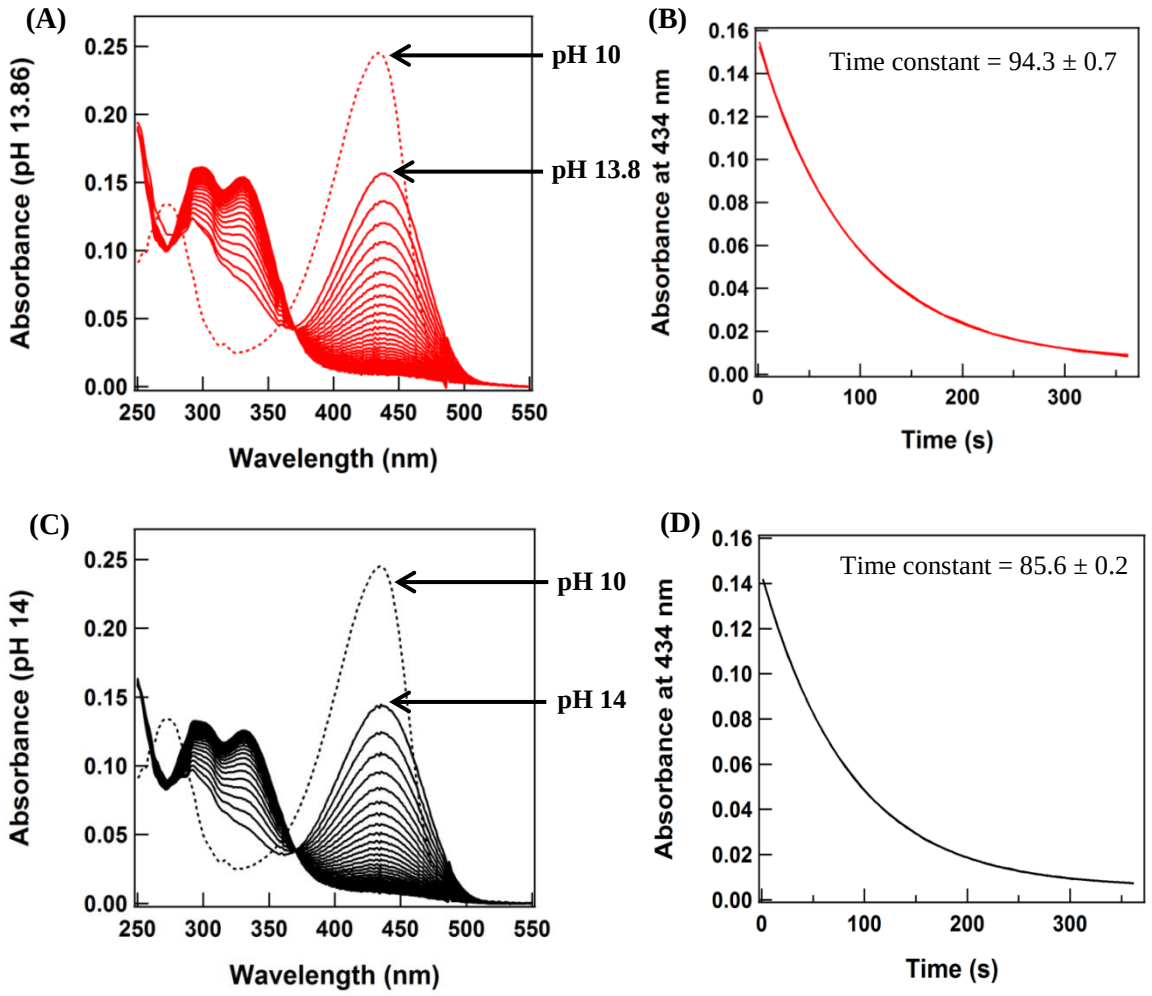


Fig. 5-8 UV/vis pH jump spectroscopy of wt Rcen PYP. The full absorbance spectra (A, C) and kinetics at the indicated wavelengths (B, D) upon exposing the protein to a jump in pH from 10.0 to 13.8 (A, B) and from pH 10.0 to 14 (C, D) are shown. The dashed spectra in A and C represent the absorbance spectra at pH 10.0 before the pH jump.

5.2.5. Photochemical activity of R. centenum PYP and its mutants:

The photochemical activity of wt Rcen PYP and its five mutants was studied. These protein samples all exhibited a photocycle induced by illumination with blue light. In line with previous reports (*75*, *101*) the light-dark difference spectra for wt Rcen PYP contain a negative peak near 434 nm due to the bleaching of the pG state and two positive peaks (Fig. 5-9A). At neutral pH, the positive peak near 340 nm can be attributed to the formation of a pB intermediate,

while the positive peak near 475 nm indicates the formation of a pR-like species. The five mutants in general exhibited very similar properties, except that for E46Q PYP the peak values were shifted to 451 and 496 nm, and that in V28F Rcen PYP difference spectrum the 340 nm peak was much more pronounced and was slightly shifted near 350 nm (probably due to higher percentage of pB formation), while a signal near 475 nm was not observed. In the M100A mutant the 475 nm peak was reduced compared to wt and most of the other mutants.

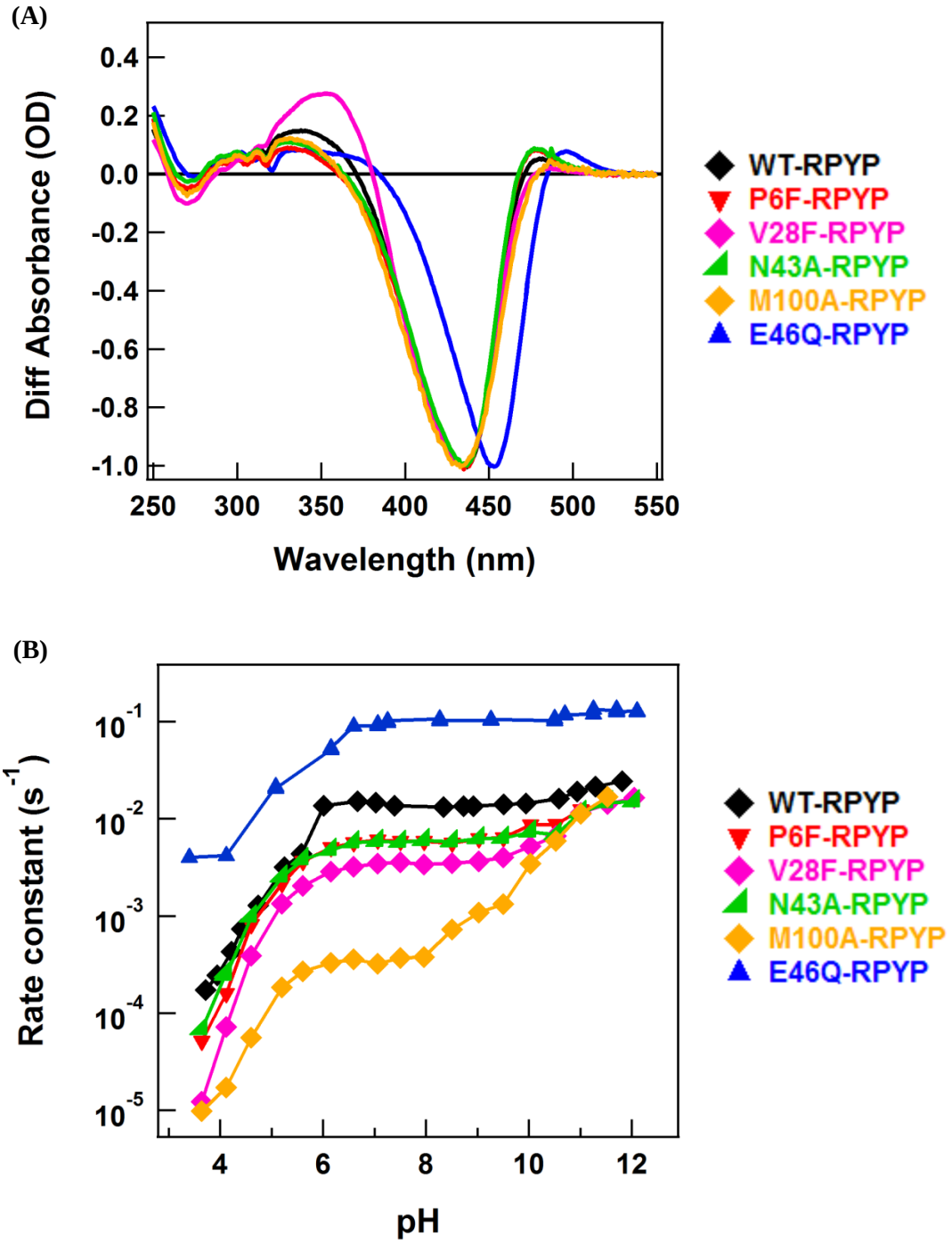


Fig. 5-9 Photochemical activity of Rcen PYP. (A) The amplitude normalized UV/vis difference absorbance spectra during pB to pG photocycle transition of wt Rcen PYP and its mutants in 10 mM Tris-HCl buffer, pH 7.5 at RT. (B) Effect of pH on kinetics of pB decay/recovery of pG state of wt Rcen PYP and the mutants in 100 mM mixed buffer at different pH values measured by UV/vis spectrophototmeter after continuous illumination with blue light.

For wt Rcen PYP at neutral pH, the pB state decayed to the initial pG state with a time constant 67 s, very similar to the previously reported value of 50 s (*101*). At neutral pH the rate of pG recovery was slowed down by a factor 2.5 for the P6F and N43A mutants, a factor 4.3 for the V28F mutant, and a factor 45 for the M100A mutant. In contrast, E46Q mutant increased the rate of pG recovery by a factor 6 times compared to wt Rcen PYP (Table 5-2). This analysis shows that four of the mutants slow down pG recovery to different degrees while one mutant (E46Q) accelerates pG recovery.

w/His tag	WT	P6F	V28F	N43A	M100A	E46Q
pB lifetime at neutral pH (s)	67	164	287	170	3000	10

Table 5-2 pB lifetimes of wt Rcen PYP and its mutants measured in mixed buffer at room temperature.

We also examined the pB decay kinetics for wt and the mutants over a pH range from 3.5 to 12.0 (Fig. 5-9B). While for wt Rcen PYP the pG recovery rate remains almost unchanged in the pH range 6-10, the rate was decreased by approximately 2 orders of magnitude upon reducing pH from 6 to 3.5. The pH dependence of pG recovery kinetics of the P6F, V28F, N43A, and E46Q mutants exhibited a very similar pattern. For the M100A mutant the effect of pH on pB decay rate was stronger and more complex. The pH dependence of the rate constants for all PYPs could not be described by the bell shaped curve used previously for the pH dependence of the pG recovery kinetics of Hhal PYP (*65*). For all PYPs it was observed that above pH ~9 the rate of pB decay was increased; this effect was the strongest in the M100A mutant and corresponded to the lower percentage of photobleaching in the photostationary state at higher pH values.

5.2.6. The pH dependence of the thermal equilibrium between the pR and pB species in Rcen PYP

The light-dark absorbance spectra of Rcen PYP exhibit positive peaks for both a pB-like and a pR-like species (Fig. 5-10A), and the pH dependence of the light-dark difference spectra (Fig. 5-10B) indicate that the relative contribution of these two species is pH dependent. At low pH the pB state predominates, while at high pH the pR intermediate predominates. These data indicate that a pH-dependent equilibrium between a pR-like state and a pB-like state exists. The pH dependence of the thermal equilibrium between the pR and pB species appears to be governed by two distinct transitions with approximate apparent pK_a values of ~4.5 and ~11.5. Above pH 12.5 the measurements were not reliable due to probable thioester hydrolysis caused by the combined effect of light and high pH.

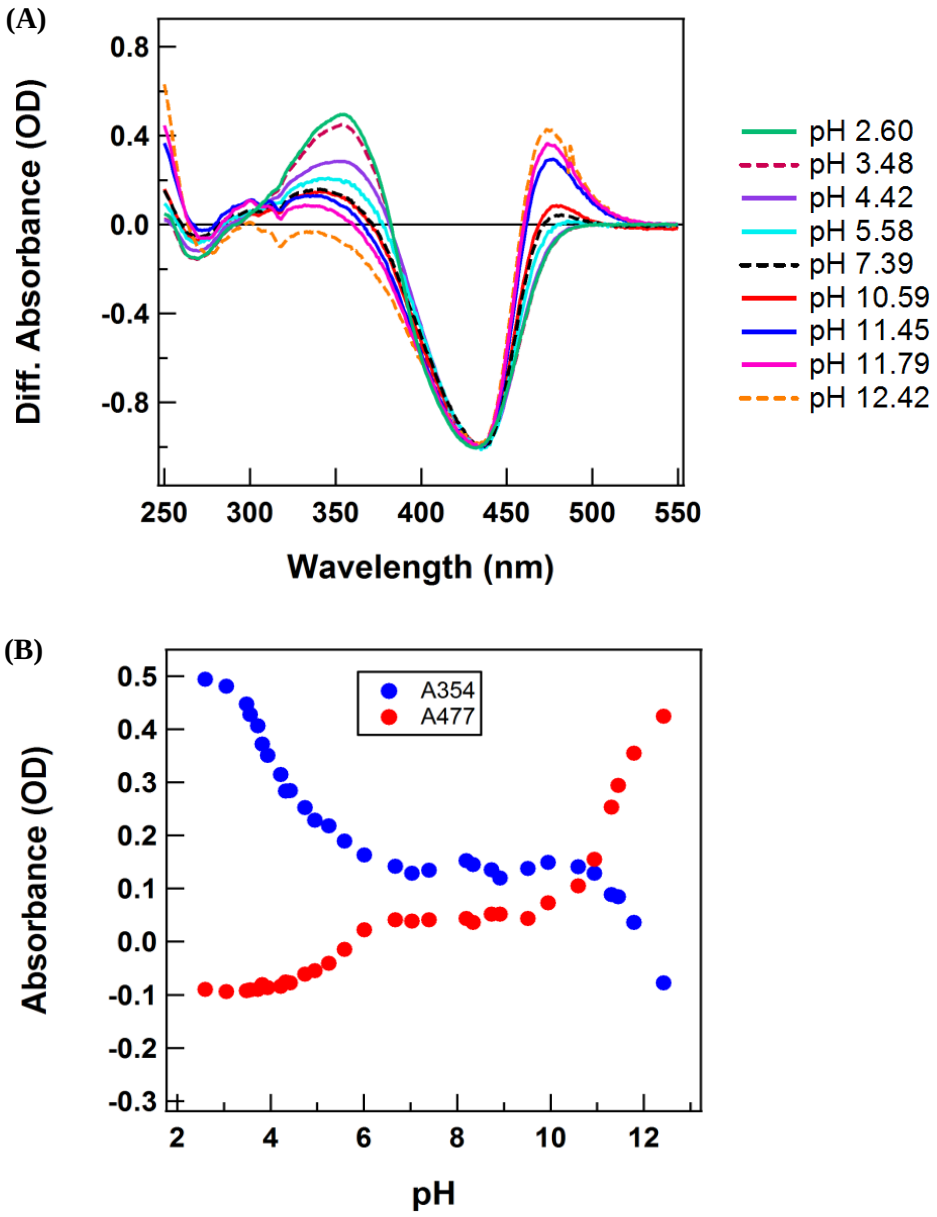


Fig. 5-10 The pH dependence of the normalized light-dark difference spectra for wt Rcen PYP. (A) Light-dark difference spectra normalized at 1 at the negative peak near 434 nm at pH values from 2.6 to 12.4. (B) The pH dependence of the signals at 354 nm (to probe pB) and 477 nm (to probe pR) in the normalized light-dark difference spectra.

A first approximation for the estimated absorbance spectra of the pB and pR states of Rcen PYP can be obtained based on modest assumptions (Fig. 5-11). Their absorbance maxima are 356 nm and 466 nm, respectively. Both species appear to have somewhat reduced maximal extinction coefficients, similar to the case of Hhal PYP. The resulting spectra indicate that at pH

2.6 the amount of pR species is negligible, while at pH 12.4 still a small amount of pB appears to be present, as indicated by the small ~350 feature in the "pure pR" spectrum. This conclusion appears to match the pattern in Fig 5-10B.

Note in Fig.5-11 that the absorbance of the pR species at 434 nm is fairly small, indicating that the normalization of the light-dark difference spectra in Fig. 5-10A is reasonable.

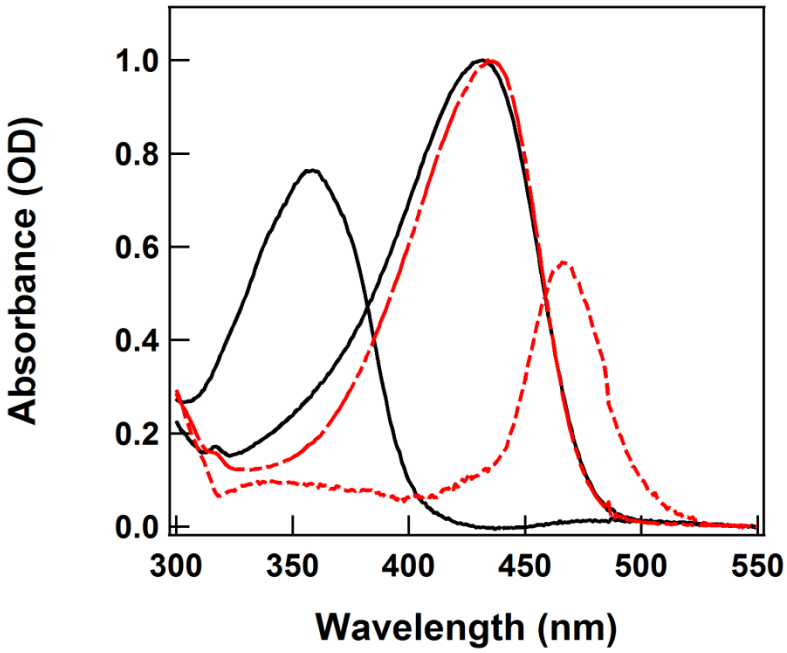


Fig. 5-11 Reconstructed pure species of the pB (black) and pR species (red dashed) in Rcen PYP. It was assumed that at pH 2.6 the contribution of pR is negligible, and at pH 12.4 the contribution of pB is negligible. The fraction cycling (amount of pG state to subtract) from the light-dark difference spectra for the pB species was estimated by making the absorbance above 400 nm near zero. The fraction cycling (amount of pG state to subtract) for the pR species was estimated by judging a reasonable spectral shape for pR (no positive or negative shoulder at the pG absorbance maximum) and a reasonable extinction coefficient relative to the pG state. The spectra for the pG dark state at 2.6 (black) and 12.4 (red dash-dot) is also indicated.

5.2.7. Folding and stability of Rcen PYP derived from denaturant titrations

We examined the stability and equilibrium unfolding of wt Rcen PYP and its mutants in guanidinium hydrochloride (Gdm-HCl) titrations of the pG dark state (Fig. 5-12). The folding of wt Hhal PYP follows two-state behavior (5) as described in previous chapters. Similarly, we found that the denaturant unfolding curves of wt Rcen PYP and its mutants could be described as

a two-state transition wherein the fully folded pG state (λ_{max} 434/451 nm) is converted to the fully denatured pG state (λ_{max} near 340 nm). Transitions for wt Rcen PYP and its mutants were isosbestic within the signal to noise of the measurements, with an isosbestic point near 372 nm (Fig. 5-12A).

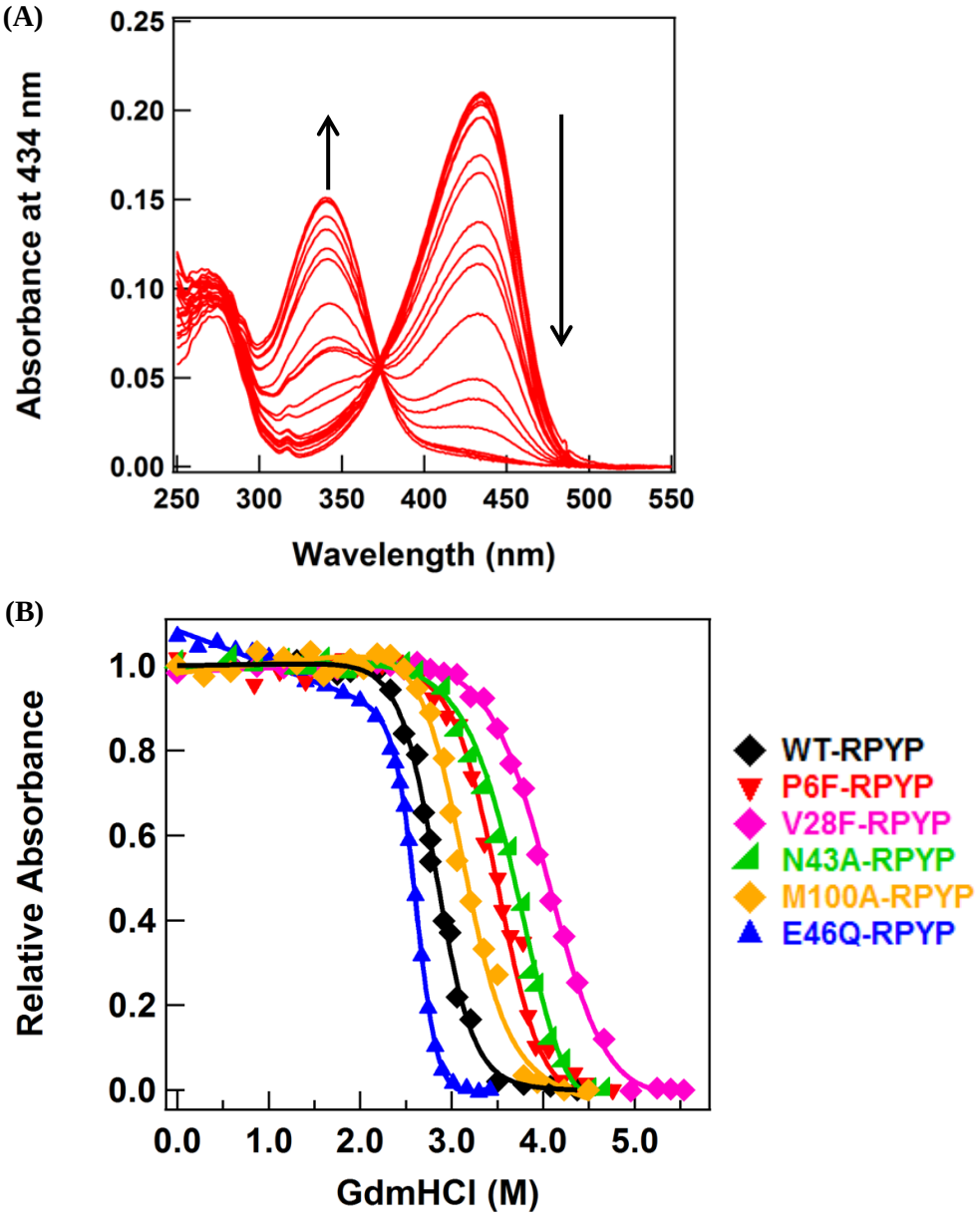


Fig. 5-12 Guanidinium hydrochloride denaturation of the pG dark state of wt Rcen PYP and its mutants, in 10 mM Tris-HCl buffer, pH 7.5 as monitored by UV/vis absorbance spectra. (A) Dependence of the absorbance spectrum of wt Rcen PYP on Gdm-HCl as an example denaturant titration. (B) The normalized titration curves were measured at 434 nm for wt, P6F, V28F, N43A, and M100A Rcen PYP and at 451 nm for E46Q Rcen PYP derived from the visible absorbance spectra.

The denaturant titration data were analyzed quantitatively to determine $\Delta G_{U,}$ the *m* value, and the resulting C_m (midpoint concentration of Gdm-HCl) where half of the PYP molecules are unfolded (Table 5-3). Mutants P6F, V28F, N43A, and M100A exhibited ΔG_U values very similar to wt Rcen PYP but the *m* values (steepness of the slope) were reduced in all mutants with a significant reduction in V28F and N43A. However, the C_m was observed to be higher in all four mutants compared to wt Rcen PYP, again with the highest values found in V28F and N43A (Table 5-3). As a result of these two compensating effects, the ΔG_U of the mutants was largely unchanged. In sharp contrast, the E46Q mutant exhibited surprisingly high values for ΔG_U and *m* value (almost twice of that of wt Rcen PYP) but a reduced C_m value (Table 5-3).

***Rcen* PYP w /His tag**	**WT**	**E46Q**	**P6F**	**V28F**	**N43A**	**M100A**
ΔG_U (kJ/mol)	**34.61 ± 2.83**	**54.94 ± 1.93**	**34.43 ± 2.49**	**30.66 ± 1.38**	**30.06 ± 2.81**	**32.79 ± 3.79**
***m* (kJ/mol/M)**	**12.23 ± 1.05**	**20.96 ± 0.77**	**9.73 ± 0.79**	**7.50 ± 0.38**	**7.67 ± 0.96**	**10.56 ± 1.32**
C_m (M)	**2.83 ± 0.01**	**2.57 ± 0.01**	**3.49 ± 0.02**	**4.04 ± 0.01**	**3.68 ± 0.02**	**3.16 ± 0.02**

Table 5-3 Effect of different point mutants on Rcen PYP by denaturant (Gdm-HCl) induced unfolding, measured at pH 7.5.

5.2.8. Light induced structural changes in Rcen PYP

The light induced conformational changes in wt Rcen PYP and the E46Q mutant were measured by steady state light induced FTIR difference spectroscopy. We analyzed the pB-pG difference spectra of wt Rcen PYP and the E46Q mutant for two specific spectral markers of light induced conformational changes in PYP; the 1726 cm^{-1} negative band assigned to COOD group of Glu46 (*32*) and the positive band at 1624 cm^{-1} produced mainly by the C=O stretching of the

protein backbone (amide I) (*181, 182*). The pB – pG FTIR difference spectrum of wt Rcen PYP showed the characteristic amide I band at 1625 cm^{-1} while the Glu46 band was downshifted from 1726 to 1723 cm^{-1}. The pB-pG difference spectrum of the E46Q mutant revealed significantly reduced amplitudes of amide I signals, which indicates reduced conformational changes during the transition from pG to pB. This result establishes that the electrostatic epicenter, (i.e. the COO^- buried charge of Glu46) that drives the large amplitude protein quake is functioning in Rcen PYP. Also, we were able to assign the important 1723 cm^{-1} signal in Rcen PYP to Glu46 since this signal was not observed in the E46Q mutant.

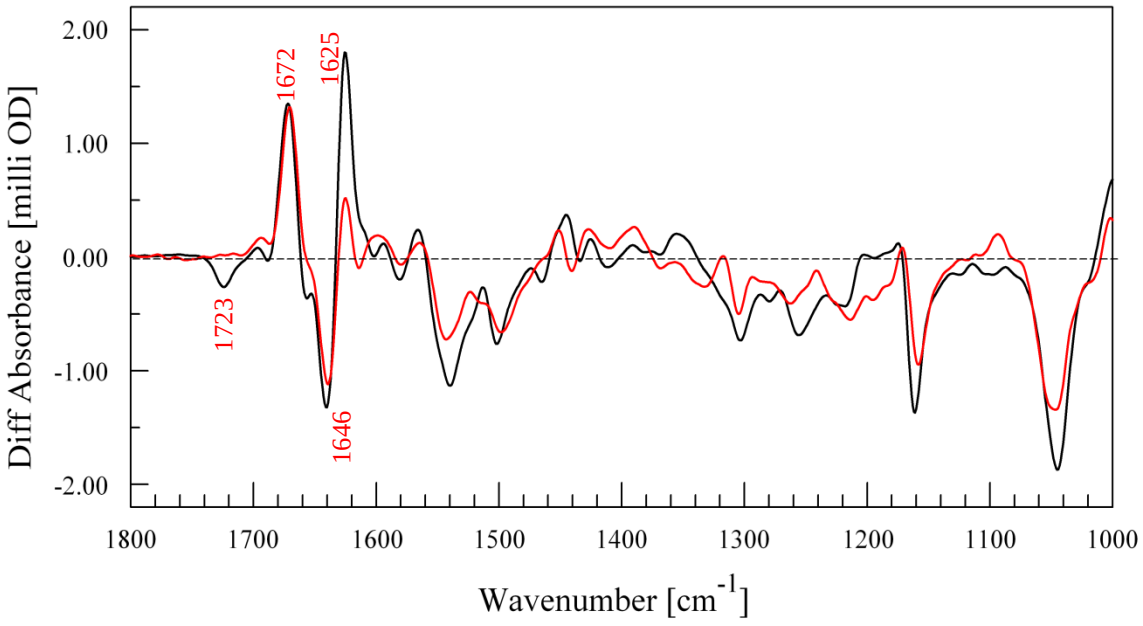


Fig. 5-13 Structural changes upon pB formation in Rcen PYP detected by time-resolved FTIR difference spectroscopy. The above panel shows the pB–pG difference spectra of wt Rcen PYP (black) and its E46Q mutant (red) at pH 7.5 in D_2O (*Data measured by Ningning Xu*).

5.3. Discussion

To study the structure-function relationship across the PYP family, we use the PYP from *Rhodospirillum centenum* as a test case. Taking advantage of the very similar 3D structure and

42% sequence identity to *H. halophila* PYP, we used the extensive knowledge and data available on the site-directed mutants on *H. halophila* PYP to predict substitutions that affect the functional properties of Rcen PYP. Ultimately, we aim to correlate the functional properties of PYP with amino acid sequences and understand the structure-function relationships in the entire PYP family of photoreceptors. While *R. centenum* PYP is a strong candidate for studying *in vivo* signaling transduction mechanisms in photoreceptors and other signaling proteins, no mutagenesis studies probing the role of functionally important residues are available in literature. In this chapter, we have reported a spectroscopic characterization of wt Rcen PYP and five different point mutants; (P6F, V28F, N43A, M100A, and E46Q) that were selected to probe different events during the photocycle of PYP. This allows the effects of each of these mutants on the functional and active site properties of Rcen PYP to be examined and the degree to which they are comparable to Hhal PYP to be determined. Analysis of these data was performed in two steps. First, we examined the different properties of wt Rcen PYP and the variation in spectroscopic properties in all five mutants. Next, we compared the differences observed in the functional properties of wt Hhal PYP and the corresponding Hhal mutants for the case of Rcen PYP.

Tuning of active site/functional properties of pG dark state

***Spectral tuning* (λ_{max})**

The absorbance maxima of the pG state of wt Rcen PYP and four mutants (P6F, V28F, N43A, and M100A) were observed to be identical at 434 nm while the E46Q mutant exhibited a 17 nm (868 cm^{-1}) red shifted spectrum with λ_{max} at 451 nm. In Hhal PYP, substitution at Glu46 (E46Q), Phe6 (F6A), and Met100 (M100A) have been reported to show a similar change in λ_{max} of Hhal PYP. F6A (*1, 202*) and M100A (*97*) have a λ_{max} at 446 nm similar to wt Hhal PYP whereas, while the E46Q mutant shows a strong 16 nm (776 cm^{-1}) redshift from 446 nm to 462 nm (*65*). In Hhal PYP, mutation of Phe28 and Asn43 resulted in small blue shifts in the

absorbance spectra with λ_{max} at 444 nm (*1*) and 441 nm (*103*) respectively, but had no effect on λ_{max} (434 nm) in Rcen PYP (Fig. 5-5A, 5-14). Thus, three out of the five substitutions in Rcen PYP have essentially the same effect as on Hhal PYP, whereas the small blue-shift caused by the N43A mutation was not observed for Rcen PYP.

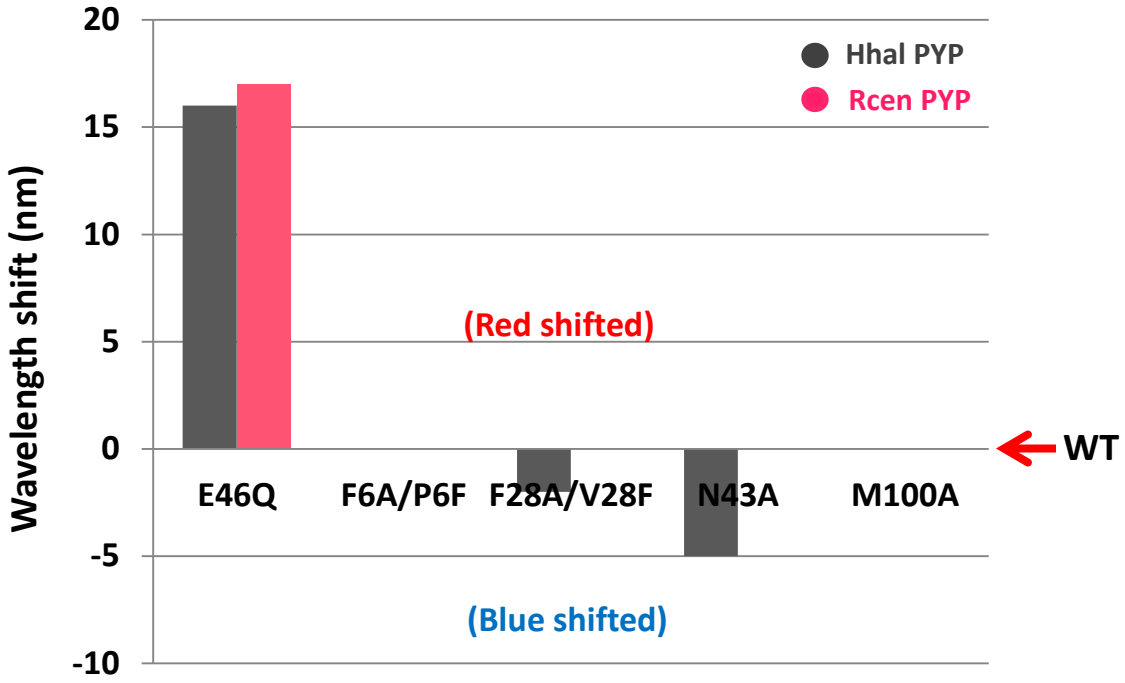


Fig. 5-14 Comparison of the effects of the respective mutants on absorbance maximum of Hhal PYP and Rcen PYP.

pK_a tuning of pCA in the pG dark state

Acid titration data shows an exceptionally low *p*CA pK_a of wt Rcen PYP (1.32) and is the lowest value reported for pK_a of any PYP. Since this transition is spectroscopically isosbestic and its *n* value is very close to 1, the most parsimonious interpretation is that this value provides a good estimate for the pK_a of the *p*CA in wt Rcen PYP. The pK_a of the *p*CA in Rcen PYP was slightly increased by the P6F, V28F, N43A, and M100A mutations and the trend of these increases was similar to Hhal PYP mutants except for Phe6, where F6A did not affect the *p*CA pK_a of Hhal PYP while in Rcen PYP P6F mutant was observed to increase the pK_a by > 0.5 pH

unit (Table 5-4). The pK_a of the *p*CA in the E46Q Rcen PYP mutant was upshifted by 1.75 pH units compared to wt Rcen PYP, similar to what is observed in Hhal PYP where the E46Q mutant also shows an upshift in pK_a of 2.7.

PYP w /His tag	WT	E46Q	P6F/F6A	V28F/F28A	N43A	M100A
Rcen PYP, pK_a	1.32	3.07	1.86	1.56	1.98	1.73
n value	0.95	1.52	1.30	0.98	1.54	1.29
Hhal PYP, pK_a	2.8 (*55*)	5.5 (*97, 102*)	2.8[a] (*1*)	2.8[a] (*1*)	4.6 (*103*)	3.8 (*97*)
n value	1.3	1	ND	ND	1.2	1

Table 5-4 Summary of the effect of respective mutants on the pK_a of Hhal PYP and Rcen PYP. [a] The pK_a values for both F6A and F28A in Hhal PYP were reported to be unchanged and if small changes in pK_a occur they would not be detected in high-throughput measurements, ND – not determined.

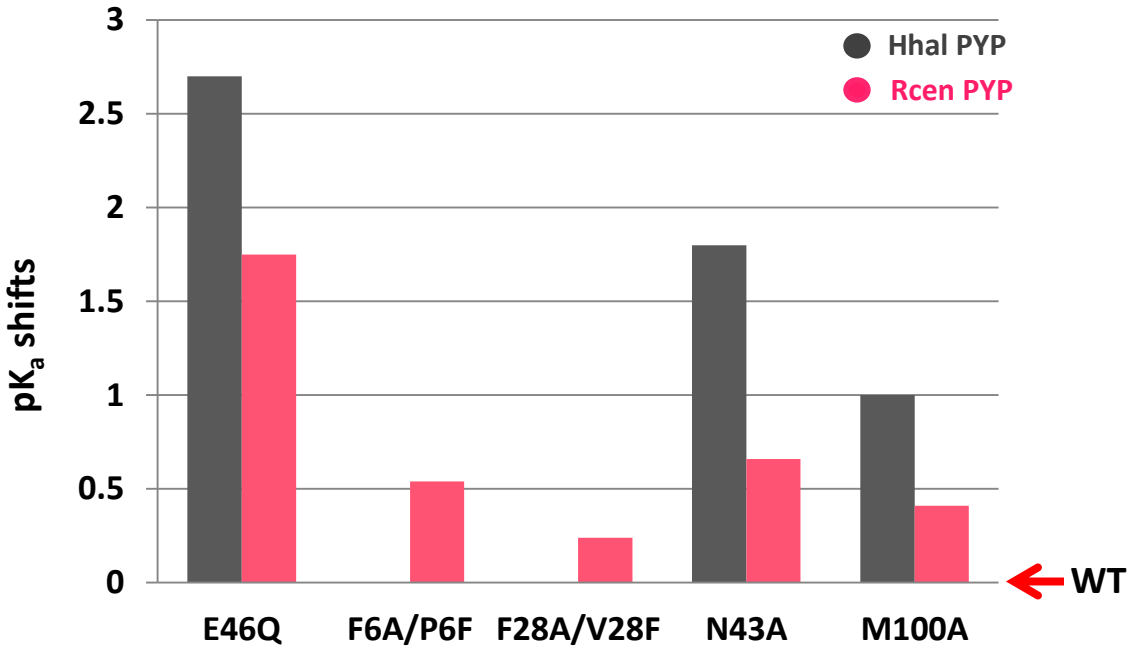


Fig. 5-15 pK_a shifts observed in different mutants of Hhal PYP and Rcen PYP.

Photoactivity of Rcen PYP and the effect of mutations on the lifetime of the light induced pB state

All five mutants expressed well in *E. coli*, retained their characteristic yellow color, and were found to be photoactive. Mutagenesis studies on Hhal PYP have revealed that substitutions at four residues; Phe6 (Pro6 in Rcen PYP), Phe28 (Val28 in Rcen PYP), Asn43, and Met100, at neutral pH slow down the photocycle rate to different degrees, while the E46Q mutation speeds up the rate of the photocycle at neutral pH (Table 5-2, Fig. 5-9B). When compared to Hhal PYP, for the substitutions at all five residues, the trends in the mutations are the same for Rcen PYP although quantitative differences exist (Fig. 5.16, Table 5.5). These results strongly support our prediction of point mutations that alter the pB lifetime of Rcen PYP based on available Hhal PYP mutant data. From this dataset, we now have four slow Rcen PYP mutants (P6F, V28F, N43A, and M100A) and one fast mutant (E46Q).

The pH dependence of the kinetics of pB decay rate was examined for Rcen PYP and its mutants. For wt Rcen PYP the dependence of rate at low pH values showed a similar trend as observed in wt Hhal PYP but exhibited a different behavior above ~pH 8. In Hhal PYP the rate is observed to decrease above ~pH 8 whereas for wt Rcen PYP above ~pH 9.5 the rate was observed to gradually increase. Analysis of the pB–pG difference spectra of wt Rcen PYP from pH 2.6 to 12.5 revealed that the species absorbing near 354 nm gradually decreases at an increased pH whereas the red shifted species absorbing near 475 nm simultaneously increases. Previous studies on wt Rcen PYP for a pH range 5.4–8 indicated that under illumination there exists a pH-dependent equilibrium between the pR-like and pB-like species (*203*). This behavior is similar to Hhal PYP which under illumination and alkaline conditions shows a pH dependent equilibrium and this equilibrium shifts towards the pR like species at high pH and at low pH it shifts towards the pB like species (*204, 205*). Here we explored in Rcen PYP over a substantially wider pH range (2.6-12.5), revealing even more pronounced shifts in this equilibrium below pH 5

and above pH 8. In addition, the light/dark difference spectra of five mutants of Rcen PYP were analyzed at neutral pH. The P6F, N43A, and E46Q mutants revealed more or less similar pR/pB equilibria compared to wt Rcen PYP. However, in M100A mutant the equilibrium was altered with a reduced amount of pR like species and for the V28F mutant the equilibrium was more significantly affected wherein the pR species was not observed, with a pronounced increase in the pB species.

PYP w/His tag	WT	P6F/F6A	V28F/F28A	N43A	M100A	E46Q
Rcen, pB lifetime at neutral pH (s)	67	164	287	170	3000	10
Hhal, pB lifetime at neutral pH (s)	0.5	6.3	9	1640	330	0.067

Table 5-5 Summary of the effect of the respective mutants on the pB lifetime of Hhal PYP and Rcen PYP.

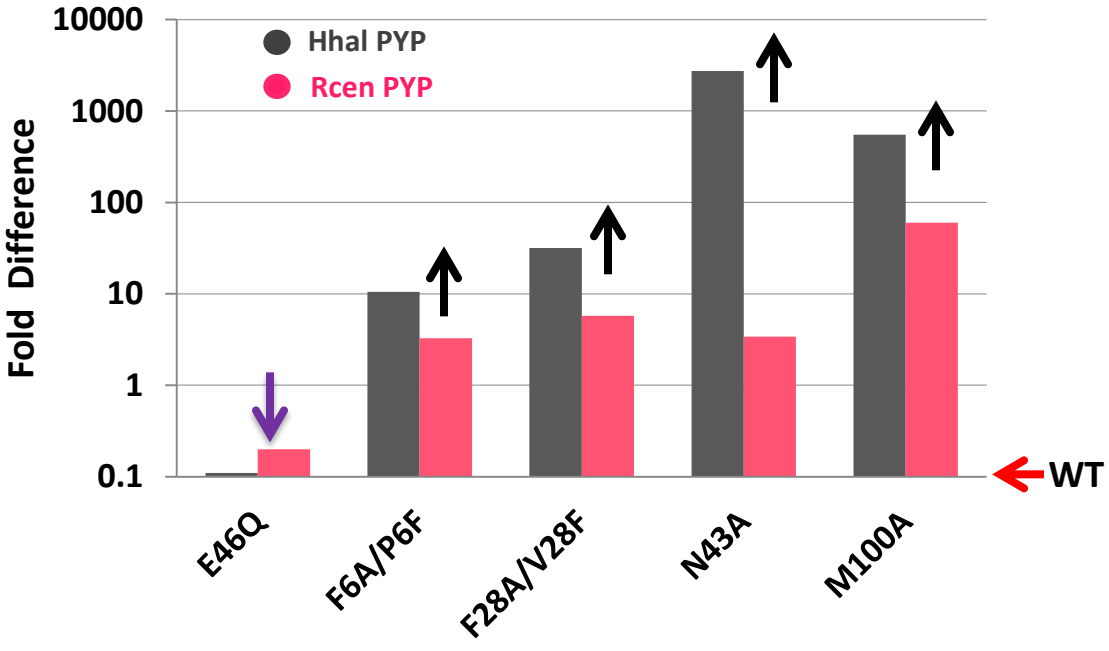


Fig. 5-16 Comparison of the effect of mutations on pB lifetime of Hhal PYP and Rcen PYP. The fold differences are depicted on a log scale and wt Rcen PYP is indicated as 1.

Effects of mutations on protein folding and stability

Denaturant titrations revealed the free energy for unfolding, ΔG_U (34.6± 2.8 kJ/mol), the denaturant *m* value (12.2 ± 1 kJ/mol/M), and the midpoint denaturant concentration, C_m (2.8 ± 0 M) of wt Rcen PYP, and effects of mutations on these factors that describe the transition from the fully folded state to the unfolded state. Except for E46Q (very high ΔG_U), all other mutants showed ΔG_U values similar to wt Rcen PYP. Unexpectedly, this property of the mutants resulted from compensating effect of the *m* and C_m values of these mutants on protein stability. E46Q was the only mutant that displayed a lower C_m than wt Rcen PYP, indicating a destabilizing effect but interestingly the steepness or the *m* value was much higher. The steepness (*m*) for all the slow mutants was reduced WT > M100A > P6F > N43A > V28F, and cancelling out this destabilizing effect we observed an increased C_m for the mutants V28F < N43A < P6F < M100A < WT (Table 5-3). In chapter 4, we have reported a similar behavior observed in I39A mutant of Hhal PYP and have proposed an explanation for the increases in denaturant *m* value. In summary, all Rcen PYP mutants tested here retain considerable stability.

Summary of the effects of substitutions at five positions on three functional properties

In order to summarize the degree to which the structure-function rules from Hhal PYP transfer to Rcen PYP, the data on three different properties of the mutants of Rcen PYP was compared to the properties of corresponding mutants in Hhal PYP. We observed qualitative agreement in 12 out of 15 cases (Table 5-6) with some disagreement in three out of 15 cases. Two mutants E46Q (faster photocycle rate) and M100A (slower photocycle rate) display a complete agreement, making them attractive candidates for future studies for probing the changes during *in vivo* signaling of PYP in *R. centenum*.

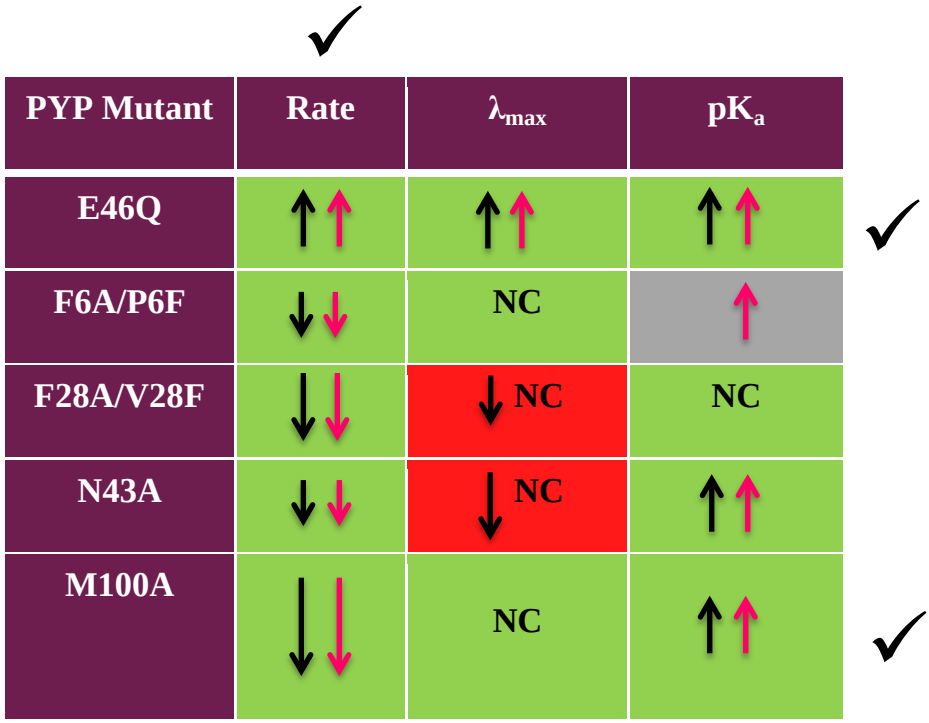


Table 5-6 Qualitative representation of the effects of mutations on three different properties of Hhal PYP (black arrows) and Rcen PYP (pink arrows), NC – No Change

The electrostatic epicenter in the PYP family and the effect of Glu46-pCA hydrogen bond on PYP active site properties

An active site proton transfer event results in protonation of *p*CA and transient ionization of the active site residue Glu46 in Hhal PYP, which triggers photoinduced large conformational changes in the pB state (*33*). The electrostatic epicenter mutant of Rcen PYP E46Q was tested in this study to probe the light induced conformational changes in Rcen PYP using rapid scan FTIR spectroscopy. Here, we report the first FTIR pB-pG difference spectra for both wt Rcen PYP and its E46Q mutant. As discussed above, the spectral markers for the protein backbone (amide I signal), the Glu46 band, and the *p*CA ring vibration mode are observed in wt Rcen PYP. In the E46Q mutant, the amide I difference signals were substantially reduced, indicating a suppression of light-induced conformational changes in this mutant. These results confirm that the

electrostatic epicenter is functional in Rcen PYP as well. Since the pB signaling state is reported to be partially unfolded, the E46Q mutant is now available to probe Rcen PYP *in vivo* and determine the degree to which partial unfolding contributes to *in vivo* signaling in PYP (See chapter 6).

FTIR difference spectroscopy is a sensitive technique that can be used to probe proton transfer reactions involving protonated carboxylic groups (*38*). The vibrational frequency of C=O stretching mode for a COOH group is sensitive to hydrogen bonding interactions. In the absence of hydrogen bonding, the C=O stretching frequency is high around 1760 cm^{-1} and with two hydrogen bonding interactions the frequency is reported to be down-shifted to 1715 cm^{-1} (*119*). The stronger the hydrogen bond is the more downshifted the C=O stretching mode (*32, 119*). The Glu46 C=O stretching mode in PYP can be used to probe the Glu46-*p*CA hydrogen bond and for Hhal PYP in pG state, this mode is reported to be at 1726 cm^{-1} (*32*). Unpublished results from our lab on Glu46 C=O stretching mode of *Salinibacter ruber* PYP was observed to be at 1719 cm^{-1} and from this study we know that for Rcen PYP this mode is at 1723 cm^{-1}. Interestingly, these data indicate an interpretable pattern in the effect of the Glu46-*p*CA hydrogen bond on two active site properties of PYP in the pG state, pG state absorbance maximum and the *p*CA pK_a. We propose an emerging theme in PYPs, in which a stronger Glu46-*p*CA hydrogen bond (detected through the frequency of the Glu46 C=O stretching mode) is associated with a blue shifted absorbance maximum and a lower *p*CA pK_a. (Table 5-7).

PYP w /His tag	Abs max pG (nm)	*p*CA pK_a in pG (*n* value)	Glu46 C=O stretching mode, pG
Hhal PYP	**446**	**2.8 (1.5)**	**1726 cm^{-1}[a]**
Rcen PYP	**434**	**~1.4 (0.95)**	**1723 cm^{-1}**
Srub PYP[b]	**432**	**1.70 (3.7)**	**1719 cm^{-1}**

[a] *Data from Xie et al. 1996;* [b] *Data from Miwa Hara et al., in preparation*

Table 5-7 Effect of Glu46–*p*CA hydrogen bond on PYP active site properties.

Active site pK_a shifts in the PYP family

In PYP the protein quake (global conformational changes) associated with the formation of the signaling state is a result of a proton transfer event involving the active site residue Glu46. It is therefore important to analyze the pK_a of the groups involved in proton transfer reaction. The pK_a of the *p*CA is ~2.8 in Hhal PYP and ~1.4 in Rcen PYP (Table 5-8), both highly down-shifted from the pK_a value of 9.0 of *p*CA in water. Unpublished results from our lab on Hhal PYP used the E46Q mutant to probe the 400 nm species formed at high pH (> 11). A rapid mixing pH jump approach was used to measure the high pH induced transitions from $A_{446} \rightarrow A_{400} \rightarrow A_{340}$ in Hhal PYP. The E46Q mutant prevented the formation of A_{400}, confirming that at high pH the active site residue Glu46 in PYP becomes deprotonated. The pK_a of this transition was determined to be 12.1 which is a strong upshift from the pK_a of Glu46 in water (4.0).

	*p*CA		Glu46	
	pK_a (n)	ΔpK_a	pK_a (n)	ΔpK_a
Hhal PYP	**2.8 (1.35)**	**- 6.2**	**12.1(3.3)**	**8.1**
Rcen PYP	**~1.4 (0.95)**	**~ -7.6**	**> 14 (nd)**	**>10**
pK_a in water : *p*CA – 9.0, Glu46 – 4.0				

Table 5-8 - Summary of pK_a shifts in the active site residues of the PYP family, nd – not determined.

Based on this information on Hhal PYP we performed a high pH titration of wt Rcen PYP and its E46Q mutant. In these experiments, Rcen PYP appears to be highly resistant to alkaline conditions up to pH ~14. Therefore, we checked the stability of wt Rcen PYP and E46Q mutant at very high pH and measured the pH dependent rate of pG decay, which was found to be almost identical for both samples. Lastly, to check the formation of the A_{400} species we performed a high pH jump measurement of wt Rcen PYP at pH values close to 14. Unexpectedly, we did not observe the formation of transient A_{400} species in these measurements. The most likely explanation for this behavior of Rcen PYP is that residue Glu46 in Rcen PYP requires even higher pH for its deprotonation, suggesting a pK_a value for Glu46 higher than 14.

Several enzymes with shifted pK_a values of catalytic groups in the active site have been reported. The highest ΔpK_a (> +8.0) has been reported for Asp96 in BR which is involved in proton transport. Unpublished studies from our lab on Hhal PYP report similar ΔpK_a (> +8.0) for Glu46 but our data on Rcen PYP indicates a remarkably high ΔpK_a (>10) for Glu46 (Table 5-8), which is the highest shift observed for a Glu/Asp in a protein. The current data offer an opportunity for further studies on the pK_a tuning and its relation to biological proton transfer.

CHAPTER VI

Conclusions

6.1. Summary

This study aims to better understand structure-function relationships at three different levels using photoactive yellow protein as a model system, with the goal of deriving insights that have broader implications for protein science: in a single protein (Hhal PYP), a protein family (the PYP family of photoreceptors), and an entire superfamily (the PAS domains). The data presented here illustrate how different approaches can be used, from bioinformatics to site-directed mutagenesis and isotope-edited FTIR spectroscopy, to probe different aspects and molecular events in the PYP photocycle, thus offering insights in the functional anatomy of proteins.

PYP has emerged as a powerful system for studying protein-ligand interactions, and functional protein dynamics and receptor activation, since its PYP photocycle involves light induced intramolecular proton transfer and subsequent protein conformational changes. The structural basis of proton transfer event within proteins is of more general interest. To help monitor the structural changes that cause proton transfer we used chemical modifications by labeling PYP with Tyr-D_4. This provides a valuable tool for ongoing and future tests of our hypothesis that the disruption of Tyr42-*p*CA hydrogen bond leads to proton transfer during the PYP photocycle.

Chapter 3 describes an extensive mass spectrometric analysis of the degree and pattern of isotope labeling of Tyr-D_4 PYP, yielding almost complete incorporation of the labeled Tyr showing no isotope dilution or isotope scrambling. This sample was then used for flash-induced microsecond time-resolved FTIR difference spectroscopy to monitor changes in Tyr hydrogen bonding. Analysis of the resulting FTIR data showed peaks that were distinctly shifted upon Tyr-D_4 labeling compared to the Tyr ring vibrational modes in unlabeled PYP. With the help of the vibrational structural marker approach developed in the Xie lab (*119*), the first indications of peak shifts that provide information on changes in Tyr hydrogen bonding were obtained. The structural interpretation of the signals that can now be assigned to Tyr side chains is in progress in the Xie lab. Thus, we successfully developed a generally applicable tool for infrared structural biology in which side chain isotope labeling can be used for the assignment of a structurally informative vibrational mode in a protein.

In chapter 4, we have explored the role of superfamily conserved residues, in this case the hydrophobic residue Ile39 from the PAS domain superfamily. Previous site directed mutagenesis studies on Hhal PYP with substitution at residue 39 showed altered functional properties, suggesting a possible role of this PAS conserved residue in signaling by a conserved allosteric switching mechanism (*1*). The data reported here on the structural changes during pB formation measured by rapid–scan FTIR spectroscopy indicate no difference between the wt PYP and its I39A mutant. However, biophysical characterization of this mutant indicates unexpected roles affecting folding cooperativity and signaling kinetics. Three distinct molecular events during pB decay can affect its rate: refolding of the protein to the native state, proton transfer to return the protonation states of the *p*CA and Glu46 to their initial state, and thermal chromophore re-isomerization. The I39A mutation probably forms a cavity in the native state, which can be proposed to hamper refolding, resulting in a longer pB lifetime. Bioinformatics analysis of side chain interactions of Ile39 across the PAS domain superfamily reveal that Ile39 forms a

conserved pattern of hydrophobic interaction with residues 31, 57, and 62 in PAS domain superfamily, indicating a conserved network of hydrophobic interactions in which the exact identity of the residues involved is less important.

Hhal PYP has been studied extensively *in vitro* by site-directed mutagenesis by several groups and a complete alanine scan of PYP from the Hoff lab (*1*) is also available. A number of mutants from these reports have been identified (from different regions of the proteins, hydrophobic core, N-terminus, and the *p*CA binding pocket) to affect different functional properties (λ_{max}, pK_a of *p*CA, and the lifetime of the pB state) of Hhal PYP as described in previous chapters. Substitutions at certain positions can affect a single property or can simultaneously alter multiple active site or functional properties of PYP. Time-resolved visible absorbance spectroscopy has been used to probe most of these functional properties of PYP. In addition to this, it has also been shown that FTIR spectroscopy allows the detection of conformational changes during the PYP photocycle (*83*). The Hhal PYP mutant, E46Q was used to experimentally establish the electrostatic epicenter driving global conformational changes in PYP upon photoactivation (*32*, *33*). Though a rich dataset on *in vitro* functional properties of Hhal PYP is available, the biological relevance of various molecular processes that occur during the PYP photocycle have not yet been experimentally connected to PYP signaling in the living bacterial cell due to the absence of genetics tools for *H. halophila*. In addition, the large body of measurements on site-directed mutants of PYP has been performed almost exclusively using Hhal PYP. Therefore, we selected the PYP from *R. centenum*, the only system currently known combining a PYP known *in vivo* function with genetic accessibility. To examine if the structure-function rules transfer from Hhal PYP to Rcen PYP, we characterized a set of Rcen PYP point mutants expected to alter different functional properties of the protein (spectral tuning (λ_{max}), pK_a of *p*CA, and pB lifetime). Analysis of these properties of all five mutants demonstrates that qualitatively the structure-function rules from Hhal PYP indeed transfer to Rcen PYP. We were

also able to use E46Q Rcen PYP to demonstrate that the electrostatic epicenter model for driving the large conformational changes during signaling is more widely applicable in the PYP family of photoreceptors (*33*). In addition, we studied the pK_a shifts of active site residues (generally associated with proton transfer events) in the PYP family, and discovered very strong shifts for the pK_a of both the *p*CA and Glu46 with ΔpK_a (>10) for Glu46 in Rcen PYP. To the best of our knowledge this is the highest shift observed for a Glu/Asp in any protein reported to date. The set of mutants reported here provides Rcen PYP variants with pB decay rates that are slowed down to different degrees and with reduced conformational changes (E46Q), opening up numerous opportunities for future *in vivo* testing of the contribution of the different molecular events in the PYP photocycle to the signal transduction mechanism in PYP inside the living cell.

6.2. Future Projections

One of the long-term goals of chapter 5 is to determine the *in vivo* signaling mechanisms of PYP in bacteria. The biophysical information derived from this thesis on *R. centenum* PYP and its mutants can be used to determine two central unresolved questions regarding the PYP photocycle: i) which photocycle intermediate functions as the signaling state in PYP? It is widely assumed that the pB state functions as the signaling state, but this has not been directly demonstrated. ii) Is light-induced partial unfolding of PYP needed for signaling? The latter aim promises to challenge the notion that two well folded interacting partners are involved in signaling by testing the suspected mechanism of partial unfolding in PYP signaling. It is known that partially unfolded and/or molten-globule like states play an important role in vitro especially in photoreceptor-mediated signaling (*5, 85*). If this mechanism of signaling involving partial unfolding of proteins is experimentally confirmed *in vivo* in *R. centenum* then it would contribute towards the understanding of several biological signal transduction mechanisms involving similar protein modules and even in unrelated signaling systems. In addition, the identification of the pB photocycle intermediate as the *in vivo* signaling state would provide physiological significance to

the large body of published biophysical data on this intermediate. A set of mutants capable of addressing (see below) the above two questions were created and its biophysical properties were characterized (Chapter 5).

Another interesting aspect of the PYP photocycle that needs to be addressed is the structural changes occurring in the N-terminal region and its importance in signaling. The role of these structural changes in the N-terminus of PYP is of more general importance in regard to the diverse PAS domain superfamily, many members of which have similar helical extensions that have been reported to exhibit conformational changes during their function (*206-208*). In case of Hhal PYP, the N-terminal region (consisting of 25 residues) is known to dissociate from the main hydrophobic core of PYP and to become floppy and unstructured (*84, 169*). *In vitro* studies on the N-terminal deletion mutant of Hhal PYP (Δ25) have revealed that it is photoactive and fully functional but has an increased pB lifetime (*82, 168, 169, 209*) and reduced light-induced structural changes indicated by reduced amide I signals in the FTIR difference spectrum of the protein (*84*). Anticipating a similar role in Rcen PYP, we constructed the N-terminal deletion mutant (Δ25) of Rcen PYP. However, the protein failed to express in the currently used *E. coli* BL21 (DE3) strain. It is suspected that this mutation probably is lethal either failing to express or undergoing rapid degradation upon reconstitution. To address this problem, future work could involve construction of partially deleted N-terminal region (for example Δ20–Δ24) of Rcen PYP. Once the photoactivity of N-terminus deleted Rcen PYP is established, it can be characterized and tested *in vitro* as was done for Hhal PYP.

Genetics tools are available for *R. centenum,* and PYP from this bacterium has been reported (*101*) to regulate the expression of the chalcone synthase gene (*chs*). To determine the effect of the five point mutants studied here (P6F, V28F, N43A, M100A, and E46Q) on the *in vivo* signaling of PYP, the Ppr deletion strain of *R. centenum* containing the *Lac*Z reporter system for *chs* expression constructed by Dr. Carl Bauer's lab can be used in future experiments. The first

step would include the introduction of these mutations in the full length *pyp* containing *ppr* gene expressed in appropriate *E. coli* strain followed by conjugation in *R. centenum* cells lacking the *ppr* gene. The *Lac*Z reporter system can be used to monitor if plasmid based expression of Ppr regulates the expression of chalcone synthase in *R. centenum* grown in different light conditions (infrared and/or blue light) using *Lac*Z assays.

In these experiments the intensity of blue light can be varied and fluency response curves for the induction of chalcone synthase by blue light can be measured. Both the maximal level of chalcone synthase and the intensity of blue light that triggers 50% of chalcone synthase induction in the mutants can be measured and compared to wt Rcen Ppr. If the pB state is the *in vivo* signaling state, an increase in its lifetime (P6F, V28F, N43A, and M100A Rcen Ppr mutants) would be expected to cause a corresponding decrease in the light intensity that causes 50% induction of the *chs* gene. If partial unfolding is required for *in vivo* signaling by PYP, light induced induction of the *chs* gene in the E46Q Ppr mutant would be greatly reduced, i.e., highly reduced signaling activity would be observed as a decrease in the maximal *chs* induction level. For the N-terminal deletion mutant of *R. centenum* Ppr, the fluency response curves for *chs* expression in different blue light intensities can be measured in a similar manner. Thus, when induced with blue light, the mutants altering only the pB lifetime are expected to show an unperturbed maximal level of *chs* expression, with the difference being in the intensity of light that causes 50% *chs* induction in different strains of *R. centenum*. However, for mutants that alter pB lifetime as well as reduce the light-induced conformational changes (E46Q mutant and the yet to be tested N-terminal deletion mutant) will show a reduced maximal level of *chs* gene expression, which will confirm their role and importance in the *in vivo* signaling mechanism of PYP.

An alternative method to measure the different levels of *chs* expression in wt Rcen Ppr and its mutants would be by using *q*RT-PCR on cells grown in different light conditions as

suggested for the *Lac*Z assays. The experiments proposed here will potentially identify which molecular events in the PYP photocycle are needed in signal transduction mechanism *in vivo*, and will help understand the functional role of photoinduced partially unfolded states of proteins that are involved in biological transduction. This line of research promises to have broader applications in studying structure-function relationships and signal transduction in a wide range of signaling proteins.

REFERENCES

1. Philip, A. F., Kumauchi, M., and Hoff, W. D. (2010) Robustness and evolvability in the functional anatomy of a PER-ARNT-SIM (PAS) domain, *P Natl Acad Sci USA 107*, 17986-17991.
2. Kumauchi, M., Hara, M. T., Stalcup, P., Xie, A. H., and Hoff, W. D. (2008) Identification of six new photoactive yellow proteins - Diversity and structure-function relationships in a bacterial blue light photoreceptor, *Photochem. Photobiol. 84*, 956-969.
3. Dobson, C. M. (2003) Protein folding and misfolding, *Nature 426*, 884-890.
4. Santoro, M. M., and Bolen, D. W. (1992) A test of the linear extrapolation of unfolding free-energy changes over an extended denaturant concentration range, *Biochemistry 31*, 4901-4907.
5. Lee, B.-C., Pandit, A., Croonquist, P. A., and Hoff, W. D. (2001) Folding and signaling share the same pathway in a photoreceptor, *Proceedings of the National Academy of Sciences 98*, 9062-9067.
6. Wang, M., Wales, T. E., and Fitzgerald, M. C. (2006) Conserved thermodynamic contributions of backbone hydrogen bonds in a protein fold, *Proc Natl Acad Sci U S A 103*, 2600-2604.

7. Brudler, R., Meyer, T. E., Genick, U. K., Devanathan, S., Woo, T. T., Millar, D. P., Gerwert, K., Cusanovich, M. A., Tollin, G., and Getzoff, E. D. (2000) Coupling of hydrogen bonding to chromophore conformation and function in photoactive yellow protein, *Biochemistry 39*, 13478-13486.

8. Reynaud, E. (2010) Protein misfolding and degenerative diseases, *Nature Education 3*, 28.

9. Dobson, C. M. (2002) Protein-misfolding diseases: Getting out of shape, *Nature 418*, 729-730.

10. Chiti, F., and Dobson, C. M. (2006) Protein misfolding, functional amyloid, and human disease, *Annu Rev Biochem 75*, 333-366.

11. Macdonald, M. E., Ambrose, C. M., Duyao, M. P., Myers, R. H., Lin, C., Srinidhi, L., Barnes, G., Taylor, S. A., James, M., Groot, N., Macfarlane, H., Jenkins, B., Anderson, M. A., Wexler, N. S., Gusella, J. F., Bates, G. P., Baxendale, S., Hummerich, H., Kirby, S., North, M., Youngman, S., Mott, R., Zehetner, G., Sedlacek, Z., Poustka, A., Frischauf, A. M., Lehrach, H., Buckler, A. J., Church, D., Doucettestamm, L., Odonovan, M. C., Ribaramirez, L., Shah, M., Stanton, V. P., Strobel, S. A., Draths, K. M., Wales, J. L., Dervan, P., Housman, D. E., Altherr, M., Shiang, R., Thompson, L., Fielder, T., Wasmuth, J. J., Tagle, D., Valdes, J., Elmer, L., Allard, M., Castilla, L., Swaroop, M., Blanchard, K., Collins, F. S., Snell, R., Holloway, T., Gillespie, K., Datson, N., Shaw, D., and Harper, P. S. (1993) A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntingtons-disease chromosomes, *Cell 72*, 971-983.

12. Bertram, L., and Tanzi, R. E. (2005) The genetic epidemiology of neurodegenerative disease, *J Clin Invest 115*, 1449-1457.

13. Detwiler, L. A. (1992) Scrapie, *Rev Sci Tech 11*, 491-537.

14. Murphy, R. M. (2002) Peptide aggregation in neurodegenerative disease, *Annu Rev Biomed Eng 4*, 155-174.

15. Clancy, S. (2008) Genetic Mutation, *Nature Education 1*, 187.

16. Pauling, L., Itano, H. A., and et al. (1949) Sickle cell anemia a molecular disease, *Science 110*, 543-548.

17. Ridge, K. D., and Palczewski, K. (2007) Visual rhodopsin sees the light: Structure and mechanism of G protein signaling, *Journal of Biological Chemistry 282*, 9297-9301.

18. Huggins, D. J., Sherman, W., and Tidor, B. (2012) Rational Approaches to Improving Selectivity in Drug Design, *Journal of Medicinal Chemistry 55*, 1424-1444.

19. Cho, K. I., Lee, K., Lee, K. H., Kim, D., and Lee, D. (2006) Specificity of molecular interactions in transient protein-protein interaction interfaces, *Proteins-Structure Function and Bioinformatics 65*, 593-606.

20. Keskin, O., and Nussinov, R. (2007) Similar binding sites and different partners: Implications to shared proteins in cellular pathways, *Structure 15*, 341-354.

21. Maerkl, S. J., and Quake, S. R. (2007) A systems approach to measuring the binding energy landscapes of transcription factors, *Science 315*, 233-237.

22. Mintseris, J., and Weng, Z. P. (2005) Structure, function, and evolution of transient and obligate protein-protein interactions, *P Natl Acad Sci USA 102*, 10930-10935.

23. Morell, M., Espargaro, A., Aviles, F. X., and Ventura, S. (2007) Detection of transient protein-protein interactions by bimolecular fluorescence complementation: The Abl-SH3 case, *Proteomics 7*, 1023-1036.

24. Nooren, I. M., and Thornton, J. M. (2003) Structural characterisation and functional significance of transient protein-protein interactions, *J Mol Biol 325*, 991-1018.

25. Sprinzak, E., Altuvia, Y., and Margalit, H. (2006) Characterization and prediction of protein-protein interactions within and between complexes, *P Natl Acad Sci USA 103*, 14718-14723.

26. Keskin, O., Gursoy, A., Ma, B., and Nussinov, R. (2008) Principles of protein-protein interactions: What are the preferred ways for proteins to interact?, *Chem Rev 108*, 1225-1244.

27. Gunner, M. R., and Alexov, E. (2000) A pragmatic approach to structure based calculation of coupled proton and electron transfer in proteins, *Bba-Bioenergetics 1458*, 63-87.

28. Ishikita, H., and Saito, K. (2014) Proton transfer reactions and hydrogen-bond networks in protein environments, *Journal of the Royal Society Interface 11*.

29. Yoo, H. Y., Boatz, J. A., Helms, V., McCammon, J. A., and Langhoff, P. W. (2001) Chromophore protonation states and the proton shuttle mechanism in green fluorescent protein: Inferences drawn from ab initio theoretical studies of chemical structures and vibrational spectra, *Journal of Physical Chemistry B 105*, 2850-2857.

30. Balashov, S. P. (2000) Protonation reactions and their coupling in bacteriorhodopsin, *Bba-Bioenergetics 1460*, 75-94.

31. Lanyi, J. K. (1998) Understanding structure and function in the light-driven proton pump bacteriorhodopsin, *J Struct Biol 124*, 164-178.

32. Xie, A. H., Hoff, W. D., Kroon, A. R., and Hellingwerf, K. J. (1996) Glu46 donates a proton to the 4-hydroxycinnamate anion chromophore during the photocycle of photoactive yellow protein, *Biochemistry 35*, 14671-14678.

33. Xie, A. H., Kelemen, L., Hendriks, J., White, B. J., Hellingwerf, K. J., and Hoff, W. D. (2001) Formation of a new buried charge drives a large-amplitude protein quake in photoreceptor activation, *Biochemistry 40*, 1510-1517.

34. Polyanichko, A. M., Wieser, H. (2007) Vibrational Circular Dichroism and its Applications to Protein Studies, In *Methods in Protein Structure and Stability Analysis: Vibrational spectroscopy* (Uversky, V., Permyakov, E, Ed.), pp 267-302, Nova Science Publishers, Inc., New York.

35. Barth, A. (2007) Infrared spectroscopy of proteins, *Bba-Bioenergetics 1767*, 1073-1101.

36. Gerwert, K. (1999) Molecular reaction mechanisms of proteins monitored by time-resolved FTIR-spectroscopy, *Biol Chem 380*, 931-935.

37. Maeda, A. (1995) Application of FTIR spectroscopy to the structural study on the function of bacteriorhodopsin, *Isr J Chem 35*, 387-400.

38. Rothschild, K. J. (1992) Ftir Difference Spectroscopy of Bacteriorhodopsin - toward a Molecular-Model, *J Bioenerg Biomembr 24*, 147-167.

39. Gennis, R. B. (2003) Some recent contributions of FTIR difference spectroscopy to the study of cytochrome oxidase, *Febs Lett 555*, 2-7.

40. Barth, A., vonGermar, F., Kreutz, W., and Mantele, W. (1996) Time-resolved infrared spectroscopy of the Ca2+-ATPase - The enzyme at work, *Journal of Biological Chemistry 271*, 30637-30646.

41. Schultz, J., Milpetz, F., Bork, P., and Ponting, C. P. (1998) SMART, a simple modular architecture research tool: identification of signaling domains, *Proc Natl Acad Sci U S A 95*, 5857-5864.

42. Bateman, A., Coin, L., Durbin, R., Finn, R. D., Hollich, V., Griffiths-Jones, S., Khanna, A., Marshall, M., Moxon, S., Sonnhammer, E. L. L., Studholme, D. J., Yeats, C., and Eddy, S. R. (2004) The Pfam protein families database, *Nucleic Acids Res 32*, D138-D141.

43. Orengo, C. A., and Thornton, J. M. (2005) Protein families and their evolution - A structural perspective, *Annu Rev Biochem 74*, 867-900.

44. 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.

45. Ponting, C. P., and Aravind, L. (1997) PAS: a multifunctional domain family comes to light, *Current Biology 7*, R674-R677.

46. Schibler, U. (1998) Circadian rhythms - New cogwheels in the clockworks, *Nature 393*, 620-621.

47. Kay, S. A. (1997) Circadian rhythms - PAS, present, and future: Clues to the origins of circadian clocks, *Science 276*, 753-754.

48. Sprenger, W. W., Hoff, W. D., Armitage, J. P., and Hellingwerf, K. J. (1993) The Eubacterium Ectothiorhodospira-Halophila Is Negatively Phototactic, with a Wavelength Dependence That Fits the Absorption-Spectrum of the Photoactive Yellow Protein, *J Bacteriol 175*, 3096-3104.

49. Semenza, G. L. (2000) Surviving ischemia: adaptive responses mediated by hypoxia-inducible factor 1, *Journal of Clinical Investigation 106*, 809-812.

50. Maxwell, P. H., Dachs, G. U., Gleadle, J. M., Nicholls, L. G., Harris, A. L., Stratford, I. J., Hankinson, O., Pugh, C. W., and Ratcliffe, P. J. (1997) Hypoxia-inducible factor-1 modulates gene expression in solid tumors and influences both angiogenesis and tumor growth, *P Natl Acad Sci USA 94*, 8104-8109.

51. Nambu, J. R., Lewis, J. O., Wharton, K. A., and Crews, S. T. (1991) The Drosophila Single-Minded Gene Encodes a Helix-Loop-Helix Protein That Acts as a Master Regulator of Cns Midline Development, *Cell 67*, 1157-1167.

52. Chen, J., Zou, A., Splawski, I., Keating, M. T., and Sanguinetti, M. C. (1999) Long QT syndrome-associated mutations in the Per-Arnt-Sim (PAS) domain of HERG potassium channels accelerate channel deactivation, *J Biol Chem 274*, 10113-10118.

53. Cusanovich, M. A., and Meyer, T. E. (2003) Photoactive yellow protein: A prototypic PAS domain sensory protein and development of a common signaling mechanism, *Biochemistry 42*, 4759-4770.

54. 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, *P Natl Acad Sci USA 95*, 5884-5890.

55. Meyer, T. E. (1985) Isolation and Characterization of Soluble Cytochromes, Ferredoxins and Other Chromophoric Proteins from the Halophilic Phototrophic Bacterium Ectothiorhodospira-Halophila, *Biochim Biophys Acta 806*, 175-183.

56. Meyer, T. E., Yakali, E., Cusanovich, M. A., and Tollin, G. (1987) Properties of a Water-Soluble, Yellow Protein Isolated from a Halophilic Phototrophic Bacterium That Has Photochemical Activity Analogous to Sensory Rhodopsin, *Biochemistry 26*, 418-423.

57. Baca, M., Borgstahl, G. E. O., Boissinot, M., Burke, P. M., Williams, D. R., Slater, K. A., and Getzoff, E. D. (1994) Complete chemical-structure of photoactive yellow protein - novel thioester-linked 4-hydroxycinnamyl chromophore and photocycle chemist, *Biochemistry 33*, 14369-14377.

58. Hoff, W. D., Dux, P., Hard, K., Devreese, B., Nugterenroodzant, I. M., Crielaard, W., Boelens, R., Kaptein, R., Vanbeeumen, J., and Hellingwerf, K. J. (1994) Thiol Ester-Linked P-Coumaric Acid as a New Photoactive Prosthetic Group in a Protein with Rhodopsin-Like Photochemistry, *Biochemistry 33*, 13959-13962.

59. Koh, M., VanDriessche, G., Samyn, B., Hoff, W. D., Meyer, T. E., Cusanovich, M. A., and VanBeeumen, J. J. (1996) Sequence evidence for strong conservation of the photoactive yellow proteins from the halophilic phototrophic bacteria Chromatium salexigens and Rhodospirillum salexigens, *Biochemistry 35*, 2526-2534.

60. Kort, R., Hoff, W. D., VanWest, M., Kroon, A. R., Hoffer, S. M., Vlieg, K. H., Crielaard, W., VanBeeumen, J. J., and Hellingwerf, K. J. (1996) The xanthopsins: A new family of eubacterial blue-light photoreceptors, *Embo J 15*, 3209-3218.

61. Meyer, T. E., Fitch, J. C., Bartsch, R. G., Tollin, G., and Cusanovich, M. A. (1990) Soluble Cytochromes and a Photoactive Yellow Protein Isolated from the Moderately Halophilic Purple Phototrophic Bacterium, Rhodospirillum-Salexigens, *Biochim Biophys Acta 1016*, 364-370.

62. Haker, A., Hendriks, J., Gensch, T., Hellingwerf, K., and Crielaard, W. (2000) Isolation, reconstitution and functional characterisation of the Rhodobacter sphaeroides photoactive yellow protein, *Febs Lett 486*, 52-56.

63. Meyer, T. E., Kyndt, J. A., Memmi, S., Moser, T., Colon-Acevedo, B., Devreese, B., and Van Beeumen, J. J. (2012) The growing family of photoactive yellow proteins and their presumed functional roles, *Photoch Photobio Sci 11*, 1495-1514.

64. Hellingwerf, K. J., Hendriks, J., and Gensch, T. (2003) Photoactive Yellow Protein, a new type of photoreceptor protein: Will this "yellow lab" bring us where we want to go?, *J Phys Chem A 107*, 1082-1094.

65. Genick, U. K., Devanathan, S., Meyer, T. E., Canestrelli, I. L., Williams, E., Cusanovich, M. A., Tollin, G., and Getzoff, E. D. (1997) Active site mutants implicate key residues for control of color and light cycle kinetics of photoactive yellow protein, *Biochemistry 36*, 8-14.

66. Lagarias, D., Wu, S.-H., and Lagarias, J. C. (1995) Atypical phytochrome gene structure in the green alga Mesotaenium caldariorum, *Plant Molecular Biology 29*, 1127-1142.

67. Borgstahl, G. E., Williams, D. R., and Getzoff, E. D. (1995) 1.4 A structure of photoactive yellow protein, a cytosolic photoreceptor: unusual fold, active site, and chromophore, *Biochemistry 34*, 6278-6287.

68. Getzoff, E. D., Gutwin, K. N., and Genick, U. K. (2003) Anticipatory active-site motions and chromophore distortion prime photoreceptor PYP for light activation, *Nat Struct Biol 10*, 663-668.

69. Van Aalten, D. M. F., Crielaard, W., Hellingwerf, K. J., and Joshua-Tor, L. (2000) Conformational substates in different crystal forms of the photoactive yellow protein - Correlation with theoretical and experimental flexibility, *Protein Sci 9*, 64-72.

70. 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.

71. Dux, P., Rubinstenn, G., Vuister, G. W., Boelens, R., Mulder, F. A. A., Hard, K., Hoff, W. D., Kroon, A. R., Crielaard, W., Hellingwerf, K. J., and Kaptein, R. (1998) Solution structure and backbone dynamics of the photoactive yellow protein, *Biochemistry 37*, 12689-12699.

72. Borgstahl, G. E. O., Williams, D. R., and Getzoff, E. D. (1995) 1.4 Angstrom Structure of Photoactive Yellow Protein, a Cytosolic Photoreceptor - Unusual Fold, Active-Site, and Chromophore, *Biochemistry 34*, 6278-6287.

73. Kim, M., Mathies, R. A., Hoff, W. D., and Hellingwerf, K. J. (1995) Resonance Raman evidence that the thioester-linked 4-hydroxycinnamyl chromophore of photoactive yellow protein is deprotonated, *Biochemistry 34*, 12669-12672.

74. Yoda, M., Houjou, H., Inoue, Y., and Sakurai, M. (2001) Spectral tuning of photoactive yellow protein. Theoretical and experimental analysis of medium effects on the absorption spectrum of the chromophore, *Journal of Physical Chemistry B 105*, 9887-9895.

75. Imamoto, Y., and Kataoka, M. (2007) Structure and photoreaction of photoactive yellow protein, a structural prototype of the PAS domain superfamily, *Photochem Photobiol 83*, 40-49.

76. Hoff, W. D., Devreese, B., Fokkens, R., NugterenRoodzant, I. M., VanBeeumen, J., Nibbering, N., and Hellingwerf, K. J. (1996) Chemical reactivity and spectroscopy of the thiol ester-linked p-coumaric acid chromophore in the photoactive yellow protein from Ectothiorhodospira halophila, *Biochemistry 35*, 1274-1281.

77. Imamoto, Y., Ito, T., Kataoka, M., and Tokunaga, F. (1995) Reconstitution Photoactive Yellow Protein from Apoprotein and P-Coumaric Acid-Derivatives, *Febs Lett 374*, 157-160.

78. Hoff, W. D., Van Stokkum, I. H., Van Ramesdonk, H. J., Van Brederode, M. E., Brouwer, A. M., Fitch, J. C., Meyer, T. E., Van Grondelle, R., and Hellingwerf, K. J. (1994) Measurement and global analysis of the absorbance changes in the photocycle of the photoactive yellow protein from Ectothiorhodospira halophila, *Biophys J 67*, 1691-1705.

79. Bogomolni, R. A., and Spudich, J. L. (1982) Identification of a 3rd-Rhodopsin-Like Pigment in Phototactic Halobacterium-Halobium, *P Natl Acad Sci-Biol 79*, 6250-6254.

80. Kort, R., Vonk, H., Xu, X., Hoff, W. D., Crielaard, W., and Hellingwerf, K. J. (1996) Evidence for trans-cis isomerization of the p-coumaric acid chromophore as the photochemical basis of the photocycle of photoactive yellow protein, *Febs Lett 382*, 73-78.

81. Zhao, J. M., Lee, H., Nome, R. A., Majid, S., Scherert, N. F., and Hoff, W. D. (2006) Single-molecule detection of structural changes during Per-Arnt-Sim (PAS) domain activation, *P Natl Acad Sci USA 103*, 11561-11566.

82. Bernard, C., Houben, K., Derix, N. M., Marks, D., van der Horst, M. A., Hellingwerf, K. J., Boelens, R., Kaptein, R., and van Nuland, N. A. (2005) The solution structure of a transient photoreceptor intermediate: Delta25 photoactive yellow protein, *Structure 13*, 953-962.

83. Hoff, W. D., Xie, A., Van Stokkum, I. H. M., Tang, X. J., Gural, J., Kroon, A. R., and Hellingwerf, K. J. (1999) Global conformational changes upon receptor stimulation in photoactive yellow protein, *Biochemistry 38*, 1009-1017.

84. Imamoto, Y., Kamikubo, H., Harigai, M., Shimizu, N., and Kataoka, M. (2002) Light-induced global conformational change of photoactive yellow protein in solution, *Biochemistry 41*, 13595-13601.

85. Lee, B. C., Croonquist, P. A., Sosnick, T. R., and Hoff, W. D. (2001) PAS domain receptor photoactive yellow protein is converted to a molten globule state upon activation, *Journal of Biological Chemistry 276*, 20821-20823.

86. Van Brederode, M. E., Hoff, W. D., Van Stokkum, I. H. M., Groot, M. L., and Hellingwerf, K. J. (1996) Protein folding thermodynamics applied to the photocycle of the photoactive yellow protein, *Biophysical Journal 71*, 365-380.

87. Rubinstenn, G., Vuister, G. W., Mulder, F. A. A., Dux, P. E., Boelens, R., Hellingwerf, K. J., and Kaptein, R. (1998) Structural and dynamic changes of photoactive yellow protein during its photocycle in solution, *Nat Struct Biol 5*, 568-570.

88. van der Horst, M. A., Stalcup, T. P., Kaledhonkar, S., Kumauchi, M., Hara, M., Xie, A., Hellingwerf, K. J., and Hoff, W. D. (2009) Locked chromophore analogs reveal that photoactive yellow protein regulates biofilm formation in the deep sea bacterium Idiomarina loihiensis, *Journal of the American Chemical Society 131*, 17443-17451.

89. Purcell, E. B., and Crosson, S. (2008) Photoregulation in prokaryotes, *Curr Opin Microbiol 11*, 168-178.

90. Hellingwerf, K. J., van der Horst, M. A., and Key, J. (2007) Photosensing in chemotrophic, non-phototrophic bacteria: let there be light sensing too, *Trends Microbiol 15*, 554-562.

91. Kroon, A. R., Hoff, W. D., Fennema, H. P., Gijzen, J., Koomen, G. J., Verhoeven, J. W., Crielaard, W., and Hellingwerf, K. J. (1996) Spectral tuning, fluorescence, and photoactivity in hybrids of photoactive yellow protein, reconstituted with native or modified chromophores, *J Biol Chem 271*, 31949-31956.

92. van der Horst, M. A., Arents, J. C., Kort, R., and Hellingwerf, K. J. (2007) Binding, tuning and mechanical function of the 4-hydroxy-cinnamic acid chromophore in photoactive yellow protein, *Photochem Photobiol Sci 6*, 571-579.

93. Yamada, H., Kumauchi, M., Hamada, N., Zheng, X. G., Park, I. H., Masuda, K., Yoshihara, K., and Tokunaga, F. (2006) Analogue chromophore study of the influence of electronic perturbation on color regulation of photoactive yellow protein, *Photochem Photobiol 82*, 1422-1425.

94. Yamato, T., Ishikura, T., Kakitani, T., Kawaguchi, K., and Watanabe, H. (2007) Spectral tuning of photoactive yellow protein, *Photochem Photobiol 83*, 323-327.

95. Hoff, W. D., Van Stokkum, I. H. M., Gural, J., and Hellingwerf, K. J. (1997) Comparison of acid denaturation and light activation in the eubacterial blue-light receptor photoactive yellow protein, *Bba-Bioenergetics 1322*, 151-162.

96. Imamoto, Y., Koshimizu, H., Mihara, K., Hisatomi, O., Mizukami, T., Tsujimoto, K., Kataoka, M., and Tokunaga, F. (2001) Roles of amino acid residues near the chromophore of photoactive yellow protein, *Biochemistry 40*, 4679-4685.

97. Meyer, T. E., Devanathan, S., Woo, T., Getzoff, E. D., Tollin, G., and Cusanovich, M. A. (2003) Site-specific mutations provide new insights into the origin of pH effects and alternative spectral forms in the photoactive yellow protein from Halorhodospira halophila, *Biochemistry 42*, 3319-3325.

98. Yoda, M., Inoue, Y., and Sakurai, M. (2003) Effect of protein environment on pK(a) shifts in the active site of photoactive yellow protein, *Journal of Physical Chemistry B 107*, 14569-14575.

99. Kyndt, J. A., Fitch, J. C., Meyer, T. E., and Cusanovich, M. A. (2005) Thermochromatium tepidum photoactive yellow protein/bacteriophytochrome/ diguanylate cyclase: Characterization of the PYP domain, *Biochemistry 44*, 4755-4764.

100. Haker, A., Hendriks, J., van Stokkum, I. H. M., Heberle, J., Hellingwerf, K. J., Crielaard, W., and Gensch, T. (2003) The two photocycles of photoactive yellow protein from Rhodobacter sphaeroides, *Journal of Biological Chemistry 278*, 8442-8451.

101. Jiang, Z. Y., Swem, L. R., Rushing, B. G., Devanathan, S., Tollin, G., and Bauer, C. E. (1999) Bacterial photoreceptor with similarity to photoactive yellow protein and plant phytochromes, *Science 285*, 406-409.

102. Mihara, K., Hisatomi, O., Imamoto, Y., Kataoka, M., and Tokunaga, F. (1997) Functional expression and site-directed mutagenesis of photoactive yellow protein, *J Biochem-Tokyo 121*, 876-880.

103. Kumauchi, M., Kaledhonkar, S., Philip, A. F., Wycoff, J., Hara, M., Li, Y., Xie, A., and Hoff, W. D. (2010) A conserved helical capping hydrogen bond in PAS domains controls signaling kinetics in the superfamily prototype photoactive yellow protein, *J Am Chem Soc 132*, 15820-15830.

104. Meyer, T. E., Tollin, G., Hazzard, J. H., and Cusanovich, M. A. (1989) Photoactive Yellow Protein from the Purple Phototrophic Bacterium, Ectothiorhodospira-Halophila - Quantum Yield of Photobleaching and Effects of Temperature, Alcohols, Glycerol, and Sucrose on Kinetics of Photobleaching and Recover, *Biophysical Journal 56*, 559-564.

105. Lee, B. C., Croonquist, P. A., and Hoff, W. D. (2001) Mimic of photocycle by a protein folding reaction in photoactive yellow protein, *J Biol Chem 276*, 44481-44487.

106. Philip, A. F., Eisenman, K. T., Papadantonakis, G. A., and Hoff, W. D. (2008) Functional Tuning of Photoactive Yellow Protein by Active Site Residue 46, *Biochemistry 47*, 13800-13810.

107. Lozier, R. H., Bogomolni, R. A., and Stoeckenius, W. (1975) Bacteriorhodopsin: a light-driven proton pump in Halobacterium halobium, *Biophys J 15*, 955-962.

108. Haupts, U., Tittor, J., and Oesterhelt, D. (1999) Closing in on bacteriorhodopsin: Progress in understanding the molecule, *Annu Rev Bioph Biom 28*, 367-399.

109. Heberle, J. (2000) Proton transfer reactions across bacteriorhodopsin and along the membrane, *Bba-Bioenergetics 1458*, 135-147.

110. Hirai, T., Subramaniam, S., and Lanyi, J. K. (2009) Structural snapshots of conformational changes in a seven-helix membrane protein: lessons from bacteriorhodopsin, *Curr Opin Struc Biol 19*, 433-439.

111. Lanyi, J. K. (2004) Bacteriorhodopsin, *Annu Rev Physiol 66*, 665-688.

112. Lanyi, J. K. (2006) Proton transfers in the bacteriorhodopsin photocycle, *Bba-Bioenergetics 1757*, 1012-1018.

113. Braun-Sand, S., Sharma, P. K., Chu, Z. T., Pisliakov, A. V., and Warshel, A. (2008) The energetics of the primary proton transfer in bacteriorhodopsin revisited: It is a sequential light-induced charge separation after all, *Bba-Bioenergetics 1777*, 441-452.

114. Warshel, A. (1979) Conversion of Light Energy to Electrostatic energy in the proton pump of Halobacterium-halobium, *Photochem Photobiol 30*, 285-290.

115. Garczarek, F., and Gerwert, K. (2006) Functional waters in intraprotein proton transfer monitored by FTIR difference spectroscopy, *Nature 439*, 109-112.

116. Lanyi, J. K. (2000) Crystallographic studies of the conformational changes that drive directional transmembrane ion movement in bacteriorhodopsin, *Bba-Bioenergetics 1459*, 339-345.

117. Wang, J. H. (1970) Directional character of proton transfer in enzyme catalysis, *P Natl Acad Sci USA 66*, 874-&.

118. Silverman, D. N., and McKenna, R. (2007) Solvent-mediated proton transfer in catalysis by carbonic anhydrase, *Acc Chem Res 40*, 669-675.

119. Nie, B., Stutzman, J., and Xie, A. (2005) A vibrational spectral maker for probing the hydrogen-bonding status of protonated Asp and Glu residues, *Biophys J 88*, 2833-2847.

120. Nie, B. (2006) Probing hydrogen bonding interactions and proton transfer in proteins, *Oklahoma State University, Stillwater, OK.*, 243.

121. Thubagere, A. J. (2008) Advanced FTIR spectroscopy of protein structural dynamics, *Oklahoma State University, Stillwater, OK.*, 142.

122. Berthomieu, C., and Hienerwadel, R. (2009) Fourier transform infrared (FTIR) spectroscopy, *Photosynth Res 101*, 157-170.

123. Kotting, C., and Gerwert, K. (2005) Proteins in action monitored by time-resolved FTIR spectroscopy, *Chemphyschem 6*, 881-888.

124. Rothschild, K. J., Braiman, M. S., He, Y. W., Marti, T., and Khorana, H. G. (1990) Vibrational spectroscopy of bacteriorhodopsin mutants. Evidence for the interaction of aspartic acid 212 with tyrosine 185 and possible role in the proton pump mechanism, *J Biol Chem 265*, 16985-16991.

125. Vogel, R., and Siebert, F. (2000) Vibrational spectroscopy as a tool for probing protein function, *Curr Opin Chem Biol 4*, 518-523.

126. Braiman, M. S., Ahl, P. L., and Rothschild, K. J. (1987) Millisecond Fourier-Transform Infrared Difference Spectra of Bacteriorhodopsin M412 Photoproduct, *Proc. Natl. Acad. Sci. U.S.A. 84*, 5221-5225.

127. Hessling, B., Herbst, J., Rammelsberg, R., and Gerwert, K. (1997) Fourier transform infrared double-flash experiments resolve bacteriorhodopsin's M-1 to M-2 transition, *Biophys. J. 73*, 2071-2080.

128. Hu, X. H., Frei, H., and Spiro, T. G. (1996) Nanosecond step-scan FTIR spectroscopy of hemoglobin: Ligand recombination and protein conformational changes, *Biochemistry 35*, 13001-13005.

129. Hage, W., Kim, M., Frei, H., and Mathies, R. A. (1996) Protein dynamics in the bacteriorhodopsin photocycle: A nanosecond step-scan FTIR investigation of the KL to L transition, *J. Phys. Chem. 100*, 16026-16033.

130. Hastings, G. (2001) Time-resolved step-scan Fourier transform infrared and visible absorption difference spectroscopy for the study of photosystem I, *Appl. Spectrosc. 55*, 894-900.

131. Sun, H., and Frei, H. (1997) Time-resolved step-scan Fourier transform infrared spectroscopy of triplet excited duroquinone in a zeolite, *J Phys Chem B 101*, 205-209.

132. Dioumaev, A. K., and Braiman, M. S. (1997) Two bathointermediates of the bacteriorhodopsin photocycle, distinguished by nanosecond time-resolved FTIR spectroscopy at room temperature, *J Phys Chem B 101*, 1655-1662.

133. Xie, A., Kelemen, L., Hendriks, J., White, B. J., Hellingwerf, K. J., and Hoff, W. D. (2001) Formation of a new buried charge drives a large-amplitude protein quake in photoreceptor activation, *Biochemistry 40*, 1510−1517.

134. Arkin, I. T. (2006) Isotope-edited IR spectroscopy for the study of membrane proteins, *Curr Opin Chem Biol 10*, 394-401.

135. Das, K. P., Choo-Smith, L. P., Petrash, J. M., and Surewicz, W. K. (1999) Insight into the secondary structure of non-native proteins bound to a molecular chaperone alpha-crystallin - An isotope-edited infrared spectroscopic study, *Journal of Biological Chemistry 274*, 33209-33212.

136. Li, T. S., Horan, T., Osslund, T., Stearns, G., and Arakawa, T. (1997) Conformational changes in G-CSF/receptor complex as investigated by isotope-edited FTIR spectroscopy, *Biochemistry 36*, 8849-8857.

137. Tatulian, S. A. (2010) Structural analysis of proteins by isotope-edited FTIR spectroscopy, *Spectrosc-Int J 24*, 37-43.

138. Warscheid, B., Brucker, S., Kallenbach, A., Meyer, H. E., Gerwert, K., and Kotting, C. (2008) Systematic approach to group-specific isotopic labeling of proteins for vibrational spectroscopy, *Vibrational Spectroscopy 48*, 28-36.

139. Zhang, M., Fabian, H., Mantsch, H. H., and Vogel, H. J. (1994) Isotope-edited Fourier transform infrared spectroscopy studies of calmodulin's interaction with its target peptides, *Biochemistry 33*, 10883-10888.

140. deJongh, H. H. J., Goormaghtigh, E., and Ruysschaert, J. M. (1997) Monitoring structural stability of trypsin inhibitor at the submolecular level by amide-proton exchange using Fourier transform infrared spectroscopy: A test case for more general application, *Biochemistry 36*, 13593-13602.

141. Muchmore, D. C., McIntosh, L. P., Russell, C. B., Anderson, D. E., and Dahlquist, F. W. (1989) Expression and nitrogen-15 labeling of proteins for proton and nitrogen-15 nuclear magnetic resonance, *Methods Enzymol 177*, 44-73.

142. Muona, M., Aranko, A. S., Raulinaitis, V., and Iwai, H. (2010) Segmental isotopic labeling of multi-domain and fusion proteins by protein trans-splicing in vivo and in vitro, *Nat Protoc 5*, 574-587.

143. Xu, R., Ayers, B., Cowburn, D., and Muir, T. W. (1999) Chemical ligation of folded recombinant proteins: Segmental isotopic labeling of domains for NMR studies, *P Natl Acad Sci USA 96*, 388-393.

144. Yamazaki, T., Otomo, T., Oda, N., Kyogoku, Y., Uegaki, K., Ito, N., Ishino, Y., and Nakamura, H. (1998) Segmental isotope labeling for protein NMR using peptide splicing, *J Am Chem Soc 120*, 5591-5592.

145. Sonar, S., Lee, C. P., Coleman, M., Patel, N., Liu, X. M., Marti, T., Khorana, H. G., Rajbhandary, U. L., and Rothschild, K. J. (1994) Site-directed isotope labeling and Ftir spectroscopy of bacteriorhodopsin, *Nat Struct Biol 1*, 512-517.

146. Yabuki, T., Kigawa, T., Dohmae, N., Takio, K., Terada, T., Ito, Y., Laue, E. D., Cooper, J. A., Kainosho, M., and Yokoyama, S. (1998) Dual amino acid-selective and site-directed stable-isotope labeling of the human c-Ha-Ras protein by cell-free synthesis, *J Biomol Nmr 11*, 295-306.

147. Dawson, P. E., and Kent, S. B. H. (2000) Synthesis of native proteins by chemical ligation, *Annu Rev Biochem 69*, 923-960.

148. Decatur, S. M., and Antonic, J. (1999) Isotope-edited infrared spectroscopy of helical peptides, *J Am Chem Soc 121*, 11914-11915.

149. Decatur, S. M. (2006) Elucidation of residue-level structure and dynamics of polypeptides via isotope-edited infrared spectroscopy, *Acc Chem Res 39*, 169-175.

150. Kim, H. W., Perez, J. A., Ferguson, S. J., and Campbell, I. D. (1990) The specific incorporation of labeled aromatic-amino-acids into proteins through growth of bacteria in the presence of glyphosate - Application to fluorotryptophan labeling to the H+-Atpase of Escherichia-coli and Nmr-studies, *Febs Lett 272*, 34-36.

151. Imamoto, Y., Mihara, K., Hisatomi, O., Kataoka, M., Tokunaga, F., Bojkova, N., and Yoshihara, K. (1997) Evidence for proton transfer from Glu-46 to the chromophore during the photocycle of photoactive yellow protein, *J Biol Chem 272*, 12905-12908.

152. Brudler, R., Rammelsberg, R., Woo, T. T., Getzoff, E. D., and Gerwert, K. (2001) Structure of the I1 early intermediate of photoactive yellow protein by FTIR spectroscopy, *Nat Struct Biol 8*, 265-270.

153. Sambrook, J., and Gething, M. J. (1989) Protein structure. Chaperones, paperones, *Nature 342*, 224-225.

154. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular cloning: a laboratory manual*, Cold Spring Harbor Laboratory Press, New York.

155. Keller, A., Nesvizhskii, A. I., Kolker, E., and Aebersold, R. (2002) Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search, *Anal Chem 74*, 5383-5392.

156. Frisch, M. J., Trucks, G. W., Schlegel, H. B., Scuseria, G. E., Robb, M. A., Cheeseman, J. R., Montgomery, J. A., Vreven, J., T., Kudin, K. N., Burant, J. C., Millam, J. M., Iyengar, S. S., Tomasi, J., Barone, V., Mennucci, B., Cossi, M., Scalmani, G., Rega, N.,

Petersson, G. A., Nakatsuji, H., Hada, M., Ehara, M., Toyota, K., Fukuda, R., Hasegawa, J., Ishida, M., Nakajima, T., Honda, Y., Kitao, O., Nakai, H., Klene, M., Li, X., Knox, J. E., Hratchian, H. P., Cross, J. B., Adamo, C., Jaramillo, J., Gomperts, R., Stratmann, R. E., Yazyev, O., Austin, A. J., Cammi, R., Pomelli, C., Ochterski, J. W., Ayala, P. Y., Morokuma, K., Voth, G. A., Salvador, P., Dannenberg, J. J., Zakrzewski, V. G., Dapprich, S., Daniels, A. D., Strain, M. C., Farkas, O., Malick, D. K., Rabuck, A. D., Raghavachari, K., Foresman, J. B., Ortiz, J. V., Cui, Q., Baboul, A. G., Clifford, S., Cioslowski, J., Stefanov, B. B., Liu, G., Liashenko, A., Piskorz, P., Komaromi, I., Martin, R. L., Fox, D. J., Keith, T., Al-Laham, M. A., Peng, C. Y., Nanayakkara, A., Challacombe, M., Gill, P. M. W., Johnson, B., Chen, W., Wong, M. W., Gonzalez, C., and Pople, J. A. (2003) *Gaussian 03, Revision A.1*, Gaussian, Inc., Pittsburgh, PA.

157. Kumauchi, M., Hamada, N., Sasaki, J., and Tokunaga, F. (2002) A role of methionine100 in facilitating PYPM-decay process in the photocycle of photoactive yellow protein, *J Biochem-Tokyo 132*, 205-210.

158. Nie, B. (2006) Probing hydrogen bonding interactions and proton transfer in proteins, p 243, Oklahoma State University, Stillwater, OK.

159. Hienerwadel, R., Boussac, A., Breton, J., Diner, B. A., and Berthomieu, C. (1997) Fourier transform infrared difference spectroscopy of photosystem II tyrosine D using site-directed mutagenesis and specific isotope labeling, *Biochemistry 36*, 14712-14723.

160. Takahashi, R., and Noguchi, T. (2007) Criteria for determining the hydrogen-bond structures of a tyrosine side chain by fourier transform infrared spectroscopy: density functional theory analyses of model hydrogen-bonded complexes of p-cresol, *J Phys Chem B 111*, 13833-13844.

161. Gutteridge, A., and Thornton, J. M. (2008) Understanding nature's catalytic toolkit, *Trends Biochem. Sci. 30*, 1153–1165.

162. Pandini, A., and Bonati, L. (2005) Conservation and specialization in PAS domain dynamics, *Protein Eng Des Sel 18*, 127-137.

163. 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.

164. 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.

165. 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.

166. Crosson, S., and Moffat, K. (2001) Structure of a flavin-binding plant photoreceptor domain: Insights into light-mediated signal transduction, *P Natl Acad Sci USA 98*, 2995-3000.

167. Mihara, K., Hisatomi, O., Imamoto, Y., Kataoka, M., and Tokunaga, F. (1997) Functional expression and site-directed mutagenesis of photoactive yellow protein, *J Biochem 121*, 876-880.

168. van der Horst, M. A., van Stokkum, I. H., Crielaard, W., and Hellingwerf, K. J. (2001) The role of the N-terminal domain of photoactive yellow protein in the transient partial unfolding during signalling state formation, *Febs Lett 497*, 26-30.

169. Harigai, M., Imamoto, Y., Kamikubo, H., Yamazaki, Y., and Kataoka, M. (2003) Role of an N-terminal loop in the secondary structural change of photoactive yellow protein, *Biochemistry 42*, 13893-13900.

170. Holm L, R. P. (2010) Dali server: conservation mapping in 3D., *Nucl. Acids Res. 38*, 5.

171. Henry, J. T., and Crosson, S. (2011) Ligand-Binding PAS Domains in a Genomic, Cellular, and Structural Context, *Annual Review of Microbiology, Vol 65 65*, 261-286.

172. Vehlow, C. S., H.; Winkelmann, M.; Duarte, J. M.; Petzold, L.; Dinse, J. and Lappe, M. (2011) CMView: Interactive aontact map visualization and analysis . *Bioinformatics 10*.

173. Guex, N., and Peitsch, M. C. (1997) SWISS-MODEL and the Swiss-PdbViewer: An environment for comparative protein modeling., *Electrophoresis 18*, 10.

174. Aloy, P., Ceulemans, H., Stark, A., and Russell, R. B. (2003) The relationship between sequence and interaction divergence in proteins, *J. Mol. Biol. 332*, 989-998.

175. Rost, B. (1999) Twilight zone of protein sequence alignments, *Protein Eng. 12*, 85-94.

176. Vogt, G., Etzold, T., and Argos, P. (1995) An assessment of amino-acid exchange matrices in aligning protein sequences – The twilight zone revisited, *J. Mol. Biol. 249*, 816-831.

177. Yan, B., Takahashi, T., Johnson, R., and Spudich, J. L. (1991) Identification of signaling states of a sensory receptor by modulation of lifetimes of stimulus-induced conformations: the case of sensory rhodopsin II, *Biochemistry 30*, 10686-10692.

178. Jung, K. H., Trivedi, V. D., and Spudich, J. L. (2003) Demonstration of a sensory rhodopsin in eubacteria, *Mol Microbiol 47*, 1513-1522.

179. Janiak-Spens, F., Sparling, J. M., Gurfinkel, M., and West, A. H. (1999) Differential stabilities of phosphorylated response regulator domains reflect functional roles of the yeast osmoregulatory SLN1 and SSK1 proteins, *J Bacteriol 181*, 411-417.

180. Thomas, S. A., Brewster, J. A., and Bourret, R. B. (2008) Two variable active site residues modulate response regulator phosphoryl group stability, *Mol Microbiol 69*, 453-465.

181. Venyaminov, S., and Kalnin, N. N. (1990) Quantitative IR spectrophotometry of peptide compounds in water (H2O) solutions. II. Amide absorption bands of polypeptides and

fibrous proteins in alpha-, beta-, and random coil conformations, *Biopolymers 30*, 1259-1271.

182. Arrondo, J. L., Muga, A., Castresana, J., and Goni, F. M. (1993) Quantitative studies of the structure of proteins in solution by Fourier-transform infrared spectroscopy, *Prog Biophys Mol Biol 59*, 23-56.

183. Matthews, B. W. (1996) Structural and genetic analysis of the folding and function of T4 lysozyme, *FASEB J 10*, 35-41.

184. Matthews, B. W. (1995) Studies on protein stability with T4 lysozyme, *Adv Protein Chem 46*, 249-278.

185. Betney, R., and McEwan, I. J. (2003) Role of conserved hydrophobic amino acids in androgen receptor AF-1 function, *J Mol Endocrinol 31*, 427-439.

186. Myers, J. K., Pace, C. N., and Scholtz, J. M. (1995) Denaturant m values and heat capacity changes: relation to changes in accessible surface areas of protein unfolding, *Protein Sci 4*, 2138-2148.

187. Lee, B. C., Kumauchi, M., and Hoff, W. D. (2010) Modulating native-like residual structure in the fully denatured state of photoactive yellow protein affects its refolding, *J Biol Chem 285*, 12579-12586.

188. Mirny, L., and Shakhnovich, E. (2001) Evolutionary conservation of the folding nucleus, *Journal of Molecular Biology 308*, 123-129.

189. Larson, S. M., Ruczinski, I., Davidson, A. R., Baker, D., and Plaxco, K. W. (2002) Residues participating in the protein folding nucleus do not exhibit preferential evolutionary conservation, *Journal of Molecular Biology 316*, 225-233.

190. Van Beeumen, J. J., Devreese, B. V., Van Bun, S. M., Hoff, W. D., Hellingwerf, K. J., Meyer, T. E., McRee, D. E., and Cusanovich, M. A. (1993) Primary structure of a photoactive yellow protein from the phototrophic bacterium Ectothiorhodospira-

halophila, with evidence for the mass and the binding-site of the chromophore, *Protein Sci 2*, 1114-1125.

191. Devanathan, S., Brudler, R., Hessling, B., Woo, T. T., Gerwert, K., Getzoff, E. D., Cusanovich, M. A., and Tollin, G. (1999) Dual photoactive species in Glu46Asp and Glu46Ala mutants of photoactive yellow protein: A pH-driven color transition, *Biochemistry 38*, 13766-13772.

192. El-Mashtoly, S. F., Unno, M., Kumauchi, M., Hamada, N., Fujiwara, K., Sasaki, J., Imamoto, Y., Kataoka, M., Tokunaga, F., and Yamauchi, S. (2004) Resonance Raman spectroscopy reveals the origin of an intermediate wavelength form in photoactive yellow protein, *Biochemistry 43*, 2279-2287.

193. Derix, N. M., Wechselberger, R. W., van der Horst, M. A., Hellingwerf, K. J., Boelens, R., Kaptein, R., and van Nuland, N. A. J. (2003) Lack of negative charge in the E46Q mutant of photoactive yellow protein prevents partial unfolding of the blue-shifted intermediate, *Biochemistry 42*, 14501-14506.

194. Mulkidjanian, A. Y. (1999) Conformationally controlled pK-switching in membrane proteins: One more mechanism specific to the enzyme catalysis?, *Febs Lett 463*, 199-204.

195. Harris, T. K., and Turner, G. J. (2002) Structural basis of perturbed pK(a) values of catalytic groups in enzyme active sites, *Iubmb Life 53*, 85-98.

196. Shinkarev VP, W. C. (1993) *In: The Photosynthetic reaction center*, Vol. 1, Academic Press, San Diego.

197. Zscherp, C., Schlesinger, R., Tittor, J., Oesterhelt, D., and Heberle, J. (1999) In situ determination of transient pK(a) changes of internal amino acids of bacteriorhodopsin by using time-resolved attenuated total reflection Fourier-transform infrared spectroscopy, *P Natl Acad Sci USA 96*, 5498-5503.

198. Forsyth, W. R., Antosiewicz, J. M., and Robertson, A. D. (2002) Empirical relationships between protein structure and carboxyl pK(a) values in proteins, *Proteins-Structure Function and Bioinformatics 48*, 388-403.

199. Martin, C. R. (1993) Structure, Function, and Regulation of the Chalcone Synthase, *Int Rev Cytol 147*, 233-284.

200. Rajagopal, S., and Moffat, K. (2003) Crystal structure of a photoactive yellow protein from a sensor histidine kinase: Conformational variability and signal transduction, *P Natl Acad Sci USA 100*, 1649-1654.

201. DeLano, W. L. (2002) PyMOL: An Open-Source Molecular Graphics Tool, DeLano Scientific, San Carlos, CA.

202. Harigai, M., Kataoka, M., and Imamoto, Y. (2006) A single CH/pi weak hydrogen bond governs stability and the photocycle of the photoactive yellow protein, *J Am Chem Soc 128*, 10646-10647.

203. Kamikubo, H., Koyama, T., Hayashi, M., Shirai, K., Yamazaki, Y., Imamoto, Y., and Kataoka, M. (2008) The photoreaction of the photoactive yellow protein domain in the light sensor histidine kinase Ppr is influenced by the C-terminal domains†, *Photochem Photobiol 84*, 895-902.

204. Imamoto, Y., Harigal, M., and Kataoka, M. (2004) Direct observation of the PH-dependent equilibrium between L-like and M intermediates of photoactive yellow protein, *Febs Lett 577*, 75-80.

205. Borucki, B., Otto, H., Joshi, C. P., Gasperi, C., Cusanovich, M. A., Devanathan, S., Tollin, G., and Heyn, M. P. (2003) pH dependence of the photocycle kinetics of the E46Q mutant of photoactive yellow protein: Protonation equilibrium between I-1 and I-2 intermediates, chromophore deprotonation by hydroxyl uptake, and protonation relaxation of the dark state, *Biochemistry 42*, 8780-8790.

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

207. Yildiz, O., Doi, M., Yujnovsky, I., Cardone, L., Berndt, A., Hennig, S., Schulze, S., Urbanke, C., Sassone-Corsi, P., and Wolf, E. (2005) Crystal structure and interactions of the PAS repeat region of the Drosophila clock protein PERIOD, *Molecular Cell 17*, 69-82.

208. 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.

209. Harigai, M., Yasuda, S., Imamoto, Y., Yoshihara, F., Tokunaga, F., and Kataoka, M. (2001) Amino acids in the N-terminal region regulate the photocycle of photoactive yellow protein, *J Biochem 130*, 51-56.

APPENDICES

This section contains additional work that I performed during my PhD project. The first half includes the preliminary characterization of I31A, D24A, and Q41A mutants of Hhal PYP that can be used to probe the functional role of residues Asp24, Ile31, and Gln41 in PYP. The latter half includes the construction of the Δ25 mutant of Rcen PYP, protein expression and reconstitution, construction of the E46Q Ppr mutant of *R. centenum*, preliminary tests for detecting *chs* gene activity in wt and the Ppr deletion strains of *R. centenum* and the associated experimental complications.

A.1 Preliminary partial characterization of I31A, D24A, and Q41A mutants of Hhal PYP

Mutagenic primers from (*1*) were used to introduce the desired mutations and the protein samples were overexpressed, reconstituted and purified as described for I39A mutant in previous chapters. The data obtained during the preliminary characterization of these mutants is discussed individually in the following sections.

Asp24 is not conserved in PAS domains but is highly conserved in PYPs. Asp24 is in the N-terminal region and forms hydrogen bonds with Ala44 in the photoactive core of PYP that are thought to be important in regulating the signaling kinetics based on the alanine mutagenesis scan of Hhal PYP.

All the measurements were made on D24A Hhal PYP with an N-terminal 6X histidine tag. The absorbance spectrum of the pG ground state of the mutant at neutral pH was identical to wt Hhal PYP and displayed a λ_{max} at 446 nm. Two properties of D24A mutant were analyzed, the pK_a and the lifetime of the pH dependence of the pB decay kinetics. Low pH titration was performed and the absorbance spectra for each pH value were recorded. From the pH titration curve (Fig. A-1B), the pK_a was calculated to be 2.8 and the *n* value was 1.38 which is similar to the values observed for wt Hhal PYP. The transition from the unprotonated to protonated *p*CA (acid denatured state absorbing around 350 nm) was isosbestic and only two species with different protonation states were observed (Fig. A-1A).

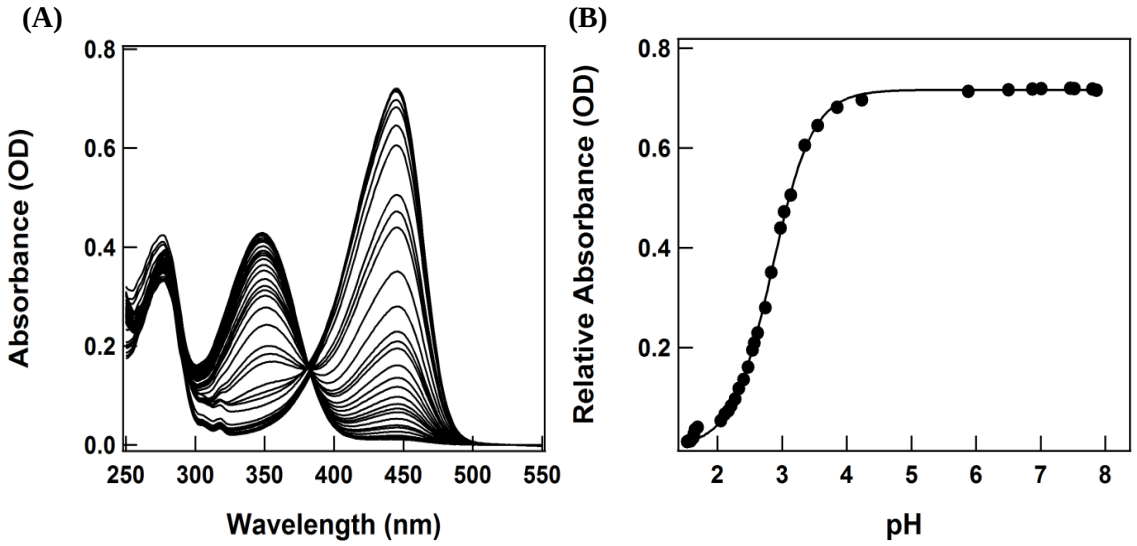


Fig. A-1 Effect of D24A mutation of pK_a of *p*CA. (A) pH dependent absorbance spectra of D24A PYP for a range of decreasing pH values from pH 8 to pH 1.5. (B) pH titration curves of D24A mutant obtained at 446 nm.

D24A showed photoactivity and the pB-pG difference spectrum at neutral pH exhibits a negative peak near 446 nm and a positive peak near 350 nm (Fig. A-2A) similar to wt Hhal PYP (see chapter 4).

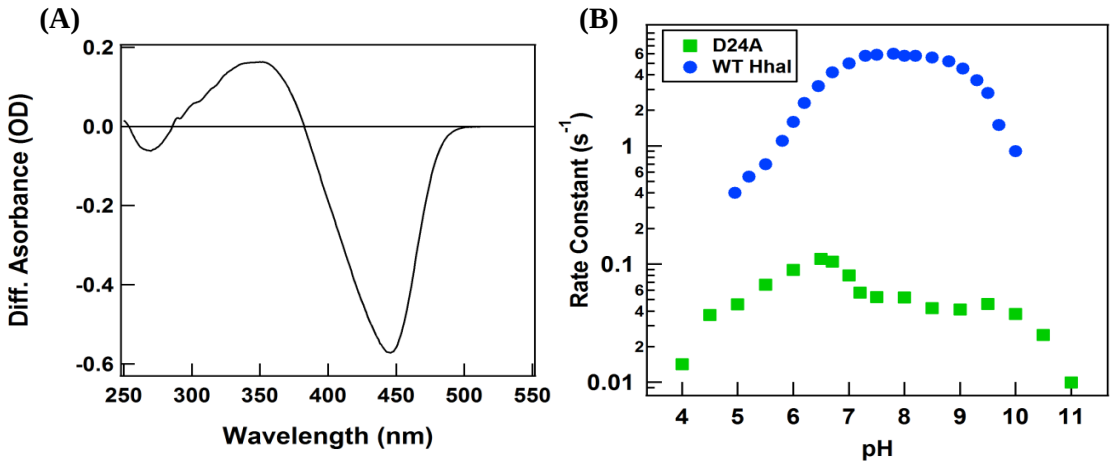


Fig. A-2 Photoactivity of D24A mutant of Hhal PYP. (A) pB-pG UV/vis absorbance difference spectrum of D24A. (B) The pH dependence of the kinetics of pB decay of wt Hhal PYP (*65*) and its D24A mutant (D24A kinetics – *measured by Josie Hilley*).

The rate of pB decay was measured for a range of pH values from pH 4 to pH 11 (Fig. A-2B). At pH 7, the time constant for the pB decay was observed to be ~15 seconds which is almost 30 times slower than wt PYP, similar to what was observed earlier for its I39A mutant. The rate was moderately affected by pH and was much slower at both the highest as well as the lowest pH value recorded for this mutant. A sharp feature similar to I39A mutant with the His tag was observed around pH 6, suggesting that protonation of the His residues modestly affect the rate of pB decay. The results obtained for D24A PYP confirm its role in regulating the pB decay kinetics and in future can be used to probe its role in the light induced structural changes using rapid scan FTIR spectroscopy.

From the alanine scan we have identified Ile31 as one of the six positions in PYP substitution at which alters the λ_{max}, pK_a, and pB lifetime (others are: G29, Y42, N43, E46, I49). Its side chain points towards phenolic O^- of *p*CA (within 5 Å) and unpublished collaborative

computational results (coarse grained molecular dynamics simulations) indicate that Ile31 is one of the few residues forming novel non-native contacts in the pB state. Due to its distinct location Ile31 could be important for the structural integrity and maintaining the stability of PYP. All measurements were made on the N-terminal 6X histidine tagged I31A protein sample at room temperature. The absorbance spectrum of the pG dark state of the I31A mutant was measured in 10 mM Tris-HCl using UV/vis spectroscopy. The λ_{max} of I31A was observed to be at 450 nm which is 4 nm red-shifted compared to wt Hhal PYP. Measurements on the pH dependence of the kinetics of pB decay of I31A Hhal PYP mutant were performed. The light-dark difference absorbance spectrum of I31A at pH 7.5 displayed a negative peak near 450 nm and a positive peak near 350 nm indicating photoconversion of the pG ground state of PYP (Fig. A-3A).

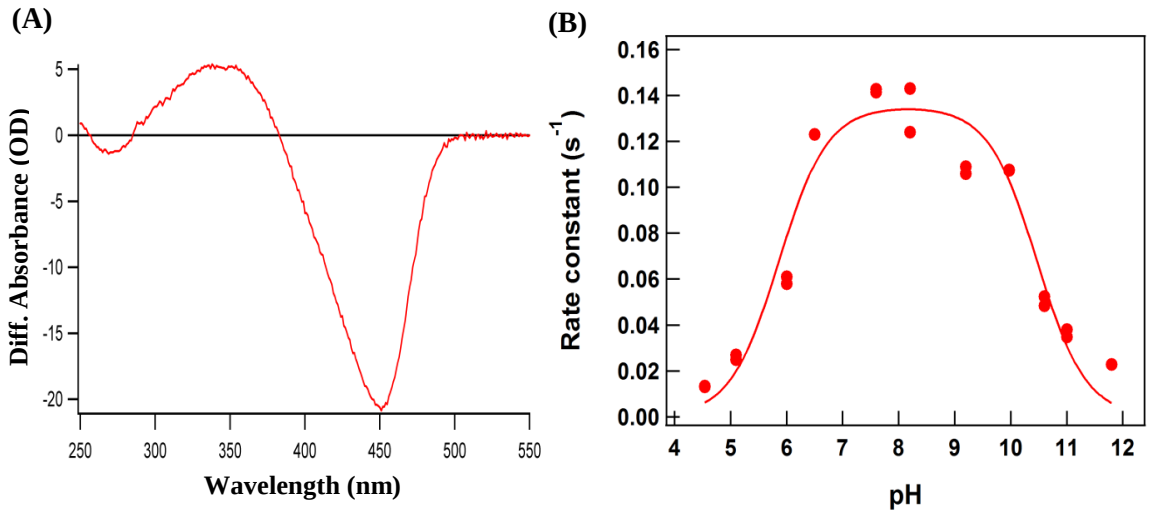


Fig. A-3 Photoactivity of I31A mutant of Hhal PYP. (A) pB-pG UV/vis absorbance difference spectrum of I31A. (B) The pH dependence of the kinetics of pB decay of the I31A mutant of Hhal PYP.

The pH dependence of the rate of pB decay was also studied from pH 4.5 to pH 12 and the data could be described using a bell-shaped fit (Fig. A-3B) as was previously reported for wt Hhal PYP. This fit for I31A PYP was governed by two apparent pK_a values (5.8 and 10.5), which

is slightly different from wt PYP, where the pK_a values were reported to be 6.4 and 9.4 (see Fig. A-2B) (*65*). At pH 7.5 the time constant for pB decay of I31A was observed to be around seven seconds, which is slower than wt PYP by a factor ~14. The rate was also moderately dependent on pH. In conclusion, the results demonstrate that the I31A mutant of PYP does affect both active site and functional properties. Thus, along with D24A, I31A is a promising candidate for probing light induced conformational changes in the pB state using FTIR difference spectroscopy.

From the alanine scan data for Hhal PYP it was identified that the Q41A mutation slows down the decay of the pB state in the PYP from *H. halophila*. In addition, in time-resolved FTIR spectroscopy measurements a band at 1689 cm^{-1} is present in the pB-pG difference spectrum, which has been used as a marker for structural changes upon pB formation (Fig. A-4) (*33*). This band could be produced by the C=O stretching of Gln41. To test this hypothesis, the Q41A PYP sample was purified and characterized with the goal of performing rapid-scan measurement by FTIR to determine if the band at 1689 cm^{-1} is absent in the Q41A mutant. This would allow the assignment of this signal to the side chain of Gln41.

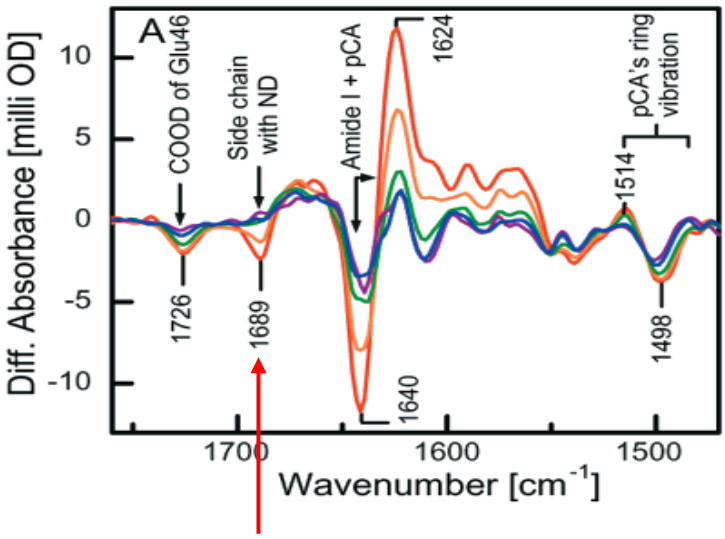


Fig. A-4 Adapted from (*33*) – Time-resolved FTIR difference spectra of wt-PYP at pH* 7 in D_2O at 25 µs *(purple), 100* µs *(blue), 400* µs *(green), 1.6* ms (orange), and 6.6 ms (red).

The Q41A mutant of PYP with an N-terminal 6X His tag was purified and used to study its absorbance spectrum and the photocycle kinetics. The absorbance spectrum of the pG dark state of Q41A PYP measured at pH 7.5 at room temperature was identical to wt PYP, with a λ_{max} at 446 nm. Using 2% SDS as explained in Chapter 2 for other PYP mutants, the molar extinction of *p*CA absorption band of Q41A was determined to be 41,902 $M^{-1}cm^{-1}$ slightly reduced compared to the value of wt PYP (45, 500 $M^{-1}cm^{-1}$). All other measurements were performed as described for I31A mutant. The UV/vis pB-pG difference spectrum was observed to be identical to wt PYP with a negative peak near 446 nm and a positive peak near 350 nm (Fig. A-5A). The photocycle kinetics for the Q41A mutant was measured over the pH range pH 3.5 to pH 11.2. The data fit the bell shaped curve described by two pK_a values at 4.3 and 10.7, with an unusual pH dependence feature around pH 6, similar to D24A and I39A mutants (Fig. A-5B). The rate of pB decay was very significantly affected by pH in Q41A PYP, and at neutral pH the pB lifetime was 3.1 s, approximately six times slower than wt PYP. Thus the data indicate only moderate reduction in pB lifetime of the Q41A mutant of PYP.

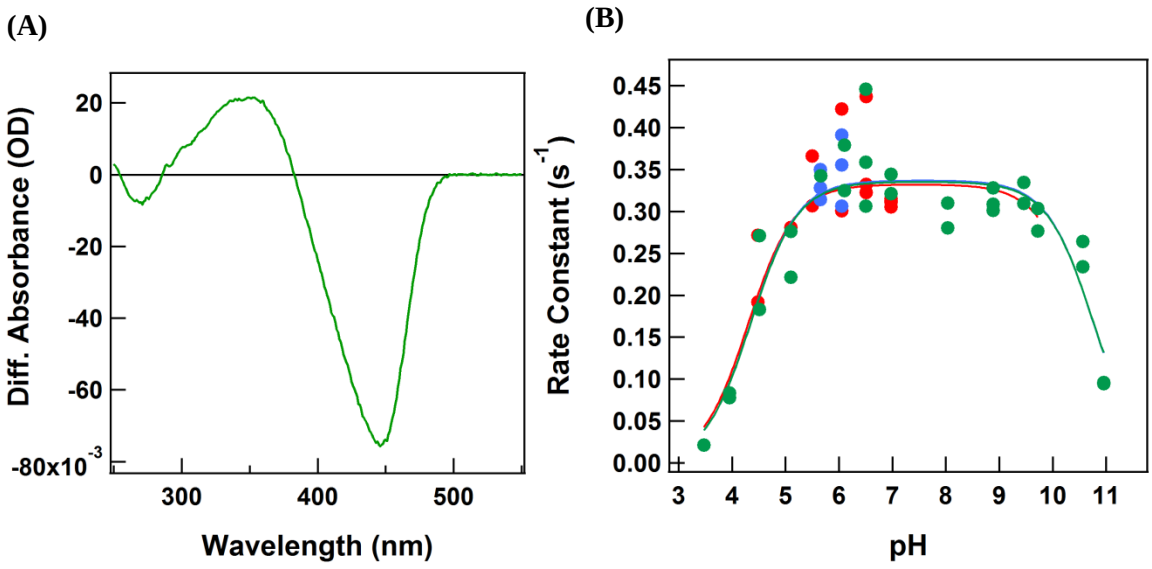


Fig. A-5 Photoactivity of Q41A mutant of Hhal PYP. (A) pB-pG UV/vis absorbance difference spectrum of Q41A. (B) The pH dependence of the kinetics of pB decay of the Q41A mutant of Hhal PYP (the red and blue solid dots are kinetic measurements that were repeated for pH 5.5 – pH 7).

To probe the 1689 cm^{-1} signal, time resolved FTIR difference spectroscopy was used. For these measurements, the Q41A PYP mutant without the 6X His tag (obtained by recloning in pET16b without the His tag) was used to avoid any interference due to histidine. The pB–pG FTIR difference spectrum of Q41A was identical to wt PYP. The structural changes occurring upon the formation of the pB state were similar in both the samples with all the important spectral markers present (see Fig. A-6). Interestingly, the 1689 cm^{-1} signal was observed in the Q41A mutant which was contradictory to our prediction. Since this signal was also observed in the N43A mutant of PYP, we believe that the most likely candidate origin of the 1689 cm^{-1} signal is the amide I of the protein backbone in a β-sheet conformation.

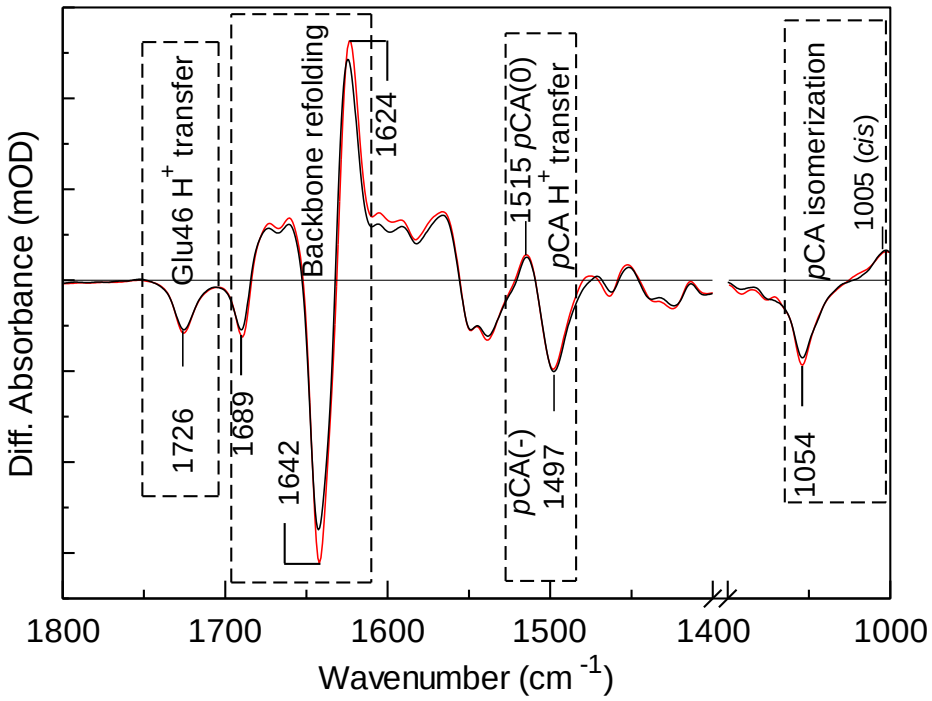


Fig. A-6 Structural changes upon pB formation in Hhal PYP detected by time-resolved FTIR difference spectroscopy. The above panel shows the pB–pG difference spectra of wt PYP (black) and its Q41A mutant (red) at pH 7.5 in D_2O (*Data measured by Sandip Kaledhonkar*).

A.2 Additional work on Rcen PYP, Rcen Ppr, and chalcone synthase gene expression

N-terminally truncated mutants of Hhal PYP have been reported to be yellow and photoactive, but with a greatly reduced rate of pB decay. We wanted to examine if N-terminal truncation of Rcen PYP had similar effects. For the construction of Δ25 Rcen PYP, The following mutagenic primer pair was used:

D25-RPYP-FW –5'-CATATGCCGGTCGGCGCCAT-3' (with *Nde*I)
RPYP – RV (w/ *Eco*RI) – See chapter 2

The Δ25 Rcen PYP was cloned in pET16B and transformed in *E. coli* BL21 (DE3). Our attempts to overexpress the protein in LB medium were unsuccessful: the protein failed to express. To check for suspected problems of misfolding or incomplete folding of the protein, different induction conditions were explored. The cells were grown overnight (~11 h) with ampicillin (50μg/mL) and then induced with 1M IPTG at 25°C (4 h and 8 h) and at 37 °C (1 h and 3 h). The cells were then spun down, lysed and reconstituted following the same protocol described in Chapter 2. The protein failed to express in all the test conditions used and the absence of PYP was confirmed by SDS-PAGE. As discussed in chapter 6, removal of the entire N-terminal region could cause the protein to be unable to fold or reconstitute, and logical next step would be to partially delete the N-terminus and then study its effects on the functioning of Rcen PYP.

Construction of E46Q Ppr

We performed experiments with the aim of testing the effects of the point mutations of Rcen PYP studied in Chapter V in the context of the full-length Ppr protein. In this work we used the following available *E. coli* strains constructed in Dr. Carl Bauer's lab:

1. BL21 (DE3)/pZJ191, which is an expression strain for the PYP domain from Ppr. pZJ191 is a KmR pET28a (+) plasmid that contains a *Nde*1-*Eco*R1 PCR generated fragment (amino acids 1-130) of the PYP domain.
2. BL21 (DE3)/pZJ137::Ppr/pEThmuO. This strain has the KmR plasmid pZJ137::Ppr which is a full length PCR amplified *ppr* gene cloned into the *Nde*I-*Eco*RI sites of pET28a (+). This strain also harbors the AmpR plasmid pEThmuO, which contains a heme oxygenase gene that converts heme to billiverdin which is the other chromophore associated with Ppr.
3. BL21 (DE3) strain has the plasmid pBBR1MCS-5, which is a gentamycin resistant plasmid that replicates in *R. centenum*.

The full length *ppr* gene is ~2.6 kb and due to its large size it is difficult to perform mutagenic PCR using the QuikChange kit. Also, due to the lack of unique restriction sites needed for the cloning of the available E46Q Rcen PYP (from chapter 3) into the full length Ppr containing pZJ137 plasmid, we used a partial fragment of Ppr (~600 bp, digested at *Nde*I and *Fse*I) to introduce the E46Q mutation. For this step, we used the pBluscript SK- plasmid which was modified to allow the introduction of the *Fse*I site by disrupting the existing *Pst*I site. After the E46Q mutant was introduced in the partial fragment of Ppr in SK- the fragment was cloned into the expression plasmid pZJ137. The next step was to introduce the full length E46Q Ppr gene in the shuttle vector pBBR1MCS-5. Multiple attempts were made to clone the gene, but the vector did not incorporate the gene. We suspect that the ligation step could be the major hurdle in this process, possibly due to large size of the gene.

Chalcone synthase gene activity in R. centenum

Available *R. centenum* strains from Dr. Carl Bauer's lab:

1. Wild type *R. centenum* containing a fusion of the chalcone synthase promoter to *Lac*Z on a suicide plasmid that has been integrated into the genome at a neutral intergenic region. This plasmid has a GmR gene and can be selected with 10 µg/mL gentamycin.
2. The Ppr deletion strain (C145) that contains the same chalcone promoter fused to *Lac*Z integrated into the chromosome.

Ideal growth conditions for R. centenum – *R. centenum* is best grown at 42°C on CENS medium. The cells do not store well at 4°C. Instead, cells can be maintained on CENS plates for few weeks at room temperature or stored for longer times at -80°C as a glycerol stock. They grow well both aerobically in the dark or anaerobically in the presence of light using 60 W incandescent bulbs as a light source.

To measure *chs* activity both the wt and the C145 mutant, cultures of *R. centenum* were grown in CENS medium at ~37–42°C in different light conditions (white, blue, and infrared) to reproduce previously reported results (*101*). Quantitative assays for β-galactosidase activity were performed on both the strains. Unexpectedly, almost no activity was observed for cells grown in all conditions.

Since the above method did not yield the expected results, a second approach based on qRT-RT-PCR was used to measure *chs* expression. This work was performed in Dr. Carl Bauer's lab at Indiana University, Bloomington. In addition to the above mentioned two strains, wt and C145 strains of *R. centenum* without a *Lac*Z fusion were also used for these experiments. The growth medium was changed to Peptone-yeast extract-soytone (PYVS) medium to replicate exact conditions used in (*101*). All strains were inoculated in PYVS medium with 1µg/mL gentamycin and were grown overnight at 37°C in white light until OD_{650} reached ~1.5 after which each strain

was subcultured to obtain 0.01OD. The subcultures were again incubated at 37°C in white light until OD_{650} reached ~0.5–0.6. Three biological replicates were used for each strain. These cells were used for RNA isolation using BIOLINE's Isolate RNA Mini Kit and the isolated RNA was used for performing qRT-RT-PCR using the SensiFAST™ SYBR Hi-ROX One-Step Kit. *bcs*A is the gene for *chs* activity and the house keeping gene *rpo*Z was used as an internal standard. The primers for each were designed using PRIMER 3,

*bcs*A - FW 5' CAG GAA CTG ACC CGC GAT 3'

RV 5' CTA CGG CGC TTT CCA GAT ACA 3'

*rpo*Z - FW 5' GAC AAC GAC AAG AAC CCG GT 3'

RV 5' TCG TCA TCC AGC AGT TCG GA3'

The quantification of the RNA levels was performed using relative quantification by the $\Delta\Delta C_t$ method to calculate the fold difference. Using this method showed ~ 0.5-0.6 fold difference in the C145 mutant as compared to wt *R. centenum* (both the samples). Although we did not obtain the expected larger difference using q-RT-RT-PCR, these experiments did show different levels of gene expression, which can be used in the future to study *chs* gene activity in optimized light conditions.

VITA

Rachana Ramila Suresh Rathod

Candidate for the Degree of

Doctor of Philosophy

Thesis: STRUCTURE–FUNCTION RELATIONSHIPS IN THE PHOTOACTIVE YELLOW PROTEIN FAMILY OF PHOTORECEPTORS

Major Field: Microbiology and Molecular Genetics

Biographical:

Education:

Completed the requirements for the Doctor of Philosophy/Education in Microbiology and Molecular Genetics at Oklahoma State University, Stillwater, Oklahoma in December, 2014.

Completed the requirements for the Master of Science in Botany at University of Mumbai, Mumbai, Maharashtra, India in 2006.

Completed the requirements for the Bachelor of Science in Botany at University of Mumbai, Mumbai, Maharashtra, India in 2004.

Experience: Teaching and Research Assistant at Department of Microbiology and Molecular Genetics, Oklahoma State University, Stillwater, OK.

Professional Memberships: American Society of Microbiology