Phytochrome Fluorescence and Folding

by

Shyamosree Bhattacharya

A dissertation submitted in partial fulfillment of

the requirements for the degree of

Doctor of Philosophy

(Chemistry)

at the

UNIVERSITY OF WISCONSIN - MADISON

2016

Date of final oral examination: 08/12/2016

The dissertation is approved by the following members of the Final Oral Committee:

Katrina T. Forest, Professor, Bacteriology

James L. Keck, Professor, Biochemistry

James C. Weisshaar, Professor, Chemistry

Silvia Cavagnero, Professor, Chemistry

Ronald T. Raines, Professor, Chemistry

Acknowledgements

This work would not have been possible without my advisor Dr. Katrina Forest. I would like to extend my thanks to my committee members Drs. James Keck, James Weisshaar, Silvia Cavagnero and Ronald Raines. I would like to thank my graduate school mentor Dr. Michele Auldridge as well as Dr. Kenneth Satyshur. Their contribution was crucial in my papers and understanding of structural biology.

I would like to take this opportunity to thank all the past and present members of the Forest laboratory. Special mention goes to Dr. Anna Baker, Katiria Gonzalez-Rivera, Peter Newhouse, Jeffrey Dwullit Smith, Neydis Moreno and Austin Becker. They have been my constant companions the last few years. Laura Franz, Drs. Ana Misic and Lorraine McLaughlin were instrumental in my initial laboratory training and I am very grateful for their continued help and support.

Lastly I would like to thank my family and friends for encouraging me through this entire process. I wouldn't be here without my parents Anup and Keya Bhattacharya and my uncle and aunt Sutirtha and Dimpi Bhattacharya. My brother Swarbhanu Bhattacharya has been a silent support these past few years along with my cousins Shouvik and Shoumik. Kelly, Kate, Jennifer, Jaritza, Paul, Caryn, Zhihui and Michael have been my pillars of support in Madison and I couldn't have done this without them.

Abstract

Use of fluorescent proteins in studying *in vivo* processes in mammalian systems requires development of near-infrared biomarkers due to clear signals unimpeded by absorption or autofluorescence of biomolecules (1). Bacteriophytochromes that use biliverdin as their chromophore have been engineered to form monomeric near-infrared biomarkers. The original design template for a fluorescent BphP was the chromophore binding domain of *Deinococcus radiodurans* (DrCBD), with a D207H substitution (2). This variant proved to be the hallmark for the next generation of phytochrome biomarkers including IFP1.4 (D207H+11 substitutions), Wi-Phy (D207HY263F) and the iRFP family based on *Rhodopseudomonas palustris* BphPs (3–6).We solved the x-ray crystal structures of IFP1.4 and several monomeric DrCBD variants, to explain the origins of fluorescence in derived BphPs (4, 7). A comparative study revealed two main themes. First, hydrophobic packing around the D-ring increases fluorescence by limiting D-ring motion. Second, while the 207th residue is critical to photochemistry it certainly does not need to be His in order to maximize fluorescence. In fact, the highest fluorescence quantum yield to date in this family belongs to a D207L variant in which waters are excluded from the chromophore vicinity (WiPhy2). Higher quantum yields and longer excited state lifetimes than in the parent suggest the loss of a de-excitation pathway via proton transfer. Continuing our structural analysis, we have turned to the formation of the deep figure of eight knot in the protein structure and uncovered the kinetic dependence of knot formation on cis-trans isomerization of Pro236 in the lasso loop of the GAF domain, which is essential for threading of the knot. We have also engineered a smaller phytochrome variant without the characteristic figure of eight knot topology. Our engineered knotless

CBD folds and binds bilin. Combining structural insights with protein engineering will help us understand assembly of holo-phytochrome and design better near-infrared fluorophores.

Table of Contents **Page Numbers**

List of Figures and Tables **Page Numbers**

Chapter I

Introduction

Near-IR fluorescent biomarkers

Visualization of molecular processes has revolutionized the way we understand life at the cellular level. Biochemical processes were difficult to monitor in live model systems before the advent of fluorescent biomarkers. Studying *in vivo* processes involved the use of toxic and radioactive materials and often interfered with the system being studied. The discovery of green fluorescent protein from the jellyfish *Aequorea victoria* has facilitated the age of molecular visualization (9, 10). Subsequently, in order to enhance fluorescence and extend or decrease the stability of the wild-type GFP at higher temperatures, a number of laboratories have modified its native structure (10–17). GFP and GFP-derivatives have become the most widely used *in vivo* protein markers for studying a plethora of different molecular processes during development. For example, the contraction of individual encardiomyocytes can be visualized based on GFP fluorescence in live embryos; GFP has also been used as a reporter for virus infections in plants (18, 19).

Despite the extensive utilization of GFP, it or any other currently available biomarkers has certain drawbacks. Most of the bioluminescent and fluorescent proteins used as molecular imaging tools are excited and emit in the visible region of the electromagnetic spectrum (Figure 1.1). The visible wavelength range is limiting in terms of *in vivo* use of GFP-based proteins because the excitation light can poorly penetrate deep tissues due to absorption by lipids and hemoglobin and scatter by water (1, 20, 21). The question arises as to whether we can make *in vivo* imaging feasible for deep tissues or not. One way to address this issue is to develop a protein-based fluorophore with excitation

maximum in the near IR region of the electromagnetic spectrum. These longer wavelengths (650-900 nm) have minimal absorbance by biomolecules and are less prone to the dissipating effects of light-scattering compared to their counterparts which are excited in the visible region of the spectrum. Light in the near IR region also transmits through deep tissues and bones. As a result, the signal to noise ratio of biomarkers that fluoresce in the >650 nm region is significantly higher than its visible spectrum counterparts (22, 23). Near-infrared fluorescent spectroscopy allows for real-time visualization of biological process like tumor biogenesis, cancer progression, effect of biomolecules and drugs (24–26). In recent years, near-infrared fluorescent spectroscopy has also been used to quantify brain amyloid depositions which can be used for diagnosis of neurological disorders like Alzheimer's and other protein misfolding diseases (27). Use of near-infrared probes for spectroscopy is a sensitive and versatile tool and has wide applications in clinical research.

There are two ways the near-infrared fluorophores can be developed, organic small molecule tracers or genetically encoded molecules (23, 28). A lot of progress has been made in the development of genetically encoded fluorophores using natural photoreceptors as starting material for fluorescent probes. Beside the GFP family of proteins, engineered variants of Light-oxygen-voltage (LOV) domain proteins, the blue-light utilizing FAD domain and cryptochromes have all been modified for fluorescent purposes (29–32). In the following sections we are going to look into bacterial phytochromes (BphPs) as a template for fluorophore design.

Introduction to Phytochromes

Phytochromes are photoreceptors that use a linear tetrapyrrole moiety as their small molecule cofactor to sense red and far red light (33, 34). The chromophore is tethered via a thioether bond to a highly conserved cysteine residue. Phytochromes are ubiquitous in the plants, bacteria and fungi (35, 36). The phytochrome family of proteins is usually dimeric and has a very modular architecture with an N-terminal photosensory core domain and a C-terminal regulatory domain. The N-terminal core can be further subdivided into the Per-ARNT-SIM (PAS) domain, the cGMP phosphodiesterase/adenylate cyclase/FhlA (GAF) domain and the phytochrome-specific (PHY) domains. The small molecule chromophore is covalently bound to the N-terminal region in BphPs and within the GAF domain in Cph and plant phytochromes (Figure 1.1A). In all cases the chromophore interacts with conserved residues in the GAF domain. The C-terminal regulatory domain is usually a Histidine Kinase in bacterial and fungal phytochromes, though other protein-protein interaction domains like diguanylate cyclase, diguanylate phosphodiesterase have also been discovered in recent years (35). These output domains are responsible for conveying light signal further downstream for cellular signalling. Phytochromes are unique in their ability to reversibly photoconvert between the red light absorbing Pr ground state and the far red light absorbing Pfr ground state which are both relatively stable. The Pr state is considered the stable ground state and photoconversion results in formation of Pfr state. The reversible photoconversion is key to phytochrome's function as a photoswitch in cellular signalling (37–39), (Figure1.1).

BphPs use Biliverdin IXα (BV) as their small chromophore, whereas plant and cyanobacterial phytochromes use PCB and PΦB respectively (40). The difference in the chromophores results in different absorbance spectra in BphPs. BphPs which are bound to BV have an absorbance maximum at 700 nm which is more red-shifted than the absorbance maxima of plant and cyanobacterial phytochromes (~630nm). Upon excitation by red-light photons, phytochrome photoconverts to a far-red light absorbing Pfr state (absorption maxima ~720 nm for BV and 650 nm for PCB and PΦB), reversibly, (41, 42). BphPs which prefer the Pfr state as their primary ground state have been discovered and they are referred to as the bathyphytochrome sub-family (42, 43). BV, PCB and PΦB are products of heme catabolism and are produced by region-specific cleavage of the heme moiety using heme oxygenase enzyme.

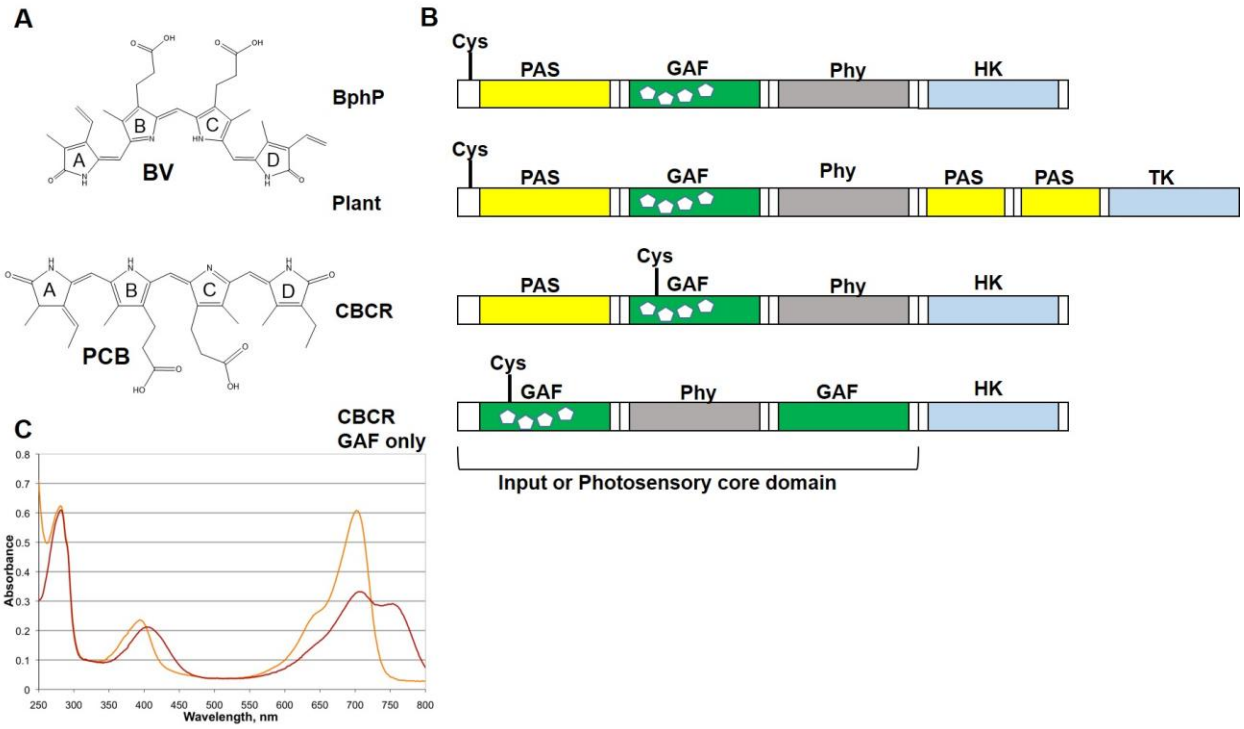
Historically, phytochromes have been studied in plants. The first phytochrome paper was published in 1959 and discussed effect of red/far red photoswitching on seed germination (44, 45). Since then, genetic studies have successfully elucidated the signalling cascades triggered by red to far-red photoconversion in plants. However due to inability to isolate and purify full-length plant phytochromes, much of the chromophore-protein interactions, both biochemical and structural, remained unknown. Fortunately, bacterial and cyanobacterial phytochrome systems proved more tenable for recombinant expression and purification and the first phytochrome structure was solved in 2005 from the bacterium *Deinococcus radiodurans* (42, 46). This work done in collaboration between the Forest and Vierstra laboratories at University of Wisconsin-Madison was an important milestone in phytochrome research and one of the seminal papers in structural biology. The CBD was crystallized and the initial structure was solved at 2.5Å resolution, which

was subsequently improved to a 1.45Å resolution structure in 2007 (47). The chromophore BV was attached to an N-terminal cysteine residue through a thioether linkage. In addition to the ligation site, several key features about the chromophore-protein interaction were revealed which were key to understanding how phytochromes function as photoreceptors. One of the most interesting features was the presence of the figure-of-eight knot which connected the PAS and the GAF domains together. The phytochrome protein structure was one of the first examples of a deeply knotted protein. A 35 amino acid residue extension upstream of the PAS domain threads through the lasso loop (generated by two β strands of the GAF domain) to form this rare structural motif with four cross-over points. Subsequent research has shown that the figure-of-eight knot in conserved among all PAS containing phytochromes (48, 49). In 2014, the crystal structure of the longer *Deinococcus radiodurans* PAS-GAF-PHY domains of BphP supported our hypothesis that the knot is responsible for imparting a degree of stiffness to the phytochrome molecule which possibly allows downstream signal transduction through angstrom level change in protein movement (50, 51). The first structure of plant phytochrome *Arabidiopsis thaliana* PhyB protein has helped characterize the knot even further. The knot in PhyB is more extensive and has a different amino acid composition in the GAF domain lasso loop region (52). The difference in the light sensing knot region would explain the differences in BphP and plant phytochrome absorbance spectra.

In addition to the figure-of-eight knot, several key interactions were responsible in maintaining the small molecule chromophore in the binding pocket and affect phytochrome photochemistry. BV is embedded in a deep binding pocket in the GAF domain in a *ZZZssa* configuration and several polar interactions with arginine, serine and

histidine residues are responsible for maintaining the Pr conformation of the chromophore (46). Photoconversion from Pr to Pfr is associated with movement of the D-ring and this bears out in the 3D structure where the D-ring is ~44° out of plane in contrast to the rest of the BV molecule (53). This strain reduces the energy required for the C15-C16 rotation that leads to the flipping of the D-ring during photoconversion (54, 55). One of the crucial H-bonding networks in the binding pocket involves a water mediated H-bond between the A, B and C rings of BV and the carbonyl backbone of the aspartate 207 residue. The aspartate 207 residue is part of a highly conserved asp-ile-pro motif and has important implications in phytochrome photochemistry (56, 57). Substitution of the aspartate 207 residue completely alters phytochrome photochemistry and in some cases results in formation of fluorescent phytochromes (2–6, 58) .

Figure 1.1 Introduction to phytochrome superfamily in plants, proteobacteria (BphP) and cyanobacteria (Cph). A. The small molecule chromophores associated with the phytochrome families. BphPs use biliverdin IXα, plants use phytochromobilin and cyano bacteria use phycocyanobilin (PCB). Structures of the linear tetrapyrroles are shown. PCB and PΦB are very similar except for the C18 ethyl side chain in PΦB. B. Domain architecture of phytochrome families from plants, cyanobacteria (Cph) and proteobacteria (BphP). The input domain consists of the PAS, GAF and Phy domains and the output domains are shown in blue. The cysteine residues on the N-terminal end of the CBD forms the thioether linkage with the A-rings of the bilin molecules and the chromophore itself is housed in a deep pocket in the GAF domain. C. Absorption spectrum of monomeric *Deinococcus radiodurans* CBD (DrCBD) bound to BV. The Pr spectrum was measured in dark. The sample was subsequently irradiated with red light at 700 nm. After ensuring sample saturation absorbance spectrum for the Pfr state was collected. The absorbance maximum for Pfr is 750 nm.


A
BV
BphP
Plant
PCB
CBCR
B
Cys
PAS
GAF
Phy
HK
PAS
GAF
Phy
PAS
PAS
TK
Cys
PAS
GAF
Phy
HK
Cys
GAF
Phy
GAF
HK
Input or Photosensory core domain
C
CBCR
GAF only
Absorbance
Wavelength, nm

Bacteriophytochrome Photochemistry

In order to understand the signalling cascades activated by light reception of phytochromes, it is essential to get a deeper understanding of the photochemistry of the bilin chromophore. Light signal is propagated via the chromophore into the protein molecule itself. The rapidly expanding three dimensional structural database has been instrumental in revealing how the photoisomerization of a C=C bond in the linear tetrapyrrole leads to changes in the protein three dimensional structure mediated via the figure-of-eight knot through the long helix into the output modules (47, 48, 50, 52, 59) (Figure 1.2A).

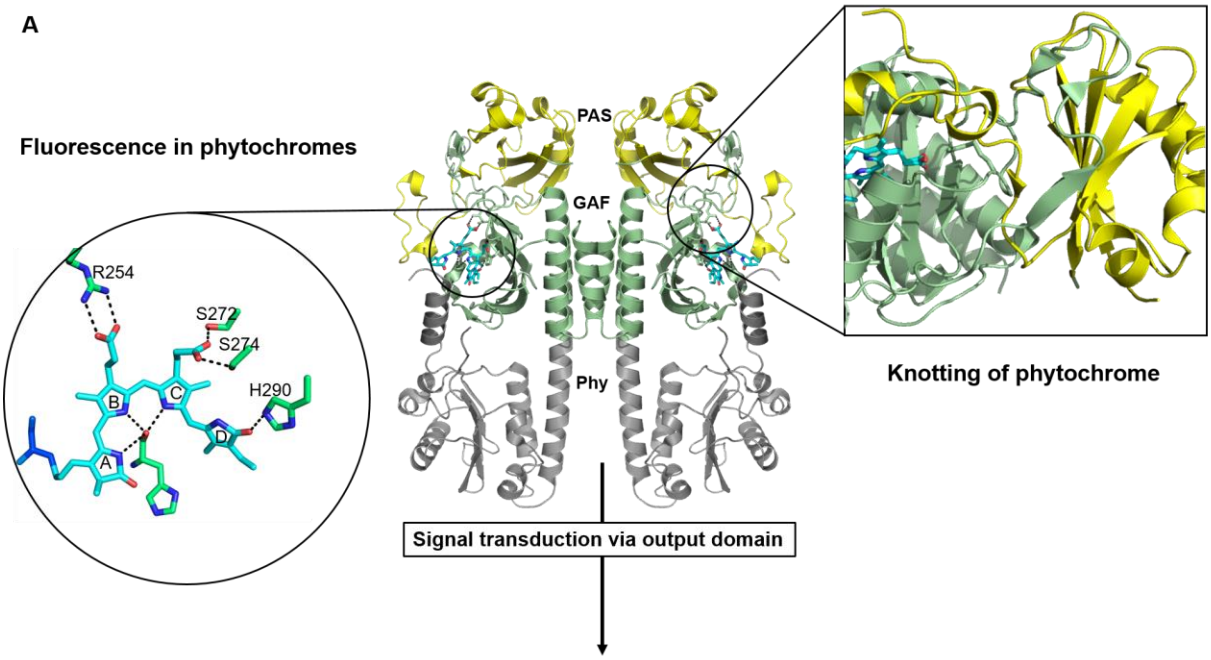
Full-length BphPs can exist in two stable ground states- Pr and Pfr which are photo-interconvertible. The Pr-Pfr Photoreversible chemistry is unique to phytochromes and is driven by conformational changes of the chromophore BV, During the Pr→ Pfr photoconversion, the chromophore BV, goes through a number of intermediate states during the Z→E isomerization of the C15=C16 bond (60, 61). In addition, intermediates are also formed during de-protonation followed by re-protonation of the chromophore. Both the isomerization and reversible protonation of BV lead to conformational changes that generate the Pfr state (48, 60, 62, 63).

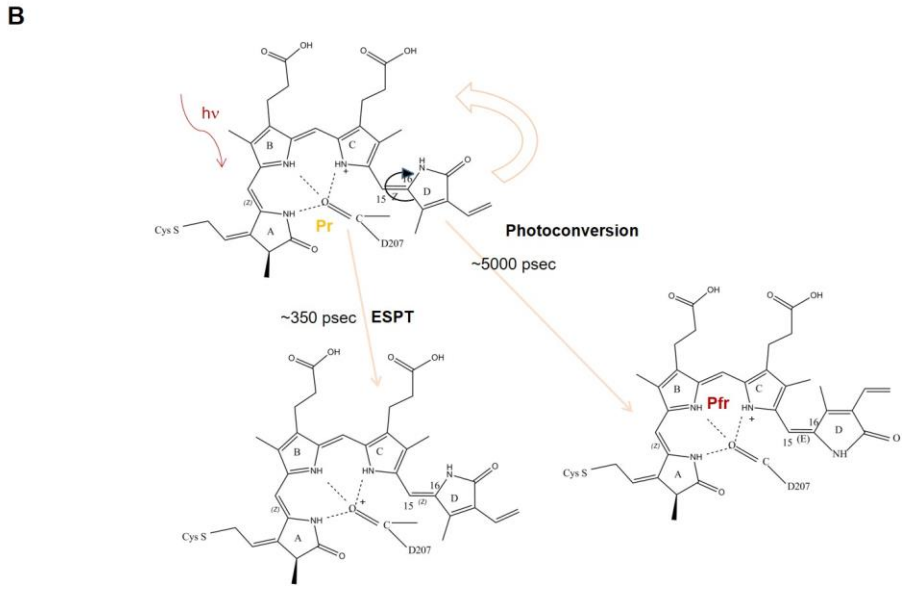
The first crystal structure of the DrCBD identified a number of protein/chromophore/water contacts that are probably important to phytochrome photochemistry (2, 46, 47). BV is locked in place by a complex lattice of static interactions and assumes in a *ZZZssa* configuration with the nitrogens in the A-C pyrrole rings coordinated with a fixed ‘pyrrole water’. The pyrrole water, in turn, forms conserved interactions with aspartate and

histidine residues (Asp207 and His260), respectively in DrBphP (Figure 1.2). This lattice is essential for proper bilin photochemistry with the pyrrole water probably serving as the immediate proton acceptor during photoconversion (57, 64). Whereas the A-C pyrrole rings are relatively co-planar, the D pyrrole ring is contorted out-of-plane by an electrostatic interaction with an invariant histidine (His290 in DrBphP) (Figure 1.2B).

Time resolved spectroscopic work done by John Kennis' laboratory in 2009 on *Rhodopseudomonas palustris* BphP2 and BphP3 has shown that reversible deprotonation of the chromophore mediated via excited state proton transfer is the dominant pathway for relaxation and accounts for 89% of the excited state population. In excited state proton transfer the delocalized proton over the BV is transferred through a polar interaction to the main chain carbonyl residue on the highly conserved aspartate 216 residue (Asp207 in *Deinococcus radiodurans*). Radiative decay to the Pr ground state or fluorescence only accounted for 5% of the de-excitation process in *R. palustris* P3 and photoconversion through rotation of the D-ring was estimated to be 6% (65). It is evident from the results that fluorescence, photoconversion and ESPT are competitive pathways used by the excited state holoprotein molecule to relax back to ground state, either Pr or Pfr. We can exploit the spectroscopic information to manipulate fluorescence of bacterial phytochromes.

Figure 1.2. Phytochrome structure and photochemistry. A. Structure of full-length phytochrome with PAS (yellow), GAF (green) and Phy (grey) domains. B. Photochemistry of BphPs. The two principal modes of relaxation are ESPT and Photoconversion. Photoconversion is reversible and significantly less efficient than non-radiative decay via ESPT.





Bacteriophytochromes as near-infrared biomarkers

As previously discussed protein-based fluorophores including the very well-known green fluorescent protein (GFP), are being extensively used as molecular biomarkers and probes. GFP protein molecules have been extensively studied and edited to make them highly fluorescent (quantum yields >60%), photostable and soluble in water, properties which make them ideal candidates to look into living organisms (12, 14, 15). Natural and mutant GFP molecules have substantially aided us in understanding cellular processes like gene expression, protein-protein interactions and movement of biomolecules. Most of the imaging advances in the fields of cellular and developmental biology can be attributed to GFP and GFP-like proteins. Variants of GFP span the spectral window from the blue region into the green (10, 66, 67). In addition to GFP, molecular probes have also been designed using phycobiliproteins and peridinin-chlorophyll complex. However due to complicated assembly of the holoprotein complexes compared to GFP, the use of the other two complexes has been limited to *in-vitro* usage (30, 68), (Figure 1.3).

In order to be used for non-invasive imaging in live mammalian model systems, we need to look beyond GFP to design a red-shifted and self-assembled fluorescent protein probe. BphPs can be used as a template for design of red-shifted fluorescent biomarkers. Due to use of BV as their small molecule chromophore, BphPs absorb at longer wavelengths in the near-infrared region of the electromagnetic spectrum at ~700 nm. BphPs have high molar absorption coefficients and apoproteins can spontaneously bind the bilin chromophore. The chromophore linkage is covalent and BV resides deep in the

chromophore binding pocket with little solvent exposure. Additionally, BV is a natural product of the heme catabolism pathway (69) and research has shown that apophytochrome molecules can use BV present in the mammalian system to form a functional holoprotein (5). The phytochrome variant P3 from *Rhodopseudomonas palustris* is naturally fluorescent. A mutational analysis of the DrCBD showed that an aspartate to histidine mutation in the 207th position led to increased red fluorescence. The 207th residue has important implications in the phytochrome photochemistry as was discussed in the previous section, due to its role in forming an extensive hydrogen bonding network with the nitrogen atoms of the A, B and C rings of the chromophore (56, 65).

The DrCBDD207H and its equivalent in other BphPs was further subjected to mutational analyses which resulted in the development of a plethora of near-infrared probes including IFP1.4, IFP 2.0, WiPhy and iRFP 670, 682, 702, 713 and 720 (3–6). These classes of BphP derived fluorophores consist of the PAS-GAF domains and contain amino acid substitutions around the chromophore BV, for amplified fluorescence and varied excitation and emission spectra. While IFP1.4 and WiPhy are monomeric and based on DrCBD template, the iRFP family of proteins is dimeric and designed using *Rhodopseudomonas* phytochromes P2 and P3 (3–6).

In addition to the BphP class of biomarkers, Clark Lagarias' research at University of California, Davis has shown that recombinant apophytochromes from higher plant species *Avena sativa,* cyanobacterium *Synechocystis* sp. PCC6803, and alga *Mesotaenium caldariorum* can bind phycoerythrobilin to form fluorescent proteins.

Phycoerythrobilin is the precursor of the phycoerythrin molecule which lacks the C15 bond that is required for Pr → Pfr state photoconversion. The resultant holoprotein is intensely fluorescent in the orange region of the spectrum and has been successfully utilized to monitor germination of *Arabidiopsis* seedlings (30) (Figure 1.3).

One of the principal caveats of using near-infrared spectroscopy is that the longer the excitation maximum of the probes, the lower the quantum yield. There exists a natural barrier to absorption at longer wavelengths in the electromagnetic spectrum, approximate at 1000 nm (70). Molecules that absorb in the near-infrared have relatively lower excited singlet states and the triplet states are lower in energy than their singlet states. The triplet states are metastable, contain two unpaired electrons and are biradical. It is hard to generalize thermal and photostabibility of all near-infrared fluorophores, however the occupation of both their singlet and triplet states makes near-infrared fluorophores reactive. Thus, near-infrared fluorophores react with a variety of solvent molecules in their immediate vicinity and in general longer wavelength absorption is related to low thermal and photochemical stability. Additionally, increased internal conversion with decreased energy difference between the S1 (singlet excited) and S0 (singlet ground) states also contributes to low fluorescence intensity at longer wavelengths. In other words, chances of non-radiative decay are higher in near-infrared fluorescent molecules than in molecules excited at shorter wavelengths due to the decreased energy difference between the ground and excited states (70).

The low QY problem is common for any molecule fluorescing in the near-infrared region, dyes, quantum dot and proteins. For example, the IR 800 dye which has a peak emission

of 800 nm has only ~10% quantum yield, whereas IR 700 which as the name suggests has a peak emission of 700 nm has a ~25% quantum yield, and fluorescein with an emission maximum in the visible region (521 nm) has a quantum yield of 91%. IR 700, IR 800 and fluorerescein are designed on the same small template with different aromaticities. We see this same trend reflected in fluorescent proteins as well. In 2010, the maximum quantum yield, of near-infrared fluorescent proteins was observed in IFP1.4, of approximately ~7% with an emission maximum of 700 nm. It is to be noted, that this emission maximum is 20 nm red shifted than DrCBD. The variant that was developed in our laboratory, WiPhy, maintained the longer emission maximum at 720 nm but had quantum yield of 5.6%. To put things into context, eGFP, which has been successfully used in countless *in-vitro* and *ex-vivo* experiments as a molecular probe, has a quantum yield of 61% (Figure 1.3). One of the biggest questions we were facing was, how far could we push the brightness of the near-infrared fluorescent phytochromes without compromising on the longer excitation and emission maxima?

IFP1.4 protein had the highest quantum yield amongst the fluorescent phytochromes at 7% but it faced several issues which limited its applications. The molecule was designed to be a monomeric version of DrCBD containing the D207H substitution. However, Dynamic Light Scattering and other biophysical analysis revealed it to be a poor monomer prone to aggregation in solution (7). Moreover, IFP 1.4 fluorescence is pH dependent and the brightness dropped by 30% when pH levels were below (5). IFP1.4 design was a mixture of rational and random mutagenesis. Shu and other scientists had targeted residues in the chromophore binding pocket to increase fluorescence, but 85% of the fluorescence improvement had resulted from the protein molecule being subjected to

rounds of saturated and random mutagenesis. The final design had 12 total mutations including the original D207H and thus we didn't know the underlying reason for the fluorescence of IFP1.4 and couldn't use the information to further improve its properties (3). Furthermore, the excitation and emission maxima were blue shifted compared to the wild-type template (684 nm and 705 nm) and the residues responsible for this shift remained unidentified.

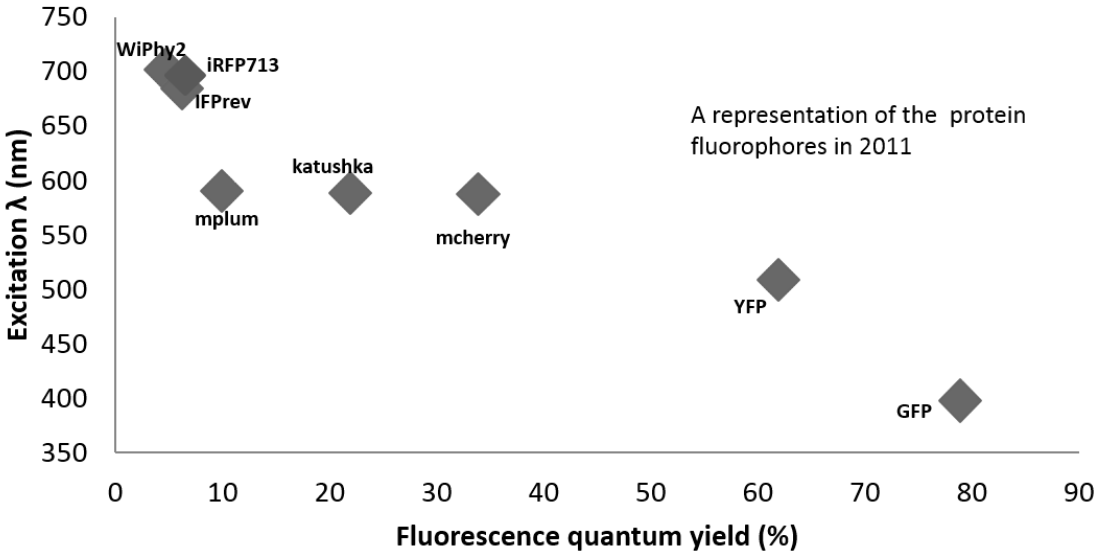
Michele Auldridge in the Forest laboratory solved the structure of WiPhy, which did aid in understanding what makes phytochrome molecules fluoresce. WiPhy has a quantum yield of 5.6% and maintained its excitation and emission maxima at 698 and 720 nm respectively. WiPhy used two amino acid substitutions D207H and Y263F to change photochemistry and three substitutions in the dimer interface to monomerize the protein. The 207th residue as previously stated is involved in non-radiative decay *via* ESPT. The D207H substitution by itself did not improve fluorescence however the Y263F variant did increase fluorescence by 60% compared to the wild-type monomeric DrCBD. Unlike its dimeric parent, DrCBD$_{mon}$ undergoes limited photoconversion, where we do see appearance of a Pfr peak at 750 nm. Substituting the non-polar phenylalanine residue for tyrosine in DrCBD$_{mon}$ (DrCBD$_{mon}$Y263F) leads to a non-photoconvertible variant. It was hypothesized that the reduced polar interactions in the chromophore binding pocket affected photoconversion (4).

The final protein in the mix is iRFP which was developed in 2011 in New York using *Rhodopseudomonas palustris* phytochrome P2 CBD, containing the same aspartate to histidine mutation as its basis. The excitation and emission maxima of this protein were

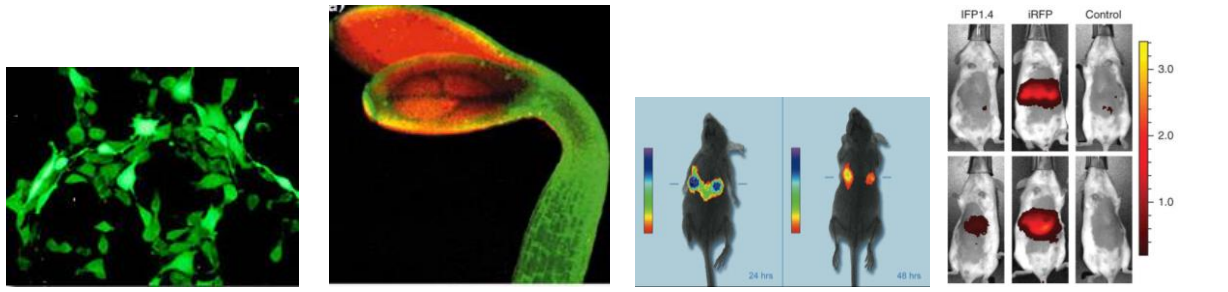
690 and 713 nm, in between IFP1.4 and WiPhy proteins. The quantum yield was 5.9%, similar to WiPhy. Unlike WiPhy or IFP 1.4, iRFP remained dimeric and contained 13 total substitutions, a combination of rational and random mutagenesis. It was reported to have higher molar extinction coefficients and photostability than IFP1.4, but in absence of analytical and structural data we didn't understand the basis of fluorescence (5).

Figure 1.3. The protein fluorophore landscape. A. A graphical representation of the protein fluorophores available for usage as molecular probes in 2011, when this work was initiated. The trend of longer excitation maxima leading to lower quantum yields is demonstrated. Rate of non-radiative decay is higher at longer excitation wavelengths which reduces the percentage of excited state population that go through fluorescence as the preferred mode of decay (3, 4, 71). B. Timeline of research in fluorescent biomarker development. Discovery of GFP was instrumental in revolutionizing non-invasive imaging techniques. The first plant based red-fluorescent proteins were discovered in 1997 and successfully used in monitoring seedling germination, shade avoidance response and growth. The field of near-infrared fluorescence spectroscopy started developing in early 2000s to exploit the ability of longer wavelengths to penetrate deep tissues and bones with minimal auto fluorescence and absorption by biomolecules. Near-infrared fluorescent phytochromes were discovered in 2008 and have been successfully used in non-invasive imaging in mice models. The figure was made from published data in the following papers (23, 30, 71, 72) .

A



B



Early 1990s — **Use of GFP in cell and tissue culture**

1997 — **PEB-plant hytochrome to monitor seed germination in oats**

2000 — **Development of near- infrared fluorescent nano-particles**

2009-12 — **Near- infrared fluorescent Bphps used in whole animal imaging**

Timeline of discovery in fluorescent biomarkers

Folding of phytochrome

The knot in the phytochrome structure plays a crucial role in the light-sensing photochemistry of the holoprotein (46). The structure of phytochrome was the second deeply knotted protein structure to be deposited at the PDB (73). Improvement in knot detecting algorithms in the PDB has led to the classification of 273 structures so far as knotted proteins (74). Is the formation of knot a spontaneous process or does the cell recruit chaperones that aid in tying the knot? Sophie Jackson and co-workers have demonstrated the spontaneous refolding of YibK, a methyl transferase with a deep trefoil knot in its structure (75, 76). Subsequent research has shown the involvement of GroEL-GroES chaperonin in efficient folding of the knot in YibK (77). In spite of the topological barrier to knot formation, the consistent ease of folding the three-dimensional protein structure into a knot inside cells suggests that folding pathways evolved to optimize for knot formation.

Todd Yeates' laboratory in UCLA proposed the involvement of a slipknot in folding of a knotted protein (78). A slipknot is an arrangement where the overall protein is unknotted but deletion of terminal regions leads the polypeptide chain to fold into itself. Yeates and colleagues also designed a knotted protein and its unknotted homolog. Their research showed that while both the knotted and unknotted versions of the polypeptide chain fold, the knotted chain folding took 20 times longer (79).

Experimental design of a protein folding study remains a significant challenge without adding the topological complexities of knot formation, which further amplify the difficulty (80). Most of the knot related research has been done on trefoil knots with three cross-

over points in the knot interface. Characterization of folding of human ubiquitin C-terminal hydrolase UCH-L3, which contains five topological crossings in its polypeptide chain shows that folding is reversible with no chaperone assistance under *in vitro* conditions (81). UCHL3 has 12 proline residues in the knot vicinity, one of prolines is a cis-proline residue. The slowest phase in refolding of UCHL3 was shown to be rate-limited by proline-isomerization events (81).

Phytochromes contain a figure of eight knot with four cross-overs. Atomic force microscopy experiments on *Deinococcus radiodurans* BphP have shown that apo and holophytochrome unfold at 47 pN and 73 pN respectively which are similar to the load borne by unknotted globular proteins of similar size and dimensions (82). While apo-phytochrome folding is reversible, once covalent attachment to BV takes place, the process becomes irreversible (82). Similar to UCHL3, 20% of the lasso loop residues in BphPs are prolines with an invariant cis-proline (Pro236 in DrBphP). In this work, we initiated the investigation the relevance of proline residues in knot formation as well as cis trans proline isomerization on protein folding. The work in chapter IV of the thesis would help explain the role of proline residues in knot folding in DrBphP.

References

1. Jobsis, F. F. (1977) Noninvasive, infrared monitoring of cerebral and myocardial oxygen sufficiency and circulatory parameters. *Science*. **198**, 1264–1267

2. Wagner, J. R., Zhang, J., von Stetten, D., Günther, M., Murgida, D. H., Mroginski, M. A., Walker, J. M., Forest, K. T., Hildebrandt, P., and Vierstra, R. D. (2008) Mutational analysis of Deinococcus radiodurans bacteriophytochrome reveals key amino acids necessary for the photochromicity and proton exchange cycle of phytochromes. *J. Biol. Chem.* **283**, 12212–12226

3. Shu, X., Royant, A., Lin, M. Z., Aguilera, T. A., Lev-Ram, V., Steinbach, P. A., and Tsien, R. Y. (2009) Mammalian expression of infrared fluorescent proteins engineered from a bacterial phytochrome. *Science*. **324**, 804–807

4. Auldridge, M. E., Satyshur, K. a, Anstrom, D. M., and Forest, K. T. (2012) Structure-guided engineering enhances a phytochrome-based infrared fluorescent protein. *J. Biol. Chem.* **287**, 7000–7009

5. Filonov, G. S., Piatkevich, K. D., Ting, L.-M., Zhang, J., Kim, K., and Verkhusha, V. V (2011) Bright and stable near-infrared fluorescent protein for in vivo imaging. *Nat. Biotechnol.* **29**, 757–761

6. Shcherbakova, D. M., and Verkhusha, V. V. (2013) Near-infrared fluorescent proteins for multicolor in vivo imaging. *Nat. Methods*. **10**, 751–754

7. Bhattacharya, S., Auldridge, M. E., Lehtivuori, H., Ihalainen, J. A., and Forest, K. T.

(2014) Origins of Fluorescence in Evolved Bacteriophytochromes. *J. Biol. Chem.* **289**, 32144–32152

8. Lehtivuori, H., Bhattacharya, S., Angenent-Mari, N. M., Satyshur, K. A., and Forest, K. T. (2015) Removal of Chromophore-Proximal Polar Atoms Decreases Water Content and Increases Fluorescence in a Near Infrared Phytofluor. *Front. Mol. Biosci.* **2**, 65–69

9. Stepanenko, O. V., Verkhusha, V. V., Kuznetsova, I. M., Uversky, V. N., and Turoverov, K. K. (2008) Fluorescent proteins as biomarkers and biosensors: throwing color lights on molecular and cellular processes. *Curr. Protein Pept. Sci.* **9**, 338–369

10. Crameri, A., Whitehorn, E. A., Tate, E., and Stemmer, W. P. (1996) Improved green fluorescent protein by molecular evolution using DNA shuffling. *Nat Biotechnol*. **14**, 315–319

11. Yoshihara, S., Shimada, T., Matsuoka, D., Zikihara, K., Kohchi, T., and Tokutomi, S. (2006) Reconstitution of Blue−Green Reversible Photoconversion of a Cyanobacterial Photoreceptor, PixJ1, in Phycocyanobilin-Producing Escherichia coli. *Biochemistry*. **45**, 3775–3784

12. Tsien, R. (1998) The green fluorescent protein. *Annu. Rev. Biochem.* **67**, 509

13. Ando, R., Hama, H., Yamamoto-Hino, M., Mizuno, H., and Miyawaki, A. (2002) An optical marker based on the UV-induced green-to-red photoconversion of a fluorescent protein. *Proc. Natl. Acad. Sci.* **99**, 12651–12656

14. Ehrig, T., O'Kane, D. J., and Prendergast, F. G. (1995) Green-fluorescent protein

mutants with altered fluorescence excitation spectra. *FEBS Lett*. **367**, 163–166

15. Reeder, P. J., Huang, Y.-M., Dordick, J. S., and Bystroff, C. (2010) A Rewired Green Fluorescent Protein: Folding and Function in a Nonsequential, Noncircular GFP Permutant. *Biochemistry*. **49**, 10773–10779

16. Hanson, G. T., McAnaney, T. B., Park, E. S., Rendell, M. E. P., Yarbrough, D. K., Chu, S., Xi, L., Boxer, S. G., Montrose, M. H., and Remington, S. J. (2002) Green Fluorescent Protein Variants as Ratiometric Dual Emission pH Sensors. 1. Structural Characterization and Preliminary Application. *Biochemistry*. **41**, 15477–15488

17. Davis, S. J., and Vierstra, R. D. (1998) Soluble, highly fluorescent variants of green fluorescent protein (GFP) for use in higher plants. *Plant Mol. Biol.* **36**, 521–528

18. Baulcombe, D. C., Chapman, S., and Cruz, S. (1995) Jellyfish green fluorescent protein as a reporter for virus infections. *Plant J.* **7**, 1045–1053

19. Peng, S.-Y., Yang, Y.-S., Chou, C.-J., Lin, K.-Y., and Wu, S.-C. (2015) Differentiation of Enhanced Green Fluorescent Protein-Labeled Mouse Amniotic Fluid-Derived Stem Cells into Cardiomyocyte-Like Beating Cells. *Acta Cardiol. Sin.* **31**, 209–214

20. Boas, D. A., Gaudette, T., Strangman, G., Cheng, X., Marota, J. J. A., and Mandeville, J. B. (2001) The Accuracy of Near Infrared Spectroscopy and Imaging during Focal Changes in Cerebral Hemodynamics. *Neuroimage*. **13**, 76–90

21. Mancini, D. M., Bolinger, L., Li, H., Kendrick, K., Chance, B., and Wilson, J. R. (1994) Validation of near-infrared spectroscopy in humans. *J. Appl. Physiol.* **77**,

2740–2747

22. Hilderbrand, S. A., and Weissleder, R. (2010) Near-infrared fluorescence: application to in vivo molecular imaging. *Curr. Opin. Chem. Biol.* **14**, 71–79

23. Luo, S., Zhang, E., Su, Y., Cheng, T., and Shi, C. (2011) A review of NIR dyes in cancer targeting and imaging. *Biomaterials*. **32**, 7127–7138

24. Ke, S., Wen, X., Gurfinkel, M., Charnsangavej, C., Wallace, S., Sevick-Muraca, E. M., and Li, C. (2003) Near-Infrared Optical Imaging of Epidermal Growth Factor Receptor in Breast Cancer Xenografts. *Cancer Res.* **63**, 7870–7875

25. Bugaj, J. E., Achilefu, S., Dorshow, R. B., and Rajagopalan, R. (2001) Novel fluorescent contrast agents for optical imaging of in vivo tumors based on a receptor-targeted dye-peptide conjugate platform. *J. Biomed. Opt.* **6**, 122–132

26. Achilefu, S., Dorshow, R. B., Bugaj, J. E., and Rajagopalan, R. (2000) Novel receptor-targeted fluorescent contrast agents for in vivo tumor imaging. *Invest. Radiol.* **35**, 479–485

27. Bongarzone, S., Staderini, M., and Bolognesi, M. L. (2014) Multitarget ligands and theranostics: sharpening the medicinal chemistry sword against prion diseases. *Future Med. Chem.* **6**, 1017–1029

28. Doerr, A. (2009) Fluorescent proteins: into the infrared. *Nat. Methods*. **6**, 482–483

29. Cashmore, A. R. (2003) Cryptochromes: Enabling plants and animals to determine circadian time. *Cell*. **114**, 537–543

30. Murphy, J. T., and Lagarias, J. C. (1997) The phytofluors: a new class of fluorescent protein probes. *Curr. Biol.* **7**, 870–876

31. Christie, J. M., Gawthorne, J., Young, G., Fraser, N. J., and Roe, A. J. (2012) LOV to BLUF: Flavoprotein Contributions to the Optogenetic Toolkit. *Mol. Plant*. **5**, 533–544

32. Pudasaini, A., El-Arab, K. K., and Zoltowski, B. D. (2015) LOV-based optogenetic devices: light-driven modules to impart photoregulated control of cellular signaling. *Front. Mol. Biosci.* **2**, 18–22

33. Möglich, A., Yang, X., Ayers, R. A., and Moffat, K. (2010) Structure and Function of Plant Photoreceptors. *Annu. Rev. Plant Biol.* **61**, 21–47

34. Rockwell, N. C., and Lagarias, J. C. (2010) A Brief History of Phytochromes. *ChemPhysChem*. **11**, 1172–1180

35. Auldridge, M. E., and Forest, K. T. (2011) Bacterial phytochromes: more than meets the light. *Crit. Rev. Biochem. Mol. Biol.* **46**, 67–88

36. Al-Sady, B., Ni, W., Kircher, S., Schaefer, E., and Quail, P. H. (2006) Photoactivated Phytochrome Induces Rapid PIF3 Phosphorylation Prior to Proteasome-Mediated Degradation. *Mol. Cell*. **23**, 439–446

37. Fankhauser, C. (2001) The Phytochromes, a Family of Red/Far-red Absorbing Photoreceptors. *J. Biol. Chem.* **276**, 11453–11456

38. Rizzini, L., Favory, J.-J., Cloix, C., Faggionato, D., O'Hara, A., Kaiserli, E., Baumeister, R., Schäfer, E., Nagy, F., Jenkins, G. I., and Ulm, R. (2011) Perception of UV-B by the Arabidopsis UVR8 protein. *Science*. **332**, 103–106

39. Quail, P. H. (2002) Phytochrome photosensory signalling networks. *Nat. Rev. Mol. Cell Biol.* **3**, 85–93

40. Kehoe, D. M., and Grossman, A. R. (1996) Similarity of a chromatic adaptation sensor to phytochrome and ethylene receptors. *Science*. **273**, 1409–1412

41. Terauchi, K., Montgomery, B. L., Grossman, A. R., Lagarias, J. C., and Kehoe, D. M. (2004) RcaE is a complementary chromatic adaptation photoreceptor required for green and red light responsiveness. *Mol. Microbiol.* **51**, 567–577

42. Davis, S. J., Vener, A. V, and Vierstra, R. D. (1999) Bacteriophytochromes: phytochrome-like photoreceptors from nonphotosynthetic eubacteria. *Science*. **286**, 2517–2520

43. Bhoo, S. H., Davis, S. J., Walker, J., Karniol, B., and Vierstra, R. D. (2001) Bacteriophytochromes are photochromic histidine kinases using a biliverdin chromophore. *Nature*. **414**, 776–779

44. Butler, W. L., Norris, K. H., Siegelman, H. W., and Hendricks, S. B. (1959) Detection, assay, and preliminary purification of the pigment controlling photoresponsive development of plants. *Proc. Natl. Acad. Sci. U. S. A.* **45**, 1703–1708

45. Blaauw-Jansen, G. (1959) The influence of red and far red light on growth and phototropism of the avena seedling. *Acta Bot. Neerl.* **8**, 1–39

46. Wagner, J. R., Brunzelle, J. S., Forest, K. T., and Vierstra, R. D. (2005) A light-sensing knot revealed by the structure of the chromophore-binding domain of phytochrome. *Nature*. **438**, 325–331

47. Wagner, J. R., Zhang, J., Brunzelle, J. S., Vierstra, R. D., and Forest, K. T. (2007) High resolution structure of Deinococcus bacteriophytochrome yields new insights

into phytochrome architecture and evolution. *J. Biol. Chem.* **282**, 12298–12309

48. Yang, X., Stojkovic, E. A., Kuk, J., and Moffat, K. (2007) Crystal structure of the chromophore binding domain of an unusual bacteriophytochrome, RpBphP3, reveals residues that modulate photoconversion. *Proc. Natl. Acad. Sci.* **104**, 12571–12576

49. Nagatani, A. (2010) Phytochrome: structural basis for its functions. *Curr. Opin. Plant Biol.* **13**, 565–570

50. Takala, H., Björling, A., Berntsson, O., Lehtivuori, H., Niebling, S., Hoernke, M., Kosheleva, I., Henning, R., Menzel, A., Ihalainen, J. A., and Westenhoff, S. (2014) Signal amplification and transduction in phytochrome photosensors. *Nature*. **509**, 245–248

51. Baker, A. W., and Forest, K. T. (2014) Structural biology: Action at a distance in a light receptor. *Nature*. **509**, 174–175

52. Burgie, E. S., Wang, T., Bussell, A. N., Walker, J. M., Li, H., and Vierstra, R. D. (2014) Crystallographic and Electron Microscopic Analyses of a Bacterial Phytochrome Reveal Local and Global Rearrangements During Photoconversion. *J. Biol. Chem.* **289**, 24573–24587

53. Yang, X., Kuk, J., and Moffat, K. (2009) Conformational differences between the Pfr and Pr states in Pseudomonas aeruginosa bacteriophytochrome. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 15639–15644

54. Bischoff, M., Hermann, G., Rentsch, S., and Strehlow, D. (2001) First Steps in the Phytochrome Phototransformation: A Comparative Femtosecond Study on the

Forward (Pr → Pfr) and Back Reaction (Pfr → Pr). *Biochemistry*. **40**, 181–186

55. Andel, F., Hasson, K. C., Gai, F., Anfinrud, P. A., and Mathies, R. A. (1997) Femtosecond time-resolved spectroscopy of the primary photochemistry of phytochrome. *Biospectroscopy*. **3**, 421–433

56. Toh, K. C., Stojkovic, E. A., van Stokkum, I. H., Moffat, K., and Kennis, J. T. (2011) Fluorescence quantum yield and photochemistry of bacteriophytochrome constructs. *Phys Chem Chem Phys*. **13**, 11985–11997

57. Dasgupta, J., Frontiera, R. R., Taylor, K. C., Lagarias, J. C., and Mathies, R. A. (2009) Ultrafast excited-state isomerization in phytochrome revealed by femtosecond stimulated Raman spectroscopy. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 1784–1789

58. Piatkevich, K. D., Subach, F. V, and Verkhusha, V. V. (2013) Engineering of bacterial phytochromes for near-infrared imaging, sensing, and light-control in mammals. *Chem. Soc. Rev.* **42**, 3441–3452

59. Yang, X., Stojković, E. A., Ozarowski, W. B., Kuk, J., Davydova, E., and Moffat, K. (2015) Light Signaling Mechanism of Two Tandem Bacteriophytochromes. *Structure*. **23**, 1179–1189

60. Rockwell, N. C., Su, Y.-S., and Lagarias, J. C. (2006) Phytochrome structure and signaling mechanisms. *Annu. Rev. Plant Biol.* **57**, 837–858

61. Kim, P. W., Freer, L. H., Rockwell, N. C., Martin, S. S., Lagarias, J. C., and Larsen, D. S. (2012) Femtosecond Photodynamics of the Red/Green Cyanobacteriochrome NpR6012g4 from *Nostoc punctiforme* . 2. Reverse Dynamics. *Biochemistry*. **51**,

619–630

62. Ulijasz, A. T., Cornilescu, G., Cornilescu, C. C., Zhang, J., Rivera, M., Markley, J. L., and Vierstra, R. D. (2010) Structural basis for the photoconversion of a phytochrome to the activated Pfr form. *Nature*. **463**, 250–254

63. Kneip, C., Hildebrandt, P., Schlamann, W., Braslavsky, S. E., Mark, F., and Schaffner, K. (1999) Protonation state and structural changes of the tetrapyrrole chromophore during the Pr/Pfr phototransformation of phytochrome: a resonance Raman spectroscopic study. *Biochemistry*. **38**, 15185–15192

64. Spillane, K. M., Dasgupta, J., and Mathies, R. A. (2012) Conformational Homogeneity and Excited-State Isomerization Dynamics of the Bilin Chromophore in Phytochrome Cph1 from Resonance Raman Intensities. *Biophys. J.* **102**, 709–717

65. Toh, K. C., Stojkovic, E. A., van Stokkum, I. H. M., Moffat, K., and Kennis, J. T. M. (2010) Proton-transfer and hydrogen-bond interactions determine fluorescence quantum yield and photochemical efficiency of bacteriophytochrome. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 9170–9175

66. Cormack, B. P., Valdivia, R. H., and Falkow, S. (1996) FACS-optimized mutants of the green fluorescent protein (GFP). *Gene*. **173**, 33–38

67. Heim, R., and Tsien, R. Y. (1996) Engineering green fluorescent protein for improved brightness, longer wavelengths and fluorescence resonance energy transfer. *Curr Biol*. **6**, 178–182

68. Afar, B., Merrill, J., and Clark, E. A. (1991) Detection of lymphocyte subsets using

three-color/single-laser flow cytometry and the fluorescent dye Peridinin chlorophyll-a protein. *J. Clin. Immunol.* **11**, 254–261

69. Nakajima, O., and Gray, C. H. (1967) Studies on haem alpha-methenyl oxygenase. Isomeric structure of formylbiliverdin, a possible precursor of biliverdin. *Biochem. J.* **104**, 20–22

70. Schäfer, F. P. (1973) Principles of dye laser operation. in *n Topics in Applied Physics Dye Lasers* (Schäfer, F. P. ed), pp. 1–89, Springer-Verlag, New York, 10.1007/3-540-51558-5_7

71. Filonov, G. S., Krumholz, A., Xia, J., Yao, J., Wang, L. V, and Verkhusha, V. V. (2012) Deep-tissue photoacoustic tomography of a genetically encoded near-infrared fluorescent probe. *Angew Chem Int Ed Engl*. **51**, 1448–1451

72. Bartz-Schmidt, K. U., Walter, P., Krott, R., Brunner, R., Esser, P., and Heimann, K. (1996) [Effects of fluorescein and indocyanine green angiography on subsequent dark adaptation and the electroretinogram]. *Klin Monbl Augenheilkd*. **208**, 224–228

73. Taylor, W. R. (2000) A deeply knotted protein structure and how it might fold. *Nature*. **406**, 916–919

74. Khatib, F., Weirauch, M. T., and Rohl, C. A. (2006) Rapid knot detection and application to protein structure prediction. *Bioinformatics*. **22**, e252–e259

75. Mallam, A. L., Morris, E. R., and Jackson, S. E. (2008) Exploring knotting mechanisms in protein folding. *Proc. Natl. Acad. Sci.* **105**, 18740–18745

76. Mallam, A. L., and Jackson, S. E. (2006) Probing Nature's Knots: The Folding Pathway of a Knotted Homodimeric Protein. *J. Mol. Biol.* **359**, 1420–1436

77. Mallam, A. L., Rogers, J. M., and Jackson, S. E. (2010) Experimental detection of knotted conformations in denatured proteins. *Proc Natl Acad Sci U S A*. **107**, 8189–8194

78. King, N. P., Yeates, E. O., and Yeates, T. O. (2007) Identification of Rare Slipknots in Proteins and Their Implications for Stability and Folding. **373**, 153–166

79. King, N. P., Jacobitz, A. W., Sawaya, M. R., Goldschmidt, L., and Yeates, T. O. (2010) Structure and folding of a designed knotted protein. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 20732–20737

80. Dill, K. A., and MacCallum, J. L. (2012) The protein-folding problem, 50 years on. **338**, 1042–1046

81. Andersson, F. I., Pina, D. G., Mallam, A. L., Blaser, G., and Jackson, S. E. (2009) Untangling the folding mechanism of the 52-knotted protein UCH-L3. **276**, 2625–2635

82. Bornschlogl, T., Anstrom, D. M., Mey, E., Dzubiella, J., Rief, M., and Forest, K. T. (2009) Tightening the knot in phytochrome by single-molecule atomic force microscopy. *Biophys J*. **96**, 1508–1514

83. von Stetten, D., Seibeck, S., Michael, N., Scheerer, P., Mroginski, M. A., Murgida, D. H., Krauss, N., Heyn, M. P., Hildebrandt, P., Borucki, B., and Lamparter, T. (2007) Highly conserved residues Asp-197 and His-250 in Agp1 phytochrome control the proton affinity of the chromophore and Pfr formation. *J. Biol. Chem.* **282**, 2116–2123

84. Anders Borg, O., and Durbeej, B. (2008) Which factors determine the acidity of the

phytochromobilin chromophore of plant phytochrome? *Phys. Chem. Chem. Phys.* **10**, 2528–2537

85. Lehtivuori, H., Bhattacharya, S., Angenent-Mari, N., Satyshur, K. A., and Forest, K. T. (2015) Removal of Chromophore-Proximal Polar Atoms Decreases Water Content and Increases Fluorescence in a Near Infrared Phytofluor. *Front. Mol. Biosci.* **2**, 65–69

86. Spillane, K. M., Dasgupta, J., Lagarias, J. C., and Mathies, R. A. (2009) Homogeneity of phytochrome Cph1 vibronic absorption revealed by resonance Raman intensity analysis. *J Am Chem Soc*. **131**, 13946–13948

87. Essen, L.-O., Mailliet, J., and Hughes, J. (2008) The structure of a complete phytochrome sensory module in the Pr ground state. *Proc. Natl. Acad. Sci.* **105**, 14709–14714

88. Yang, X., Kuk, J., and Moffat, K. (2008) Crystal structure of Pseudomonas aeruginosa bacteriophytochrome: photoconversion and signal transduction. *Proc. Natl. Acad. Sci.* **105**, 14715–14720

89. von Stetten, D., Gunther, M., Scheerer, P., Murgida, D. H., Mroginski, M. A., Krauss, N., Lamparter, T., Zhang, J., Anstrom, D. M., Vierstra, R. D., Forest, K. T., and Hildebrandt, P. (2008) Chromophore heterogeneity and photoconversion in phytochrome crystals and solution studied by resonance Raman spectroscopy. *Angew Chem Int Ed Engl*. **47**, 4753–4755

90. Dasgupta, J., Frontiera, R. R., Taylor, K. C., Lagarias, J. C., and Mathies, R. A. (2009) Ultrafast excited-state isomerization in phytochrome revealed by

femtosecond stimulated Raman spectroscopy. *Proc. Natl. Acad. Sci.* **106**, 1784–1789

91. Foerstendorf, H., Mummert, E., Sch??fer, E., Scheer, H., and Siebert, F. (1996) Fourier-transform infrared spectroscopy of phytochrome: Difference spectra of the intermediates of the photoreactions. *Biochemistry*. **35**, 10793–10799

92. Toh, K. C., van Stokkum, I. H., Hendriks, J., Alexandre, M. T., Arents, J. C., Perez, M. A., van Grondelle, R., Hellingwerf, K. J., and Kennis, J. T. (2008) On the signaling mechanism and the absence of photoreversibility in the AppA BLUF domain. *Biophys J*. **95**, 312–321

93. Fischer, A. J., and Lagarias, J. C. (2004) Harnessing phytochrome's glowing potential. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 17334–17339

94. Lehtivuori, H., Rissanen, I., Takala, H., Bamford, J., Tkachenko, N. V., and Ihalainen, J. A. (2013) Fluorescence Properties of the Chromophore-Binding Domain of Bacteriophytochrome from *Deinococcus radiodurans*. *J. Phys. Chem. B*. **117**, 11049–11057

95. Otwinowski, Z., and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326

96. McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007) Phaser crystallographic software. *J Appl Crystallogr*. **40**, 658–674

97. Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F., and Vagin, A. A. (2011) REFMAC5 for the

refinement of macromolecular crystal structures. *Acta Crystallogr. D. Biol. Crystallogr.* **67**, 355–367

98. Hahn, J., Strauss, H. M., Landgraf, F. T., Gimenèz, H. F., Lochnit, G., Schmieder, P., and Hughes, J. (2006) Probing protein-chromophore interactions in Cph1 phytochrome by mutagenesis. *FEBS J.* **273**, 1415–1429

99. Jamroz, M., Niemyska, W., Rawdon, E. J., Stasiak, A., Millett, K. C., Sułkowski, P., and Sulkowska, J. I. (2015) KnotProt: a database of proteins with knots and slipknots. *Nucleic Acids Res.* **43**, D306–D314

100. Lai, Y.-L., Yen, S.-C., Yu, S.-H., and Hwang, J.-K. (2007) pKNOT: the protein KNOT web server. *Nucleic Acids Res.* **35**, W420–W424

101. Millett, K. C., Rawdon, E. J., Stasiak, A., and Sułkowska, J. I. (2013) Identifying knots in proteins. *Biochem. Soc. Trans.* **41**, 533–537

102. Burgie, E. S., and Vierstra, R. D. (2014) Phytochromes: An Atomic Perspective on Photoactivation and Signaling. *Plant Cell Online*. **26**, 4568–4583

103. Rockwell, N. C., Martin, S. S., Gulevich, A. G., and Lagarias, J. C. (2012) Phycoviolobilin Formation and Spectral Tuning in the DXCF Cyanobacteriochrome Subfamily. *Biochemistry*. **51**, 1449–1463

104. Sineshchekov, V., Mailliet, J., Psakis, G., Feilke, K., Kopycki, J., Zeidler, M., Essen, L.-O., and Hughes, J. (2014) Tyrosine 263 in cyanobacterial phytochrome Cph1 optimizes photochemistry at the prelumi-R→lumi-R step. *Photochem. Photobiol.* **90**, 786–795

105. Narikawa, R., Fukushima, Y., Ishizuka, T., Itoh, S., and Ikeuchi, M. (2008) A Novel

Photoactive GAF Domain of Cyanobacteriochrome AnPixJ That Shows Reversible Green/Red Photoconversion. *J. Mol. Biol.* **380**, 844–855

106. Narikawa, R., Ishizuka, T., Muraki, N., Shiba, T., Kurisu, G., and Ikeuchi, M. (2013) Structures of cyanobacteriochromes from phototaxis regulators AnPixJ and TePixJ reveal general and specific photoconversion mechanism. *Proc. Natl. Acad. Sci.* **110**, 918–923

107. Narikawa, R., Nakajima, T., Aono, Y., Fushimi, K., Enomoto, G., Ni-Ni-Win, Itoh, S., Sato, M., Ikeuchi, M., Narikawa, R., Enomoto, G., Ni, W. N., Fushimi, K., and Ikeuchi, M. (2015) A biliverdin-binding cyanobacteriochrome from the chlorophyll d–bearing cyanobacterium Acaryochloris marina. *Sci. Rep.* **5**, 7950–7975

108. Zhang, J., Wu, X.-J., Wang, Z.-B., Chen, Y., Wang, X., Zhou, M., Scheer, H., and Zhao, K.-H. (2010) Fused-Gene Approach to Photoswitchable and Fluorescent Biliproteins. *Angew. Chemie Int. Ed.* **49**, 5456–5458

109. Rockwell, N. C., Martin, S. S., and Lagarias, J. C. (2012) Red/Green Cyanobacteriochromes: Sensors of Color and Power. *Biochemistry*. **51**, 9667–9677

110. Pal, D., and Chakrabarti, P. (1999) Cis peptide bonds in proteins: residues involved, their conformations, interactions and locations. *J. Mol. Biol.* **294**, 271–288

111. Joseph, A. P., Srinivasan, N., and de Brevern, A. G. (2012) Cis-trans peptide variations in structurally similar proteins. *Amino Acids*. **43**, 1369–1381

112. MacArthur, M. W., and Thornton, J. M. (1991) Influence of proline residues on protein conformation. *J. Mol. Biol.* **218**, 397–412

113. Fischer, G., and Schmid, F. X. (1990) The mechanism of protein folding.

Implications of in vitro refolding models for de novo protein folding and translocation in the cell. *Biochemistry*. **29**, 2205–2212

114. Hesterkamp, T., and Bukau, B. (1996) Identification of the prolyl isomerase domain of Escherichia coli trigger factor. *FEBS Lett.* **385**, 67–71

115. Kramer, G., Patzelt, H., Rauch, T., Kurz, T. A., Vorderwülbecke, S., Bukau, B., and Deuerling, E. (2004) Trigger Factor Peptidyl-prolyl cis/trans Isomerase Activity Is Not Essential for the Folding of Cytosolic Proteins in Escherichia coli. *J. Biol. Chem.* **279**, 14165–14170

116. Stoller, G., Rücknagel, K. P., Nierhaus, K. H., Schmid, F. X., Fischer, G., and Rahfeld, J. U. (1995) A ribosome-associated peptidyl-prolyl cis/trans isomerase identified as the trigger factor. *EMBO J.* **14**, 4939–4948

117. Kramer, G., Rauch, T., Rist, W., Vorderwülbecke, S., Patzelt, H., Schulze-Specking, A., Ban, N., Deuerling, E., and Bukau, B. (2002) L23 protein functions as a chaperone docking site on the ribosome. *Nature*. **419**, 171–174

118. Gupta, R., Lakshmipathy, S. K., Chang, H.-C., Etchells, S. A., and Hartl, F. U. (2010) Trigger factor lacking the PPIase domain can enhance the folding of eukaryotic multi-domain proteins in Escherichia coli. *FEBS Lett.* **584**, 3620–3624

119. Ferbitz, L., Maier, T., Patzelt, H., Bukau, B., Deuerling, E., and Ban, N. (2004) Trigger factor in complex with the ribosome forms a molecular cradle for nascent proteins. *Nature*. **431**, 590–596

120. Thomason, L. C., Costantino, N., Court, D. L., Thomason, L. C., Costantino, N., and Court, D. L. (2007) *E. coli* Genome Manipulation by P1 Transduction. in *Current*

Protocols in Molecular Biology, pp. 1.17.1–1.17.8, John Wiley & Sons, Inc., Hoboken, NJ, USA, 10.1002/0471142727.mb0117s79

121. Burgie, E. S., Walker, J. M., Phillips, G. N., and Vierstra, R. D. (2013) A Photo-Labile Thioether Linkage to Phycoviolobilin Provides the Foundation for the Blue/Green Photocycles in DXCF-Cyanobacteriochromes. *Structure*. **21**, 88–97

122. Frey, U. H., Bachmann, H. S., Peters, J., and Siffert, W. (2008) PCR-amplification of GC-rich regions: “slowdown PCR.” *Nat. Protoc.* **3**, 1312–1317

123. Gibson, D. G., Young, L., Chuang, R.-Y., Venter, J. C., Hutchison, C. A., and Smith, H. O. (2009) Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat Meth*. **6**, 343–345

124. Schneider, C. A., Rasband, W. S., and Eliceiri, K. W. (2012) NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods*. **9**, 671–675

125. Rumyantsev, K. A., Shcherbakova, D. M., Zakharova, N. I., Emelyanov, A. V., Turoverov, K. K., and Verkhusha, V. V. (2015) Minimal domain of bacterial phytochrome required for chromophore binding and fluorescence. *Sci. Rep.* **5**, 18348–18352

126. Mallam, A. L. (2009) How does a knotted protein fold? *FEBS J.* **276**, 365–375

127. Sułkowska, J. I., Noel, J. K., Ramírez-Sarmiento, Esar A, Rawdon, E. J., Millett, K. C., and Onuchic, J. N. (2013) Knotting pathways in proteins. *Biochem. Soc. Trans.* **41**, 523–527

128. Mallam, A. L., and Jackson, S. E. (2011) Knot formation in newly translated proteins is spontaneous and accelerated by chaperonins. *Nat. Chem. Biol.* **8**, 147–153

129. Yang, X., Stojković, E. A., Ozarowski, W. B., Kuk, J., Davydova, E., and Moffat, K. (2015) Light Signaling Mechanism of Two Tandem Bacteriophytochromes. *Structure*. **23**, 1179–1189

130. Yang, X., Ren, Z., Kuk, J., and Moffat, K. (2011) Temperature-scan cryocrystallography reveals reaction intermediates in bacteriophytochrome. *Nature*. **479**, 428–432

131. Yang, X., Stojkovic, E. A., Kuk, J., and Moffat, K. (2007) Crystal structure of the chromophore binding domain of an unusual bacteriophytochrome, RpBphP3, reveals residues that modulate photoconversion. *Proc Natl Acad Sci U S A*. **104**, 12571–12576

132. Burgie, E. S., Bussell, A. N., Walker, J. M., Dubiel, K., and Vierstra, R. D. (2014) Crystal structure of the photosensing module from a red/far-red light-absorbing plant phytochrome. *Proc. Natl. Acad. Sci.* **111**, 10179–10184

133. Yu, D., Gustafson, W. C., Han, C., Lafaye, C., Noirclerc-Savoye, M., Ge, W.-P., Thayer, D. A., Huang, H., Kornberg, T. B., Royant, A., Jan, L. Y., Jan, Y. N., Weiss, W. A., and Shu, X. (2014) An improved monomeric infrared fluorescent protein for neuronal and tumour brain imaging. *Nat. Commun.* **5**, 3626–3632

134. Giraud, E., Zappa, S., Vuillet, L., Adriano, J.-M., Hannibal, L., Fardoux, J., Berthomieu, C., Bouyer, P., Pignol, D., and Verméglio, A. (2005) A new type of bacteriophytochrome acts in tandem with a classical bacteriophytochrome to control the antennae synthesis in Rhodopseudomonas palustris. *J. Biol. Chem.* **280**, 32389–32397

CHAPTER II

Origin of fluorescence in IFP1.4

This chapter has been previously published in the Journal of Biological Chemistry

Bhattacharya, S., Auldridge, M. E., Lehtivuori, H., Ihalainen, J. a, and Forest, K. T. (2014) Origins of Fluorescence in Evolved Bacteriophytochromes. J. Biol. Chem. 289, 32144–32152

Abstract

Use of fluorescent proteins to study *in vivo* processes in mammals requires near-infrared (NIR) biomarkers that exploit the ability of light in this range to penetrate tissue. Bacteriophytochromes (BphPs) are photoreceptors whose biological function is to couple absorbance of NIR light by a biliverdin chromophore housed within the protein to photoisomerization, protein conformational changes and signal transduction. BphPs have been engineered to form monomeric NIR fluorophores, including IFP1.4, Wi-Phy, and iRFP, initially by replacement of Asp207 by His. This position was suggestive because its main chain carbonyl is within hydrogen bonding distance of pyrrole ring nitrogens in biliverdin, thus potentially functioning as a crucial proton sink during photoconversion. In order to explain the origin of fluorescence in these phytofluors, we solved the crystal structures of IFP1.4 and a comparison non-fluorescent monomeric phytochrome. Met186 and Val288 in IFP1.4 are responsible for the formation of a more tightly-packed hydrophobic hub around the biliverdin D ring. Met186 is also responsible for the blueshift of IFP1.4's absorbance maximum relative to the parent BphP. The structure of IFP1.4 revealed decreased structural heterogeneity and a contraction of two surface regions as direct consequences of side chain substitutions. Unexpectedly, IFP1.4 with Asp207 reinstalled (IFP_{rev}) has a higher fluorescence quantum yield (9%) than any NIR phytofluor published to date. In agreement, fluorescence lifetime measurements confirm the exceptionally long excited state lifetimes, up to 815 ps, in IFP1.4 and IFP_{rev}. Our research has helped delineate the origin of fluorescence in engineered BphPs, and will facilitate the wide-spread adoption of phytofluors as biomarkers.

Introduction

INTRODUCTION

Visualization of molecular processes has revolutionized the way we understand life at the cellular level. Biochemical pathways were difficult to monitor in live model systems before the advent of fluorescent biomarkers (1-4). Studying *in vivo* characteristics involved the use of toxic and radioactive materials and often interfered with the system being studied. Based on Dr. Frans Jöbsis' research on tissue oxygenation (5-7), near-infrared (NIR) spectroscopy has been developed to be a highly feasible monitoring device for cellular mechanisms (8). Compared to visible light, NIR light is less scattered by skin, bone or other organs, and is hardly absorbed by biomolecules or water. Development of fluorescent biomarkers with NIR capabilities is therefore crucial to the development of *in vivo* imaging techniques (6,7). Bacteriophytochromes (BPhPs) are ideally suited for this purpose for several reasons including NIR absorbance wavelength maxima, availability of the biliverdin IXα (BV) chromophore in mammalian cells, and the existence of two relatively stable photoreversible ground states (8-12). The absorption spectrum of the phytochrome class of sensory proteins is tuned by the interactions between the protein and linear tetrapyrrole chromophore, BV, which is an intermediate in normal mammalian heme catabolism. Phytochromes are covalently linked to BV through a thioether bond to a conserved cysteine side chain; phytochromes are autocatalytic BV lyases. Available fluorescent BphPs to date, including the very recently described IFP2.0, are most successful for imaging when BV is supplemented or generated by an accessory heme oxygenase (9,10,13). Nonetheless, the bioavailability of BV and the lack of accessory enzymes needed to load the chromophore make BphPs an attractive design template for

a mammalian NIR fluorescent biomarker. Moreover, the photoreversible nature of phytochromes adds another dimension of utility to these proteins as photoactivatable fluorophores (14,15).

Phytochromes have a modular structure with the PAS, GAF and PHY domains making up an input photosensing module (12). In many BphPs, this input module is followed C-terminally by a histidine kinase output module. Despite the thioether cysteine location N-terminal to the PAS domain, the GAF domain houses the BV chromophore (16). These two domains thus comprise the chromophore-binding domain (CBD). Absorption of a red light photon (λ=700 nm) by the chromophore in the Pr ground state promotes an isomerization of the BV C15=C16 double bond. During photoconversion, the chromophore becomes transiently deprotonated. The BV D-ring movements lead to changes in the orientation and protonation state of several amino acid sidechains in vicinity of the BV, which lead to local protein secondary structural changes in the PHY domain (17). Ultimately, the large-scale movements of the output domain form the far-red light absorbing Pfr state (17-22). Through extensive structure-guided amino acid substitutions in BphPs, it was discovered that truncation of the polypeptide chain as well as variations in the amino acid residues surrounding the chromophore significantly reduced the photoconvertability of these proteins (8,23,24). Absorption of red light excites the molecule, but due either to an inability to isomerize the chromophore, or a disruption in the normal proton shuttle, photoconversion is blocked. In some cases, fluorescence quantum yield increases (8-10). The utility of phytochrome-based fluorophores was explored as early as 2004 by Lagarias and coworkers, who introduced the apt term phytofluor (25). The archetypal BphP phytofluor carried a structure-based substitution of

histidine for the invariant aspartic acid whose main chain carbonyl interacts with the BV A ring (23). Four rounds of directed or randomized *in vitro* evolution of this DrCBD-D207H led to IFP1.4 (10), which carries 12 additional substitutions (M54V, G119A, V186M, L195M, H196Q, I208T, A288V, Y307E, L311K, L314G, V318R, and T135I).

Wild-type DrBphP is a dimer stabilized by a six-helix bundle interface contributed by GAF domain helices (26). In order to increase BphP's utility as a fluorophore, residues in this GAF dimer interface have been rationally mutated in efforts to create a monomer (8,10). Interestingly, in the case of IFP1.4 the latter rounds of random mutagenesis with selection for brighter variants led to additional changes in the dimer interface. Size exclusion chromatography demonstrated that the final IFP1.4 forms a mixed population of monomer as well as higher order oligomers (9). Our dynamic light scattering of IFP1.4 also demonstrated protein aggregation (data not shown), and recent additional directed evolution of IFP1.4 focused on stabilization of the monomeric form (13). DrCBD itself was also independently monomerized to form $DrCBD_{mon}$ by designed interruption of three favorable hydrophobic interactions in the dimer interface. Analytical ultracentrifugation confirmed the monomeric nature of $DrCBD_{mon}$ (8). Addition of D207H and a Y263F substitution both increased fluorescence of $DrCBD_{mon}$, leading to the Wi-Phy phytofluor (8).

IFP1.4 has an absorbance maximum at 684 nm, and fluorescence excitation and emission maxima at 684 nm and 708 nm, respectively, which are blue-shifted from their respective wavelengths of the wild-type DrCBD. It was demonstrated to function as a near-infrared biomarker in live mice (10). However, presently available near-infrared biomarkers based on the Asp207His variation, still have a number of properties which

need to be improved, including but not limited to a propensity to photobleach and relatively low quantum yields (14).

In this paper we report the three dimensional structures of IFP1.4 and $DrCBD_{mon}$ (8). These structures further the collective understanding of the origin of fluorescence in BphPs. Additionally, we engineered an IFP1.4-D207 revertant, IFP_{rev}, and present a detailed analysis of its spectroscopic properties. *In vitro*, IFP_{rev} is the brightest BphP-derived phytofluor to date. By measuring the fluorescence lifetime profiles of these variants we are able to distinguish site-selectively the contributions of individual amino acids to the excited state properties, and thus, to the fluorescence quenching mechanism of BV within the DrCBD ligand pocket.

Results

Crystal Structure

Overall In order to understand the characteristic fluorescent behavior of IFP1.4, a high-resolution crystal structure of the protein was solved and refined against 1.6 Å resolution data. The structure was compared to the 1.7 Å resolution structure of $DrCBD_{mon}$, also obtained in this study (Table I) and to other variants of DrCBD (8,16,26).

The overall architecture of IFP1.4 and $DrCBD_{mon}$, with canonical PAS and GAF domains and a distinctive figure-of-eight knot, is identical to previously published DrCBD variant structures (Figure 1A). In both, N-terminal residues, C-terminal histidines and the PAS-GAF linker are disordered, and in IFP1.4 there is an additional break in the main chain density at a connecting loop (107-108). Electron density is unambiguous for eleven of twelve expected amino acid substitutions (T135I falls within the disordered linker). Alignment of the three-dimensional structures showed that there is an overall RMSD of 0.38 Å between IFP1.4 and $DrCBD_{mon}$ for all Cα atoms. The integrity of the chromophore-binding pocket is maintained in IFP1.4 and the D-ring tilt relative to B and C rings is within the range seen in other structures (Suppl. Table 1). Polar interactions with the B and C rings are conserved; the B-ring propionate forms a salt bridge with Arg254 and hydrogen-bonds with Tyr216 and Ser257, while the C-ring propionate hydrogen-bonds with His260, Ser272 and Ser274 (Figure 1A, 1B). The plane of the His260 imidazole also forms a perfect Van der Waals packing platform for the planar B and C rings and the methine bridge that connects them.

Table 2.1. Data Collection and Refinement Statistics

	IFP1.4 (408G)	$DrCBD_{mon}$ (4IJG)
Data Collection		
Cell Dimensions:		
a, b, c (Å)	96.2, 53.3, 66.8	94.6,54.8,69.8
α β γ (°)	90.0, 90.6, 90.0	90.0, 91.8, 90.0
Resolution (Å)*	30.0 - 1.62 (1.65-1.62)	25.0 -1.70 (1.75-1.70)
Mosaicity (°)	0.34	nr
R_{sym}	0.068 (0.22)	0.069 (0.21)
I/σ	29.7 (8.8)	15.5
Completeness (%)	93.7 (99.9)	95.2 (77.0)
No. reflections	38,000 (2711)	35,551 (2232)
Redundancy	7.1 (6.6)	3.5 (2.7)
Wilson B ($Å^2$)	16.5	20.5
Refinement		
Resolution (Å)	24.0-1.65 (1.70-1.65)	24.0-1.70 (1.74-1.70)
R_{work} / R_{free} (%)	18.5 (19.3)/ 21.1 (21.9)	20.6 (21.3) / 22.7 (23)
No. atoms		
Protein	2374	2370
Ligand	86	67
Water	227	211
B-factor ($Å^2$)		

Protein	20.9	25.2
Ligand	10.8	7.2
Water	26.1	29.9
Rmsd		
Bond lengths (Å)	0.010	0.007
Bond angles (°)	1.6	1.5
Ramachandran (%favored/allowed)	99.0/1.0	99.7/0.3

*Values in parentheses represent the highest resolution shell.

D-ring packing and Wavelength Shift The most striking change in the structure is the emergent tightly-packed hydrophobic hub in the chromophore-binding pocket of IFP1.4, formed by the novel Val288 and Met186 side chains, the original Met174 residue, the D-ring, and the methyl group of the C-ring. This hub locks the D-ring in place and likely prevents relaxation through photoconversion (Figure 2A). Other amino acids undergo subtle shifts (tenths of an Å) between the two structures and contribute to the packing around the D-ring. These include Tyr263 and Met267. Additional amino acids, whose positions are not significantly altered between the structures, further stabilize the D-ring position. These include Tyr176, Phe203 and Phe198 (Figure 2A).

To determine the relative effects of amino acid changes within the D-ring hub on IFP1.4 fluorescence, we installed the A288V substitution in both DrCBD$_{mon}$ and DrCBD-Y263F and found increased fluorescence in both cases. V186M, on the other hand, had the striking effect of blueshifting the 700 nm absorption and excitation maxima of DrCBD$_{mon}$ to 692 nm (Figure 2.2B, Table 2.2) without any positive impact on fluorescence quantum yield (Table 2.2).

Figure 2.1. Structure of IFP1.4. A) Architecture of IFP1.4 is almost identical to previously published DrCBD structures. A protein knot keeps the PAS domain and N-terminal residues (lime) packed against the GAF domain (mint). BV (cyan), covalently attached at Cys24, is nestled in the GAF domain. B. Polar contacts to BV (dotted lines) are unchanged while new and formerly observed hydrophobic contacts with the chromophore (starbursts; green for IFP1.4, blue for $DrCBD_{mon}$) stabilize the D ring. In particular, Met186 and Val288 are absent in the wild-type counterpart $DrCBD_{mon}$.

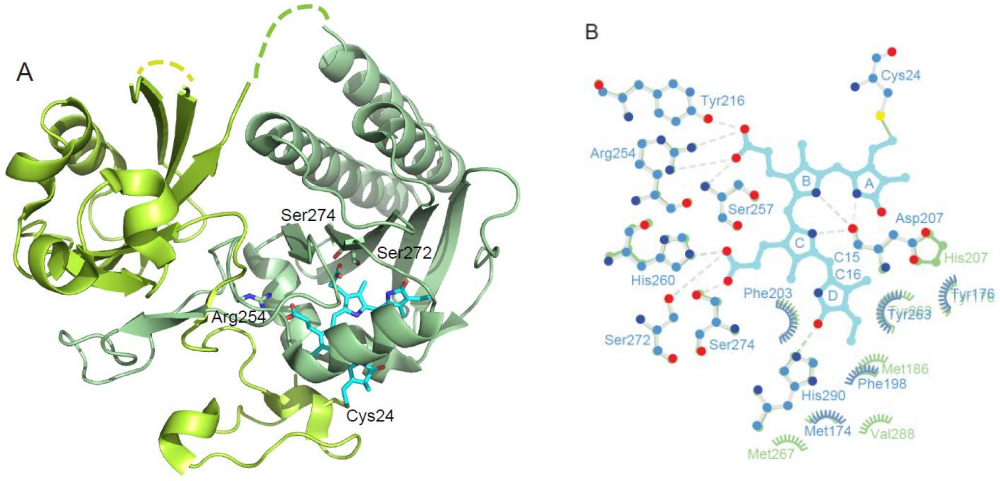
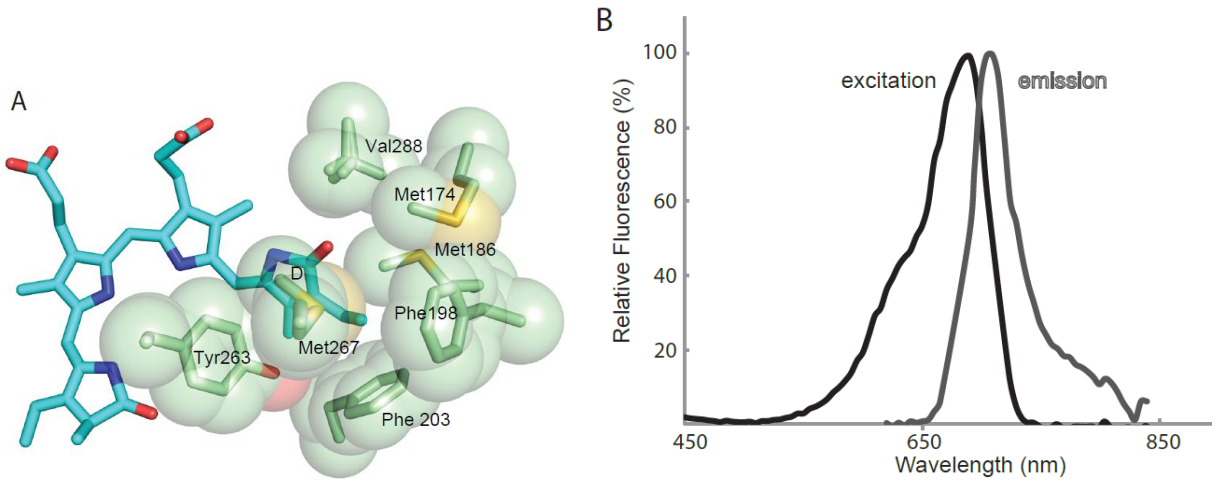
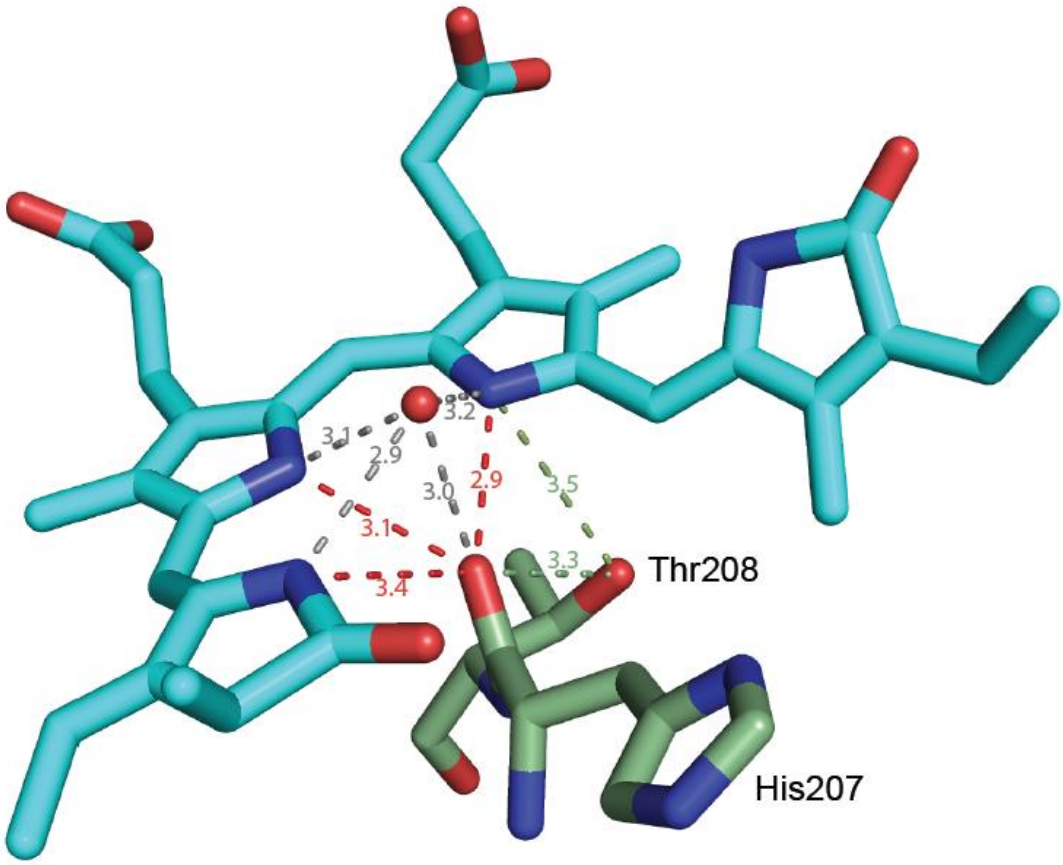


Figure 2.2. The evolved hydrophobic hub leads to excitation wavelength shift. A) Hydrophobic interactions of conserved and novel (Met186, Val188) restrict the D ring. B) A single amino acid change of Val186 to Met is sufficient to blue-shift the excitation wavelength maximum to 692 from 701. Scans shown are from 5-nm step size excitation (monitored at 713 nm) and emission (excited at 645 nm) scans.



Polar Chromophore Water Interactions The Asp-Ile-Pro (DIP) motif is a highly conserved structural motif found in canonical phytochromes. The main chain carbonyl of Asp207 has been implicated as a proton sink in both the proton release and uptake required for normal photoconversion as well as during non-radiative decay via excited state proton transfer, due to its proximity to all three protonated ring nitrogens and the highly ordered pyrrole water (Fig. 2.3) (23,24,27,28). The water-mediated hydrogen-bonding network also extends between the pyrrole ring nitrogens of the chromophore and the N**δ**1 of His260 (Fig. 2.3A). The partial negative charge of the carbonyl oxygen of Asp207 and the N**δ**1 of His260 would aid in stabilizing the protonated chromophore (28,29). In IFP1.4, amino acids 207 and 208 have been altered to His and Thr, respectively. The impact of His207 has been previously discussed and could be on positioning of Tyr263 and/or interactions with the D-ring hydroxyl in the Pfr form (8). Additionally, in IFP1.4 the novel hydroxyl group of Thr208 makes a hydrogen-bond to the main chain carbonyl of residue 207 (Figure 3B). This polar contact potentially affects the excited state proton transfer pathway. The direct interaction between Thr208 and the chromophore is weak, consisting only of a long (3.5 Å) hydrogen-bond between the Thr -OH and the C-ring nitrogen. In DrCBDmon a second water molecule is found in the binding pocket (Fig. 2.3B), in space made available by the repositioning of Tyr263 to the most distal position it has been observed.

Figure 2.3. Network of polar contacts involving BV ring nitrogens. The carbonyl backbone of His207 forms direct polar contacts (red) with the A, B and C ring nitrogens of BV, all of which are protonated. The pyrrole-water is a central part of this network, interacting with each of the ring nitrogens and the carbonyl. The novel H-bond donor Thr208 may reduce the likelihood of proton transfer from the chromophore to the 207 carboxylate.

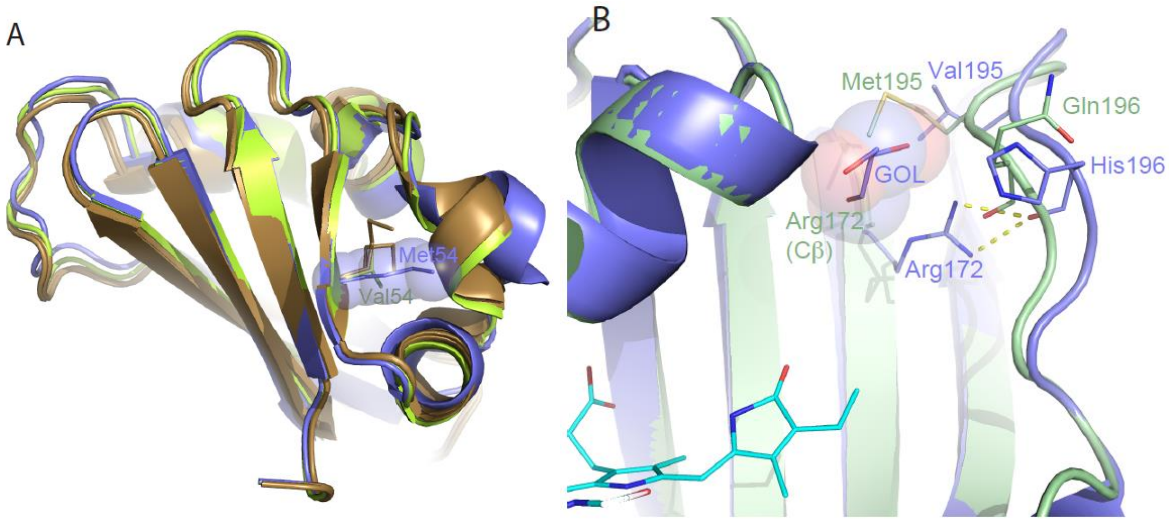


Increased rigidity and compactness The M54V substitution was isolated in the context of the dimeric phytochrome in the second stage of the *in vitro* evolution of IFP1.4 and provided a 32% increase in fluorescence quantum yield (10). By aligning seven DrCBD structures we are able to conclude that the parental Met at position 54 can adopt one of three rotamers, placing the Cγ of the side chain in one of two positions with the consequence that the short PAS domain helix α2 can also occupy one of two positions (Figure 4A). The Val substitution in IFP1.4 mimics only one of these two, and thus promotes a more compact and less heterogeneous structure (Figure 4A).

Limiting conformational alternatives of the α2 helix only indirectly improves fluorescence quantum yield in IFP1.4. We found upon installing the M54V substitution in $DrCBD_{mon}$ there was no increase in brightness (Table 2.2). Thus increased rigidity of the protein decreases the chance of thermal dissipation of energy and leads to increased fluorescence only in a phytochrome which has a higher than background probability of nonradiative decay.

We discovered two alternative main chain configurations for the loop region between residues 192 and 201 containing the Met195 and the Gln196 variants. In the more clearly defined main chain alternative, this region is approximately 1.5 Å closer to the chromophore in comparison to $DrCBD_{mon}$ (Figure 2.4B). Met195 and Gln196 are thus a second example of a change that increases phytofluor fluorescence by stabilizing a single rigid protein conformation.

Figure 2.4. Contraction of surface residues in IFP1.4. A. Packing of the Met54 pushes PAS domain β2 outwards in some DrCBD variants (WiPhy 3s7q shown in blue), but can adopt variable positions in the family of DrCBD structures (dimeric DrCBD, 2O9C, and high resolution DrCBD D207H, 3S7O, shown in sand). In fluorescent IFP1.4 (lime), the Val54 side chain cannot accomplish this displacement, so β2 adopts the more compact position closer to β2. B. Substitution V195M in IFP1.4 (mint) pulls residues 195-197 of the GAF domain inward, closing a surface cleft seen in $DrCBD_{mon}$ (blue) and displacing glycerol (GOL) as well as disrupting the interaction between 196 main chain and Arg172 side chain.



Time-resolved Fluorescence

Excited state lifetime measurements provide information about fluorescence properties of phytofluors independent on the concentration of the sample. The fluorescence decay properties of BV molecules in the binding pocket were studied by a photon counting method with excitation wavelength of 660 nm and monitoring wavelength of 714 nm (Figure 5). The fluorescence decays were fitted with either mono- or biexponential functions (30) to obtain the excited state lifetimes of particular DrCBD constructs. We found two categories; in the first, the decay profile of the non-fluorescent monomer required two exponential components, whereas in the case of fluorescent variants Wi-Phy, DrCBD-Y263F, IFP1.4 and IFPrev, monoexponential fits were sufficient to describe the fluorescence decay (Table 2.2). Thus, more decay components are present in the first category than in the second. As predicted, the excited state lifetime measurements yield longer lifetimes, up to 815 ps, for the variants with the highest independently measured fluorescence quantum yields (Table 2.2). However, we did not find a linear dependence between the excited state lifetimes and fluorescence quantum yields across the entire family of phytochrome variant proteins measured in this study. The lack of such a correlation indicates a variation in the radiative lifetimes which apparently depends on the chromophore environment created by individual amino acid substitutions.

Figure 2.5. Fluorescence decays of dark-adapted phytochromes. Emission decays of purified IFP1.4, DrCBD$_{mon}$ and DrCBD$_{mon}$ variants (in 30 mM Tris, pH 8.0) excited at 660 nm and monitored at 714 nm. IRF is the instrument response function. Solid lines show the multiexponential fit of the data.

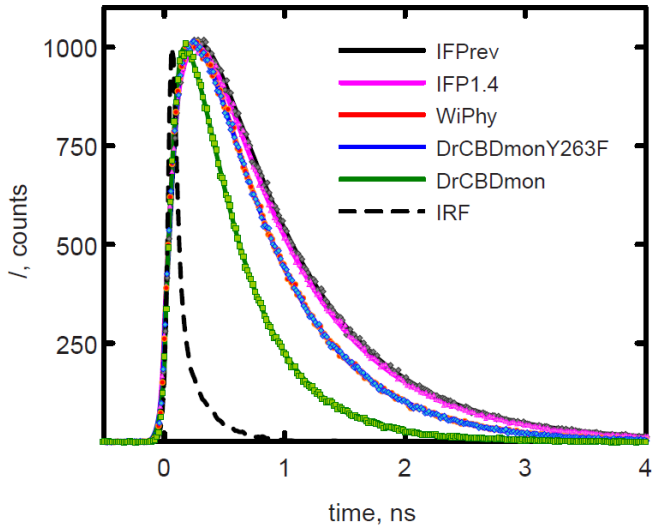


Table 2.2: Quantum yield measurements and fluorescence lifetimes

Protein Variant	Abs. max, (nm)	Em. max, (nm)	Φ[a] (%)	Lifetime[b] (ps)	
				τ_1	τ_2
$DrCBD_{mon}$	698	718	2.9±0.1	390±30 (77%)	620±70 (33%)
$DrCBD_{mon}$Y263F	701	722	4.0±0.1	-	670±10
WiPhy	701	722	6.3±0.2	-	670±10
IFP1.4	684	708	7.0±0.3	-	800±10
$IFP1.4_{rev}$	684	708	8.7±0.1	-	815±11

[a] Fluorescence quantum yield measurements as determined by the absolute quantum yield method.

Discussion

Taking advantage of time-resolved and steady state fluorescence spectroscopy as well as the time-averaged technique of protein crystallography we present evidence for three pathways by which fluorescence quantum yield is increased in BphP-based phytofluors. In the first, the chromophore is rigidified by van der Waals packing interactions with neighboring side chains, particularly in the D-ring. The excited state chromophore is then prevented from C15=C16 double bond rotation and fluorescence quantum yield is increased (Figure 2). This interpretation agrees perfectly with the recent computational and spectroscopic results on fluorescence in the RpBphP2-derived phytofluors (31). In the second mechanism, changes to the local network of polar interactions between solvent, protein and BV lessen the likelihood for excited state proton transfer thus increasing fluorescence quantum yield (Figure 3). The third mechanism is less direct but also minimizes non-radiative decay. It involves the structural repositioning of amino acids not in direct contact with the chromophore. Such repositioning can have the effect of lessening structural heterogeneity and/or increasing rigidity of the protein thus decreasing thermal relaxation (Figure 4a). It also has the consequence of reducing solvent and ion access to the chromophore-binding pocket. For example, residues 195-196 are pulled inward 1.5 Å in IFP1.4 compared to DrCBDmon, excluding a solvent glycerol (Figure 4b).

One consequence of the selection scheme used to optimize IFP1.4 was that the procedure optimized for fluorescence based on FACS sorting using a shorter excitation wavelength than the absorption maximum of DrCBD (10). It is thus not surprising that the resulting phytofluor has an excitation maximum of 684 nm. We have shown that a single

change (V186M) in the hydrophobic environment around the BV D-ring was sufficient to cause blue-shifting (Fig. 1), and moreover that there is no gain of fluorescence intensity when this side chain is added to DrCBDmonY263F (Table 2). This is a welcome result given the impetus to find NIR phytofluors with high quantum yield and long wavelength optima.

On replacement of Asp207 by His, DrCBD exhibits near-IR fluorescent capabilities (8,10,23). Recent work has demonstrated that reversal of the His residue to Asp in the DIP motif had no impact on the fluorescent quantum yield of the DrCBDmon-Y263F variant (8). Rhodopseudomonas palustris BphP3 is naturally fluorescent with an Asp at this position (24,32,33). We thus studied the solution behaviors of DrCBD variants with either Asp or His at position 207 and discovered that IFP1.4-D207 (IFP1.4rev) is ~30% more fluorescent than its progenitor with a quantum yield of nearly 9%, in comparison to the 7% of IFP1.4 (Table 2). The underlying reason for the increased quantum yield of IFPrev is not obvious. pKa differences between the His and Asp side chains could have differential affects in the subtly different chromophore environments of the now multiply-described NIR phytofluors and warrant further investigation. Alternatively, the steric blocking of the second water in the center of the BV by the interaction of Tyr263 with His207 may be the relevant variable. This interaction has been noted previously (8). We found IFP1.4 and Wi-Phy showed higher variability in quantum yield measurements between experimental techniques and investigators than IFP1.4rev or DrCBDmon, which were consistent. We concluded that the presence of His207 position leads to sensitivity to even slight changes in pH, ion concentration, illumination flux, or other possible experimental variations. In any case, our steady state and time-resolved fluorescence

measurements clearly indicate the His207 residue is not necessary for enhanced fluorescence of BphPs.

Typically, DrCBD wild type samples show biexponential fluorescence decay profiles (30), suggesting either a heterogeneous population of molecules in the ground state or a second fluorescent species that appears in the excited state of the BV. The fluorescence lifetime of a molecule is extremely sensitive to local environmental changes. Such heterogeneity can be manifest on very small distance and time scales, for example slight repositioning of the pyrrole water and/or a small rotation of the D-ring. Comparing DrCBDmon constructs to more rigid IFPs reveals some interesting behavior. Quantum yields and fluorescence lifetimes are increased (Table 2). Any disturbance of the π-system, e.g. blocking the torsional vibrations of the C15=C16 double bond of the methine bridge between rings C and D, causes these values to increase. This can be explained by the increase in the σπ-interaction or/and the number and strength of the hydrogen-bonds between the protein and D-ring. Spectroscopic studies confirmed the spatial observations of X-ray crystallography. The excited state lifetimes for fluorescent DrCBD derivatives are consistently longer and monoexponential compared with the value found for CBDs from D. radiodurans (30) and R. palustris (33). However, Cph1 lifetimes have been reported as long as 1.8 nsec at room temperature (34,35). The longer decay lifetime and monoexponential decay behavior in the case of the DrCBDmonY263F variant indicates greater rigidity compared to the wild type and/or that the wild type lacks any type of fluorescent photoproduct. In the case of IFP1.4 and IFPrev, the time constants are longer than observed in any other Dr phytochrome variants (33). The considerably longer lifetimes together with the blue-shifted absorption and emission wavelengths of the IFP1.4

proteins compared to the other DrCBD systems arise from the hydrophobic hub and thus confirm the hub forms in solution as well as in a crystalline environment.

Both IFP1.4 and Wi-Phy contain the D207H mutation. Interestingly, pairwise comparison of the corresponding constructs with and without D207H replacement (IFP1.4 vs. IFP1.4rev and Wi-Phy vs. DrCBDmonY263F) reveals similar excited state lifetimes yet different fluorescent quantum yields. Such an effect suggests changes in the quenching mechanism between members of each matched pair. In agreement with our crystal structures, small changes in the charge distribution in the BV binding pocket may lead to variation, for example in the access of solvent ions to the binding pocket, and therefore change the nature of the quenching pathway.

In all of the phytofluors studied here, the quantum yield for bona fide photoconversion is essentially zero ((8) and data not shown), yet the fluorescence quantum yield does not exceed 10%. This discrepancy indicates a large contribution for non-radiative decay processes of the excited state of the BV molecule. Our structural and spectroscopic data indicate there are several avenues for further improvement in the desirable properties of NIR biomarkers based on bacterial phytochromes. High temporal and spatial resolution will be needed to understand and minimize the pathways for excited state proton transfer and static and dynamic quenching mechanisms.

Experimental Methods

Cloning

Standard PCR procedures were used to perform all DNA engineering. QuickChangeTM (Stratagene, La Jolla, CA) method was used to make point mutations. IFP1.4 in pBAD was shared by Prof. Roger Tsien and Dr. Xiakun Shu (UCSD). The following primers were used to transfer the IFP1.4 gene into the pET21a vector: GAAAT AATTT TGTTTA ACTTT AAGAA GGAGA TATAC ATATG GCCCG GGACC CGTTG C and CAGTG GTGGT GGTGG TGGTG CTCGA GCGCT TCCTT GCGTT GAACT TGGC. The D207 variant of IFP1.4 was created using the following primers: GTTTC CCGGC TAGCG ATACC CCGGC and CCTGC GCCGG GGTAT CGCTA GCCGG. The DrCBDmon/DrCBDmonY263F M54V, V186M and A288V variants were created using the following primers respectively:

GCGAGGTGCTCCAGGTGAGCCTCAACGC and
GCGTTGAGGCTCACCTGGAGCACCTCGC

CGCCACCGGCGAAATGATTGCCGAGGCC and
GGCCTCGGCAATCATTTCGCCGGTGGCG

GGGCCTGATCGTGTGCCACCACCAGAC and
GTCTGGTGGTGGCACACGATCAGGCCC

Clones were verified using DNA sequencing at the University of Wisconsin-Madison DNA Biotechnology Center.

Protein Purification

The constructs bearing DrCBD and IFP variants were transformed into BL21 (DE3) expression cells and grown at 37 °C in LB-media in the presence of 0.1 mg/ml ampicillin. At an O.D of 0.6, the cells were induced with isopropyl β-D-1-thiogalactopyranoside. The cells were harvested after 4 hours by centrifugation at 5000 X g and subsequently resuspended in lysis buffer (25 mM Tris buffer (pH 8.0), 50 mM NaCl). In addition the lysis buffer contained 5 mM imidazole for the DrCBD variants. Cells were sonicated and centrifuged at 40000 X g. Cell lysate was incubated with 200µl of 20 mM BV in the dark for an hour. For IFP variants approximately double the amount of BV was added to the lysate followed by an overnight incubation in the dark. Proteins were affinity purified under green light using Ni2+ -NTA resin (Qiagen, Valencia, CA). Further purification was performed using a hydrophobic interacting phenyl sepharose column (GE Life Sciences) in order to separate apo- and holo-phytochrome. The purified protein was subjected to absorption scans to ascertain chromophore binding (25) as well as shifts in maximum absorption of the holoprotein.

Spectroscopy

UV-Vis absorption scans were collected on a Beckman Coulter DU640B spectrophotometer using 1 nm steps from 280 to 800 nm.

Fluorescence excitation and emission scans were collected on a TecanTM Infinite M1000 monochromator-based plate reader set to a 5 nm bandwidth, with samples in Greiner FLUOTRAC 200 96-well flat-bottom black plates. Emission scans were monitored between 550 and 800 nm with excitation at 676 nm for IFP1.4 and IFPrev and 696 nm for

all other DrCBDmon variants. Excitation scans were collected between 350 and 800 nm with emission monitored at 705 nm for IFP1.4 and IFPrev and 728 nm for all other variants. There was a 15 or 18 nm gap in data collection for all scans at wavelengths where the excitation and emission wavelengths coincided. The step size used for the scans was 5 nm with an additional 2 nm step size scan across the peak for newly reported variants (IFPrev, V186M, A288V, Y263F/A288V, M54V, Y263F/V186M). All samples had an absorbance of 0.25 OD at their absorbance maxima. Emission spectra were collected for this 0.25 OD stock as well as three dilutions, in three replicates. Integrated values under each emission curve were used to calculate the fluorescence quantum yields of the variants by comparison to Cy-5 standard dye, which has a quantum yield of 0.27 in PBS (8,36).

Absolute quantum yields were determined using the integrating sphere method in a Hamamatsu QuantaurusTM instrument (37). Three 2.5 ml replicates of each sample were prepared in 30 mM Tris buffer, pH 8.0 at an absorbance of 0.1 at 700 nm for DrCBDmon and Wi-Phy (684 nm for IFP1.4 and IFPrev). Fluorescence was excited at a wavelength of 640 nm and emission spectra were integrated up to 850 nm.

Time-resolved fluorescence

Fluorescence decays of the samples in the sub-nanosecond and nanosecond time scales were measured using a commercial PicoQuant HydraHarp 400 time-correlated single photon counting (TCSPC) data acquisition system. The excitation source comprised of a PicoQuant PDL 800-D pulsed diode laser driver with 660 nm (spectral FWHM 2 nm) diode laser head LDH-P-C-660. The repetition rate of the excitation pulses was set to 40 MHz

in all measurements and the output power of the laser was 1.14 mW for 660 nm excitation. The bandpass filters were used to detect the emission at 714 nm with a single photon avalanche photodiode (SPAD, MPD-1CTC). Time resolution of the experiment was determined to be approximately 60 ps (FWHM of the instrument response function). The data were fitted with a sum of two or three exponentials, as in (30), to obtain fluorescence lifetimes. In addition to the fluorescence decay components, a fast rise component of about 20 ps was needed to obtain satisfactory fits at early time points. This rise component is omitted in Table 2.

Crystallization and data collection Proteins were concentrated to approximately 20 mg/ml in 30 mM Tris, pH 8.0 and crystallized by hanging drop vapor diffusion with drops containing a 1:1 mixture of protein and reservoir solutions. IFP1.4 crystals formed when reservoir solutions contained 0.1 M sodium acetate, pH 4.6, 5.6% PEG 4000, whereas DrCBDmon crystallized with 25% PEG 400, 0.1 M phosphate citrate, pH 4.2. Crystals were cryoprotected in 12% v/v glycerol in reservoir solution. Data were collected at LS-CAT beam line 21-ID-D at the Advanced Photon Source (Argonne, IL). The resulting datasets were integrated and scaled using HKL2000 (38). DrCBD (PDB code: 2O9C) (26) was used as a search model for molecular replacement using Phaser (29). The models for both IFP1.4 and DrCBDmon were built using Coot (39). The structures were refined using REFMAC5 (40) from the CCP4 software package (41). The coordinates and the structure factors for IFP1.4 and DrCBDmon can be accessed from the Protein Data Bank (42) using the codes 4O8G and 4IJG.

The Sride server (43) was used to analyze hydrophobic core packing. Structure figures were generated with Pymol (44).

References

1. Ying, B. W., Fourmy, D., and Yoshizawa, S. (2007) Substitution of the use of radioactivity by fluorescence for biochemical studies of RNA. *Rna* **13**, 2042-2050
2. Shimomura, O., Johnson, F. H., and Saiga, Y. (1962) Extraction, purification and properties of aequorin, a bioluminescent protein from the luminous hydromedusan, Aequorea. *J Cell Comp Physiol* **59**, 223-239
3. Heim, R., Prasher, D. C., and Tsien, R. Y. (1994) Wavelength mutations and posttranslational autoxidation of green fluorescent protein. *Proc Natl Acad Sci* **91**, 12501-12504
4. Stepanenko, O. V., Verkhusha, V. V., Kuznetsova, I. M., Uversky, V. N., and Turoverov, K. K. (2008) Fluorescent proteins as biomarkers and biosensors: throwing color lights on molecular and cellular processes. *Curr Protein Pept Sci* **9**, 338-369
5. Auldridge, M. E., Satyshur, K. A., Anstrom, D. M., and Forest, K. T. (2012) Structure-guided Engineering Enhances a Phytochrome-based Infrared Fluorescent Protein. *J Biol Chem* **287**, 7000-7009
6. Jobsis, F. F. (1977) Noninvasive, infrared monitoring of cerebral and myocardial oxygen sufficiency and circulatory parameters. *Science* **198**, 1264-1267
7. Jobsis, F. F. (1977) Non-invasive, infra-red monitoring of cerebral O2 sufficiency, bloodvolume, HbO2-Hb shifts and bloodflow. *Acta Neurol Scand Suppl* **64**, 452-453

8. Filonov, G. S., Piatkevich, K. D., Ting, L. M., Zhang, J., Kim, K., and Verkhusha, V. V. (2011) Bright and stable near-infrared fluorescent protein for in vivo imaging. *Nat Biotechnol* **29**, 757-761

9. Shu, X., Royant, A., Lin, M. Z., Aguilera, T. A., Lev-Ram, V., Steinbach, P. A., and Tsien, R. Y. (2009) Mammalian expression of infrared fluorescent proteins engineered from a bacterial phytochrome. *Science* **324**, 804-807

10. Davis, S. J., Vener, A. V., and Vierstra, R. D. (1999) Bacteriophytochromes: phytochrome-like photoreceptors from nonphotosynthetic eubacteria. *Science* **286**, 2517-2520

11. Auldridge, M. E., and Forest, K. T. (2011) Bacterial phytochromes: more than meets the light. *Crit Rev Biochem Mol Biol* **46**, 67-88

12. Yu, D., Gustafson, W. C., Han, C., Lafaye, C., Noirclerc-Savoye, M., Ge, W. P., Thayer, D. A., Huang, H., Kornberg, T. B., Royant, A., Jan, L. Y., Jan, Y. N., Weiss, W. A., and Shu, X. (2014) An improved monomeric infrared fluorescent protein for neuronal and tumour brain imaging. *Nat Commun* **5**, 3626-3632

13. Piatkevich, K. D., Subach, F. V., and Verkhusha, V. V. (2013) Far-red light photoactivatable near-infrared fluorescent proteins engineered from a bacterial phytochrome. *Nat Commun* **4**, 2153-2167

14. Zhang, J., Wu, X. J., Wang, Z. B., Chen, Y., Wang, X., Zhou, M., Scheer, H., and Zhao, K. H. (2010) Fused-gene approach to photoswitchable and fluorescent biliproteins. *Angew Chem Int Ed Engl* **49**, 5456-5458

15. Wagner, J. R., Brunzelle, J. S., Forest, K. T., and Vierstra, R. D. (2005) A light-sensing knot revealed by the structure of the chromophore-binding domain of phytochrome. *Nature* **438**, 325-331

16. Song, C., Psakis, G., Lang, C., Mailliet, J., Gärtner, W., Hughes, J., and Matysik, J. (2011) Two ground state isoforms and a chromophore D-ring photoflip triggering extensive intramolecular changes in a canonical phytochrome. *Proc Natl Acad Sci* **108**, 3842-3847

17. Dasgupta, J., Frontiera, R. R., Taylor, K. C., Lagarias, J. C., and Mathies, R. A. (2009) Ultrafast excited-state isomerization in phytochrome revealed by femtosecond stimulated Raman spectroscopy. *Proc Natl Acad Sci* **106**, 1784-1789

18. Yang, X., Ren, Z., Kuk, J., and Moffat, K. (2011) Temperature-scan cryocrystallography reveals reaction intermediates in bacteriophytochrome. *Nature* **479**, 428-432

19. Foerstendorf, H., Mummert, E., Schafer, E., Scheer, H., and Siebert, F. (1996) Fourier-transform infrared spectroscopy of phytochrome: difference spectra of the intermediates of the photoreactions. *Biochemistry* **35**, 10793-10799

20. Rockwell, N. C., Su, Y. S., and Lagarias, J. C. (2006) Phytochrome structure and signalling mechanisms. *Annu Rev Plant Biol* **57**, 837-858

21. Takala, H., Bjorling, A., Berntsson, O., Lehtivuori, H., Niebling, S., Hoernke, M., Kosheleva, I., Henning, R., Menzel, A., Ihalainen, J. A., and Westenhoff, S. (2014) Signal amplification and transduction in phytochrome photosensors. *Nature* **509**, 245-248

22. Wagner, J. R., Zhang, J., von Stetten, D., Gunther, M., Murgida, D. H., Mroginski, M. A., Walker, J. M., Forest, K. T., Hildebrandt, P., and Vierstra, R. D. (2008) Mutational analysis of Deinococcus radiodurans bacteriophytochrome reveals key amino acids necessary for the photochromicity and proton exchange cycle of phytochromes. *J Biol Chem* **283**, 12212-12226

23. Auldridge, M. E., Satyshur, K. A., Anstrom, D. M., and Forest, K. T. (2012) Structure-guided Engineering Enhances a Phytochrome-based Infrared Fluorescent Protein. *J Biol Chem* **287**, 7000-7009

24. Toh, K. C., Stojkovic, E. A., van Stokkum, I. H., Moffat, K., and Kennis, J. T. (2010) Proton-transfer and hydrogen-bond interactions determine fluorescence quantum yield and photochemical efficiency of bacteriophytochrome. *Proc Natl Acad Sci* **107**, 9170-9175

25. Fischer, A. J., and Lagarias, J. C. (2004) Harnessing phytochrome's glowing potential. *Proc Natl Acad Sci* **101**, 17334-17339

26. Wagner, J. R., Zhang, J., Brunzelle, J. S., Vierstra, R. D., and Forest, K. T. (2007) High resolution structure of Deinococcus bacteriophytochrome yields new insights into phytochrome architecture and evolution. *J Biol Chem* **282**, 12298-12309

27. Borucki, B., von Stetten, D., Seibeck, S., Lamparter, T., Michael, N., Mroginski, M. A., Otto, H., Murgida, D. H., Heyn, M. P., and Hildebrandt, P. (2005) Light-induced proton release of phytochrome is coupled to the transient deprotonation of the tetrapyrrole chromophore. *J. Biol. Chem.* 280, 34358–34364

28. von Stetten, D., Seibeck, S., Michael, N., Scheerer, P., Mroginski, M. A., Murgida, D. H., Krauss, N., Heyn, M. P., Hildebrandt, P., Borucki, B., and Lamparter, T.

(2007) Highly conserved residues Asp-197 and His-250 in Agp1 phytochrome control the proton affinity of the chromophore and Pfr formation. *J Biol Chem* **282**, 2116-2123

29. Lehtivuori, H., Rissanen, I., Takala, H., Bamford, J., Tkachenko, N. V., and Ihalainen, J. A. (2013) Fluorescence properties of the chromophore-binding domain of bacteriophytochrome from Deinococcus radiodurans. *J Phys Chem B* **117**, 11049-11057
30. Yang, X., Stojkovic, E. A., Kuk, J., and Moffat, K. (2007) Crystal structure of the chromophore binding domain of an unusual bacteriophytochrome, RpBphP3, reveals residues that modulate photoconversion. *Proc Natl Acad Sci* **104**, 12571-12576
31. Toh, K. C., Stojkovic, E. A., van Stokkum, I. H., Moffat, K., and Kennis, J. T. (2010) Proton-transfer and hydrogen-bond interactions determine fluorescence quantum yield and photochemical efficiency of bacteriophytochrome. *Proc Natl Acad Sci* **107**, 9170-9175
32. Toh, K. C., Stojkovic, E. A., van Stokkum, I. H., Moffat, K., and Kennis, J. T. (2011) Fluorescence quantum yield and photochemistry of bacteriophytochrome constructs. *Phys Chem Chem Phys* **13**, 11985-11997
33. Lackowicz, J. (2006) *The Principles of Fluorescence Spectroscopy*, Plenum Press, New York
34. Otwinowski, Z., and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode. **276**, 307–326

35. Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr* **60**, 2126-2132

36. Murshudov, G. N., Skubak, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F., and Vagin, A. A. (2011) REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr* **D67**, 355-367

37. McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007) Phaser crystallographic software. *J Appl Crystallogr* **40**, 658-674

38. Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N., and Bourne, P. E. (2000) The Protein Data Bank. *Nucleic Acids Res* **28**, 235-242

39. Magyar, C., Gromiha, M. M., Pujadas, G., Tusn??dy, G. E., and Simon, I. (2005) SRide: A server for identifying stabilizing residues in proteins. *Nucleic Acids Res.* 10.1093/nar/gki409

40. **Schrödinger, L.** The PyMOL Molecular Graphics System, Version 1.3.

CHAPTER III

Mutational Analyses of *Deinococcus radiodurans* bacteriophytochromes discovers key residues to augment fluorescence

Partially adapted from 2 papers:

1. Bhattacharya, S., Auldridge, M. E., Lehtivuori, H., Ihalainen, J. a, and Forest, K. T. (2014) Origins of Fluorescence in Evolved Bacteriophytochromes. J. Biol. Chem. 289, 32144–32152

2. Lehtivuori, H., Bhattacharya, S., Angenent-Mari, N., Satyshur, K. A., & Forest, K. T. (2015). Reducing Polar Interactions Near the Chromophore Increases Near-infrared Fluorescence While Blue-shifting Absorbance Maximum in a Phytochrome-based Biomarker. Frontiers of Molecular Biosciences, vol. 2, pg. 65.

D207L was variant was made by Dr. Lehtivuori, protein purification and crystallization and vitrification was done by me. I assisted Dr. Lehtivuori with spectroscopic data collection on D207L. Kenneth Satyshur solved the structure. All other variants were made, purified and characterized by me. Three paragraphs in the results section of crystal structure of WiPhy2 were taken from (85)

Abstract

Based on our structural information from IFP1.4 and WiPhy fluorescent variants in comparison to the non-fluorescent photoconverting $DrCBD_{mon}$, we came up with a set of residues in the chromophore binding pocket which could lead to brighter variants. These residues were categorized into two general sets a. substitutions around the D-ring which would prevent photoconversion by increasing prevent Z→E isomerization around the C15-C16 bond, b. substitutions of and around the D207 residue. The carbonyl backbone of the 207^{th} residue forms an extended H-bonding network with BV A, B and C rings and, the pyrrole water. Removing polar interactions would decrease deprotonation of the BV, thus reducing photoconversion and ESPT. Substitutions around the BV D-ring and D207 were able to independently improve fluorescence in $DrCBD_{mon}$ background. The A288V substitution increased fluorescent quantum yield by 60% compared to $DrCBD_{mon}$ template. Additionally, we discovered that the V186M variant blue shifted the $DrCBD_{mon}$ excitation maximum by 8 nm. The brightest variant was $DrCBD_{mon}$D207LY263F (WiPhy2) at a fluorescent quantum yield of 8.7. We solved the crystal structure of WiPhy2 at 1.54 Å resolution. The structure demonstrates that a non-polar substitution in the 207^{th} position reduces the number of solvent molecules in the binding pocket. Our study reinforces the importance of D207 residue in the photocycle and shows a third mode of improving fluorescence by decreasing solvent content in the binding pocket.

Introduction

Editing the protein environment around the chromophore alters the spectroscopic properties of the bacterial phytochromes. The mechanism by which bacterial phytochromes act as red/far red light switches in plants and microorganisms lies in key interactions between the bilin chromophore and the apoprotein (Figure 3.1A) that promote chromophore ligation and photoconversion between the spectrally distinct red light (R) absorbing Pr and far-red light (FR) absorbing Pfr forms (2, 64, 86). BphPs are dimeric in nature with each monomer having a covalently attached BV to a conserved cysteine residue on the N-terminal end of the polypeptide chain (39, 60). The input module in prototypical BphPs consist of PAS and the GAF domains which are responsible for chromophore attachment and formation of the binding pocket, followed by the Phytochrome associated (Phy) domain, which is required for complete photoconversion from Pr→Pfr state as well as for forming a stable Pfr state (87, 88). An output module follows the photosensory core domain, in BphPs, this is usually a histidine kinase domain (35). The chromophore BV is protonated in the Pr ground state (89). Upon exposure to red light photon, the Pr ground state is excited to a Pr* excited state with a C15-C16 bond angle strain. The Z→E isomerization of the C15-C16 bond leads to formation of the LumiR* excited state (55, 90). The first product in the relaxation pathway is LumiR. It takes microseconds for the LumiR to transform into the MetaRa photoproduct. Deprotonation of the BV at this point leads to transformation of the MetaRa intermediate into the MetaRc intermediate. The proton from the chromophore is taken up either by the pyrrole water or the oxygen atom of the carbonyl backbone of the residue next to the pyrrole water (91). The MetaRc intermediate reuptakes the proton lost in the previous

step and combined with changes in the polypeptide chain conformation, leads to the formation of the far-red absorbing Pfr state (Figure 3.1B).

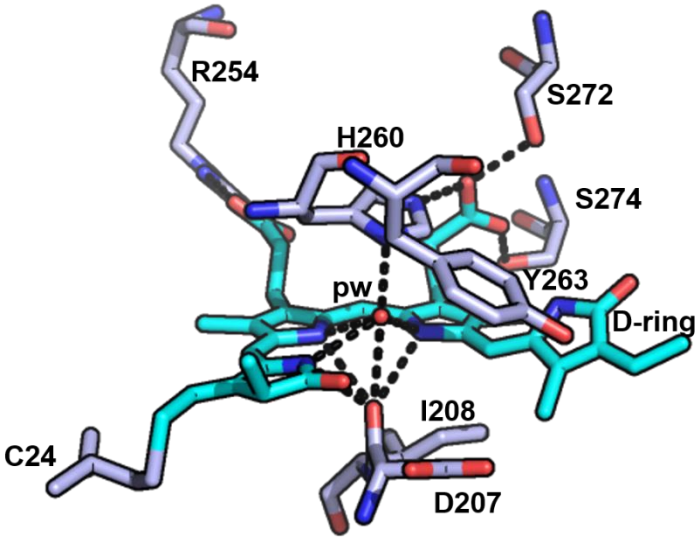
The PAS and the GAF domains together are defined as the chromophore binding domain (CBD) and DrCBD has a stable Pr state but in absence of the Phy domain, photoconversion to Pfr is impaired (2, 46). After the structure of DrCBD was solved, first to a resolution of 2.5 Å (46), then to 1.45 Å by engineering better crystal packing (47), Wagner and his co-workers made site-specific substitutions in the chromophore binding pocket to understand the role of positionally important amino acids in defining photochromicity of DrCBD. Phytochromes have evolved to be photoreceptors, when variations in amino acids in the binding pocket disrupted the ability of the chromophore to rearrange itself, flux of photons through non-traditional and non-productive relaxation pathways like fluorescence increased. The discovery of the red fluorescence associated D207H substitution in DrCBD (2) led to the development of bacterial phytofluors IFP1.4, WiPhy and iRFP (3–5). Auldridge *et al.* monomerized DrCBD to construct fluorescent WiPhy variant. Upon monomerization of DrCBD, the $DrCBD_{mon}$ variant displayed limited photoconversion ability. It was proposed that absence of the dimerization packing constraint led to greater freedom of movement of the polypeptide chain which allowed for photoconversion to occur (4). WiPhy contains two mutations in the binding pocket, D207H and Y263F that contribute to its fluorescence quantum yield of 5.6%. Auldridge *et al*. also demonstrated that both D207H and Y263F were independently contributing to the improved fluorescence of $DrCBD_{mon}$. The two substitutions destabilized the BV conformation that results due to the Z→E isomerization and leads to the LumiR* state

collapsing back to the Pr state. Additionally, D207H is the residue closest to the pyrrole water and the oxygen atom on the carbonyl backbone of D207 is proposed to be the putative proton sink (56, 92). Changing the H-bonding network in this region decreases the rate of BV deprotonation between the MetaRa and MetaRc states of the photocycle and increases ESPT (4), Figure 1B.

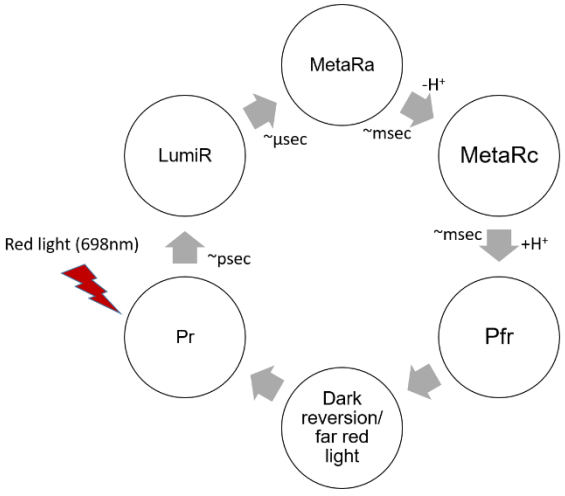
Our structural comparison of IFP1.4 and $DrCBD_{mon}$ helped us identify the residues that contributed to increased fluorescence of IFP1.4. These included the hydrophobic changes in the binding pocket around the D-ring, A288V and V186M substitutions, and the M54V substitution in the PAS domain which we hypothesize prevents non-radiative decay. The other observation from IFP1.4 was the increased polar interaction between BV chromophore and the 207-208th residues. In wild-type $DrCBD_{mon}$ these residues are aspartate and isoleucine. IFP1.4 had histidine and threonine in the 207th and 208th positions respectively. We established in chapter II that histidine to aspartate reversion in the IFP1.4 molecule improves fluorescence quantum yield by 11%. Here we probed the effect of site specific substitutions from our structural knowledge of WiPhy and IFP1.4 to further increase phytochrome fluorescence.

Figure 3.1. Structure and photochemistry of the binding pocket. A. 3D- position of amino acid residues in the BV binding pocket of $DrCBD_{mon}$. Key polar interactions are displayed in black dashed lines. BV is attached via a thioether linkage to the cysteine 24 residue. pw stand for the highly conserved pyrrole water which is involved in deprotonation and subsequent reprotonation of BV (PDB 4IJG, (7)). B. Graphic representation of the Pr→Pfr photocycle, with relative timescales.

A



B



Results

Previous work has used site-directed mutagenesis, both random and saturated to generate fluorescent variants of bacteriophytochromes. We leveraged our high-resolution structures of WiPhy [PDB ID 3S7Q (Auldridge et al, 2012) and IFP1.4 (7) to pin point residues that are photochemically important for increasing fluorescence. These amino acid residues are in the chromophore binding domain (PAS-GAF) and are either in close contact with the chromophore or around the lasso loop. The substitutions can be classified into:

a. Introducing Hydrophobic residues
b. Removing Ionic residues

To this end, we identified key residues from the structures of IFP1.4 and WiPhy and transferred them to wild-type monomeric DrCBD backgrounds. Interpretation of the mutational studies was mostly limited to expression, chromophore assembly and spectral characterization. The results of the variants that expressed and bound BV and allowed for spectroscopy data collection are collated in table 1 and 2.

We did solve the structure of one of our substitutions, the $DrCBD_{mon}$D207LY263F variant and we will be discussing the implications of the structure as well.

Table 3.1. Spectroscopic and photochemical properties of 15 $DrCBD_{mon}$ variants

Protein Variant	**Protein expression**	**BV attachment**	**Photo-conversion**	**Reference**
$DrCBD_{mon}$	++	YES	YES	(4, 7)
$DrCBD_{mon}$Y263F	++	YES	NO	(4)
$DrCBD_{mon}$D207HY263F (Wi-Phy)	++	YES	NO	(7)
IFP1.4	++	YES	NO	(7)
$IFP1.4_{rev}$	++	YES	NO	(7)
$DrCBD_{mon}$A288V	++	YES	NO	(7)
$DrCBD_{mon}$Y263F/A288V	++	YES	NO	
$DrCBD_{mon}$V186M	++	YES	NO	(7)
$DrCBD_{mon}$M54V	++	YES	NO	
$DrCBD_{mon}$D207LY263F (WiPhy2)	++	YES	NO	(85)
$DrCBD_{mon}$D207L	++	YES	NO	
$DrCBD_{mon}$D207LY263FA288V	++	YES	NO	
$DrCBD_{mon}$I208T	++	YES	YES	
$DrCBD_{mon}$ΔD207	+-	NO	N/A	
$DrCBD_{mon}$ΔDI207-208	+-	NO	N/A	
$DrCBD_{mon}$ΔDIP207-208-209	--	NO	N/A	

Table 3.2. Fluorescence characteristics of $DrCBD_{mon}$ variants

Protein Variant	**Abs. max, (nm)**	**Excit. max, (nm)**	**Em. max, (nm)**	**Φ[a] (%)**	**ε (M^{-1}/cm^{-1})**	**Bright-ness (%)**	**Lifetime (ps)**	
							τ_1	**τ_2**
$DrCBD_{mon}$	698	698	718	2.9±0.1	133, 199	100	390±30	620±70
$DrCBD_{mon}$Y263F	701	700	722	4.0±0.1	128, 250	132	-	670±10
Wi-Phy	701	700	722	6.3±0.2	117, 947	192	-	670±10
IFP1.4	684	685	708	7.0±0.3	130, 533	236	-	800±10
$IFP1.4_{rev}$	684	685	708	8.7±0.1	131, 473	296	-	815±11
$DrCBD_{mon}$A288V	701	700	720	4.2±0.2	126, 566	138	ND	ND
$DrCBD_{mon}$Y263F/A288V	701	700	722	4.6±0.1	130, 210	155	ND	ND
$DrCBD_{mon}$V186M	690	692	708	3.2±0.1	135, 619	102	ND	ND
$DrCBD_{mon}$M54V	701	700	720	3.0±0.1	91,000	101	ND	ND
$DrCBD_{mon}$D207LY263F	696	698	719	8.7±0.5	125,636	285		780±30
$DrCBD_{mon}$D207L	697	698	722	7.0±0.5	119, 816	194		650±30
$DrCBD_{mon}$D207LY263FA288V	696		719	8.8	127,266	287		ND
$DrCBD_{mon}$I208T	698		718	3.1	130, 417	98		ND

Hydrophobic Residues in the chromophore binding pocket

Photoconversion from the Pr to Pfr states is mediated by *Z→E* isomerization of the C=C double bond which leads to flipping of the D-ring. The D-ring is ~44° out of plane with respect to the A, B and C rings of BV and this orientation of the D-ring is maintained by an ionic interaction with the histidine 290 residue in *Deinococcus radiodurans* BphP (46). In our IFP1.4 structure we observed that the D-ring interaction with the histidine 290 residue was impeded due to change in ring orientation. This change in orientation was mediated by the introduction of a few new residues in amino acid residues in the region. We hypothesized that the alanine 288 to valine and the valine 188 to methionine residues were the positionally significant substitutions in IFP1.4 which altered the positions of the conserved tyrosine 263, methionine 267 and 174 and phenylalanine 198 and 203 residues by tenths of angstroms and brought them closer to the chromophore and helped in locking the D-ring in place.

We introduced the Ala288Val and the Val186Met mutations separately in monomeric DrCBD backgrounds to probe their individual effects on photochemistry. Both variants expressed and were able to covalently bind BV covalently and produced normal Pr spectra. Neither variant photoconverted to Pfr upon irradiation with red light. The Ala288Val variant was the one that had a higher fluorescence signal than $DrCBD_{mon}$. The excitation maxima of both variants were 700nm and the emission spectrum peaked at 720 nm. Both the variants were tested for photostability. Neither variant photobleaches, there was ~5% loss in emission signal upon light exposure.

Changes in and around Asp 207

The aspartate 207 residue is one of the crucial residues that defines the photochemistry of the phytochrome. It is part of the highly conserved Asp-Ile-Pro motif. The main-chain carbonyl backbone of the aspartate residue is involved in an extended H-bonding network with the pyrrole nitrogens of the BV A, B and C rings. Additionally, it is also part of the polar interaction with the central pyrrole water. The aspartate side-chain forms a H-bond with the Tyr 263 hydroxyl group. Mutational analyses done previously had demonstrated that a Asp207His substitution leads to amplified red fluorescence in DrCBD (Wagner 2008). This substitution was the cornerstone of subsequent research in generating near-infrared fluorescent biomarkers using BphP as template including WiPhy, IFP 1.4 and iRFPs (3, 4, 71)

The structure of WiPhy revealed that limiting the polar interaction around the 207th residue by introducing a nonpolar residue (Tyr263Phe) increased fluorescence. Additionally, the structure of IFP 1.4 didn't suggest a reason for the presence of the histidine in the 207th position. Upon reverting the histidine back to an aspartate residue (IFPrev or IFPD207) had an 11 % higher quantum yield. We further probed the effect of the 207th position by deleting the Asp 207 residue. The deletion of the 207th residue had a deleterious effect on protein expression and the variant was mostly present as inclusion bodies in SDS-PAGE.

IFP1.4 had a threonine in position 208 which increased the lattice like H-bonding network around the pyrrole water. We wanted to test the effect of the substitution on a wild-type template and introduced it on DrCBD$_{mon}$ and found no improvement in fluorescent

capabilities. This agrees with our proposed model of role of H-bonding around the pyrrole water being crucial for deprotonation of BV.

We decided to introduce a non-polar substitution of residue 207. We introduced a leucine residue in this position, as it is the only non-polar residue that most closely resembles the structure of aspartate. We chose to substitute Leu because it is the nonpolar side chain whose structure most closely mimics that of the native Asp. The size of Leu should prevent adventitious binding of Protoporphyrin IXα (PPIXα), which interacts covalently with H207-carrying variants (2, 52, 93, 94). The evidence for PPIXα binding includes the fact that fluorescence spectroscopy of the D207A apoprotein assembled with BV detected two fluorescent species, one matching the absorption and emission spectra of incorporated PPIXα, and a second matching those for BV ((93, 94).We generated two variants the $DrCBD_{mon}$D207L and the $DrCBD_{mon}$D207LY263F (WiPhy2). We did a detailed spectral characterization of the two variants and solved the three-dimensional structure of WiPhy2. The crystallographic structure verified that the L207 side-chain had 60% less water molecules, only 4 water molecules including the conserved pyrrole water. In contract the high-resolution three-dimensional structure of the DrCBD molecule had 9 water molecules. The distance cut-off for the analysis was 5Å.

Both D207L and D207LY263F variants have an absorption maximum of ~697 nm (Pr), Both variants also had a shoulder at 650 nm. The 650 nm shoulder is more pronounced in both the leucine variants than other $DrCBD_{mon}$ derivatives. Neither variant

photoconverted upon red light irradiation, however there was an appearance of a shoulder at 730 nm which demonstrates either an incomplete photocycle or a different kind of photocycle than other variants.

We further quantified the fluorescence spectra of the leucine variants. The emission maxima of the D207L and D207LY263F variants are 722 and 719 nm respectively. D207LY263F had a quantum yield of 8.7 which is approximately 24% higher than just D207L and 45% higher than WiPhy (D207HY263F). In fact, this is one of the highest fluorescence quantum yields we have observed in fluorescent phytochromes. The time-resolved lifetime measurements matched the results of the fluorescence spectra. Using excitation wavelength of 660 nm and monitoring wavelength of 720 nm, the excitation decay properties of BV molecules in the binding pocket can be studied. The excited state decay can be described by monoexponential components, with time constants of 650 ± 30 ps for $DrCBD_{mon}$D207L and 780 ± 30 ps for WiPhy2. Analysis of the shape of the fluorescence spectra of the variants in contrast to the histidine variants of $DrCBD_{mon}$ also reveals that the leucine variants do not bind protoporphyrin IX. Instead the fluorescence spectra of WiPhy2 is due to covalent attachment to BV only.

Crystal Structure of WiPhy2

To gain insight into the fluorescent nature of WiPhy2, a 1.3 Å resolution crystal structure was obtained (Table 3.3). There were no significant changes to the overall structure of WiPhy2 compared to $DrCBD_{mon}$ (RMSD 0.82 Å over all 296 shared Cα atoms including mobile loop regions; Figure 3.2A). The BV chromophore is well-ordered with no evidence of a break in electron density for the cysteine connection to the A-ring.

The most obvious result from this new structure is the confirmation of our hypothesis: waters are less abundant around the L207 side chain than has been seen in other high-resolution structures containing either Asp or His at this position. For D207L, within 5 Å of any Leu atom there are only four waters, including the pyrrole water with strong interactions to BV nitrogen atoms in A-, B-, and C-rings. The closest of these waters to any Leu side chain atom is 3.8 Å. All four water positions are conserved in the water network of $DrCBD_{mon}$ (PDB ID: 4IJG,(7)). The second and third form a path from the pyrrole water to the solvent, whereas the fourth is located under the residue 207 side chain and forms a H-bond with Y176. Of course, there are no H-bonds from any of these waters to the L207 side chain in WiPhy2 (Figure 3.2B).

This paucity of solvent molecules can be strongly contrasted with the native sequence found in the monomer structure, which holds nine waters (Figure 3.2C). These waters permit an extensive H-bonding network of 16 different pair wise interactions of 3.8 Å or

less between any two atoms in the set containing all atoms in residue 207, the –OH of Y263, the –OH of the A-ring, and these waters.

The introduction of His at position 207 (PDB ID: 3S7O; (4)) diminishes this H-bonding network somewhat, with seven waters and nine remaining H-bonds. For IFP1.4 (PDB ID: 4O8G,), which also has the polar His at position 207 but introduces the nonpolar Phe263, there remain eight waters in the 5 Å cutoff window from the 207 side chain and 9 H-bonds. Thus, we can conclude that the identity of the residue at position 207 and not at position 263 has the greatest effect on this water network near the “mouth” leading from BV to the solvent. Among the four structures compared here, it is notable that only in $DrCBD_{mon}$ which is known to weakly photoconvert (4) is there a second water located between the chromophore and this outlet.

Table 3.3. X-ray data collection and structure determination statistics

Data Collection	**Rotating Anode - Bruker**	**LS-CAT ID-D**
Wavelength, Å	1.5418	0.9785
Resolution*, Å	47.60-1.50 (1.60-1.50)	23.8-1.30 (1.35-1.30)
Space Group	C2	C2
Unit Cell (a, b, c (Å))	94.8, 55.1, 69.8	94.4, 53.3, 65.7
(α, β, γ (°))	90.0, 91.9, 90.0	90.0, 90.91, 90.0
Completeness, %	96.0 (91.2)	96.5 (94.3)
# Unique Reflections/# Unique Reflections Observed	57159/54864	80665/77741
Redundancy	2.9 (2.0)	7.7 (7.7)
<I/σI>	14.69 (4.83)	33.2(1.0)
Wilson B value, $Å^2$	10.8	10.5
R_{sym}†, %	5.0 (17.9)	5.2 (35.3)
Refinement		
Resolution, Å	25.0-1.50 (1.54-1.51)	23.22-1.30(1.33-1.30)
R_{work}/R_{free}, ‡ %	16.0/19.5 (18.8/27.5)	14.4/16.0 (15.1/17.5)
Rmsd		
Bonds, Å	0.007	0.006
Angles, °	1.391	1.314
Ramachandran statistics, %		

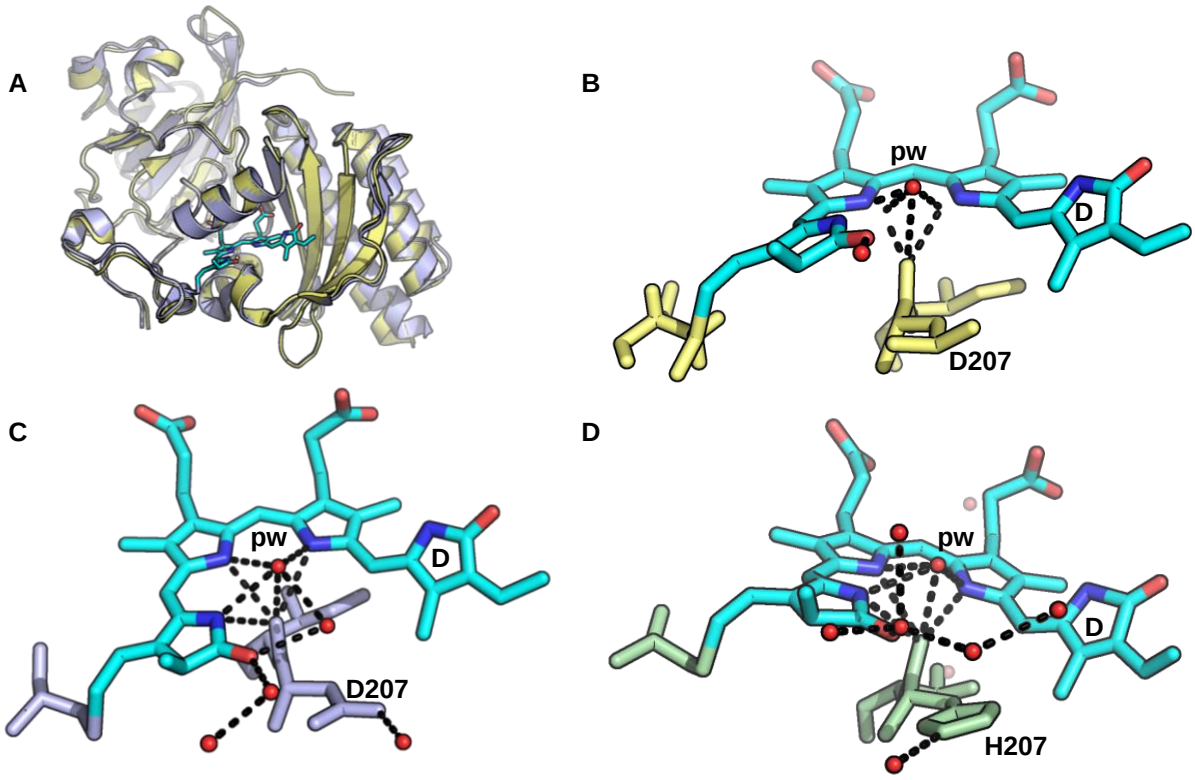
Most favored	---	
Allowed	98.6	98.0
Generously allowed	1.4	2.0
Disallowed	0	0
# atoms		
protein	2457	2551
chromophore	86	86
water	288	231
<B factor>, Å²		
protein	18.5	16.5
chromophore	9.1	7.8
water	28.4	26.1

* The highest resolution bin is indicated in parentheses.

† $R_{sym} = \Sigma\Sigma j|Ij - \langle I \rangle|\Sigma Ij$, where Ij is the intensity measurement for reflection j and $\langle I \rangle$ is the mean intensity for multiply recorded reflections.

‡ $R_{work}/R_{free} = \Sigma||F_{obs}| - |F_{calc}||/|F_{obs}|$, where the working and free R factors are calculated by using the working and free reflection sets, respectively. For the R_{free}, 5% of the total reflections were held aside throughout refinement.

Figure 3.2. WiPhy2 structural assessment. A. Overall fold of WiPhy2 (in pale yellow, PDB 4Z1W) is identical to DrCBD$_{mon}$ (4IJG, light blue). B. Analysing solvent content in binding pocket of WiPhy2. BV binding pocket in WiPhy2 displays fewer water molecules compared to any other DrCBD variant crystal structure, pyrrole water is still intact but the no. H-bonds. C. DrCBD$_{mon}$ has an extensive H-bonding network mediated by the pyrrole water. C. 3D structure of IFP1.4 (PDB 4O8G, pale green) shows a similarly extensive H-bonding network.



Materials and Methods

Cloning

Standard PCR procedures were performed using QuickChange mutagenesis (, La Jolla, CA) All DrCBD mutations were made the pET21a plasmid that encoded the $DrCBD_{mon}$ gene with N-terminal T7 tag and a C-terminal 6X His tag. The primers made for mutagenesis are listed in the table.

Correct sequences of clones were verified using DNA sequencing at the University of Wisconsin- Madison Biotechnology Center.

Protein expression and purification

The $DrCBD_{mon}$ variants were purified via Ni-NTA affinity chromatography followed by hydrophobic interaction chromatography using phenyl sepharose resin to separate the holoprotein from apo. *E. coli* cells expressing the $DrCBD_{mon}$ proteins were disrupted by sonication in 30 mM Tris-HCl (pH 8.0), 100 mM NaCl, 5 mM imidazole (lysis buffer). The supernatant was separated from debris by 45 min centrifugation at 18,000x g. The supernatants were passed through a Ni-NTA column which had been pre-equilibrated with lysis buffer. The bound proteins were eluted

with 30 mM Tris-HCl (pH 8.0), 100 mM NaCl and 250 mM imidazole. The protein sample was then loaded on to a phenyl sepharose HIC column to separate the holoprotein fractions from apoprotein.

Spectroscopic Analysis

Protein samples were diluted in 30 mM Tris, pH 8.0 for all spectral measurements (O. D700~0.1-0.25). Absorption measurements were carried out on a Beckman Coulter DU640B spectrophotometer. The proteins were incubated and all data collected in dark and at room temperature. For Pr to Pfr photoconversion assay, samples were irradiated with red light for 15 minutes using a FOSTER ACE light source and light at 700 nm was filtered out with a fitted filter (Andover Corp., Salem). The sample distance was adjusted so irradiances of 150 µmol m-2 s-1 reached the samples.

Fluorescence excitation and emission scans were collected on a Tecan™ Infinite M1000 monochromator-based plate reader set to a 5-nm bandwidth, with samples in Greiner FLUOTRAC 200 96-well flat-bottom black plates. Emission scans were monitored between 550 and 800 nm with excitation at 676 nm for IFP1.4 and IFPrev and 696 nm for all other $DrCBD_{mon}$ variants. Excitation scans were collected between 350 and 800 nm with emission monitored at 705 nm for IFP1.4 and IFPrev and 728 nm for all other variants. There was a 15- or 18-nm gap in data collection for all scans at wavelengths where the excitation and emission wavelengths coincided. The step size used for the scans was 5 nm with an additional 2-nm step size scan across the peak for newly reported variants (IFPrev, V186M, A288V, Y263F/A288V, M54V, Y263F/V186M). All samples had an absorbance of 0.25 optical density at their absorbance maxima. Emission spectra were collected for this 0.25 optical density stock as well as three dilutions in three replicates. Integrated values under each emission curve were used to calculate the fluorescence

quantum yields of the variants by comparison to Cy-5 standard dye, which has a quantum yield of 0.27 in PBS.

The fluorescence quantum yields of DrCBD$_{mon}$D207L and WiPhy2 were determined relative to two reference fluorophores with known quantum yields (Eaton, 1988). Cy5-N-hydroxysuccinimidyl ester (ΦCy5 = 0.27) (Lumiprobe) dissolved in phosphate-buffered saline and Nile Blue perchlorate (ΦNileBlue = 0.27) (Sigma Aldrich) in acidic ethanol (0.5% (v/v) 0.1 M HCl in ethanol) were used as a fluorescence quantum yield standards (Mujumdar et al., 1993; Sens and Drexhage, 1981).

Crystallization and structure determination

Proteins were concentrated to ~20 mg/ml in 30 mM Tris, pH 8.0, and crystallized by hanging drop vapor diffusion with drops containing a 1:1 mixture of protein and reservoir solutions. WiPhy2 crystals formed when reservoir solutions contained 20% PEG400 and 0.1 M phosphate citrate buffer at pH 4.0. The crystal used for data collection at LS-CAT was soaked for 5 min in a cryoprotectant of 20% glycerol in mother liquor before vitrification. Data were collected at LS-CAT beam line 21-ID-D at the Advanced Photon Source (Argonne, IL). The resulting datasets were integrated and scaled using HKL2000 (95). IFP 1.4 (PDB code 4O8G) was used as a search model for molecular replacement using Phaser (96). The models were built using Coot. The structure was refined using REFMAC5 from the CCP4 software package (97). The coordinates and the structure factors for high-resolution and low-resolution structures can be accessed from the Protein Data Bank.

Discussion

We used the high-resolution crystal structures of fluorescent BphPs IFP 1.4 and WiPhy to select the most important residues that might improve fluorescence by inhibiting Pr → Pfr photoconversion and reversible deprotonation of the chromophore (7, 35). We studied 11 mutations to the chromophore binding pocket using UV-vis absorption spectra, fluorescence excitation an emission spectra and crystallographic structure. The results showed that fluorescence was not significantly affected by single point mutations. All the variants had a stable Pr state and single mutations prevented photoconversion to Pfr. The Pfr conformation has been shown to be highly susceptible to amino acid substitutions due the specificity of residues involved in the final step in photoconversion, Meta Rc to Pfr.

Two of our variants failed to express and these were the deletions made to the conserved Asp-Ile-Pro motif. The DI and DIP deletion variants failed to express, thus demonstrating the importance of the motif in overall protein folding. The D207 deletion did show some expression, however the protein did not assemble BV and most of the protein formed inclusion bodies on SDS-PAGE. Collectively these results indicate the DIP (207-209) residues need to be precisely and correctly positioned for overall protein folding and bilin binding. All other variants including new hydrophobic residues around the D-ring and ionic changes around the Asp 207 residue expressed and bound BV covalently.

The next point to consider in analyzing the point mutations was the absorbance spectra of the holoprotein. Studies have shown that the Pr conformation of the chromophore is usually not affected by single amino acid substitutions (83, 98). The only exception we

found in our mutational analyses was the Val186Met substitution which blue shifted the absorbance maximum to 690 nm. This mutation arose from the ensemble mutations in IFP1.4 and is partially responsible for the blue shifted excitation and emission spectrum of IFP1.4. The leucine variants $DrCBD_{mon}$D207L and WiPhy2 have a maximum absorption at 696 nm with no photoconversion to Pfr upon red light irradiation. However, a pronounced shoulder at 730 nm appears in both variants upon red light irradiation. This is possibly due to the incomplete photocycle, the final step in Pr to Pfr conversion requires the extended H-bonding network mediated by the polar aspartate 207^{th} residue.

Our mutational analyses have shown that there are two ways we can manipulate and improve fluorescence properties of phytochrome variants. Our research showed that IFPrev had the highest excited state lifetime of 870 ps among the phytochrome variants at room temperature. The longer the excited state lifetime of the variants, the higher the chances are of flux of photons through fluorescence. There are two mechanisms by which we could increase the excited state lifetimes of the variants. Immobilization of the D-ring by hydrogen bonding or hydrophobic packing is the first mechanism which has been discussed in chapter II. The A288V substitution increases fluorescence in the wild-type $DrCBD_{mon}$ template by this mechanism. Introduction of the bulky side chain around the D-ring prevents Z→E isomerization of the C15-C16 double bond between the C and D rings of the chromophore. On the other hand, the leucine variants (D207L and WiPhy2) had excited state lifetimes that are longer than WiPhy at 650 and 780 ps. This agrees with the second mechanism of improving fluorescence properties, by decreasing the number of solvent molecules in the bilin binding pocket, which drastically reduces the

extensive hydrogen bonding network mediated via the pyrrole water, pyrrole nitrogen atoms on BV and polar side-chains on the protein molecule. Thus both photoconversion and excited state proton transfer is curtailed leading to increase in fluorescence. Our previous fluorescent variant structures show an abundance of water molecules in the DrCBD molecule (4, 7). WiPhy2 had fewer waters due to the presence of the non-polar Leu residue. While all the crystal structures of DrCBD molecule are in the same space group, with nearly identical crystallization conditions, differences do exist in vitrification conditions. So while relative solvent molecule content between the molecules are valid we cannot absolutely state the number of water molecules. In addition to the water molecules, the other reason for improved fluorescence properties of WiPhy2 is the discouragement in binding PPIXα in the chromophore binding pocket instead of BV. Phytochrome molecules bound to PPIXα do not fluoresce. PPIXα is cyclized version of phytochrome chromophores and fits well into the binding pocket of DrBphP and also some cyanobacterial variants (93). and the emission maxima of phytochrome bound to PPIXα at 660 nm upon being excited with light at 600-650 nm (94). D207L substitution removed PPIXα binding from DrCBD, and we obtained the emission spectra of BV upon excitation in the 600-660 nm region.

We decided to test the effect of trapping the D-ring in place on our WiPhy2 variant to see if there was any change in fluorescence. We introduced the Ala288Val substitution which has been shown to independently improve fluorescence in $DrCBD_{mon}$ background. The $DrCBD_{mon}$D207LY263FA288V triple substitution had identical photochemical and fluoresce spectrum when compared to WiPhy2. This leads us to suppose we have

reached the maximum brightness possible at the long excitation maximum of 697 nm. Further structural substitutions might make slight incremental improvements in fluorescence quantum yield; however non-radiative decay pathways like ESPT are 1000 times faster (65) than radiative processes. Near-infrared phytofluors suffer from high rates of internal conversion. However, the goal of engineering the fluorescent variants was non-invasive imaging in mammalian systems. With the advantage of being able to penetrate through tissues and bones, the signal from phytofluors are higher than brighter proteins that are excited in the visible region of the spectrum. The high fluorescent quantum yields of WiPhy2 combined with long excitation and emission maximum as well as pH and photostability make it a great candidate for an *in-vivo* biomarker.

References

1. Jobsis, F. F. (1977) Noninvasive, infrared monitoring of cerebral and myocardial oxygen sufficiency and circulatory parameters. *Science*. **198**, 1264–1267
2. Wagner, J. R., Zhang, J., von Stetten, D., Günther, M., Murgida, D. H., Mroginski, M. A., Walker, J. M., Forest, K. T., Hildebrandt, P., and Vierstra, R. D. (2008) Mutational analysis of Deinococcus radiodurans bacteriophytochrome reveals key amino acids necessary for the photochromicity and proton exchange cycle of phytochromes. *J. Biol. Chem.* **283**, 12212–12226
3. Shu, X., Royant, A., Lin, M. Z., Aguilera, T. A., Lev-Ram, V., Steinbach, P. A., and Tsien, R. Y. (2009) Mammalian expression of infrared fluorescent proteins engineered from a bacterial phytochrome. *Science*. **324**, 804–807
4. Auldridge, M. E., Satyshur, K. a, Anstrom, D. M., and Forest, K. T. (2012) Structure-guided engineering enhances a phytochrome-based infrared fluorescent protein. *J. Biol. Chem.* **287**, 7000–7009
5. Filonov, G. S., Piatkevich, K. D., Ting, L.-M., Zhang, J., Kim, K., and Verkhusha, V. V (2011) Bright and stable near-infrared fluorescent protein for in vivo imaging. *Nat. Biotechnol.* **29**, 757–761
6. Shcherbakova, D. M., and Verkhusha, V. V. (2013) Near-infrared fluorescent proteins for multicolor in vivo imaging. *Nat. Methods*. **10**, 751–754
7. Bhattacharya, S., Auldridge, M. E., Lehtivuori, H., Ihalainen, J. A., and Forest, K. T. (2014) Origins of Fluorescence in Evolved Bacteriophytochromes. *J. Biol. Chem.* **289**, 32144–32152

8. Lehtivuori, H., Bhattacharya, S., Angenent-Mari, N. M., Satyshur, K. A., and Forest, K. T. (2015) Removal of Chromophore-Proximal Polar Atoms Decreases Water Content and Increases Fluorescence in a Near Infrared Phytofluor. *Front. Mol. Biosci.* **2**, 65–69

9. Stepanenko, O. V., Verkhusha, V. V., Kuznetsova, I. M., Uversky, V. N., and Turoverov, K. K. (2008) Fluorescent proteins as biomarkers and biosensors: throwing color lights on molecular and cellular processes. *Curr. Protein Pept. Sci.* **9**, 338–369

10. Crameri, A., Whitehorn, E. A., Tate, E., and Stemmer, W. P. (1996) Improved green fluorescent protein by molecular evolution using DNA shuffling. *Nat Biotechnol*. **14**, 315–319

11. Yoshihara, S., Shimada, T., Matsuoka, D., Zikihara, K., Kohchi, T., and Tokutomi, S. (2006) Reconstitution of Blue−Green Reversible Photoconversion of a Cyanobacterial Photoreceptor, PixJ1, in Phycocyanobilin-Producing Escherichia coli. *Biochemistry*. **45**, 3775–3784

12. Tsien, R. (1998) The green fluorescent protein. *Annu. Rev. Biochem.* **67**, 509

13. Ando, R., Hama, H., Yamamoto-Hino, M., Mizuno, H., and Miyawaki, A. (2002) An optical marker based on the UV-induced green-to-red photoconversion of a fluorescent protein. *Proc. Natl. Acad. Sci.* **99**, 12651–12656

14. Ehrig, T., O'Kane, D. J., and Prendergast, F. G. (1995) Green-fluorescent protein mutants with altered fluorescence excitation spectra. *FEBS Lett*. **367**, 163–166

15. Reeder, P. J., Huang, Y.-M., Dordick, J. S., and Bystroff, C. (2010) A Rewired Green Fluorescent Protein: Folding and Function in a Nonsequential, Noncircular

GFP Permutant. *Biochemistry*. **49**, 10773–10779

16. Hanson, G. T., McAnaney, T. B., Park, E. S., Rendell, M. E. P., Yarbrough, D. K., Chu, S., Xi, L., Boxer, S. G., Montrose, M. H., and Remington, S. J. (2002) Green Fluorescent Protein Variants as Ratiometric Dual Emission pH Sensors. 1. Structural Characterization and Preliminary Application. *Biochemistry*. **41**, 15477–15488
17. Davis, S. J., and Vierstra, R. D. (1998) Soluble, highly fluorescent variants of green fluorescent protein (GFP) for use in higher plants. *Plant Mol. Biol.* **36**, 521–528
18. Baulcombe, D. C., Chapman, S., and Cruz, S. (1995) Jellyfish green fluorescent protein as a reporter for virus infections. *Plant J.* **7**, 1045–1053
19. Peng, S.-Y., Yang, Y.-S., Chou, C.-J., Lin, K.-Y., and Wu, S.-C. (2015) Differentiation of Enhanced Green Fluorescent Protein-Labeled Mouse Amniotic Fluid-Derived Stem Cells into Cardiomyocyte-Like Beating Cells. *Acta Cardiol. Sin.* **31**, 209–214
20. Boas, D. A., Gaudette, T., Strangman, G., Cheng, X., Marota, J. J. A., and Mandeville, J. B. (2001) The Accuracy of Near Infrared Spectroscopy and Imaging during Focal Changes in Cerebral Hemodynamics. *Neuroimage*. **13**, 76–90
21. Mancini, D. M., Bolinger, L., Li, H., Kendrick, K., Chance, B., and Wilson, J. R. (1994) Validation of near-infrared spectroscopy in humans. *J. Appl. Physiol.* **77**, 2740–2747
22. Hilderbrand, S. A., and Weissleder, R. (2010) Near-infrared fluorescence: application to in vivo molecular imaging. *Curr. Opin. Chem. Biol.* **14**, 71–79
23. Luo, S., Zhang, E., Su, Y., Cheng, T., and Shi, C. (2011) A review of NIR dyes in

cancer targeting and imaging. *Biomaterials*. **32**, 7127–7138

24. Ke, S., Wen, X., Gurfinkel, M., Charnsangavej, C., Wallace, S., Sevick-Muraca, E. M., and Li, C. (2003) Near-Infrared Optical Imaging of Epidermal Growth Factor Receptor in Breast Cancer Xenografts. *Cancer Res.* **63**, 7870–7875
25. Bugaj, J. E., Achilefu, S., Dorshow, R. B., and Rajagopalan, R. (2001) Novel fluorescent contrast agents for optical imaging of in vivo tumors based on a receptor-targeted dye-peptide conjugate platform. *J. Biomed. Opt.* **6**, 122–132
26. Achilefu, S., Dorshow, R. B., Bugaj, J. E., and Rajagopalan, R. (2000) Novel receptor-targeted fluorescent contrast agents for in vivo tumor imaging. *Invest. Radiol.* **35**, 479–485
27. Bongarzone, S., Staderini, M., and Bolognesi, M. L. (2014) Multitarget ligands and theranostics: sharpening the medicinal chemistry sword against prion diseases. *Future Med. Chem.* **6**, 1017–1029
28. Doerr, A. (2009) Fluorescent proteins: into the infrared. *Nat. Methods*. **6**, 482–483
29. Cashmore, A. R. (2003) Cryptochromes: Enabling plants and animals to determine circadian time. *Cell*. **114**, 537–543
30. Murphy, J. T., and Lagarias, J. C. (1997) The phytofluors: a new class of fluorescent protein probes. *Curr. Biol.* **7**, 870–876
31. Christie, J. M., Gawthorne, J., Young, G., Fraser, N. J., and Roe, A. J. (2012) LOV to BLUF: Flavoprotein Contributions to the Optogenetic Toolkit. *Mol. Plant*. **5**, 533–544
32. Pudasaini, A., El-Arab, K. K., and Zoltowski, B. D. (2015) LOV-based optogenetic devices: light-driven modules to impart photoregulated control of cellular signaling.

Front. Mol. Biosci. **2**, 18–22

33. Möglich, A., Yang, X., Ayers, R. A., and Moffat, K. (2010) Structure and Function of Plant Photoreceptors. *Annu. Rev. Plant Biol.* **61**, 21–47
34. Rockwell, N. C., and Lagarias, J. C. (2010) A Brief History of Phytochromes. *ChemPhysChem*. **11**, 1172–1180
35. Auldridge, M. E., and Forest, K. T. (2011) Bacterial phytochromes: more than meets the light. *Crit. Rev. Biochem. Mol. Biol.* **46**, 67–88
36. Al-Sady, B., Ni, W., Kircher, S., Schaefer, E., and Quail, P. H. (2006) Photoactivated Phytochrome Induces Rapid PIF3 Phosphorylation Prior to Proteasome-Mediated Degradation. *Mol. Cell*. **23**, 439–446
37. Fankhauser, C. (2001) The Phytochromes, a Family of Red/Far-red Absorbing Photoreceptors. *J. Biol. Chem.* **276**, 11453–11456
38. Rizzini, L., Favory, J.-J., Cloix, C., Faggionato, D., O'Hara, A., Kaiserli, E., Baumeister, R., Schäfer, E., Nagy, F., Jenkins, G. I., and Ulm, R. (2011) Perception of UV-B by the Arabidopsis UVR8 protein. *Science*. **332**, 103–106
39. Quail, P. H. (2002) Phytochrome photosensory signalling networks. *Nat. Rev. Mol. Cell Biol.* **3**, 85–93
40. Kehoe, D. M., and Grossman, A. R. (1996) Similarity of a chromatic adaptation sensor to phytochrome and ethylene receptors. *Science*. **273**, 1409–1412
41. Terauchi, K., Montgomery, B. L., Grossman, A. R., Lagarias, J. C., and Kehoe, D. M. (2004) RcaE is a complementary chromatic adaptation photoreceptor required for green and red light responsiveness. *Mol. Microbiol.* **51**, 567–577
42. Davis, S. J., Vener, A. V, and Vierstra, R. D. (1999) Bacteriophytochromes:

phytochrome-like photoreceptors from nonphotosynthetic eubacteria. *Science*. **286**, 2517–2520

43. Bhoo, S. H., Davis, S. J., Walker, J., Karniol, B., and Vierstra, R. D. (2001) Bacteriophytochromes are photochromic histidine kinases using a biliverdin chromophore. *Nature*. **414**, 776–779
44. Butler, W. L., Norris, K. H., Siegelman, H. W., and Hendricks, S. B. (1959) Detection, assay, and preliminary purification of the pigment controlling photoresponsive development of plants. *Proc. Natl. Acad. Sci. U. S. A.* **45**, 1703–1708
45. Blaauw-Jansen, G. (1959) The influence of red and far red light on growth and phototropism of the avena seedling. *Acta Bot. Neerl.* **8**, 1–39
46. Wagner, J. R., Brunzelle, J. S., Forest, K. T., and Vierstra, R. D. (2005) A light-sensing knot revealed by the structure of the chromophore-binding domain of phytochrome. *Nature*. **438**, 325–331
47. Wagner, J. R., Zhang, J., Brunzelle, J. S., Vierstra, R. D., and Forest, K. T. (2007) High resolution structure of Deinococcus bacteriophytochrome yields new insights into phytochrome architecture and evolution. *J. Biol. Chem.* **282**, 12298–12309
48. Yang, X., Stojkovic, E. A., Kuk, J., and Moffat, K. (2007) Crystal structure of the chromophore binding domain of an unusual bacteriophytochrome, RpBphP3, reveals residues that modulate photoconversion. *Proc. Natl. Acad. Sci.* **104**, 12571–12576
49. Nagatani, A. (2010) Phytochrome: structural basis for its functions. *Curr. Opin. Plant Biol.* **13**, 565–570

50. Takala, H., Björling, A., Berntsson, O., Lehtivuori, H., Niebling, S., Hoernke, M., Kosheleva, I., Henning, R., Menzel, A., Ihalainen, J. A., and Westenhoff, S. (2014) Signal amplification and transduction in phytochrome photosensors. *Nature*. **509**, 245–248

51. Baker, A. W., and Forest, K. T. (2014) Structural biology: Action at a distance in a light receptor. *Nature*. **509**, 174–175

52. Burgie, E. S., Wang, T., Bussell, A. N., Walker, J. M., Li, H., and Vierstra, R. D. (2014) Crystallographic and Electron Microscopic Analyses of a Bacterial Phytochrome Reveal Local and Global Rearrangements During Photoconversion. *J. Biol. Chem.* **289**, 24573–24587

53. Yang, X., Kuk, J., and Moffat, K. (2009) Conformational differences between the Pfr and Pr states in Pseudomonas aeruginosa bacteriophytochrome. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 15639–15644

54. Bischoff, M., Hermann, G., Rentsch, S., and Strehlow, D. (2001) First Steps in the Phytochrome Phototransformation: A Comparative Femtosecond Study on the Forward (Pr → Pfr) and Back Reaction (Pfr → Pr). *Biochemistry*. **40**, 181–186

55. Andel, F., Hasson, K. C., Gai, F., Anfinrud, P. A., and Mathies, R. A. (1997) Femtosecond time-resolved spectroscopy of the primary photochemistry of phytochrome. *Biospectroscopy*. **3**, 421–433

56. Toh, K. C., Stojkovic, E. A., van Stokkum, I. H., Moffat, K., and Kennis, J. T. (2011) Fluorescence quantum yield and photochemistry of bacteriophytochrome constructs. *Phys Chem Chem Phys*. **13**, 11985–11997

57. Dasgupta, J., Frontiera, R. R., Taylor, K. C., Lagarias, J. C., and Mathies, R. A.

(2009) Ultrafast excited-state isomerization in phytochrome revealed by femtosecond stimulated Raman spectroscopy. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 1784–1789

58. Piatkevich, K. D., Subach, F. V, and Verkhusha, V. V. (2013) Engineering of bacterial phytochromes for near-infrared imaging, sensing, and light-control in mammals. *Chem. Soc. Rev.* **42**, 3441–3452

59. Yang, X., Stojković, E. A., Ozarowski, W. B., Kuk, J., Davydova, E., and Moffat, K. (2015) Light Signaling Mechanism of Two Tandem Bacteriophytochromes. *Structure*. **23**, 1179–1189

60. Rockwell, N. C., Su, Y.-S., and Lagarias, J. C. (2006) Phytochrome structure and signaling mechanisms. *Annu. Rev. Plant Biol.* **57**, 837–858

61. Kim, P. W., Freer, L. H., Rockwell, N. C., Martin, S. S., Lagarias, J. C., and Larsen, D. S. (2012) Femtosecond Photodynamics of the Red/Green Cyanobacteriochrome NpR6012g4 from *Nostoc punctiforme* . 2. Reverse Dynamics. *Biochemistry*. **51**, 619–630

62. Ulijasz, A. T., Cornilescu, G., Cornilescu, C. C., Zhang, J., Rivera, M., Markley, J. L., and Vierstra, R. D. (2010) Structural basis for the photoconversion of a phytochrome to the activated Pfr form. *Nature*. **463**, 250–254

63. Kneip, C., Hildebrandt, P., Schlamann, W., Braslavsky, S. E., Mark, F., and Schaffner, K. (1999) Protonation state and structural changes of the tetrapyrrole chromophore during the Pr/Pfr phototransformation of phytochrome: a resonance Raman spectroscopic study. *Biochemistry*. **38**, 15185–15192

64. Spillane, K. M., Dasgupta, J., and Mathies, R. A. (2012) Conformational

Homogeneity and Excited-State Isomerization Dynamics of the Bilin Chromophore in Phytochrome Cph1 from Resonance Raman Intensities. *Biophys. J.* **102**, 709–717

65. Toh, K. C., Stojkovic, E. A., van Stokkum, I. H. M., Moffat, K., and Kennis, J. T. M. (2010) Proton-transfer and hydrogen-bond interactions determine fluorescence quantum yield and photochemical efficiency of bacteriophytochrome. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 9170–9175
66. Cormack, B. P., Valdivia, R. H., and Falkow, S. (1996) FACS-optimized mutants of the green fluorescent protein (GFP). *Gene*. **173**, 33–38
67. Heim, R., and Tsien, R. Y. (1996) Engineering green fluorescent protein for improved brightness, longer wavelengths and fluorescence resonance energy transfer. *Curr Biol*. **6**, 178–182
68. Afar, B., Merrill, J., and Clark, E. A. (1991) Detection of lymphocyte subsets using three-color/single-laser flow cytometry and the fluorescent dye Peridinin chlorophyll-a protein. *J. Clin. Immunol.* **11**, 254–261
69. Nakajima, O., and Gray, C. H. (1967) Studies on haem alpha-methenyl oxygenase. Isomeric structure of formylbiliverdin, a possible precursor of biliverdin. *Biochem. J.* **104**, 20–22
70. Schäfer, F. P. (1973) Principles of dye laser operation. in *n Topics in Applied Physics Dye Lasers* (Schäfer, F. P. ed), pp. 1–89, Springer-Verlag, New York, 10.1007/3-540-51558-5_7
71. Filonov, G. S., Krumholz, A., Xia, J., Yao, J., Wang, L. V, and Verkhusha, V. V. (2012) Deep-tissue photoacoustic tomography of a genetically encoded near-

infrared fluorescent probe. *Angew Chem Int Ed Engl*. **51**, 1448–1451

72. Bartz-Schmidt, K. U., Walter, P., Krott, R., Brunner, R., Esser, P., and Heimann, K. (1996) [Effects of fluorescein and indocyanine green angiography on subsequent dark adaptation and the electroretinogram]. *Klin Monbl Augenheilkd*. **208**, 224–228
73. Taylor, W. R. (2000) A deeply knotted protein structure and how it might fold. *Nature*. **406**, 916–919
74. Khatib, F., Weirauch, M. T., and Rohl, C. A. (2006) Rapid knot detection and application to protein structure prediction. *Bioinformatics*. **22**, e252–e259
75. Mallam, A. L., Morris, E. R., and Jackson, S. E. (2008) Exploring knotting mechanisms in protein folding. *Proc. Natl. Acad. Sci.* **105**, 18740–18745
76. Mallam, A. L., and Jackson, S. E. (2006) Probing Nature's Knots: The Folding Pathway of a Knotted Homodimeric Protein. *J. Mol. Biol.* **359**, 1420–1436
77. Mallam, A. L., Rogers, J. M., and Jackson, S. E. (2010) Experimental detection of knotted conformations in denatured proteins. *Proc Natl Acad Sci U S A*. **107**, 8189–8194
78. King, N. P., Yeates, E. O., and Yeates, T. O. (2007) Identification of Rare Slipknots in Proteins and Their Implications for Stability and Folding. **373**, 153–166
79. King, N. P., Jacobitz, A. W., Sawaya, M. R., Goldschmidt, L., and Yeates, T. O. (2010) Structure and folding of a designed knotted protein. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 20732–20737
80. Dill, K. A., and MacCallum, J. L. (2012) The protein-folding problem, 50 years on. **338**, 1042–1046
81. Andersson, F. I., Pina, D. G., Mallam, A. L., Blaser, G., and Jackson, S. E. (2009)

Untangling the folding mechanism of the 52-knotted protein UCH-L3. **276**, 2625–2635

82. Bornschlogl, T., Anstrom, D. M., Mey, E., Dzubiella, J., Rief, M., and Forest, K. T. (2009) Tightening the knot in phytochrome by single-molecule atomic force microscopy. *Biophys J.* **96**, 1508–1514

83. von Stetten, D., Seibeck, S., Michael, N., Scheerer, P., Mroginski, M. A., Murgida, D. H., Krauss, N., Heyn, M. P., Hildebrandt, P., Borucki, B., and Lamparter, T. (2007) Highly conserved residues Asp-197 and His-250 in Agp1 phytochrome control the proton affinity of the chromophore and Pfr formation. *J. Biol. Chem.* **282**, 2116–2123

84. Anders Borg, O., and Durbeej, B. (2008) Which factors determine the acidity of the phytochromobilin chromophore of plant phytochrome? *Phys. Chem. Chem. Phys.* **10**, 2528–2537

85. Lehtivuori, H., Bhattacharya, S., Angenent-Mari, N., Satyshur, K. A., and Forest, K. T. (2015) Removal of Chromophore-Proximal Polar Atoms Decreases Water Content and Increases Fluorescence in a Near Infrared Phytofluor. *Front. Mol. Biosci.* **2**, 65–69

86. Spillane, K. M., Dasgupta, J., Lagarias, J. C., and Mathies, R. A. (2009) Homogeneity of phytochrome Cph1 vibronic absorption revealed by resonance Raman intensity analysis. *J Am Chem Soc*. **131**, 13946–13948

87. Essen, L.-O., Mailliet, J., and Hughes, J. (2008) The structure of a complete phytochrome sensory module in the Pr ground state. *Proc. Natl. Acad. Sci.* **105**, 14709–14714

88. Yang, X., Kuk, J., and Moffat, K. (2008) Crystal structure of Pseudomonas aeruginosa bacteriophytochrome: photoconversion and signal transduction. *Proc. Natl. Acad. Sci.* **105**, 14715–14720

89. von Stetten, D., Gunther, M., Scheerer, P., Murgida, D. H., Mroginski, M. A., Krauss, N., Lamparter, T., Zhang, J., Anstrom, D. M., Vierstra, R. D., Forest, K. T., and Hildebrandt, P. (2008) Chromophore heterogeneity and photoconversion in phytochrome crystals and solution studied by resonance Raman spectroscopy. *Angew Chem Int Ed Engl*. **47**, 4753–4755

90. Dasgupta, J., Frontiera, R. R., Taylor, K. C., Lagarias, J. C., and Mathies, R. A. (2009) Ultrafast excited-state isomerization in phytochrome revealed by femtosecond stimulated Raman spectroscopy. *Proc. Natl. Acad. Sci.* **106**, 1784–1789

91. Foerstendorf, H., Mummert, E., Sch??fer, E., Scheer, H., and Siebert, F. (1996) Fourier-transform infrared spectroscopy of phytochrome: Difference spectra of the intermediates of the photoreactions. *Biochemistry*. **35**, 10793–10799

92. Toh, K. C., van Stokkum, I. H., Hendriks, J., Alexandre, M. T., Arents, J. C., Perez, M. A., van Grondelle, R., Hellingwerf, K. J., and Kennis, J. T. (2008) On the signaling mechanism and the absence of photoreversibility in the AppA BLUF domain. *Biophys J*. **95**, 312–321

93. Fischer, A. J., and Lagarias, J. C. (2004) Harnessing phytochrome's glowing potential. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 17334–17339

94. Lehtivuori, H., Rissanen, I., Takala, H., Bamford, J., Tkachenko, N. V., and Ihalainen, J. A. (2013) Fluorescence Properties of the Chromophore-Binding

Domain of Bacteriophytochrome from *Deinococcus radiodurans*. *J. Phys. Chem. B*. **117**, 11049–11057

95. Otwinowski, Z., and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326

96. McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007) Phaser crystallographic software. *J Appl Crystallogr*. **40**, 658–674

97. Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F., and Vagin, A. A. (2011) REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. D. Biol. Crystallogr.* **67**, 355–367

98. Hahn, J., Strauss, H. M., Landgraf, F. T., Gimenèz, H. F., Lochnit, G., Schmieder, P., and Hughes, J. (2006) Probing protein-chromophore interactions in Cph1 phytochrome by mutagenesis. *FEBS J.* **273**, 1415–1429

99. Jamroz, M., Niemyska, W., Rawdon, E. J., Stasiak, A., Millett, K. C., Sułkowski, P., and Sulkowska, J. I. (2015) KnotProt: a database of proteins with knots and slipknots. *Nucleic Acids Res.* **43**, D306–D314

100. Lai, Y.-L., Yen, S.-C., Yu, S.-H., and Hwang, J.-K. (2007) pKNOT: the protein KNOT web server. *Nucleic Acids Res.* **35**, W420–W424

101. Millett, K. C., Rawdon, E. J., Stasiak, A., and Sułkowska, J. I. (2013) Identifying knots in proteins. *Biochem. Soc. Trans.* **41**, 533–537

102. Burgie, E. S., and Vierstra, R. D. (2014) Phytochromes: An Atomic Perspective on Photoactivation and Signaling. *Plant Cell Online*. **26**, 4568–4583

103. Rockwell, N. C., Martin, S. S., Gulevich, A. G., and Lagarias, J. C. (2012) Phycoviolobilin Formation and Spectral Tuning in the DXCF Cyanobacteriochrome Subfamily. *Biochemistry*. **51**, 1449–1463

104. Sineshchekov, V., Mailliet, J., Psakis, G., Feilke, K., Kopycki, J., Zeidler, M., Essen, L.-O., and Hughes, J. (2014) Tyrosine 263 in cyanobacterial phytochrome Cph1 optimizes photochemistry at the prelumi-R→lumi-R step. *Photochem. Photobiol.* **90**, 786–795

105. Narikawa, R., Fukushima, Y., Ishizuka, T., Itoh, S., and Ikeuchi, M. (2008) A Novel Photoactive GAF Domain of Cyanobacteriochrome AnPixJ That Shows Reversible Green/Red Photoconversion. *J. Mol. Biol.* **380**, 844–855

106. Narikawa, R., Ishizuka, T., Muraki, N., Shiba, T., Kurisu, G., and Ikeuchi, M. (2013) Structures of cyanobacteriochromes from phototaxis regulators AnPixJ and TePixJ reveal general and specific photoconversion mechanism. *Proc. Natl. Acad. Sci.* **110**, 918–923

107. Narikawa, R., Nakajima, T., Aono, Y., Fushimi, K., Enomoto, G., Ni-Ni-Win, Itoh, S., Sato, M., Ikeuchi, M., Narikawa, R., Enomoto, G., Ni, W. N., Fushimi, K., and Ikeuchi, M. (2015) A biliverdin-binding cyanobacteriochrome from the chlorophyll d–bearing cyanobacterium Acaryochloris marina. *Sci. Rep.* **5**, 7950–7975

108. Zhang, J., Wu, X.-J., Wang, Z.-B., Chen, Y., Wang, X., Zhou, M., Scheer, H., and Zhao, K.-H. (2010) Fused-Gene Approach to Photoswitchable and Fluorescent Biliproteins. *Angew. Chemie Int. Ed.* **49**, 5456–5458

109. Rockwell, N. C., Martin, S. S., and Lagarias, J. C. (2012) Red/Green Cyanobacteriochromes: Sensors of Color and Power. *Biochemistry*. **51**, 9667–9677

110. Pal, D., and Chakrabarti, P. (1999) Cis peptide bonds in proteins: residues involved, their conformations, interactions and locations. *J. Mol. Biol.* **294**, 271–288

111. Joseph, A. P., Srinivasan, N., and de Brevern, A. G. (2012) Cis-trans peptide variations in structurally similar proteins. *Amino Acids*. **43**, 1369–1381

112. MacArthur, M. W., and Thornton, J. M. (1991) Influence of proline residues on protein conformation. *J. Mol. Biol.* **218**, 397–412

113. Fischer, G., and Schmid, F. X. (1990) The mechanism of protein folding. Implications of in vitro refolding models for de novo protein folding and translocation in the cell. *Biochemistry*. **29**, 2205–2212

114. Hesterkamp, T., and Bukau, B. (1996) Identification of the prolyl isomerase domain of Escherichia coli trigger factor. *FEBS Lett.* **385**, 67–71

115. Kramer, G., Patzelt, H., Rauch, T., Kurz, T. A., Vorderwülbecke, S., Bukau, B., and Deuerling, E. (2004) Trigger Factor Peptidyl-prolyl cis/trans Isomerase Activity Is Not Essential for the Folding of Cytosolic Proteins in Escherichia coli. *J. Biol. Chem.* **279**, 14165–14170

116. Stoller, G., Rücknagel, K. P., Nierhaus, K. H., Schmid, F. X., Fischer, G., and Rahfeld, J. U. (1995) A ribosome-associated peptidyl-prolyl cis/trans isomerase identified as the trigger factor. *EMBO J.* **14**, 4939–4948

117. Kramer, G., Rauch, T., Rist, W., Vorderwülbecke, S., Patzelt, H., Schulze-Specking, A., Ban, N., Deuerling, E., and Bukau, B. (2002) L23 protein functions as a chaperone docking site on the ribosome. *Nature*. **419**, 171–174

118. Gupta, R., Lakshmipathy, S. K., Chang, H.-C., Etchells, S. A., and Hartl, F. U. (2010) Trigger factor lacking the PPIase domain can enhance the folding of

eukaryotic multi-domain proteins in Escherichia coli. *FEBS Lett.* **584**, 3620–3624

119. Ferbitz, L., Maier, T., Patzelt, H., Bukau, B., Deuerling, E., and Ban, N. (2004) Trigger factor in complex with the ribosome forms a molecular cradle for nascent proteins. *Nature*. **431**, 590–596
120. Thomason, L. C., Costantino, N., Court, D. L., Thomason, L. C., Costantino, N., and Court, D. L. (2007) *E. coli* Genome Manipulation by P1 Transduction. in *Current Protocols in Molecular Biology*, pp. 1.17.1–1.17.8, John Wiley & Sons, Inc., Hoboken, NJ, USA, 10.1002/0471142727.mb0117s79
121. Burgie, E. S., Walker, J. M., Phillips, G. N., and Vierstra, R. D. (2013) A Photo-Labile Thioether Linkage to Phycoviolobilin Provides the Foundation for the Blue/Green Photocycles in DXCF-Cyanobacteriochromes. *Structure*. **21**, 88–97
122. Frey, U. H., Bachmann, H. S., Peters, J., and Siffert, W. (2008) PCR-amplification of GC-rich regions: “slowdown PCR.” *Nat. Protoc.* **3**, 1312–1317
123. Gibson, D. G., Young, L., Chuang, R.-Y., Venter, J. C., Hutchison, C. A., and Smith, H. O. (2009) Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat Meth*. **6**, 343–345
124. Schneider, C. A., Rasband, W. S., and Eliceiri, K. W. (2012) NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods*. **9**, 671–675
125. Rumyantsev, K. A., Shcherbakova, D. M., Zakharova, N. I., Emelyanov, A. V., Turoverov, K. K., and Verkhusha, V. V. (2015) Minimal domain of bacterial phytochrome required for chromophore binding and fluorescence. *Sci. Rep.* **5**, 18348–18352
126. Mallam, A. L. (2009) How does a knotted protein fold? *FEBS J.* **276**, 365–375

127. Sułkowska, J. I., Noel, J. K., Ramírez-Sarmiento, Esar A, Rawdon, E. J., Millett, K. C., and Onuchic, J. N. (2013) Knotting pathways in proteins. *Biochem. Soc. Trans.* **41**, 523–527

128. Mallam, A. L., and Jackson, S. E. (2011) Knot formation in newly translated proteins is spontaneous and accelerated by chaperonins. *Nat. Chem. Biol.* **8**, 147–153

129. Yang, X., Stojković, E. A., Ozarowski, W. B., Kuk, J., Davydova, E., and Moffat, K. (2015) Light Signaling Mechanism of Two Tandem Bacteriophytochromes. *Structure*. **23**, 1179–1189

130. Yang, X., Ren, Z., Kuk, J., and Moffat, K. (2011) Temperature-scan cryocrystallography reveals reaction intermediates in bacteriophytochrome. *Nature*. **479**, 428–432

131. Yang, X., Stojkovic, E. A., Kuk, J., and Moffat, K. (2007) Crystal structure of the chromophore binding domain of an unusual bacteriophytochrome, RpBphP3, reveals residues that modulate photoconversion. *Proc Natl Acad Sci U S A*. **104**, 12571–12576

132. Burgie, E. S., Bussell, A. N., Walker, J. M., Dubiel, K., and Vierstra, R. D. (2014) Crystal structure of the photosensing module from a red/far-red light-absorbing plant phytochrome. *Proc. Natl. Acad. Sci.* **111**, 10179–10184

133. Yu, D., Gustafson, W. C., Han, C., Lafaye, C., Noirclerc-Savoye, M., Ge, W.-P., Thayer, D. A., Huang, H., Kornberg, T. B., Royant, A., Jan, L. Y., Jan, Y. N., Weiss, W. A., and Shu, X. (2014) An improved monomeric infrared fluorescent protein for neuronal and tumour brain imaging. *Nat. Commun.* **5**, 3626–3632

134. Giraud, E., Zappa, S., Vuillet, L., Adriano, J.-M., Hannibal, L., Fardoux, J.,

Berthomieu, C., Bouyer, P., Pignol, D., and Verméglio, A. (2005) A new type of bacteriophytochrome acts in tandem with a classical bacteriophytochrome to control the antennae synthesis in Rhodopseudomonas palustris. *J. Biol. Chem.* **280**, 32389–32397

CHAPTER IV

Knot Formation in Phytochrome

All work in this chapter was carried out by me and is yet to be published. Erik Jessen helped me to design the P1 phage transduction experiment and Katiria Gonzalez Rivera assisted me with Gibson assembly.

Abstract

The structure of *Deinococcus radiodurans* phytochrome contains a chromophore binding domain that includes a figure of eight knot with four crossover points. DrCBD was the second deeply knotted protein structures to be to be deposited in the PDB. The knot is crucial for Pr→ Pfr photoconversion but not for fluorescence. The importance of the knot might lie in signal transduction through the photosensory core to the output histidine kinase domain. We wanted to probe the importance the knot in maintaining the integrity of the protein molecule itself as well as understand the folding of the figure-of-eight knot.

We deleted the entire knot by constructing two DrCBDknotless variants. Both variants were able to express and fold recombinantly in *E. coli*, and displayed limited BV attachment in its binding pocket. Additionally, the high (approximately 20%) concentration of proline in the knot region suggested a role for peptidyl-prolyl isomerases in knot formation. We chose the trigger factor chaperone associated peptidyl-prolyl isomerase to test our hypothesis. Our results indicate that absence of TF-PPIase activity slows folded protein accumulation in $DrCBD_{mon}$ and $DrCBD_{mon}$P236G variant as measured by protein expression. The high number of proline residues is also significant since a P236-245-247-G triple substitution had no recombinant expression. Taken together, our study expands our understanding of phytochrome knot folding and may be broadly applicable to deeply knotted proteins.

Introduction

Proteins have to fold into a compact structure in order to be functional. For decades proteins were assumed to be knotless; the topological barrier to knot formation was considered to be too high. The first deeply knotted protein structure was deposited in the Protein Data Bank in 2000 (73) and since then structural biology has helped identify a number of different knotted folds in proteins. Approximately 2% of total protein structures in the protein data base contain knots (74, 99). Most of these knots are trefoil knots in which one loop threads into another without affecting the rest of the protein chain. Protein structure prediction software has introduced algorithms that can now predict knots in protein sequences using sequence similarity to the structures deposited in the PDB (99–101).

Even among the set of known knotted proteins, phytochromes are unusually complex with their four crossovers, and are also deeply knotted, with at least 34 amino acids threaded through the lasso loop (in DrBphP). Although folding of apo-phytochrome is reversible, once the protein covalently incorporates BV, the folding process is irreversible (82). Our previous research using atomic force microscopy had shown that under load apo phytochrome unfolds at ~47pN and holoprotein bound to BV unfolded at 73pN. These numbers are comparable to unknotted proteins of similar dimensions (82). No additional mechanical stability was associated with knotting of the protein, which suggests that mechanical stability is not the reason for evolutionary conservation of the knot in all PAS domain containing phytochromes. We noticed that although the lasso loop has no conserved sequence motifs, all bacterial phytochromes have a high proline concentration across this 24-amino acid sequence (Fig. 1). Studies have shown that in YibK, which

contains a deep trefoil not with three crossovers, proline isomerization is one of the rate determining steps (77).

In this study we tested two hypotheses. First we tested the role of the knot on overall protein folding and chromophore assembly by constructing a knotless variant of DrCBD. The crystallographic three-dimensional structure of the chromophore binding domain in *Deinococcus radiodurans* identified several unique features about the structural integration of the PAS and GAF domains (46, 47). One of the most important discoveries was the figure-of-eight knot that interconnects the PAS and GAF domains. This knot in question is a rare fold with 35 amino acid chain threading through a lasso loop with 4 cross-over points (Figure 4.1A). We believe that the knot formation occurs even before the chromophore is bound to the protein molecule (35, 102). Subsequent structural work has shown that the figure-of-eight knot is a characteristic of all phytochromes that contain the PAS domain, including plant phytochromes and BphPs. The importance of the knot in signal transduction could lie in its ability to give the photosensory core domain a degree of flexibility that allows downstream channeling of the light signal.

We designed knot deletion of a monomeric DrCBD variant to better understand two phenomena. Firstly, if the figure-of-eight knot has been evolutionarily conserved for photoconversion, a knotless deletion should not be able to photoconvert. Deletion of the knot would pave the way for deleting the entire PAS domain itself while still maintaining BV incorporation. On the other hand, we also wanted to understand the importance of the knot in protein folding itself. If the knot is crucial for the structural integrity of the

phytochrome protein molecule itself, then designing a one-domain near-infrared fluorescent phytochrome would be near impossible.

GAF-only chromophore binding domains of phytochromes already exist in nature. Cyanobacterial phytochromes which use phycocyanobilin as their small molecule chromophore and can go through the reversible photocycle with only the GAF domain as its photosensory core (103–105). The structure of the phytochrome from *Thermosynechococcus elongatus* and *Anabena spp* PCC7120 was solved in 2013 (106). The GAF domains of CBCRs can fluoresce and GAF-only CBCR which can bind BV have also been designed (107–109).

We synthesized and characterized two knotless DrCBD variants with a deletion of 231-255 and 231-253 residues. Both variants were expressed as soluble protein and had limited chromophore binding, however the ~10% of protein that bound BV was able to orient it correctly in the binding pocket, as demonstrated by a small absorbance peak at 700 nm. This is a significant result as it shows the importance of knot lies in holoprotein formation, rather than apo-protein structure.

Our second hypothesis revolved around the predominance of proline residues in the primary amino acid sequence of the lasso loop region. Approximately 20% of the residues in the loop are proline with the P236 residue invariantly adopting a cis-conformation in all DrCBD structures we have solved so far (4, 7, 46, 47, 85). Cis prolines are not uncommon in protein structures, approximately 5% of proline residues in PDB have been found to be in cis-conformation in contrast to an overall probability of <1% for other amino acid residues(110–112). Finding cis P236 as an invariant residue within the lasso loop

suggested it as key to knot formation either for protein structure or function. We propose that the trans to cis isomerization of the P236 residue within the lasso loop keeps it extended for the period of time that is necessary to form the knot. Absence of the P236 residue would have an effect on folding of the knot. In order to test the relevance of the cis proline residue as well as the high concentration of proline residue in this region, we constructed P236G mutation and P236-245-247G triple mutation on $DrCBD_{mon}$. The in variant cis-proline substitution P236G had slower accumulation time upon recombinant expression in *E. coli*. But upon complete protein expression, the $DrCBD_{mon}$P236G variant was able to incorporate BV and had a normal Pr absorbance spectrum identical to $DrCBD_{mon}$. Thus the cis-proline is affecting the kinetics of the protein expression. The P236-245-247G substitution did not express. Thus the number of proline residues is important for knot formation.

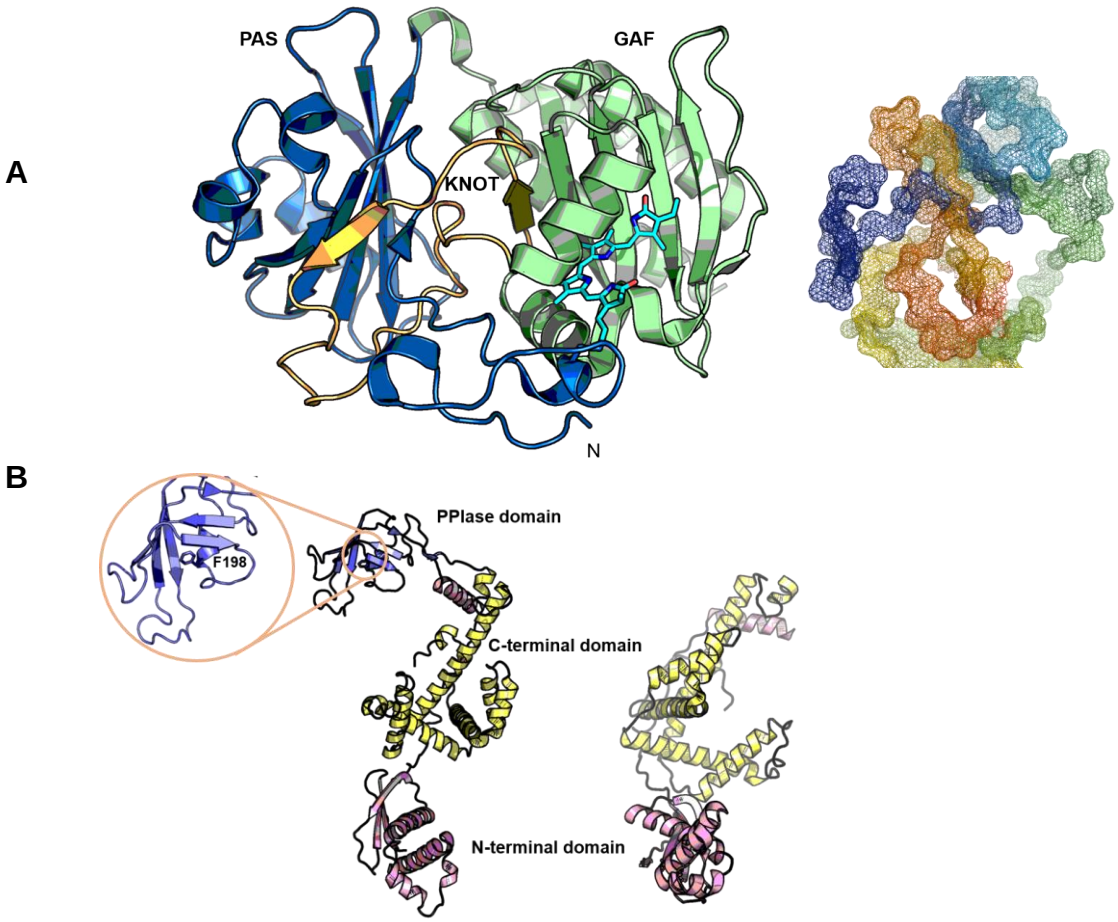
Cis-trans isomerization of peptide bonds are catalyzed by peptidyl-prolyl isomerase enzymes in cells (113). If trans to cis isomerization of the P236 residue is key to folding of the knot, then in absence of peptidyl-prolyl isomerase would negatively impact folding of the knot and thus soluble protein expression. Within cells there are numerous prolyl isomerases, but we decided to investigate the peptidyl-prolyl isomerase activity (PPIase) associated with the general protein chaperone trigger factor because of its low specificity for the sequence around the proline, its ubiquity in all cells, and its ribosome association (114, 115).

Trigger factor is a chaperone with prolyl isomerase activity (116); it has been shown that a F198A mutation in the PPIase domain completely impairs the peptidyl-prolyl isomerase activity but leaves the chaperonin function intact (115). Trigger factor has been found to

interact with nascent polypeptide chains as they exit the ribosome and helps in co-translational protein folding (114, 117), Figure 4.4.1 A. Abolishing the PPIase functionality by either introducing F198A mutation or deleting the entire PPIase domain did not affect the folding of *E. coli* cytosolic proteins and in fact enhances co-translational folding of eukaryotic multidomain proteins expressed in *E. coli* (115, 118)(Figure 4.1B).

We compared folding of native and knotless DrCBD proteins in wild-type *E. coli*, trigger factor mutant (ΔTF), trigger factor with the F198A mutation which impairs PPIase, and intact trigger factor backgrounds. We found increased protein accumulation times (~2 fold higher) in impaired PPIase backgrounds compared to intact trigger factor in $DrCBD_{mon}$ and $DrCBD_{mon}$P236G. There was no difference in knotless protein expression in either background. Taken together the steady state kinetic results suggest that cis-trans proline isomerization is involved in folding of the knotted phytochrome molecule.

Figure 4.1. Structure of DrCBD and trigger factor. A. Structure of DrCBD with PAS domain in blue and GAF in green connected by the figure of eight knot, PDB ID 2O9C. The crossover points displayed is adapted from pKnot server developed by Lai *et al.* published in NAR in 2007 (100). B. Three major domain of trigger factor. The PPIase domain is on the C-terminal end connected to the rest of the molecule with a long linker. Deletion of the entire domain or a F198A mutation retains chaperonin activity of trigger factor while completely abolishing PPiase activity, PDB ID 1W26 (115, 118, 119).



Results

Trigger Factor mutation transferred to *E. coli* BL21(DE3) cell line

The trigger factor deletion in *E. coli* MC4100 background was obtained as genomic DNA from Bernd Bukau's lab in Heidelberg, Germany and the deletion was originally in the *E. coli* MC4100 and linked to kanamycin resistance (115). The deletion was successfully transduced into *E. coli* BL21(DE3) cells for optimized expression of pET vectors, using the P1 phage transduction method (120).

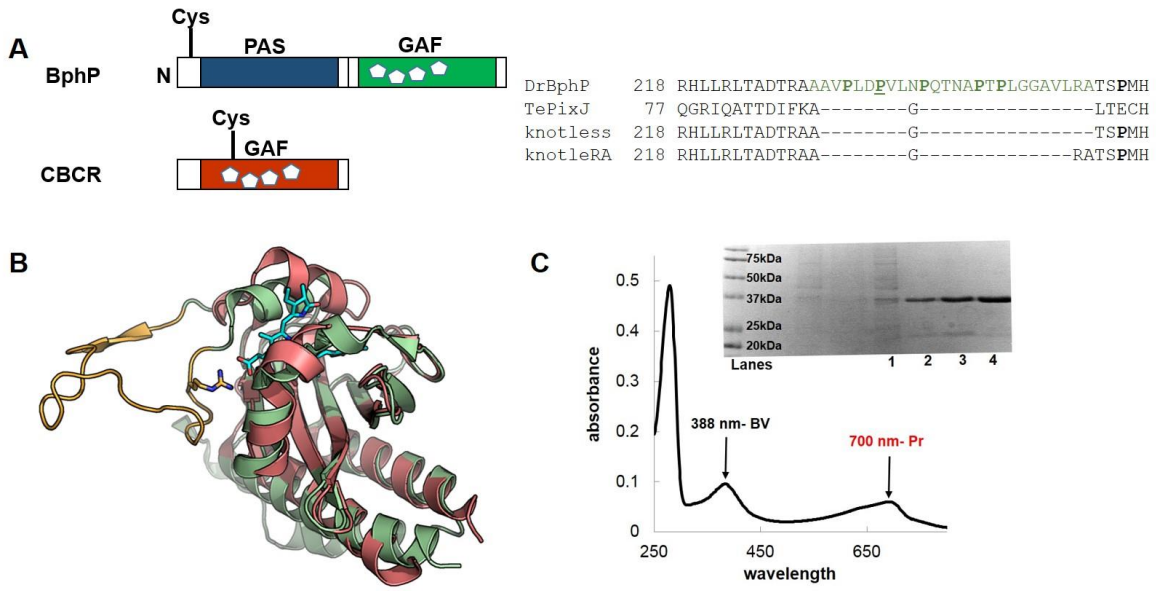
Design of Knotless phytochrome variants

The initial design of the knotless phytochrome was designed on the superposition of the TePixJ ((121), Figure 4.2 A and B). Our first design had a deletion of residues 231-255 which removes Arg254 and its interaction with the BV B ring propionate side chain. The knotless CBD was ordered as a synthetic gene from DNA2.0. I expressed and purified the DrCBDknotless variant recombinantly in *E. coli* expression systems. Optimal expression of DrCBDknotless occurred after 16 hours of incubation post induction with IPTG, which is three times longer than DrCBD knotted variant expression. Additionally, I demonstrated that the DrCBDknotless protein also binds biliverdin, to a limited extent Figure 4.2C).

The original knotless design closely matched TePixJ structure and deleted the entire loop that did not match the tertiary structure which included the R254 residue. Previous

research had shown the importance of the R254 residue in BV incorporation. In various R254 substitutions, BV binding was hindered and in case of a glutamine substitution, completely inhibited (2). We hypothesized that re-insertion of the arginine residue in our DrCBDknotless variant might improve its capacity to bind BV. Using the synthetic DrCBDknotless gene as template, I engineered a DrCBDknotless-RA variant in which the R254 and A255 residues were added to the original deletion. The new construct thus carried a knot deletion from 231-253. This knotless-RA variant behaves identically to our knotless Δ231-255 variant. We could readily express and purify both knotless variants but binding to BV is limited. BV binding ability of the protein was tested by either: a. co-expressing heme oxygenase 1 enzyme or b. exogenous addition of BVIXα. Knotless variants displayed low 388 and 698 nm absorption (Figure 4.2C). Analysis of the absorbance spectra shows that the specific absorbance of knotless variants (absorbance at 700 nm/280 nm) was 10 fold lower than wild-type DrCBD, 10% of total protein was bound to BV but the limited assembly of BV was in the correct orientation. This was a significant improvement since an interim deletion construct of 8 residues 231-239, had no expression in *E. coli* (data not shown).

Figure 4.2. Design and characterization of DrCBDknotless variants. A. CBCR can function as a reversible photoreceptor with only a GAF domain as its photosensory core domain. The Cysteine residue which forms the thioether linkage with the chromophore is part of the GAF domain in CBCR. In proteobacterial phytochromes the cysteine residue is one the N-terminus and the extension from the N-terminus into the GAF domain was deleted to construct the knotless variant. B. Overlay of DrCBD GAF domain (PDB2OC, purple) with TePixJ (pdb, cyan). The region highlighted in green is the lasso loop (231-255) that was deleted. The R254 residue hanging from this extension has been shown to be important for BV binding in DrCBD due to a salt-bridge interaction with the B-ring propionate. C. Absorbance spectra of knotless variant shows absorption at 388 nm (BV) and 698 nm Pr peak. The SDS-PAGE gel shows that DrCBDknotless folds and runs at 35 kDa.

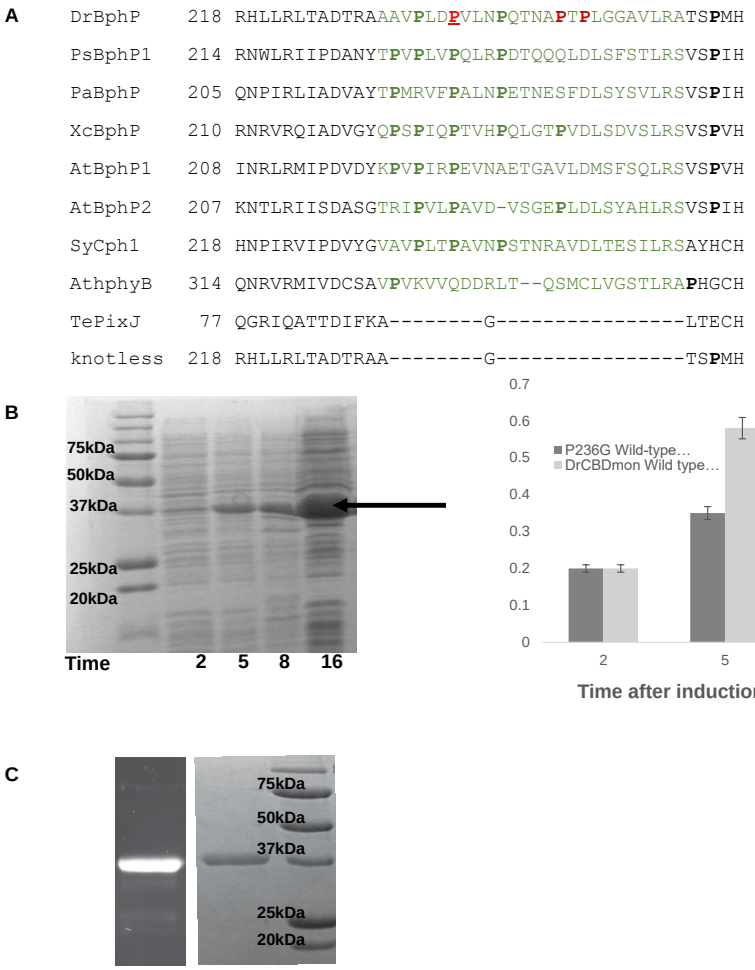


Effect of proline rich lasso loop on folding of Phytochrome

Two percent of the total number of protein structures deposited in the PDB have a knot. Among the set of proteins that are known to be knotted, phytochrome is unique, with a high concentration of Proline residues in the knot region of its primary amino acid sequence (Figure 4.3A). 20% of the residues in the knot region are Prolines and in DrBphP the Pro236 residue exists as a cis-isomer. The cis-proline residue has remained invariant across all the structures of DrCBD that we have solved so far. We hypothesized that the slow step of trans to cis isomerization of proline within this loop is the rate determining step of knot formation.

In order to probe the importance of proline residues in knot formation we created a $DrCBD_{mon}$P236G variant and $DrCBD_{mon}$P236-245-247-G triple variant. Due to the high GC content and complexity of the DrCBD gene, especially in this region, we had to utilize a variety of methods including a modified PCR described by Frey *et al.*, which depends on slowing down ramp rate of the thermocycler (122). The triple variant was created in two fragments which were then assembled via the Gibson method (123). The P236G variant had a slower expression/accumulation time in *E. coli* compared to $DrCBD_{mon}$ (Figure4.3B). However, upon expression and purification $DrCBD_{mon}$P236G was able to incorporate BV and had a Pr absorbance spectrum identical to $DrCBD_{mon}$ (Figure 4.3C). The triple proline to glycine mutant, while successfully cloned, failed to express in *E. coli.*

Figure 4.3. Proline mutant design and characterization. A. Alignment of Lasso Loop region of BphPs, our knotless CBD (see Fig. 2), and a CBCR. Species are *Dr Deinococcus radiodurans; Ps Pseudomonas syringae; Pa Pseudomonas aeruginosa; Xc Xanthamonus campestris; At Agrobacterium tumefaciens; Sy Synechococcus; Ath Arabadopsis thaliana; Te Thermococcus elongatus*. BphP and Phy numbering is based on full-length native proteins; for *Te*PixJ on the GAF domain only. Lasso loop residues are in green. The proline mutations we studied are highlighted in red, P236, P245, P247. B. Expression profile of P236GThe comparative protein yields are displayed in the bar chart. C. P236G is able to bind BV covalently as displayed by the bright Zn signal.



Effect of trigger factor peptidyl-prolyl isomerase activity on folding of phytochrome

Next we tested expression of $DrCBD_{mon}$ and DrCBDP236G variants in ΔTF-*E. coli* BL21(DE3) strain. We observed a global change in protein expression such that the overall expression was lower under identical expression conditions. We compared the overall protein production to our positive control where ΔTF-*E. coli* BL21(DE3) was complemented with a plasmid that carried the TF gene. Overall protein production in ΔTF-*E. coli* BL21(DE3)+TF was similar to wild-type yields. Both $DrCBD_{mon}$ and $DrCBD_{mon}$P236G were overexpressed in absence of TF and in complemented *E. coli* strains. There was a 50% difference in yield of $DrCBD_{mon}$ variants between the 2 strains, with overall yields being much lower in ΔTF (Figure 4.4.A).

Next we probed the effect of the PPiase functionality of the trigger factor on the variants. In this experiment we complemented the ΔTF-*E. coli* BL21(DE3) strain with a plasmid bearing the trigger factor gene containing a mutation that deactivates PPIase activity. Overall protein expression and overexpression of $DrCBD_{mon}$ variants was 25% lower in comparison to complementation of intact trigger factor function, as quantified a=by absorbance 280 measurements. The cells were induced with IPTG for 4 hours for expression testing. Expression tests were carried out in two biological templates of $DrCBD_{mon}$ and $DrCBD_{mon}$P236G in three technical replicates. The overall protein levels in The ΔTF-*E. coli* BL21(DE3)+TF was used as a positive control for protein expression.

In the course of our experiments we also discovered a time dependence in protein production for $DrCBD_{mon}$ and $DrCBD_{mon}$P236G when comparing the three strains, ΔTF,

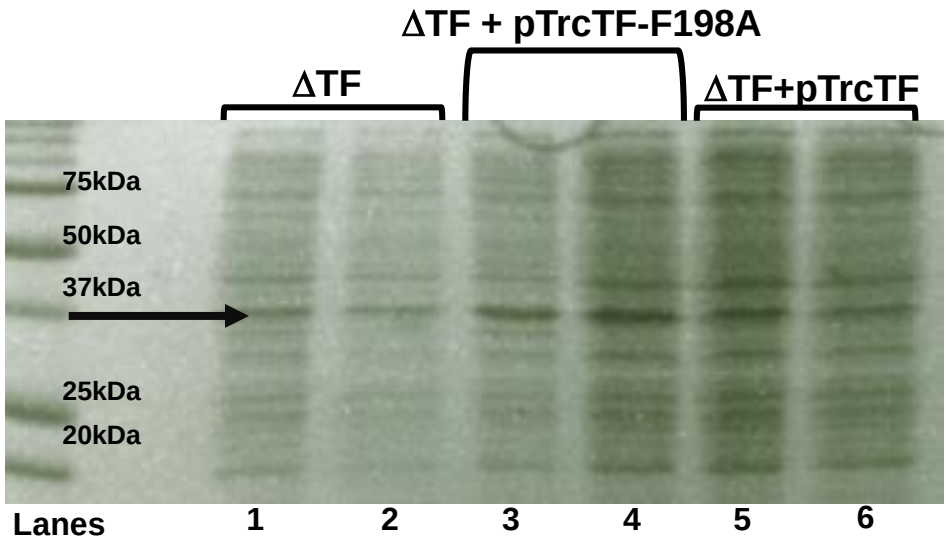
(TF-ΔPPiase) and reverse complemented TF. In both the trigger factor deletion and PPiase impaired strains, maximum protein expression was observed after overnight induction with IPTG at 22°C. Whereas in the strain where full functionality of trigger factor was restored via complementation, optimal protein expression was identical to our standard pipeline, 5 hours of induction at 22°C. The only exception was knotless variants which had a slower accumulation time in wild-type background. Presence or absence of PPIase activity did not make a difference in knotless expression. There were no qualitative differences in protein production. All variants were able to incorporate BV and had their characteristic Pr absorbance spectrum. The time course experiment is summarized in the graphs is Figure 4.4.B.

Figure 4.4. $DrCBD_{mon}$ variants expression in TF mutant strains. A. Expression tests on $DrCBD_{mon}$ and $DrCBD_{mon}$P236G. Lanes 1-6 were expressed in ΔTF, ΔTF complemented with trigger factor with no PPIase functionality and, ΔTF complemented with intact trigger factor. Lanes 1, 3 and 5 were $DrCBD_{mon}$P236G samples and lanes 2, 4 and 6 were $DrCBD_{mon}$. C. Time dependent expression of $DrCBD_{mon}$ and $DrCBD_{mon}$P236G. Protein expression levels were quantified by UV-Vis absorbance in mg/ml. We observed that there was a linear increase in folded protein accumulation which peaked at ~0.7 mg/ml and after the maximum, protein concentrations reached saturation and curve became an asymptote. We used the following equation:

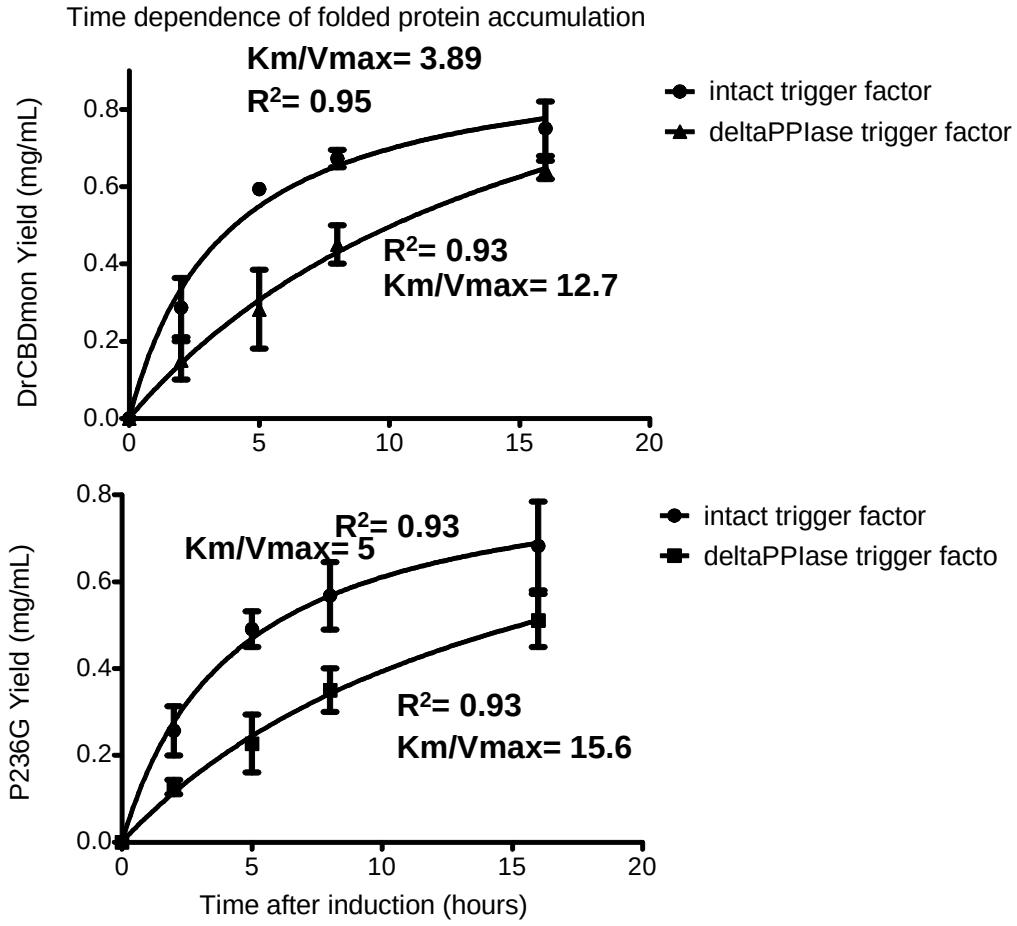
concentration of protein= $\frac{maximum\ protein\ concentration.\ [time\ after\ induction]}{time\ after\ induction\ at \frac{1}{2} of\ Pmax+\ time\ after\ induction}$.

We observed a linear increase in protein production in PPIase impaired conditions. In presence of fully functional trigger factor, protein production plateaued at the 5-hour mark.

A


ΔTF + pTrcTF-F198A
ΔTF
ΔTF+pTrcTF
75kDa
50kDa
37kDa
25kDa
20kDa
Lanes
1
2
3
4
5
6

B


Time dependence of folded protein accumulation
Km/Vmax= 3.89
R²= 0.95
R²= 0.93
Km/Vmax= 12.7
intact trigger factor
deltaPPIase trigger factor
DrCBDmon Yield (mg/mL)
0.0
0.2
0.4
0.6
0.8
0
5
10
15
20
R²= 0.93
Km/Vmax= 5
R²= 0.93
Km/Vmax= 15.6
intact trigger factor
deltaPPIase trigger facto
P236G Yield (mg/mL)
Time after induction (hours)

Materials and methods

P1 phage transduction

The MC4100 strain carrying the trigger factor deletion lined to kanamycin resistance and pTrc plasmids carrying the trigger factor and trigger factor F198A genes were shared with us by Professor Bernd Bukau at Zentrum fur Molekulare Biologie, Heidelburg, Germany. The phage is first grown on MC4100 strain carrying ΔTF deletion, and the resulting phage lysate was used to infect a second recipient *E. coli* BL21 (DE3) strain (120). MC4100 donor strain was grown in 5 ml of LB containing 100 mM $CaCl_2$ for 30 minutes. 100 µl of KL300 phage stock was added to the 5 ml of donor culture and grown at 37°C for 4 hours until lysis. The stock solution was filtered with 0.2 µM filter into a 15 ml falcon tube. For the actual transduction 100 µl of E. coli BL21(DE3) strain was mixed with 100 µl of the phage stock. The resulting mix was grown at 37°C for 20 minutes. The mixture was then transferred to a falcontube and 1 ml of LB and 200 µl of 0.325 M NaPPi was added to the falcon tube and grown at 37°C for 1 hour. 100 µl of the resulting solution was plated on LB+ 50mM Kanamycin+0.325M NaPPi plate. Sequences were verified using Sanger sequencing at the University of Wisconsin-Madison DNA sequencing facility using the following primers:

ΔTF-BL21 FP: 5’ ACAGATGAGATGGTCAGACTAAACT 3’

ΔTF-BL21 RP: 3’ CCGTTTCTGTAAGAAGGAGAAAACT 5’

Cloning

The list of primers is attached at the end of the document

The DrCBDknotless construct was ordered as a synthetic gene from DNA2.0 on pD454 plasmid. DrCBDknotless-RA variant was constructed on the knotless plasmid using the Knotless forward and reverse primers.

We developed a “step-up” PCR procedure to make the variant. The annealing temperature was increased from 60-75 degrees, over 30 cycles using Phusion polymerase (Thermo Scientific).

Due to the high GC content and complexity of the DrCBD gene, we used the aforementioned annealing temperature “step-up” method and combined it with slowing down the ramp rate of the thermocycler to 1.5 °C/sec to construct the $DrCBD_{mon}$P236G mutation using the P236G forward and reverse primers using Pfu turbo enzyme (Agilent technologies).

The P236G template was used to install the P245-247G mutations. Again, site-directed mutagenesis protocols didn’t work thus we used Gibson assembly to for the final P236-245-247G triple mutant. The gene was constructed in two parts. The first fragment was amplified using Gibson 1 and 2 primers. The second gene fragment was made using Gibson 2 and 3 primers.

The gene fragments were assembled using the Gibson assembly protocol originally described by Daniel Gibson and co-workers in 2009 (123) The PCR product was purified and DpnI digested. 7.5µl of Gibson mix and 2.5µl of PCR product was mixed in together

and the mixture was incubated on ice for 5 minutes. The enzyme was heat inactivated and the product transformed in to competent E. coli top10 cells.

The Gibson reaction mix was made in house using the following protocol:

Ingredients	**Amount needed for 120 reactions**
Peg 8000 Biograde (500g)	1.5g
NAD HPLC (250 mg)	20mg
1M Tris HCl pH 7.5 Bio grade (500 ml)	3mL
T5 exonuclease (100U)	6.4U
Phusion HF	40 U
Taq DNA ligase (10,000U)	6,400U

Expression and protein production

The pET21a-DrCBD$_{mon}$-Pro236Gly and pET21a-DrCBD$_{mon}$ were co-transformed into ΔTF *Escherichia coli (E. coli)* BL21 (DE3) (Invitrogen) cells for expression tests. with plasmid to perform *in vivo* protein expression and chromophore assembly. DrCBD$_{mon}$, knotless and P236G variants were co-transformed with either the pTrc TF(F198A) or pTrc (TF).

Protein Expression The constructs bearing DrCBDknotless and DrCBDknotless-RA variants were transformed into BL21 (DE3) expression cells and grown at 37 °C in LB-media in the presence of 0.1mg/ml ampicillin. At an O.D of 0.6, the cells were induced with isopropyl β-D-1-thiogalactopyranoside. The cells were harvested after 2 hours, 5

hours, 8 hours and overnight incubation at 22° C. Cells were centrifugation at 5000 X g and subsequently resuspended in lysis buffer (25mM Tris buffer (pH 8.0), 50 mM NaCl, 5mM imidazole). Cells were sonicated at 45 % amplitude, 10 seconds on and 10 seconds off for 15 minutes. Lysate was separated out by centrifugation at 40000 X g. Cell lysate was run on SDS-PAGE gels to compare protein expression levels.

Cell lysate was incubated with 20µl of 20mM BV in the dark for an hour. Proteins were affinity purified under green light using Ni^{2+} -NTA resin (Qiagen, Valencia, CA) in small batches. The purified protein was subjected to absorption scans to ascertain chromophore binding (93) as well as shifts in maximum absorption of the holoprotein.

UV-vis spectroscopy measurements were performed on a PerkinElmer Life Sciences (Waltham, MA) Lambda 650 spectrophotometer (PerkinElmer Life Sciences). Both UV-vis measurements using molar extinction coefficient of $DrCBD_{mon}$ and ImageJ software to analyze SDS-PAGE gels were used to quantify protein amounts (124).

Discussion

The figure-of-eight knot has been evolutionarily conserved in all PAS containing phytochromes. Such conservation suggests that the BphP knot provides a selective advantage, but the nature of any such advantage has yet to be identified. Previous atomic force measurement study done in collaboration Matthias Rief's laboratory had shown that approximately ~47pN of force is required to unfold apo-phytochrome. The amount of force is comparable to other unknotted proteins; thus the figure-of-eight knot does not confer additional mechanical stability to the protein (82).

Fully functional GAF-only phytochromes from the CBCR family exist in nature, thus the functional aspect of the presence of knot in BphPs has also not been clear. Minimal domain GAF only near-infrared fluorescent protein design has demonstrated that bacterial phytochromes can function and bind BV without the knot (125). In this study we hoped to understand the implications of knot formation in the folding pathway of the protein in three ways, a. by deleting the lasso loop that is not a part of GAF only phytochromes, b. by testing proline mutations in the lasso loop to understand their significance and c. testing our variants in presence and absence of PPIase activity of trigger factor chaperone.

Our designed knotless variant carried a deletion from residues 231 to 255 which closely matched the sequence of GAF only CBCR TePixJ (Figure 4.2A & B). On our way to the 14 residue deletion, we had originally constructed an eight residue deletion on DrCBD. This variant had no expression in *E. coli*. We hypothesized that this could be due to the short loop that prevented the protein from knot formation and folding and removal of the entire loop might have a less deleterious effect. Our designed knotless variant with the

fourteen residue deletion was able to readily express and purify and moreover we detected incorporation of BV. BV binding was inefficient with a low absorbance maximum at 700 nm (~90% lower absorbance than wild-type) (Figure 4.2C). Protein accumulation was much slower than wild-type and optimal expression was seen 16 hours after induction. The 10% of protein that did bind to BV was able to position it correctly in the chromophore binding pocket as evidenced by the Pr absorbance peak at 700 nm (Figure 4.2C). The knotless design included the R254 residue which is involved in a salt-bridge interaction with B-ring propionate of BV and previous mutational analysis of the knot had revealed that absence of R254 had impaired BV binding (2). We engineered arginine back into the DrCBDknotless. DrCBDknotless-RA variant had identical expression and purification conditions compared to DrCBDknotless. There was no change in the absorbance spectrum of the RA version, with a low 388 nm (BV absorbance) and 700 nm (holoprotein absorbance) peaks.

It was the lack of knot that inhibited BV incorporation while the apo-protein could still fold and express. Due to the deletion of the 13 residues preceding the arginine residue, the arginine in the loop was no longer in the correct position or orientation to be in polar contact with BV and thus couldn't stabilize BV interaction in the binding pocket. We conclude that while the knotted protein may be the most stable, phytochrome is able to fold in the absence of the knot but it is stuck in a folding dead end.

The knotless variants were also expressed in ΔTF, ΔTF strains complemented with trigger factor lacking PPiase functionality. Multivariable expression of knotless variants revealed that overnight induction was optimal for maximum protein production. Production of knotless variant was relatively affected by PPiase functionality. We propose that the

peptidyl-prolyl activity of the trigger factor is involved in knot formation, thus in absence of the knot the protein does not need trigger factor for optimal folding. This result agrees with our original hypothesis that trigger factor contributes to BphP knot formation and explains the relative insensitivity of knotless variants to trigger factor.

We also examined how the folding pathway of knotted phytochrome is dependent on characteristics of the proline-rich lasso loop. We proposed that threading of the polypeptide chain through the unstructured proline-rich lasso loop (possibly forming a slip knot intermediate) was the rate determining step for figure-of-eight knot formation, following which proline isomerization, catalyzed by prolyl cis-trans isomerase, promotes collapse of the loop (126–128).

Approximately 20% of the residues in the lasso loop region in bacteriophytochromes are prolines. In particular, the Pro236 residue in all bacteriophytochromes structures solved *Deinococcus radiodurans*, *Rhodopseudomonas palustris* P2 and P3, *Pseudomonas aeruginosa* takes a cis orientation (4, 7, 47, 85, 129–131). We made a Pro236Gly mutant to test if this residue was essential to knot folding. We expected that this variant would cause significant changes to the lasso loop which would affect folding of the figure-of-eight knot and thus have a global effect on protein folding and BV incorporation. Our designed $DrCBD_{mon}$P236G variant showed slower accumulation of folded protein compared to $DrCBD_{mon}$. Upon purification using our standard protein production pipeline was able to bind BV and had a normal Pr absorption spectra (Figure 4.3B & C). This leads us to think that presence of the cis-proline might be relevant for kinetic efficiency of knot formation. Presence of the glycine was sufficient for the polypeptide chain to pass through the lasso loop and still form the knot. It is possible that the proline has been conserve this

position due to the degree of flexibility it allows the lasso loop, in which case substitution with a glycine which is a small and flexible amino acid residue did not affect the structural integrity of the lasso loop and thus the protein was still able to fold and assemble BV Further *in vitro* experiment need to be conducted to verify the hypothesis. The phytochrome B structure from *Arabidiopsis thaliana* has a very different composition of knot residues with prolines in the knot lasso present in the GAF domain. The plant phytochrome knot has four cross-over points and is has a longer extension out of the GAF domain (132). The proline residues are thus applicable to BphPs only.

We also constructed a Pro236-245-247Gly triple substitution on the $DrCBD_{mon}$ background. The proline to glycine triple variant failed to express. This was an interesting discovery as it highlights the concentration effect of the proline residues on knot formation. The number of proline residues are essential to form the loop structure. When we decreased the number of proline from 5 to 2, presence of the glycine is not enough to keep the loop extended for knot formation to happen. This would lead to collapse of the entire knot region and unfolding of the entire molecule.

During testing expression of $DrCBD_{mon}$ and $DrCBD_{mon}$P236G variants in trigger factor and PPIase deficient strains of *E. coli*, we noticed a time variability in folded protein accumulation times. The proteins had low expression levels in ΔTF background which was expected because of the loss of a major chaperone. However, in comparing complementation by TF-PPiase and TF plasmids, we observed slower accumulation of protein in absence of PPIase (Fig 4.4 C). We saw a 2-fold slower protein accumulation in TF-PPIase background for $DrCBD_{mon}$. The protein concentration increased and reached asymptotic levels which were fitted to the curve using the following equation:

concentration of protein= $\frac{maximum\ protein\ concentration.\ [time\ after\ induction]}{time\ after\ induction\ at\frac{1}{2}of\ Pmax+\ time\ after\ induction}$.

In TF-PPIase background, the curve never plateaued. This result suggests that absence of trigger factor peptidyl-prolyl isomerase activity has a kinetic effect on protein folding. It is possible in absence of the trigger factor another peptidyl-prolyl isomerase is being recruited by the ribosome to fold apo-phytochrome. The slower expression of P236G variant coupled with its ability to fold and assemble BV with a Pr spectrum, suggests that the role of the proline has to do with the kinetics of knot formation. Our time dependent analysis of protein production shows that production of phytochrome protein slowed down in absence of PPIase activity in ΔTF *E. coli* cells. The formation of the knot requires the chaperone as well as PPIase functionality of Trigger Factor. The process of folding a knot is complex and possibly involves multiple chaperones. We need to explore the possibility that another chaperone, in particular the GroEL-GroES chaperonin, may be a key player in phytochrome folding.

References

1. Jobsis, F. F. (1977) Noninvasive, infrared monitoring of cerebral and myocardial oxygen sufficiency and circulatory parameters. *Science*. **198**, 1264–1267

2. Wagner, J. R., Zhang, J., von Stetten, D., Günther, M., Murgida, D. H., Mroginski, M. A., Walker, J. M., Forest, K. T., Hildebrandt, P., and Vierstra, R. D. (2008) Mutational analysis of Deinococcus radiodurans bacteriophytochrome reveals key amino acids necessary for the photochromicity and proton exchange cycle of phytochromes. *J. Biol. Chem.* **283**, 12212–12226

3. Shu, X., Royant, A., Lin, M. Z., Aguilera, T. A., Lev-Ram, V., Steinbach, P. A., and Tsien, R. Y. (2009) Mammalian expression of infrared fluorescent proteins engineered from a bacterial phytochrome. *Science*. **324**, 804–807

4. Auldridge, M. E., Satyshur, K. a, Anstrom, D. M., and Forest, K. T. (2012) Structure-guided engineering enhances a phytochrome-based infrared fluorescent protein. *J. Biol. Chem.* **287**, 7000–7009

5. Filonov, G. S., Piatkevich, K. D., Ting, L.-M., Zhang, J., Kim, K., and Verkhusha, V. V (2011) Bright and stable near-infrared fluorescent protein for in vivo imaging. *Nat. Biotechnol.* **29**, 757–761

6. Shcherbakova, D. M., and Verkhusha, V. V. (2013) Near-infrared fluorescent proteins for multicolor in vivo imaging. *Nat. Methods*. **10**, 751–754

7. Bhattacharya, S., Auldridge, M. E., Lehtivuori, H., Ihalainen, J. A., and Forest, K. T.

(2014) Origins of Fluorescence in Evolved Bacteriophytochromes. *J. Biol. Chem.* **289**, 32144–32152

8. Lehtivuori, H., Bhattacharya, S., Angenent-Mari, N. M., Satyshur, K. A., and Forest, K. T. (2015) Removal of Chromophore-Proximal Polar Atoms Decreases Water Content and Increases Fluorescence in a Near Infrared Phytofluor. *Front. Mol. Biosci.* **2**, 65–69

9. Stepanenko, O. V., Verkhusha, V. V., Kuznetsova, I. M., Uversky, V. N., and Turoverov, K. K. (2008) Fluorescent proteins as biomarkers and biosensors: throwing color lights on molecular and cellular processes. *Curr. Protein Pept. Sci.* **9**, 338–369

10. Crameri, A., Whitehorn, E. A., Tate, E., and Stemmer, W. P. (1996) Improved green fluorescent protein by molecular evolution using DNA shuffling. *Nat Biotechnol*. **14**, 315–319

11. Yoshihara, S., Shimada, T., Matsuoka, D., Zikihara, K., Kohchi, T., and Tokutomi, S. (2006) Reconstitution of Blue−Green Reversible Photoconversion of a Cyanobacterial Photoreceptor, PixJ1, in Phycocyanobilin-Producing Escherichia coli. *Biochemistry*. **45**, 3775–3784

12. Tsien, R. (1998) The green fluorescent protein. *Annu. Rev. Biochem.* **67**, 509

13. Ando, R., Hama, H., Yamamoto-Hino, M., Mizuno, H., and Miyawaki, A. (2002) An optical marker based on the UV-induced green-to-red photoconversion of a fluorescent protein. *Proc. Natl. Acad. Sci.* **99**, 12651–12656

14. Ehrig, T., O'Kane, D. J., and Prendergast, F. G. (1995) Green-fluorescent protein mutants with altered fluorescence excitation spectra. *FEBS Lett*. **367**, 163–166

15. Reeder, P. J., Huang, Y.-M., Dordick, J. S., and Bystroff, C. (2010) A Rewired Green Fluorescent Protein: Folding and Function in a Nonsequential, Noncircular GFP Permutant. *Biochemistry*. **49**, 10773–10779

16. Hanson, G. T., McAnaney, T. B., Park, E. S., Rendell, M. E. P., Yarbrough, D. K., Chu, S., Xi, L., Boxer, S. G., Montrose, M. H., and Remington, S. J. (2002) Green Fluorescent Protein Variants as Ratiometric Dual Emission pH Sensors. 1. Structural Characterization and Preliminary Application. *Biochemistry*. **41**, 15477–15488

17. Davis, S. J., and Vierstra, R. D. (1998) Soluble, highly fluorescent variants of green fluorescent protein (GFP) for use in higher plants. *Plant Mol. Biol.* **36**, 521–528

18. Baulcombe, D. C., Chapman, S., and Cruz, S. (1995) Jellyfish green fluorescent protein as a reporter for virus infections. *Plant J.* **7**, 1045–1053

19. Peng, S.-Y., Yang, Y.-S., Chou, C.-J., Lin, K.-Y., and Wu, S.-C. (2015) Differentiation of Enhanced Green Fluorescent Protein-Labeled Mouse Amniotic Fluid-Derived Stem Cells into Cardiomyocyte-Like Beating Cells. *Acta Cardiol. Sin.* **31**, 209–214

20. Boas, D. A., Gaudette, T., Strangman, G., Cheng, X., Marota, J. J. A., and Mandeville, J. B. (2001) The Accuracy of Near Infrared Spectroscopy and Imaging during Focal Changes in Cerebral Hemodynamics. *Neuroimage*. **13**, 76–90

21. Mancini, D. M., Bolinger, L., Li, H., Kendrick, K., Chance, B., and Wilson, J. R. (1994) Validation of near-infrared spectroscopy in humans. *J. Appl. Physiol.* **77**, 2740–2747

22. Hilderbrand, S. A., and Weissleder, R. (2010) Near-infrared fluorescence: application to in vivo molecular imaging. *Curr. Opin. Chem. Biol.* **14**, 71–79

23. Luo, S., Zhang, E., Su, Y., Cheng, T., and Shi, C. (2011) A review of NIR dyes in cancer targeting and imaging. *Biomaterials*. **32**, 7127–7138

24. Ke, S., Wen, X., Gurfinkel, M., Charnsangavej, C., Wallace, S., Sevick-Muraca, E. M., and Li, C. (2003) Near-Infrared Optical Imaging of Epidermal Growth Factor Receptor in Breast Cancer Xenografts. *Cancer Res.* **63**, 7870–7875

25. Bugaj, J. E., Achilefu, S., Dorshow, R. B., and Rajagopalan, R. (2001) Novel fluorescent contrast agents for optical imaging of in vivo tumors based on a receptor-targeted dye-peptide conjugate platform. *J. Biomed. Opt.* **6**, 122–132

26. Achilefu, S., Dorshow, R. B., Bugaj, J. E., and Rajagopalan, R. (2000) Novel receptor-targeted fluorescent contrast agents for in vivo tumor imaging. *Invest. Radiol.* **35**, 479–485

27. Bongarzone, S., Staderini, M., and Bolognesi, M. L. (2014) Multitarget ligands and theranostics: sharpening the medicinal chemistry sword against prion diseases. *Future Med. Chem.* **6**, 1017–1029

28. Doerr, A. (2009) Fluorescent proteins: into the infrared. *Nat. Methods*. **6**, 482–483

29. Cashmore, A. R. (2003) Cryptochromes: Enabling plants and animals to determine

circadian time. *Cell*. **114**, 537–543

30. Murphy, J. T., and Lagarias, J. C. (1997) The phytofluors: a new class of fluorescent protein probes. *Curr. Biol.* **7**, 870–876

31. Christie, J. M., Gawthorne, J., Young, G., Fraser, N. J., and Roe, A. J. (2012) LOV to BLUF: Flavoprotein Contributions to the Optogenetic Toolkit. *Mol. Plant*. **5**, 533–544

32. Pudasaini, A., El-Arab, K. K., and Zoltowski, B. D. (2015) LOV-based optogenetic devices: light-driven modules to impart photoregulated control of cellular signaling. *Front. Mol. Biosci.* **2**, 18–22

33. Möglich, A., Yang, X., Ayers, R. A., and Moffat, K. (2010) Structure and Function of Plant Photoreceptors. *Annu. Rev. Plant Biol.* **61**, 21–47

34. Rockwell, N. C., and Lagarias, J. C. (2010) A Brief History of Phytochromes. *ChemPhysChem*. **11**, 1172–1180

35. Auldridge, M. E., and Forest, K. T. (2011) Bacterial phytochromes: more than meets the light. *Crit. Rev. Biochem. Mol. Biol.* **46**, 67–88

36. Al-Sady, B., Ni, W., Kircher, S., Schaefer, E., and Quail, P. H. (2006) Photoactivated Phytochrome Induces Rapid PIF3 Phosphorylation Prior to Proteasome-Mediated Degradation. *Mol. Cell*. **23**, 439–446

37. Fankhauser, C. (2001) The Phytochromes, a Family of Red/Far-red Absorbing Photoreceptors. *J. Biol. Chem.* **276**, 11453–11456

38. Rizzini, L., Favory, J.-J., Cloix, C., Faggionato, D., O'Hara, A., Kaiserli, E.,

Baumeister, R., Schäfer, E., Nagy, F., Jenkins, G. I., and Ulm, R. (2011) Perception of UV-B by the Arabidopsis UVR8 protein. *Science*. **332**, 103–106

39. Quail, P. H. (2002) Phytochrome photosensory signalling networks. *Nat. Rev. Mol. Cell Biol.* **3**, 85–93

40. Kehoe, D. M., and Grossman, A. R. (1996) Similarity of a chromatic adaptation sensor to phytochrome and ethylene receptors. *Science*. **273**, 1409–1412

41. Terauchi, K., Montgomery, B. L., Grossman, A. R., Lagarias, J. C., and Kehoe, D. M. (2004) RcaE is a complementary chromatic adaptation photoreceptor required for green and red light responsiveness. *Mol. Microbiol.* **51**, 567–577

42. Davis, S. J., Vener, A. V, and Vierstra, R. D. (1999) Bacteriophytochromes: phytochrome-like photoreceptors from nonphotosynthetic eubacteria. *Science*. **286**, 2517–2520

43. Bhoo, S. H., Davis, S. J., Walker, J., Karniol, B., and Vierstra, R. D. (2001) Bacteriophytochromes are photochromic histidine kinases using a biliverdin chromophore. *Nature*. **414**, 776–779

44. Butler, W. L., Norris, K. H., Siegelman, H. W., and Hendricks, S. B. (1959) Detection, assay, and preliminary purification of the pigment controlling photoresponsive development of plants. *Proc. Natl. Acad. Sci. U. S. A.* **45**, 1703–1708

45. Blaauw-Jansen, G. (1959) The influence of red and far red light on growth and phototropism of the avena seedling. *Acta Bot. Neerl.* **8**, 1–39

46. Wagner, J. R., Brunzelle, J. S., Forest, K. T., and Vierstra, R. D. (2005) A light-sensing knot revealed by the structure of the chromophore-binding domain of phytochrome. *Nature*. **438**, 325–331

47. Wagner, J. R., Zhang, J., Brunzelle, J. S., Vierstra, R. D., and Forest, K. T. (2007) High resolution structure of Deinococcus bacteriophytochrome yields new insights into phytochrome architecture and evolution. *J. Biol. Chem.* **282**, 12298–12309

48. Yang, X., Stojkovic, E. A., Kuk, J., and Moffat, K. (2007) Crystal structure of the chromophore binding domain of an unusual bacteriophytochrome, RpBphP3, reveals residues that modulate photoconversion. *Proc. Natl. Acad. Sci.* **104**, 12571–12576

49. Nagatani, A. (2010) Phytochrome: structural basis for its functions. *Curr. Opin. Plant Biol.* **13**, 565–570

50. Takala, H., Björling, A., Berntsson, O., Lehtivuori, H., Niebling, S., Hoernke, M., Kosheleva, I., Henning, R., Menzel, A., Ihalainen, J. A., and Westenhoff, S. (2014) Signal amplification and transduction in phytochrome photosensors. *Nature*. **509**, 245–248

51. Baker, A. W., and Forest, K. T. (2014) Structural biology: Action at a distance in a light receptor. *Nature*. **509**, 174–175

52. Burgie, E. S., Wang, T., Bussell, A. N., Walker, J. M., Li, H., and Vierstra, R. D. (2014) Crystallographic and Electron Microscopic Analyses of a Bacterial Phytochrome Reveal Local and Global Rearrangements During Photoconversion. *J. Biol. Chem.* **289**, 24573–24587

53. Yang, X., Kuk, J., and Moffat, K. (2009) Conformational differences between the Pfr and Pr states in Pseudomonas aeruginosa bacteriophytochrome. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 15639–15644

54. Bischoff, M., Hermann, G., Rentsch, S., and Strehlow, D. (2001) First Steps in the Phytochrome Phototransformation: A Comparative Femtosecond Study on the Forward (Pr → Pfr) and Back Reaction (Pfr → Pr). *Biochemistry*. **40**, 181–186

55. Andel, F., Hasson, K. C., Gai, F., Anfinrud, P. A., and Mathies, R. A. (1997) Femtosecond time-resolved spectroscopy of the primary photochemistry of phytochrome. *Biospectroscopy*. **3**, 421–433

56. Toh, K. C., Stojkovic, E. A., van Stokkum, I. H., Moffat, K., and Kennis, J. T. (2011) Fluorescence quantum yield and photochemistry of bacteriophytochrome constructs. *Phys Chem Chem Phys*. **13**, 11985–11997

57. Dasgupta, J., Frontiera, R. R., Taylor, K. C., Lagarias, J. C., and Mathies, R. A. (2009) Ultrafast excited-state isomerization in phytochrome revealed by femtosecond stimulated Raman spectroscopy. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 1784–1789

58. Piatkevich, K. D., Subach, F. V, and Verkhusha, V. V. (2013) Engineering of bacterial phytochromes for near-infrared imaging, sensing, and light-control in mammals. *Chem. Soc. Rev.* **42**, 3441–3452

59. Yang, X., Stojković, E. A., Ozarowski, W. B., Kuk, J., Davydova, E., and Moffat, K. (2015) Light Signaling Mechanism of Two Tandem Bacteriophytochromes. *Structure*. **23**, 1179–1189

60. Rockwell, N. C., Su, Y.-S., and Lagarias, J. C. (2006) Phytochrome structure and signaling mechanisms. *Annu. Rev. Plant Biol.* **57**, 837–858

61. Kim, P. W., Freer, L. H., Rockwell, N. C., Martin, S. S., Lagarias, J. C., and Larsen, D. S. (2012) Femtosecond Photodynamics of the Red/Green Cyanobacteriochrome NpR6012g4 from *Nostoc punctiforme* . 2. Reverse Dynamics. *Biochemistry*. **51**, 619–630

62. Ulijasz, A. T., Cornilescu, G., Cornilescu, C. C., Zhang, J., Rivera, M., Markley, J. L., and Vierstra, R. D. (2010) Structural basis for the photoconversion of a phytochrome to the activated Pfr form. *Nature*. **463**, 250–254

63. Kneip, C., Hildebrandt, P., Schlamann, W., Braslavsky, S. E., Mark, F., and Schaffner, K. (1999) Protonation state and structural changes of the tetrapyrrole chromophore during the Pr/Pfr phototransformation of phytochrome: a resonance Raman spectroscopic study. *Biochemistry*. **38**, 15185–15192

64. Spillane, K. M., Dasgupta, J., and Mathies, R. A. (2012) Conformational Homogeneity and Excited-State Isomerization Dynamics of the Bilin Chromophore in Phytochrome Cph1 from Resonance Raman Intensities. *Biophys. J.* **102**, 709–717

65. Toh, K. C., Stojkovic, E. A., van Stokkum, I. H. M., Moffat, K., and Kennis, J. T. M. (2010) Proton-transfer and hydrogen-bond interactions determine fluorescence quantum yield and photochemical efficiency of bacteriophytochrome. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 9170–9175

66. Cormack, B. P., Valdivia, R. H., and Falkow, S. (1996) FACS-optimized mutants of

the green fluorescent protein (GFP). *Gene*. **173**, 33–38

67. Heim, R., and Tsien, R. Y. (1996) Engineering green fluorescent protein for improved brightness, longer wavelengths and fluorescence resonance energy transfer. *Curr Biol*. **6**, 178–182

68. Afar, B., Merrill, J., and Clark, E. A. (1991) Detection of lymphocyte subsets using three-color/single-laser flow cytometry and the fluorescent dye Peridinin chlorophyll-a protein. *J. Clin. Immunol.* **11**, 254–261

69. Nakajima, O., and Gray, C. H. (1967) Studies on haem alpha-methenyl oxygenase. Isomeric structure of formylbiliverdin, a possible precursor of biliverdin. *Biochem. J.* **104**, 20–22

70. Schäfer, F. P. (1973) Principles of dye laser operation. in *n Topics in Applied Physics Dye Lasers* (Schäfer, F. P. ed), pp. 1–89, Springer-Verlag, New York, 10.1007/3-540-51558-5_7

71. Filonov, G. S., Krumholz, A., Xia, J., Yao, J., Wang, L. V, and Verkhusha, V. V. (2012) Deep-tissue photoacoustic tomography of a genetically encoded near-infrared fluorescent probe. *Angew Chem Int Ed Engl*. **51**, 1448–1451

72. Bartz-Schmidt, K. U., Walter, P., Krott, R., Brunner, R., Esser, P., and Heimann, K. (1996) [Effects of fluorescein and indocyanine green angiography on subsequent dark adaptation and the electroretinogram]. *Klin Monbl Augenheilkd*. **208**, 224–228

73. Taylor, W. R. (2000) A deeply knotted protein structure and how it might fold. *Nature*. **406**, 916–919

74. Khatib, F., Weirauch, M. T., and Rohl, C. A. (2006) Rapid knot detection and application to protein structure prediction. *Bioinformatics*. **22**, e252–e259

75. Mallam, A. L., Morris, E. R., and Jackson, S. E. (2008) Exploring knotting mechanisms in protein folding. *Proc. Natl. Acad. Sci.* **105**, 18740–18745

76. Mallam, A. L., and Jackson, S. E. (2006) Probing Nature's Knots: The Folding Pathway of a Knotted Homodimeric Protein. *J. Mol. Biol.* **359**, 1420–1436

77. Mallam, A. L., Rogers, J. M., and Jackson, S. E. (2010) Experimental detection of knotted conformations in denatured proteins. *Proc Natl Acad Sci U S A*. **107**, 8189–8194

78. King, N. P., Yeates, E. O., and Yeates, T. O. (2007) Identification of Rare Slipknots in Proteins and Their Implications for Stability and Folding. **373**, 153–166

79. King, N. P., Jacobitz, A. W., Sawaya, M. R., Goldschmidt, L., and Yeates, T. O. (2010) Structure and folding of a designed knotted protein. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 20732–20737

80. Dill, K. A., and MacCallum, J. L. (2012) The protein-folding problem, 50 years on. **338**, 1042–1046

81. Andersson, F. I., Pina, D. G., Mallam, A. L., Blaser, G., and Jackson, S. E. (2009) Untangling the folding mechanism of the 52-knotted protein UCH-L3. **276**, 2625–2635

82. Bornschlogl, T., Anstrom, D. M., Mey, E., Dzubiella, J., Rief, M., and Forest, K. T. (2009) Tightening the knot in phytochrome by single-molecule atomic force

microscopy. *Biophys J.* **96**, 1508–1514

83. von Stetten, D., Seibeck, S., Michael, N., Scheerer, P., Mroginski, M. A., Murgida, D. H., Krauss, N., Heyn, M. P., Hildebrandt, P., Borucki, B., and Lamparter, T. (2007) Highly conserved residues Asp-197 and His-250 in Agp1 phytochrome control the proton affinity of the chromophore and Pfr formation. *J. Biol. Chem.* **282**, 2116–2123

84. Anders Borg, O., and Durbeej, B. (2008) Which factors determine the acidity of the phytochromobilin chromophore of plant phytochrome? *Phys. Chem. Chem. Phys.* **10**, 2528–2537

85. Lehtivuori, H., Bhattacharya, S., Angenent-Mari, N., Satyshur, K. A., and Forest, K. T. (2015) Removal of Chromophore-Proximal Polar Atoms Decreases Water Content and Increases Fluorescence in a Near Infrared Phytofluor. *Front. Mol. Biosci.* **2**, 65–69

86. Spillane, K. M., Dasgupta, J., Lagarias, J. C., and Mathies, R. A. (2009) Homogeneity of phytochrome Cph1 vibronic absorption revealed by resonance Raman intensity analysis. *J Am Chem Soc*. **131**, 13946–13948

87. Essen, L.-O., Mailliet, J., and Hughes, J. (2008) The structure of a complete phytochrome sensory module in the Pr ground state. *Proc. Natl. Acad. Sci.* **105**, 14709–14714

88. Yang, X., Kuk, J., and Moffat, K. (2008) Crystal structure of Pseudomonas aeruginosa bacteriophytochrome: photoconversion and signal transduction. *Proc. Natl. Acad. Sci.* **105**, 14715–14720

89. von Stetten, D., Gunther, M., Scheerer, P., Murgida, D. H., Mroginski, M. A., Krauss, N., Lamparter, T., Zhang, J., Anstrom, D. M., Vierstra, R. D., Forest, K. T., and Hildebrandt, P. (2008) Chromophore heterogeneity and photoconversion in phytochrome crystals and solution studied by resonance Raman spectroscopy. *Angew Chem Int Ed Engl*. **47**, 4753–4755

90. Dasgupta, J., Frontiera, R. R., Taylor, K. C., Lagarias, J. C., and Mathies, R. A. (2009) Ultrafast excited-state isomerization in phytochrome revealed by femtosecond stimulated Raman spectroscopy. *Proc. Natl. Acad. Sci.* **106**, 1784–1789

91. Foerstendorf, H., Mummert, E., Sch??fer, E., Scheer, H., and Siebert, F. (1996) Fourier-transform infrared spectroscopy of phytochrome: Difference spectra of the intermediates of the photoreactions. *Biochemistry*. **35**, 10793–10799

92. Toh, K. C., van Stokkum, I. H., Hendriks, J., Alexandre, M. T., Arents, J. C., Perez, M. A., van Grondelle, R., Hellingwerf, K. J., and Kennis, J. T. (2008) On the signaling mechanism and the absence of photoreversibility in the AppA BLUF domain. *Biophys J*. **95**, 312–321

93. Fischer, A. J., and Lagarias, J. C. (2004) Harnessing phytochrome's glowing potential. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 17334–17339

94. Lehtivuori, H., Rissanen, I., Takala, H., Bamford, J., Tkachenko, N. V., and Ihalainen, J. A. (2013) Fluorescence Properties of the Chromophore-Binding Domain of Bacteriophytochrome from *Deinococcus radiodurans*. *J. Phys. Chem. B*. **117**, 11049–11057

95. Otwinowski, Z., and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326

96. McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007) Phaser crystallographic software. *J Appl Crystallogr*. **40**, 658–674

97. Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F., and Vagin, A. A. (2011) REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. D. Biol. Crystallogr.* **67**, 355–367

98. Hahn, J., Strauss, H. M., Landgraf, F. T., Gimenèz, H. F., Lochnit, G., Schmieder, P., and Hughes, J. (2006) Probing protein-chromophore interactions in Cph1 phytochrome by mutagenesis. *FEBS J.* **273**, 1415–1429

99. Jamroz, M., Niemyska, W., Rawdon, E. J., Stasiak, A., Millett, K. C., Sułkowski, P., and Sulkowska, J. I. (2015) KnotProt: a database of proteins with knots and slipknots. *Nucleic Acids Res.* **43**, D306–D314

100. Lai, Y.-L., Yen, S.-C., Yu, S.-H., and Hwang, J.-K. (2007) pKNOT: the protein KNOT web server. *Nucleic Acids Res.* **35**, W420–W424

101. Millett, K. C., Rawdon, E. J., Stasiak, A., and Sułkowska, J. I. (2013) Identifying knots in proteins. *Biochem. Soc. Trans.* **41**, 533–537

102. Burgie, E. S., and Vierstra, R. D. (2014) Phytochromes: An Atomic Perspective on Photoactivation and Signaling. *Plant Cell Online*. **26**, 4568–4583

103. Rockwell, N. C., Martin, S. S., Gulevich, A. G., and Lagarias, J. C. (2012) Phycoviolobilin Formation and Spectral Tuning in the DXCF Cyanobacteriochrome Subfamily. *Biochemistry*. **51**, 1449–1463

104. Sineshchekov, V., Mailliet, J., Psakis, G., Feilke, K., Kopycki, J., Zeidler, M., Essen, L.-O., and Hughes, J. (2014) Tyrosine 263 in cyanobacterial phytochrome Cph1 optimizes photochemistry at the prelumi-R→lumi-R step. *Photochem. Photobiol.* **90**, 786–795

105. Narikawa, R., Fukushima, Y., Ishizuka, T., Itoh, S., and Ikeuchi, M. (2008) A Novel Photoactive GAF Domain of Cyanobacteriochrome AnPixJ That Shows Reversible Green/Red Photoconversion. *J. Mol. Biol.* **380**, 844–855

106. Narikawa, R., Ishizuka, T., Muraki, N., Shiba, T., Kurisu, G., and Ikeuchi, M. (2013) Structures of cyanobacteriochromes from phototaxis regulators AnPixJ and TePixJ reveal general and specific photoconversion mechanism. *Proc. Natl. Acad. Sci.* **110**, 918–923

107. Narikawa, R., Nakajima, T., Aono, Y., Fushimi, K., Enomoto, G., Ni-Ni-Win, Itoh, S., Sato, M., Ikeuchi, M., Narikawa, R., Enomoto, G., Ni, W. N., Fushimi, K., and Ikeuchi, M. (2015) A biliverdin-binding cyanobacteriochrome from the chlorophyll d–bearing cyanobacterium Acaryochloris marina. *Sci. Rep.* **5**, 7950–7975

108. Zhang, J., Wu, X.-J., Wang, Z.-B., Chen, Y., Wang, X., Zhou, M., Scheer, H., and Zhao, K.-H. (2010) Fused-Gene Approach to Photoswitchable and Fluorescent Biliproteins. *Angew. Chemie Int. Ed.* **49**, 5456–5458

109. Rockwell, N. C., Martin, S. S., and Lagarias, J. C. (2012) Red/Green

Cyanobacteriochromes: Sensors of Color and Power. *Biochemistry*. **51**, 9667–9677

110. Pal, D., and Chakrabarti, P. (1999) Cis peptide bonds in proteins: residues involved, their conformations, interactions and locations. *J. Mol. Biol.* **294**, 271–288

111. Joseph, A. P., Srinivasan, N., and de Brevern, A. G. (2012) Cis-trans peptide variations in structurally similar proteins. *Amino Acids*. **43**, 1369–1381

112. MacArthur, M. W., and Thornton, J. M. (1991) Influence of proline residues on protein conformation. *J. Mol. Biol.* **218**, 397–412

113. Fischer, G., and Schmid, F. X. (1990) The mechanism of protein folding. Implications of in vitro refolding models for de novo protein folding and translocation in the cell. *Biochemistry*. **29**, 2205–2212

114. Hesterkamp, T., and Bukau, B. (1996) Identification of the prolyl isomerase domain of Escherichia coli trigger factor. *FEBS Lett.* **385**, 67–71

115. Kramer, G., Patzelt, H., Rauch, T., Kurz, T. A., Vorderwülbecke, S., Bukau, B., and Deuerling, E. (2004) Trigger Factor Peptidyl-prolyl cis/trans Isomerase Activity Is Not Essential for the Folding of Cytosolic Proteins in Escherichia coli. *J. Biol. Chem.* **279**, 14165–14170

116. Stoller, G., Rücknagel, K. P., Nierhaus, K. H., Schmid, F. X., Fischer, G., and Rahfeld, J. U. (1995) A ribosome-associated peptidyl-prolyl cis/trans isomerase identified as the trigger factor. *EMBO J.* **14**, 4939–4948

117. Kramer, G., Rauch, T., Rist, W., Vorderwülbecke, S., Patzelt, H., Schulze-Specking, A., Ban, N., Deuerling, E., and Bukau, B. (2002) L23 protein functions as

a chaperone docking site on the ribosome. *Nature*. **419**, 171–174

118. Gupta, R., Lakshmipathy, S. K., Chang, H.-C., Etchells, S. A., and Hartl, F. U. (2010) Trigger factor lacking the PPIase domain can enhance the folding of eukaryotic multi-domain proteins in Escherichia coli. *FEBS Lett.* **584**, 3620–3624

119. Ferbitz, L., Maier, T., Patzelt, H., Bukau, B., Deuerling, E., and Ban, N. (2004) Trigger factor in complex with the ribosome forms a molecular cradle for nascent proteins. *Nature*. **431**, 590–596

120. Thomason, L. C., Costantino, N., Court, D. L., Thomason, L. C., Costantino, N., and Court, D. L. (2007) *E. coli* Genome Manipulation by P1 Transduction. in *Current Protocols in Molecular Biology*, pp. 1.17.1–1.17.8, John Wiley & Sons, Inc., Hoboken, NJ, USA, 10.1002/0471142727.mb0117s79

121. Burgie, E. S., Walker, J. M., Phillips, G. N., and Vierstra, R. D. (2013) A Photo-Labile Thioether Linkage to Phycoviolobilin Provides the Foundation for the Blue/Green Photocycles in DXCF-Cyanobacteriochromes. *Structure*. **21**, 88–97

122. Frey, U. H., Bachmann, H. S., Peters, J., and Siffert, W. (2008) PCR-amplification of GC-rich regions: “slowdown PCR.” *Nat. Protoc.* **3**, 1312–1317

123. Gibson, D. G., Young, L., Chuang, R.-Y., Venter, J. C., Hutchison, C. A., and Smith, H. O. (2009) Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat Meth*. **6**, 343–345

124. Schneider, C. A., Rasband, W. S., and Eliceiri, K. W. (2012) NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods*. **9**, 671–675

125. Rumyantsev, K. A., Shcherbakova, D. M., Zakharova, N. I., Emelyanov, A. V., Turoverov, K. K., and Verkhusha, V. V. (2015) Minimal domain of bacterial phytochrome required for chromophore binding and fluorescence. *Sci. Rep.* **5**, 18348–18352

126. Mallam, A. L. (2009) How does a knotted protein fold? *FEBS J.* **276**, 365–375

127. Sułkowska, J. I., Noel, J. K., Ramírez-Sarmiento, Esar A, Rawdon, E. J., Millett, K. C., and Onuchic, J. N. (2013) Knotting pathways in proteins. *Biochem. Soc. Trans.* **41**, 523–527

128. Mallam, A. L., and Jackson, S. E. (2011) Knot formation in newly translated proteins is spontaneous and accelerated by chaperonins. *Nat. Chem. Biol.* **8**, 147–153

129. Yang, X., Stojković, E. A., Ozarowski, W. B., Kuk, J., Davydova, E., and Moffat, K. (2015) Light Signaling Mechanism of Two Tandem Bacteriophytochromes. *Structure*. **23**, 1179–1189

130. Yang, X., Ren, Z., Kuk, J., and Moffat, K. (2011) Temperature-scan cryocrystallography reveals reaction intermediates in bacteriophytochrome. *Nature*. **479**, 428–432

131. Yang, X., Stojkovic, E. A., Kuk, J., and Moffat, K. (2007) Crystal structure of the chromophore binding domain of an unusual bacteriophytochrome, RpBphP3, reveals residues that modulate photoconversion. *Proc Natl Acad Sci U S A*. **104**, 12571–12576

132. Burgie, E. S., Bussell, A. N., Walker, J. M., Dubiel, K., and Vierstra, R. D. (2014)

Crystal structure of the photosensing module from a red/far-red light-absorbing plant phytochrome. *Proc. Natl. Acad. Sci.* **111**, 10179–10184

133. Yu, D., Gustafson, W. C., Han, C., Lafaye, C., Noirclerc-Savoye, M., Ge, W.-P., Thayer, D. A., Huang, H., Kornberg, T. B., Royant, A., Jan, L. Y., Jan, Y. N., Weiss, W. A., and Shu, X. (2014) An improved monomeric infrared fluorescent protein for neuronal and tumour brain imaging. *Nat. Commun.* **5**, 3626–3632

134. Giraud, E., Zappa, S., Vuillet, L., Adriano, J.-M., Hannibal, L., Fardoux, J., Berthomieu, C., Bouyer, P., Pignol, D., and Verméglio, A. (2005) A new type of bacteriophytochrome acts in tandem with a classical bacteriophytochrome to control the antennae synthesis in Rhodopseudomonas palustris. *J. Biol. Chem.* **280**, 32389–32397

CHAPTER V

Concluding Remarks and Future Directions

Engineering fluorescent phytochromes

Fluorescence competes against both Pr→Pfr photoconversion and non-radiative decay in the form of ESPT as a means of returning bacteriophytochromes to the ground state. Excited state lifetime measurements show that 89% of the excited state relaxation happens via non-radiative decay/ESPT, with only 5% of the excited state relaxation happening via photoconversion (56). Thus, inhibiting ESPT would have a more dramatic effect on increasing fluorescence and thus making this family of fluorescent proteins more suitable for use as *in vivo* biomarkers.

Through this work we have shown that we are able to suppress photoconversion through site directed substitutions around the chromophore binding pocket which leads to improvement in fluorescence. ESPT is harder to manipulate, but removing polar interactions from the BV binding pocket (Asp207 variants) have also improved fluorescence abilities in phytochrome variants (Chapter III, WiPhy2) (4, 7, 85). When we started this project the D207H substitution was supposed to be the fundamental basis of BphP fluorescence. Our work has helped reevaluate that fundamental assumption. D207H substitution did augment fluorescence in WiPhy but in IFP1.4, with a background of 12 other amino acid substitutions, a reversion of the histidine 207 to the wild-type aspartate residue actually helped improve fluorescence quantum yield by 11%. Our 1.65 Å X-ray structure of IFP1.4 showed that hydrophobic residues around the BV D-ring were principally responsible for the increased fluorescence. When we introduced one of the substitutions, A288V, on wild-type $DrCBD_{mon}$ background we saw an improved quantum yield of 4.6% (7).

The WiPhy2 variant that we built has a quantum yield of 8.5% with a long excitation and emission maxima of 700 and 720 nm. This is one of the highest quantum yields in longer excitation wavelengths amongst the phytofluors. The maximum fluorescence among any of the variants of phytofluors is ~9% at excitation wavelengths above 670 nm. I installed the A288V mutation in the WiPhy2 background to see if the brightness could be further manipulated and found no improvement. The iRFP713 and 720 proteins have a quantum yield of ~6%. iRFP670 which is 50 nm blue shifted a two-fold higher quantum yield. It is possible that we have maximized fluorescence quantum yield of phytochromes. Non-radiative decay at longer wavelengths will not allow further increase in brightness.

Phytofluors have been successfully used in *in-vivo* studies. Because of the efficient tissue penetration of excitation wavelength in the near-infrared region, the signal from IFP1.4, iRFPs are significantly higher than GFP, even with the 6-fold lower brightness. Photostability, lack of cytotoxicity and the insensitivity to environmental fluctuations and ability to maintain signal over long timespans are important characteristics of a biomarker in non-invasive imaging. Phytofluors have been successfully tested in non-invasive imaging of brain, tumor progression and epi-fluorescence microscopy (6, 133). Transgenic mice encoding iRFPs have been successfully created, thus proving that phytofluors are non-cytotoxic and able to express in mammalian tissues.

Moving forward, I believe the next frontier in engineering phytochromes is going to be to make smaller holophytochromes, which still bind BV and fluorescence in the near-infrared. Verkhusha and co-workers have made a minimal domain GAF-FP using directed evolution technique on the template *Rhodopseudomonas palustris* BphP1. At a molecular weight of 20kDa, GAF-FP is half the size of usual phytofluors, has a fluorescence quantum yield of 7.3, but the excitation maximum is blue shifted by 50 nm at 635 nm. GAF-FP is able to bind two chromophores, BV and PCB (125).

I propose two groups of experiments with WiPhy2. We should test WiPhy2 in mammalian systems. It has a high brightness at long excitation maximum, is a monomer (IFP1.4 and 2.0 dimerize at high concentrations), doesn't photobleach, and brightness does not fluctuate with pH changes. Additionally, we should engineer a smaller GAF only version of WiPhy2 while maintaining its long excitation and emission maxima. An advantage of smaller near-infrared phytofluor would help us use it in viral systems which are much smaller to report cellular processes and protein-protein interactions. Recent advances have also been made in engineering of small GAF only near-infrared fluorescent phytochromes.

The Knot Problem

Knotting in proteins stabilizes the protein molecule. If we think about knots, the reason sailors tie complicated knots with multiple cross-overs, or why we knot shoe-laces, is because directed energy is required to untie a knot. In our work in evaluating knot formation in phytochromes, our results have demonstrated the efficiency of folding the figure of eight knot containing DrCBD is 2.5 fold higher than the knotless variant. The relative ease of knotted phytochrome accumulation in cells is consistent with the hypothesis that that folding pathways evolved to optimize for knot formation. We have also shown that cis-trans proline isomerization plays a role in the phytochrome knot. In absence of PPIase activity of trigger factor, folded protein accumulation was slower and never reached saturation.

We need to further evaluate and understand our results by conducting *in-vitro* experiments using the cell-free system which would give us a more precise control over the components, the protein and the chaperone being the principal ones. We can explore additional possibilities in the cell-free system including the involvement GroEL-GroES chaperonin which has been found to be involved in folding of deep trefoil knots (77). It would also be helpful to understand both photostability and thermal stability of the variants we engineered for probing knotting in DrCBD.

Final Remarks

My original goal when I started working on this project was to engineer a brighter phytofluor. For successful engineering of any protein, its mechanism of action. For the phytochrome system, I needed to understand what makes phytofluors prefer fluorescence as a relaxation pathway, and investigate how phytochromes fold into a complex knotted topology. My thesis work has helped explain the structural basis of fluorescence in phytochromes and led us to rewrite the fundamental assumption of the D207H substitution being key to fluorescence. We have solved two detailed three dimensional structures of near-infrared fluorescent variants which have shown that we can manipulate phytochrome photochemistry by either adding bulkier residues around the D-ring of the chromophore or by decreasing the number of waters near the chromophore. Based on the structural information, we have generated a library of fluorescent variants that support our structural results. Moving forward, structure guided phytofluor design would definitely be key to further augment the capabilities of phytochromes as biomarkers. Finally, I have also shown that the TF chaperone and PPIase activity associated with TF are key players in folding of the figure of eight knot. My results demonstrate that the figure of eight knot formation contributes to kinetic stability of the phytochrome.

References

1. Jobsis, F. F. (1977) Noninvasive, infrared monitoring of cerebral and myocardial oxygen sufficiency and circulatory parameters. *Science*. **198**, 1264–1267

2. Wagner, J. R., Zhang, J., von Stetten, D., Günther, M., Murgida, D. H., Mroginski, M. A., Walker, J. M., Forest, K. T., Hildebrandt, P., and Vierstra, R. D. (2008) Mutational analysis of Deinococcus radiodurans bacteriophytochrome reveals key amino acids necessary for the photochromicity and proton exchange cycle of phytochromes. *J. Biol. Chem.* **283**, 12212–12226

3. Shu, X., Royant, A., Lin, M. Z., Aguilera, T. A., Lev-Ram, V., Steinbach, P. A., and Tsien, R. Y. (2009) Mammalian expression of infrared fluorescent proteins engineered from a bacterial phytochrome. *Science*. **324**, 804–807

4. Auldridge, M. E., Satyshur, K. a, Anstrom, D. M., and Forest, K. T. (2012) Structure-guided engineering enhances a phytochrome-based infrared fluorescent protein. *J. Biol. Chem.* **287**, 7000–7009

5. Filonov, G. S., Piatkevich, K. D., Ting, L.-M., Zhang, J., Kim, K., and Verkhusha, V. V (2011) Bright and stable near-infrared fluorescent protein for in vivo imaging. *Nat. Biotechnol.* **29**, 757–761

6. Shcherbakova, D. M., and Verkhusha, V. V. (2013) Near-infrared fluorescent proteins for multicolor in vivo imaging. *Nat. Methods*. **10**, 751–754

7. Bhattacharya, S., Auldridge, M. E., Lehtivuori, H., Ihalainen, J. A., and Forest, K. T.

(2014) Origins of Fluorescence in Evolved Bacteriophytochromes. *J. Biol. Chem.* **289**, 32144–32152

8. Lehtivuori, H., Bhattacharya, S., Angenent-Mari, N. M., Satyshur, K. A., and Forest, K. T. (2015) Removal of Chromophore-Proximal Polar Atoms Decreases Water Content and Increases Fluorescence in a Near Infrared Phytofluor. *Front. Mol. Biosci.* **2**, 65–69

9. Stepanenko, O. V., Verkhusha, V. V., Kuznetsova, I. M., Uversky, V. N., and Turoverov, K. K. (2008) Fluorescent proteins as biomarkers and biosensors: throwing color lights on molecular and cellular processes. *Curr. Protein Pept. Sci.* **9**, 338–369

10. Crameri, A., Whitehorn, E. A., Tate, E., and Stemmer, W. P. (1996) Improved green fluorescent protein by molecular evolution using DNA shuffling. *Nat Biotechnol*. **14**, 315–319

11. Yoshihara, S., Shimada, T., Matsuoka, D., Zikihara, K., Kohchi, T., and Tokutomi, S. (2006) Reconstitution of Blue−Green Reversible Photoconversion of a Cyanobacterial Photoreceptor, PixJ1, in Phycocyanobilin-Producing Escherichia coli. *Biochemistry*. **45**, 3775–3784

12. Tsien, R. (1998) The green fluorescent protein. *Annu. Rev. Biochem.* **67**, 509

13. Ando, R., Hama, H., Yamamoto-Hino, M., Mizuno, H., and Miyawaki, A. (2002) An optical marker based on the UV-induced green-to-red photoconversion of a fluorescent protein. *Proc. Natl. Acad. Sci.* **99**, 12651–12656

14. Ehrig, T., O'Kane, D. J., and Prendergast, F. G. (1995) Green-fluorescent protein mutants with altered fluorescence excitation spectra. *FEBS Lett*. **367**, 163–166

15. Reeder, P. J., Huang, Y.-M., Dordick, J. S., and Bystroff, C. (2010) A Rewired Green Fluorescent Protein: Folding and Function in a Nonsequential, Noncircular GFP Permutant. *Biochemistry*. **49**, 10773–10779

16. Hanson, G. T., McAnaney, T. B., Park, E. S., Rendell, M. E. P., Yarbrough, D. K., Chu, S., Xi, L., Boxer, S. G., Montrose, M. H., and Remington, S. J. (2002) Green Fluorescent Protein Variants as Ratiometric Dual Emission pH Sensors. 1. Structural Characterization and Preliminary Application. *Biochemistry*. **41**, 15477–15488

17. Davis, S. J., and Vierstra, R. D. (1998) Soluble, highly fluorescent variants of green fluorescent protein (GFP) for use in higher plants. *Plant Mol. Biol.* **36**, 521–528

18. Baulcombe, D. C., Chapman, S., and Cruz, S. (1995) Jellyfish green fluorescent protein as a reporter for virus infections. *Plant J.* **7**, 1045–1053

19. Peng, S.-Y., Yang, Y.-S., Chou, C.-J., Lin, K.-Y., and Wu, S.-C. (2015) Differentiation of Enhanced Green Fluorescent Protein-Labeled Mouse Amniotic Fluid-Derived Stem Cells into Cardiomyocyte-Like Beating Cells. *Acta Cardiol. Sin.* **31**, 209–214

20. Boas, D. A., Gaudette, T., Strangman, G., Cheng, X., Marota, J. J. A., and Mandeville, J. B. (2001) The Accuracy of Near Infrared Spectroscopy and Imaging during Focal Changes in Cerebral Hemodynamics. *Neuroimage*. **13**, 76–90

21. Mancini, D. M., Bolinger, L., Li, H., Kendrick, K., Chance, B., and Wilson, J. R. (1994) Validation of near-infrared spectroscopy in humans. *J. Appl. Physiol.* **77**, 2740–2747

22. Hilderbrand, S. A., and Weissleder, R. (2010) Near-infrared fluorescence: application to in vivo molecular imaging. *Curr. Opin. Chem. Biol.* **14**, 71–79

23. Luo, S., Zhang, E., Su, Y., Cheng, T., and Shi, C. (2011) A review of NIR dyes in cancer targeting and imaging. *Biomaterials*. **32**, 7127–7138

24. Ke, S., Wen, X., Gurfinkel, M., Charnsangavej, C., Wallace, S., Sevick-Muraca, E. M., and Li, C. (2003) Near-Infrared Optical Imaging of Epidermal Growth Factor Receptor in Breast Cancer Xenografts. *Cancer Res.* **63**, 7870–7875

25. Bugaj, J. E., Achilefu, S., Dorshow, R. B., and Rajagopalan, R. (2001) Novel fluorescent contrast agents for optical imaging of in vivo tumors based on a receptor-targeted dye-peptide conjugate platform. *J. Biomed. Opt.* **6**, 122–132

26. Achilefu, S., Dorshow, R. B., Bugaj, J. E., and Rajagopalan, R. (2000) Novel receptor-targeted fluorescent contrast agents for in vivo tumor imaging. *Invest. Radiol.* **35**, 479–485

27. Bongarzone, S., Staderini, M., and Bolognesi, M. L. (2014) Multitarget ligands and theranostics: sharpening the medicinal chemistry sword against prion diseases. *Future Med. Chem.* **6**, 1017–1029

28. Doerr, A. (2009) Fluorescent proteins: into the infrared. *Nat. Methods*. **6**, 482–483

29. Cashmore, A. R. (2003) Cryptochromes: Enabling plants and animals to determine

circadian time. *Cell*. **114**, 537–543

30. Murphy, J. T., and Lagarias, J. C. (1997) The phytofluors: a new class of fluorescent protein probes. *Curr. Biol.* **7**, 870–876

31. Christie, J. M., Gawthorne, J., Young, G., Fraser, N. J., and Roe, A. J. (2012) LOV to BLUF: Flavoprotein Contributions to the Optogenetic Toolkit. *Mol. Plant*. **5**, 533–544

32. Pudasaini, A., El-Arab, K. K., and Zoltowski, B. D. (2015) LOV-based optogenetic devices: light-driven modules to impart photoregulated control of cellular signaling. *Front. Mol. Biosci.* **2**, 18–22

33. Möglich, A., Yang, X., Ayers, R. A., and Moffat, K. (2010) Structure and Function of Plant Photoreceptors. *Annu. Rev. Plant Biol.* **61**, 21–47

34. Rockwell, N. C., and Lagarias, J. C. (2010) A Brief History of Phytochromes. *ChemPhysChem*. **11**, 1172–1180

35. Auldridge, M. E., and Forest, K. T. (2011) Bacterial phytochromes: more than meets the light. *Crit. Rev. Biochem. Mol. Biol.* **46**, 67–88

36. Al-Sady, B., Ni, W., Kircher, S., Schaefer, E., and Quail, P. H. (2006) Photoactivated Phytochrome Induces Rapid PIF3 Phosphorylation Prior to Proteasome-Mediated Degradation. *Mol. Cell*. **23**, 439–446

37. Fankhauser, C. (2001) The Phytochromes, a Family of Red/Far-red Absorbing Photoreceptors. *J. Biol. Chem.* **276**, 11453–11456

38. Rizzini, L., Favory, J.-J., Cloix, C., Faggionato, D., O'Hara, A., Kaiserli, E.,

Baumeister, R., Schäfer, E., Nagy, F., Jenkins, G. I., and Ulm, R. (2011) Perception of UV-B by the Arabidopsis UVR8 protein. *Science*. **332**, 103–106

39. Quail, P. H. (2002) Phytochrome photosensory signalling networks. *Nat. Rev. Mol. Cell Biol.* **3**, 85–93

40. Kehoe, D. M., and Grossman, A. R. (1996) Similarity of a chromatic adaptation sensor to phytochrome and ethylene receptors. *Science*. **273**, 1409–1412

41. Terauchi, K., Montgomery, B. L., Grossman, A. R., Lagarias, J. C., and Kehoe, D. M. (2004) RcaE is a complementary chromatic adaptation photoreceptor required for green and red light responsiveness. *Mol. Microbiol.* **51**, 567–577

42. Davis, S. J., Vener, A. V, and Vierstra, R. D. (1999) Bacteriophytochromes: phytochrome-like photoreceptors from nonphotosynthetic eubacteria. *Science*. **286**, 2517–2520

43. Bhoo, S. H., Davis, S. J., Walker, J., Karniol, B., and Vierstra, R. D. (2001) Bacteriophytochromes are photochromic histidine kinases using a biliverdin chromophore. *Nature*. **414**, 776–779

44. Butler, W. L., Norris, K. H., Siegelman, H. W., and Hendricks, S. B. (1959) Detection, assay, and preliminary purification of the pigment controlling photoresponsive development of plants. *Proc. Natl. Acad. Sci. U. S. A.* **45**, 1703–1708

45. Blaauw-Jansen, G. (1959) The influence of red and far red light on growth and phototropism of the avena seedling. *Acta Bot. Neerl.* **8**, 1–39

46. Wagner, J. R., Brunzelle, J. S., Forest, K. T., and Vierstra, R. D. (2005) A light-sensing knot revealed by the structure of the chromophore-binding domain of phytochrome. *Nature*. **438**, 325–331

47. Wagner, J. R., Zhang, J., Brunzelle, J. S., Vierstra, R. D., and Forest, K. T. (2007) High resolution structure of Deinococcus bacteriophytochrome yields new insights into phytochrome architecture and evolution. *J. Biol. Chem.* **282**, 12298–12309

48. Yang, X., Stojkovic, E. A., Kuk, J., and Moffat, K. (2007) Crystal structure of the chromophore binding domain of an unusual bacteriophytochrome, RpBphP3, reveals residues that modulate photoconversion. *Proc. Natl. Acad. Sci.* **104**, 12571–12576

49. Nagatani, A. (2010) Phytochrome: structural basis for its functions. *Curr. Opin. Plant Biol.* **13**, 565–570

50. Takala, H., Björling, A., Berntsson, O., Lehtivuori, H., Niebling, S., Hoernke, M., Kosheleva, I., Henning, R., Menzel, A., Ihalainen, J. A., and Westenhoff, S. (2014) Signal amplification and transduction in phytochrome photosensors. *Nature*. **509**, 245–248

51. Baker, A. W., and Forest, K. T. (2014) Structural biology: Action at a distance in a light receptor. *Nature*. **509**, 174–175

52. Burgie, E. S., Wang, T., Bussell, A. N., Walker, J. M., Li, H., and Vierstra, R. D. (2014) Crystallographic and Electron Microscopic Analyses of a Bacterial Phytochrome Reveal Local and Global Rearrangements During Photoconversion. *J. Biol. Chem.* **289**, 24573–24587

53. Yang, X., Kuk, J., and Moffat, K. (2009) Conformational differences between the Pfr and Pr states in Pseudomonas aeruginosa bacteriophytochrome. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 15639–15644

54. Bischoff, M., Hermann, G., Rentsch, S., and Strehlow, D. (2001) First Steps in the Phytochrome Phototransformation: A Comparative Femtosecond Study on the Forward (Pr → Pfr) and Back Reaction (Pfr → Pr). *Biochemistry*. **40**, 181–186

55. Andel, F., Hasson, K. C., Gai, F., Anfinrud, P. A., and Mathies, R. A. (1997) Femtosecond time-resolved spectroscopy of the primary photochemistry of phytochrome. *Biospectroscopy*. **3**, 421–433

56. Toh, K. C., Stojkovic, E. A., van Stokkum, I. H., Moffat, K., and Kennis, J. T. (2011) Fluorescence quantum yield and photochemistry of bacteriophytochrome constructs. *Phys Chem Chem Phys*. **13**, 11985–11997

57. Dasgupta, J., Frontiera, R. R., Taylor, K. C., Lagarias, J. C., and Mathies, R. A. (2009) Ultrafast excited-state isomerization in phytochrome revealed by femtosecond stimulated Raman spectroscopy. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 1784–1789

58. Piatkevich, K. D., Subach, F. V, and Verkhusha, V. V. (2013) Engineering of bacterial phytochromes for near-infrared imaging, sensing, and light-control in mammals. *Chem. Soc. Rev.* **42**, 3441–3452

59. Yang, X., Stojković, E. A., Ozarowski, W. B., Kuk, J., Davydova, E., and Moffat, K. (2015) Light Signaling Mechanism of Two Tandem Bacteriophytochromes. *Structure*. **23**, 1179–1189

60. Rockwell, N. C., Su, Y.-S., and Lagarias, J. C. (2006) Phytochrome structure and signaling mechanisms. *Annu. Rev. Plant Biol.* **57**, 837–858

61. Kim, P. W., Freer, L. H., Rockwell, N. C., Martin, S. S., Lagarias, J. C., and Larsen, D. S. (2012) Femtosecond Photodynamics of the Red/Green Cyanobacteriochrome NpR6012g4 from *Nostoc punctiforme* . 2. Reverse Dynamics. *Biochemistry*. **51**, 619–630

62. Ulijasz, A. T., Cornilescu, G., Cornilescu, C. C., Zhang, J., Rivera, M., Markley, J. L., and Vierstra, R. D. (2010) Structural basis for the photoconversion of a phytochrome to the activated Pfr form. *Nature*. **463**, 250–254

63. Kneip, C., Hildebrandt, P., Schlamann, W., Braslavsky, S. E., Mark, F., and Schaffner, K. (1999) Protonation state and structural changes of the tetrapyrrole chromophore during the Pr/Pfr phototransformation of phytochrome: a resonance Raman spectroscopic study. *Biochemistry*. **38**, 15185–15192

64. Spillane, K. M., Dasgupta, J., and Mathies, R. A. (2012) Conformational Homogeneity and Excited-State Isomerization Dynamics of the Bilin Chromophore in Phytochrome Cph1 from Resonance Raman Intensities. *Biophys. J.* **102**, 709–717

65. Toh, K. C., Stojkovic, E. A., van Stokkum, I. H. M., Moffat, K., and Kennis, J. T. M. (2010) Proton-transfer and hydrogen-bond interactions determine fluorescence quantum yield and photochemical efficiency of bacteriophytochrome. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 9170–9175

66. Cormack, B. P., Valdivia, R. H., and Falkow, S. (1996) FACS-optimized mutants of

the green fluorescent protein (GFP). *Gene*. **173**, 33–38

67. Heim, R., and Tsien, R. Y. (1996) Engineering green fluorescent protein for improved brightness, longer wavelengths and fluorescence resonance energy transfer. *Curr Biol*. **6**, 178–182

68. Afar, B., Merrill, J., and Clark, E. A. (1991) Detection of lymphocyte subsets using three-color/single-laser flow cytometry and the fluorescent dye Peridinin chlorophyll-a protein. *J. Clin. Immunol.* **11**, 254–261

69. Nakajima, O., and Gray, C. H. (1967) Studies on haem alpha-methenyl oxygenase. Isomeric structure of formylbiliverdin, a possible precursor of biliverdin. *Biochem. J.* **104**, 20–22

70. Schäfer, F. P. (1973) Principles of dye laser operation. in *n Topics in Applied Physics Dye Lasers* (Schäfer, F. P. ed), pp. 1–89, Springer-Verlag, New York, 10.1007/3-540-51558-5_7

71. Filonov, G. S., Krumholz, A., Xia, J., Yao, J., Wang, L. V, and Verkhusha, V. V. (2012) Deep-tissue photoacoustic tomography of a genetically encoded near-infrared fluorescent probe. *Angew Chem Int Ed Engl*. **51**, 1448–1451

72. Bartz-Schmidt, K. U., Walter, P., Krott, R., Brunner, R., Esser, P., and Heimann, K. (1996) [Effects of fluorescein and indocyanine green angiography on subsequent dark adaptation and the electroretinogram]. *Klin Monbl Augenheilkd*. **208**, 224–228

73. Taylor, W. R. (2000) A deeply knotted protein structure and how it might fold. *Nature*. **406**, 916–919

74. Khatib, F., Weirauch, M. T., and Rohl, C. A. (2006) Rapid knot detection and application to protein structure prediction. *Bioinformatics*. **22**, e252–e259

75. Mallam, A. L., Morris, E. R., and Jackson, S. E. (2008) Exploring knotting mechanisms in protein folding. *Proc. Natl. Acad. Sci.* **105**, 18740–18745

76. Mallam, A. L., and Jackson, S. E. (2006) Probing Nature's Knots: The Folding Pathway of a Knotted Homodimeric Protein. *J. Mol. Biol.* **359**, 1420–1436

77. Mallam, A. L., Rogers, J. M., and Jackson, S. E. (2010) Experimental detection of knotted conformations in denatured proteins. *Proc Natl Acad Sci U S A*. **107**, 8189–8194

78. King, N. P., Yeates, E. O., and Yeates, T. O. (2007) Identification of Rare Slipknots in Proteins and Their Implications for Stability and Folding. **373**, 153–166

79. King, N. P., Jacobitz, A. W., Sawaya, M. R., Goldschmidt, L., and Yeates, T. O. (2010) Structure and folding of a designed knotted protein. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 20732–20737

80. Dill, K. A., and MacCallum, J. L. (2012) The protein-folding problem, 50 years on. **338**, 1042–1046

81. Andersson, F. I., Pina, D. G., Mallam, A. L., Blaser, G., and Jackson, S. E. (2009) Untangling the folding mechanism of the 52-knotted protein UCH-L3. **276**, 2625–2635

82. Bornschlogl, T., Anstrom, D. M., Mey, E., Dzubiella, J., Rief, M., and Forest, K. T. (2009) Tightening the knot in phytochrome by single-molecule atomic force

microscopy. *Biophys J.* **96**, 1508–1514

83. von Stetten, D., Seibeck, S., Michael, N., Scheerer, P., Mroginski, M. A., Murgida, D. H., Krauss, N., Heyn, M. P., Hildebrandt, P., Borucki, B., and Lamparter, T. (2007) Highly conserved residues Asp-197 and His-250 in Agp1 phytochrome control the proton affinity of the chromophore and Pfr formation. *J. Biol. Chem.* **282**, 2116–2123

84. Anders Borg, O., and Durbeej, B. (2008) Which factors determine the acidity of the phytochromobilin chromophore of plant phytochrome? *Phys. Chem. Chem. Phys.* **10**, 2528–2537

85. Lehtivuori, H., Bhattacharya, S., Angenent-Mari, N., Satyshur, K. A., and Forest, K. T. (2015) Removal of Chromophore-Proximal Polar Atoms Decreases Water Content and Increases Fluorescence in a Near Infrared Phytofluor. *Front. Mol. Biosci.* **2**, 65–69

86. Spillane, K. M., Dasgupta, J., Lagarias, J. C., and Mathies, R. A. (2009) Homogeneity of phytochrome Cph1 vibronic absorption revealed by resonance Raman intensity analysis. *J Am Chem Soc*. **131**, 13946–13948

87. Essen, L.-O., Mailliet, J., and Hughes, J. (2008) The structure of a complete phytochrome sensory module in the Pr ground state. *Proc. Natl. Acad. Sci.* **105**, 14709–14714

88. Yang, X., Kuk, J., and Moffat, K. (2008) Crystal structure of Pseudomonas aeruginosa bacteriophytochrome: photoconversion and signal transduction. *Proc. Natl. Acad. Sci.* **105**, 14715–14720

89. von Stetten, D., Gunther, M., Scheerer, P., Murgida, D. H., Mroginski, M. A., Krauss, N., Lamparter, T., Zhang, J., Anstrom, D. M., Vierstra, R. D., Forest, K. T., and Hildebrandt, P. (2008) Chromophore heterogeneity and photoconversion in phytochrome crystals and solution studied by resonance Raman spectroscopy. *Angew Chem Int Ed Engl*. **47**, 4753–4755

90. Dasgupta, J., Frontiera, R. R., Taylor, K. C., Lagarias, J. C., and Mathies, R. A. (2009) Ultrafast excited-state isomerization in phytochrome revealed by femtosecond stimulated Raman spectroscopy. *Proc. Natl. Acad. Sci.* **106**, 1784–1789

91. Foerstendorf, H., Mummert, E., Sch??fer, E., Scheer, H., and Siebert, F. (1996) Fourier-transform infrared spectroscopy of phytochrome: Difference spectra of the intermediates of the photoreactions. *Biochemistry*. **35**, 10793–10799

92. Toh, K. C., van Stokkum, I. H., Hendriks, J., Alexandre, M. T., Arents, J. C., Perez, M. A., van Grondelle, R., Hellingwerf, K. J., and Kennis, J. T. (2008) On the signaling mechanism and the absence of photoreversibility in the AppA BLUF domain. *Biophys J*. **95**, 312–321

93. Fischer, A. J., and Lagarias, J. C. (2004) Harnessing phytochrome's glowing potential. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 17334–17339

94. Lehtivuori, H., Rissanen, I., Takala, H., Bamford, J., Tkachenko, N. V., and Ihalainen, J. A. (2013) Fluorescence Properties of the Chromophore-Binding Domain of Bacteriophytochrome from *Deinococcus radiodurans*. *J. Phys. Chem. B*. **117**, 11049–11057

95. Otwinowski, Z., and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326

96. McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007) Phaser crystallographic software. *J Appl Crystallogr*. **40**, 658–674

97. Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F., and Vagin, A. A. (2011) REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. D. Biol. Crystallogr.* **67**, 355–367

98. Hahn, J., Strauss, H. M., Landgraf, F. T., Gimenèz, H. F., Lochnit, G., Schmieder, P., and Hughes, J. (2006) Probing protein-chromophore interactions in Cph1 phytochrome by mutagenesis. *FEBS J.* **273**, 1415–1429

99. Jamroz, M., Niemyska, W., Rawdon, E. J., Stasiak, A., Millett, K. C., Sułkowski, P., and Sulkowska, J. I. (2015) KnotProt: a database of proteins with knots and slipknots. *Nucleic Acids Res.* **43**, D306–D314

100. Lai, Y.-L., Yen, S.-C., Yu, S.-H., and Hwang, J.-K. (2007) pKNOT: the protein KNOT web server. *Nucleic Acids Res.* **35**, W420–W424

101. Millett, K. C., Rawdon, E. J., Stasiak, A., and Sułkowska, J. I. (2013) Identifying knots in proteins. *Biochem. Soc. Trans.* **41**, 533–537

102. Burgie, E. S., and Vierstra, R. D. (2014) Phytochromes: An Atomic Perspective on Photoactivation and Signaling. *Plant Cell Online*. **26**, 4568–4583

103. Rockwell, N. C., Martin, S. S., Gulevich, A. G., and Lagarias, J. C. (2012) Phycoviolobilin Formation and Spectral Tuning in the DXCF Cyanobacteriochrome Subfamily. *Biochemistry*. **51**, 1449–1463

104. Sineshchekov, V., Mailliet, J., Psakis, G., Feilke, K., Kopycki, J., Zeidler, M., Essen, L.-O., and Hughes, J. (2014) Tyrosine 263 in cyanobacterial phytochrome Cph1 optimizes photochemistry at the prelumi-R→lumi-R step. *Photochem. Photobiol.* **90**, 786–795

105. Narikawa, R., Fukushima, Y., Ishizuka, T., Itoh, S., and Ikeuchi, M. (2008) A Novel Photoactive GAF Domain of Cyanobacteriochrome AnPixJ That Shows Reversible Green/Red Photoconversion. *J. Mol. Biol.* **380**, 844–855

106. Narikawa, R., Ishizuka, T., Muraki, N., Shiba, T., Kurisu, G., and Ikeuchi, M. (2013) Structures of cyanobacteriochromes from phototaxis regulators AnPixJ and TePixJ reveal general and specific photoconversion mechanism. *Proc. Natl. Acad. Sci.* **110**, 918–923

107. Narikawa, R., Nakajima, T., Aono, Y., Fushimi, K., Enomoto, G., Ni-Ni-Win, Itoh, S., Sato, M., Ikeuchi, M., Narikawa, R., Enomoto, G., Ni, W. N., Fushimi, K., and Ikeuchi, M. (2015) A biliverdin-binding cyanobacteriochrome from the chlorophyll d–bearing cyanobacterium Acaryochloris marina. *Sci. Rep.* **5**, 7950–7975

108. Zhang, J., Wu, X.-J., Wang, Z.-B., Chen, Y., Wang, X., Zhou, M., Scheer, H., and Zhao, K.-H. (2010) Fused-Gene Approach to Photoswitchable and Fluorescent Biliproteins. *Angew. Chemie Int. Ed.* **49**, 5456–5458

109. Rockwell, N. C., Martin, S. S., and Lagarias, J. C. (2012) Red/Green

Cyanobacteriochromes: Sensors of Color and Power. *Biochemistry*. **51**, 9667–9677

110. Pal, D., and Chakrabarti, P. (1999) Cis peptide bonds in proteins: residues involved, their conformations, interactions and locations. *J. Mol. Biol.* **294**, 271–288

111. Joseph, A. P., Srinivasan, N., and de Brevern, A. G. (2012) Cis-trans peptide variations in structurally similar proteins. *Amino Acids*. **43**, 1369–1381

112. MacArthur, M. W., and Thornton, J. M. (1991) Influence of proline residues on protein conformation. *J. Mol. Biol.* **218**, 397–412

113. Fischer, G., and Schmid, F. X. (1990) The mechanism of protein folding. Implications of in vitro refolding models for de novo protein folding and translocation in the cell. *Biochemistry*. **29**, 2205–2212

114. Hesterkamp, T., and Bukau, B. (1996) Identification of the prolyl isomerase domain of Escherichia coli trigger factor. *FEBS Lett.* **385**, 67–71

115. Kramer, G., Patzelt, H., Rauch, T., Kurz, T. A., Vorderwülbecke, S., Bukau, B., and Deuerling, E. (2004) Trigger Factor Peptidyl-prolyl cis/trans Isomerase Activity Is Not Essential for the Folding of Cytosolic Proteins in Escherichia coli. *J. Biol. Chem.* **279**, 14165–14170

116. Stoller, G., Rücknagel, K. P., Nierhaus, K. H., Schmid, F. X., Fischer, G., and Rahfeld, J. U. (1995) A ribosome-associated peptidyl-prolyl cis/trans isomerase identified as the trigger factor. *EMBO J.* **14**, 4939–4948

117. Kramer, G., Rauch, T., Rist, W., Vorderwülbecke, S., Patzelt, H., Schulze-Specking, A., Ban, N., Deuerling, E., and Bukau, B. (2002) L23 protein functions as

a chaperone docking site on the ribosome. *Nature*. **419**, 171–174

118. Gupta, R., Lakshmipathy, S. K., Chang, H.-C., Etchells, S. A., and Hartl, F. U. (2010) Trigger factor lacking the PPIase domain can enhance the folding of eukaryotic multi-domain proteins in Escherichia coli. *FEBS Lett.* **584**, 3620–3624

119. Ferbitz, L., Maier, T., Patzelt, H., Bukau, B., Deuerling, E., and Ban, N. (2004) Trigger factor in complex with the ribosome forms a molecular cradle for nascent proteins. *Nature*. **431**, 590–596

120. Thomason, L. C., Costantino, N., Court, D. L., Thomason, L. C., Costantino, N., and Court, D. L. (2007) *E. coli* Genome Manipulation by P1 Transduction. in *Current Protocols in Molecular Biology*, pp. 1.17.1–1.17.8, John Wiley & Sons, Inc., Hoboken, NJ, USA, 10.1002/0471142727.mb0117s79

121. Burgie, E. S., Walker, J. M., Phillips, G. N., and Vierstra, R. D. (2013) A Photo-Labile Thioether Linkage to Phycoviolobilin Provides the Foundation for the Blue/Green Photocycles in DXCF-Cyanobacteriochromes. *Structure*. **21**, 88–97

122. Frey, U. H., Bachmann, H. S., Peters, J., and Siffert, W. (2008) PCR-amplification of GC-rich regions: “slowdown PCR.” *Nat. Protoc.* **3**, 1312–1317

123. Gibson, D. G., Young, L., Chuang, R.-Y., Venter, J. C., Hutchison, C. A., and Smith, H. O. (2009) Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat Meth*. **6**, 343–345

124. Schneider, C. A., Rasband, W. S., and Eliceiri, K. W. (2012) NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods*. **9**, 671–675

125. Rumyantsev, K. A., Shcherbakova, D. M., Zakharova, N. I., Emelyanov, A. V., Turoverov, K. K., and Verkhusha, V. V. (2015) Minimal domain of bacterial phytochrome required for chromophore binding and fluorescence. *Sci. Rep.* **5**, 18348–18352

126. Mallam, A. L. (2009) How does a knotted protein fold? *FEBS J.* **276**, 365–375

127. Sułkowska, J. I., Noel, J. K., Ramírez-Sarmiento, Esar A, Rawdon, E. J., Millett, K. C., and Onuchic, J. N. (2013) Knotting pathways in proteins. *Biochem. Soc. Trans.* **41**, 523–527

128. Mallam, A. L., and Jackson, S. E. (2011) Knot formation in newly translated proteins is spontaneous and accelerated by chaperonins. *Nat. Chem. Biol.* **8**, 147–153

129. Yang, X., Stojković, E. A., Ozarowski, W. B., Kuk, J., Davydova, E., and Moffat, K. (2015) Light Signaling Mechanism of Two Tandem Bacteriophytochromes. *Structure*. **23**, 1179–1189

130. Yang, X., Ren, Z., Kuk, J., and Moffat, K. (2011) Temperature-scan cryocrystallography reveals reaction intermediates in bacteriophytochrome. *Nature*. **479**, 428–432

131. Yang, X., Stojkovic, E. A., Kuk, J., and Moffat, K. (2007) Crystal structure of the chromophore binding domain of an unusual bacteriophytochrome, RpBphP3, reveals residues that modulate photoconversion. *Proc Natl Acad Sci U S A*. **104**, 12571–12576

132. Burgie, E. S., Bussell, A. N., Walker, J. M., Dubiel, K., and Vierstra, R. D. (2014)

Crystal structure of the photosensing module from a red/far-red light-absorbing plant phytochrome. *Proc. Natl. Acad. Sci.* **111**, 10179–10184

133. Yu, D., Gustafson, W. C., Han, C., Lafaye, C., Noirclerc-Savoye, M., Ge, W.-P., Thayer, D. A., Huang, H., Kornberg, T. B., Royant, A., Jan, L. Y., Jan, Y. N., Weiss, W. A., and Shu, X. (2014) An improved monomeric infrared fluorescent protein for neuronal and tumour brain imaging. *Nat. Commun.* **5**, 3626–3632

134. Giraud, E., Zappa, S., Vuillet, L., Adriano, J.-M., Hannibal, L., Fardoux, J., Berthomieu, C., Bouyer, P., Pignol, D., and Verméglio, A. (2005) A new type of bacteriophytochrome acts in tandem with a classical bacteriophytochrome to control the antennae synthesis in Rhodopseudomonas palustris. *J. Biol. Chem.* **280**, 32389–32397

APPENDIX I

iRFP Purification and crystallization studies

Abstract

The iRFP family of proteins has been used as fluorescent biomarkers on account of their versatile extinction and emission maxima (660-700) nm, photostability and lack of cytotoxicity. They are constructed from two *Rhodopseudomonas palustris* bacteriophytochromes P2 and P3. P3 is a naturally fluorescent BphP. We aim to understand the structural properties that make the iRFP proteins fluoresce.

Introduction

Rhodopseudomonas palustris has a total of six bacteriophytochromes encoded in its genome with varying spectral characteristics. *R. palustris* BphP1 (*RpBphP1*) from strain CEA001, RpBphP5, and RpBphP6 from strain CGA009 favor the Pfr ground state (134). These phytochromes are termed bathy phytochromes because the absorption maxima of the bathochromic shift in their ground states. RpBphP1, RpBphP5 and RpBphP6 undergo dark conversion from Pr to the favored Pfr ground state (131).

In addition to the above bathy phytochromes, RpBphP3, which has a Pr ground state, photoconverts to the near red Pnr form with a blue-shifted absorption maximum at 645nm. RpBphP3 and RpBphP2 genes are contiguous and are located near the genes encoding the light harvesting complex 4 in the organism and have been implicated in regulation of LH4 synthesis (Toh et al., 2011; Yang et al., 2015). Lastly, RpBphP4 strains Ha2 and BisB5 photoconvert into a long-lived MetaR form. MetaRa and MetaRc are intermediates in the photoconversion from Pr to Pfr of prototypical phytochromes. However, RpBphP4 from R. palustris CGA009 lacks the biliverdin-binding cysteine and does not bind a chromophore.

RpBphP2 was the first of the *Rhodopseudomonas palustris* be engineered into a near-infrared fluorescent protein. This iRFP was modeled on the contemporary *Deinococcus radiodurans* phytochrome which had been shown to be fluorescent when a D207H (D202H in RpBphP2) substitution was introduced. Research by Filonov *et al.* designed the first iRFP (iRFP713) by subjecting truncated RpBphP2-PAS-GAF-D202H gene to three rounds of random mutagenesis and a final round of saturation mutagenesis (5). This model was further improved in 2013 by Shcherbakova *et al.* to design iRFP 670, 682, and

720 based on RpBphP6 as template and iRFP702 used RpBphP2. The number 670, 682, 702, 713 and 720 refer to the emission maxima of the variants (6).

Our research has since demonstrated that presence of the histidine in the 207th position has no significant effect in increasing the fluorescence quantum yields of bacterial phytochromes in *Deinococcus radiodurans* (4, 7). Crystal structures of near-IR phytochromes WiPhy, WiPhy2 and IFP1.4 revealed that increased hydrophobic packing of residues around the D-ring of the BV chromophore combined with absence of polar interactions in the binding pocket is responsible for the higher brightness of the engineered phytofluors (4, 7, 85).

The iRFP variants share the same basic PAS-GAF architecture with the DrCBD variants but they are dimeric and sequences have been significantly altered using site specific and saturated mutagenesis. We would like to investigate structural qualities of iRFP variants in order to further our insight about fluorescence in bacteriophytochromes and delineate the structural reasons behind it which we hope to map onto our spectroscopic information.

In order to express functional holoprotein in mammalian systems gene encoding iRFP is expressed simultaneously with a gene encoding heme oxygenase. The *hmuo* gene is from *Brazyrhizobium sp.* Expression of heme oxygenase along with iRFP *in vitro* results in the formation of functional phytofluor with chromophore bound. *hmuO* gene encoding *Bradyrhizobium* ORS278 in pWA21-AvrIINotI plasmid. In the presence of a suitable electron donor the enzyme heme oxygenase (HO,1 EC 1.14.99.3) catalyzes the oxidative degradation of heme to yield equimolar amounts of biliverdin (BV), iron, and CO.

Single turnover experiments of human HO-1 have revealed that the release of BV from HO-1 is the rate-limiting step in the conversion from heme to BV. In the presence of BV, this step is accelerated and one of the earlier electron transfer steps becomes rate-limiting. The data presented here indicate that the bacterial phytochrome BphP from *P. aeruginosa* whose gene is encoded in the same operon as *bphO* is able to interact with the BphO·BV complex and facilitates the release of BV. The BV is "pulled" out of the HO and forms a stable covalent adduct with BphP, which was confirmed by difference spectroscopy and zinc blotting. The goal of this chapter is to understand the basis of the altered spectroscopic properties and improved brightness of iRFPs. To that end we aim to solve the crystal structures of all the variants of iRFP. The crystal structure of RpBphP2 has previously been solved to 3.5 Å resolution (59), and we propose to use this structure as a model for molecular replacement to solve iRFP structures.

Materials and methods

Cloning: An initial clone of all five variants of iRFP and HO1 in pBAD and was sent by Dr. Verkhusha. The iRFP had a N-terminal His tag and an enterokinase recognition site in the C-terminal end. A new construct of iRFPs were made by QuickChange mutagenesis (Stratagene, La Jolla, CA) using an existing pET21a plasmid with N-terminal T7 and C-terminal hexahistidine tags (Auldridge et al., 2012).

FOR_PBAD: 5’ AGCTGGAATTCATCCATGAGCCGGGACCCG 3’

REV_PBAD: 5’AGCTCGGATCCGTGGTGGTGGTGGTG 3’

The sequences of clones were verified using DNA sequencing at the University of Wisconsin-Madison Biotechnology Center.

Protein Purification: Constructs bearing IRFP variants were grown at 37 °C in RM media supplemented with0.1 mg/ml ampicillin and kanamycin and 20% rhamnose to induce hemeoxygenase expression. At OD_{600} 0.5, iRFP expression was induced with 20% arabinose at 37 °C for 12 hours. Temperature was lowered to 18°C for 24 hours. Cells were then harvested by centrifugation at 5000 x g for 30 min, resuspended in lysis buffer (25 mM Tris buffer, pH 8.0, 50 mM NaCl, 5 mM imidazole), and lysed by sonication at an intensity of 45%, 10 seconds on and 10 seconds off. The supernatant was separated by centrifugation at 40,000 x g for 30 min, Proteins were affinity-purified under green light using nickel-nitrilotriacetic acid resin (Qiagen, Valencia, CA) and eluted with 25mM Tris, pH 8.0, 50 mM NaCl and 250mM imidazole. Further purification was carried out by size-exclusion chromatography using the Sephacryl-100 column (GE Life Sciences). Final

storage buffer of the protein is 30 mM Tris·HCl, pH 8.0. Samples were stored at 4°C and concentrated to 20 mg/ml for crystallization.

The pET21a constructs bearing iRFP variants were transformed into BL21 (DE3) expression cells and grown at 37 °C in LB-media in the presence of 0.1mg/ml ampicillin. At an O.D of 0.6, the cells were induced with isopropyl β-D-1-thiogalactopyranoside. The cells were harvested after 4 hours by centrifugation at 5000 X g and subsequently resuspended in lysis buffer (25mM Tris buffer (pH 8.0), 50 mM NaCl). In addition, the lysis buffer contained 5 mM imidazole for the DrCBD variants. Cells were sonicated and centrifuged at 40000 X g. Cell lysate was incubated with 200µl of 20mM BV in the dark for an hour. Proteins were affinity purified under green light using Ni^{2+}-NTA resin (Qiagen, Valencia, CA). Further purification was performed using a size exclusion Sephacryl-100 column (GE Life Sciences). The purified protein was subjected to absorption scans to ascertain chromophore binding (*25*) as well as shifts in maximum absorption of the holoprotein.

Mass Spectroscopy: We analyzed the Time of Flight LC-MS spectra of whole iRFP proteins at the University of Wisconsin-Madison Mass Spectrometry facility.

Crystal Trials: We have crystallized iRFP 670 by the hanging drop method of vapor diffusion with several conditions from the JCSG+ crystallization suite. However, we have been unable to replicate the results from the screen. Most of the conditions in the Cryos suite remained clear.

Results

We have successfully purified multiple iRFP variants including iRFP670 and 682 variants (pBAD vector, N-terminal His tag), co-expressed with Heme oxygenase (HO1) using a combination of affinity and size-exclusion columns. We also observed the consistent appearance of two bands on our SDS-PAGE gels even after multiple purification steps with both Ni-NTA and size-exclusion columns. The two bands could be two different polymeric states of iRFP. The molecular weight of both HO and iRFP monomer is ~32kDa which agrees the appearance of two bands in our SDS-PAGE gels (Figure 1). We need to verify the identity of the second band using mass spectroscopy.

Upon affinity purification, we consistently observed the appearance of 2 peaks after affinity purification through Ni-NTA column (Figure 1). This allowed us to separate the peak with higher specific absorbance (peak 2, Figure 1) from the lower (see table 1). Growth of cell culture in dark does not affect this finding.

For crystallization, the iRFP was concentrated to 10 and 20 mg/ml and crystal trials were conducted using the JCSG+ suite. Multiple initial hits were obtained with both 10 and 20 mg/ml concentrations (table 2). However, further optimization of the conditions did not yield crystals. Some of our reproduction issues with crystals have been resolved by optimizing the long-term storage condition of the protein to maintain <10% polydispersity. However, the reproducibility issues are also directly related to changes in BV binding due to ambient light conditions in which we are purifying the complex and setting up and

storing crystal trials. Even slight changes in complex formation could result in enough fluctuations to destabilize crystal lattice formation.

In addition to the JCSG+ screen, we also set up a Cryos suite screen with 10 mg/ml protein. Protein remained soluble in most conditions and we have followed up with increasing the protein concentration to 20 mg/ml in order to increase chances of crystallization.

In addition to purifying phytochrome from pBAD vector, we have cloned iRFP670 and 682 in pET21A vector without the extended EK recognition site and C-terminal His tag. We are also purifying and crystallizing the phytochrome protein without inducing HO and compensating by addition of exogenous biliverdin to the cell lysate. The entire purification and crystal trial is being conducted under dark (green light only) conditions to reduce experimental error and improve reproducibility of protein purification and crystal trials.

Figure 1. Large Scale purification results of iRFP variants. A. Single peak and single band on SDS-PAGE from pET21a vectors. B. Two peaks appear after affinity purification peak1 660/280:1.5/1, peak 2 660/280: 2.2/1. Further purification and crystal trials were conducted on peak 2 fractions. Purification was done in ambient light conditions.

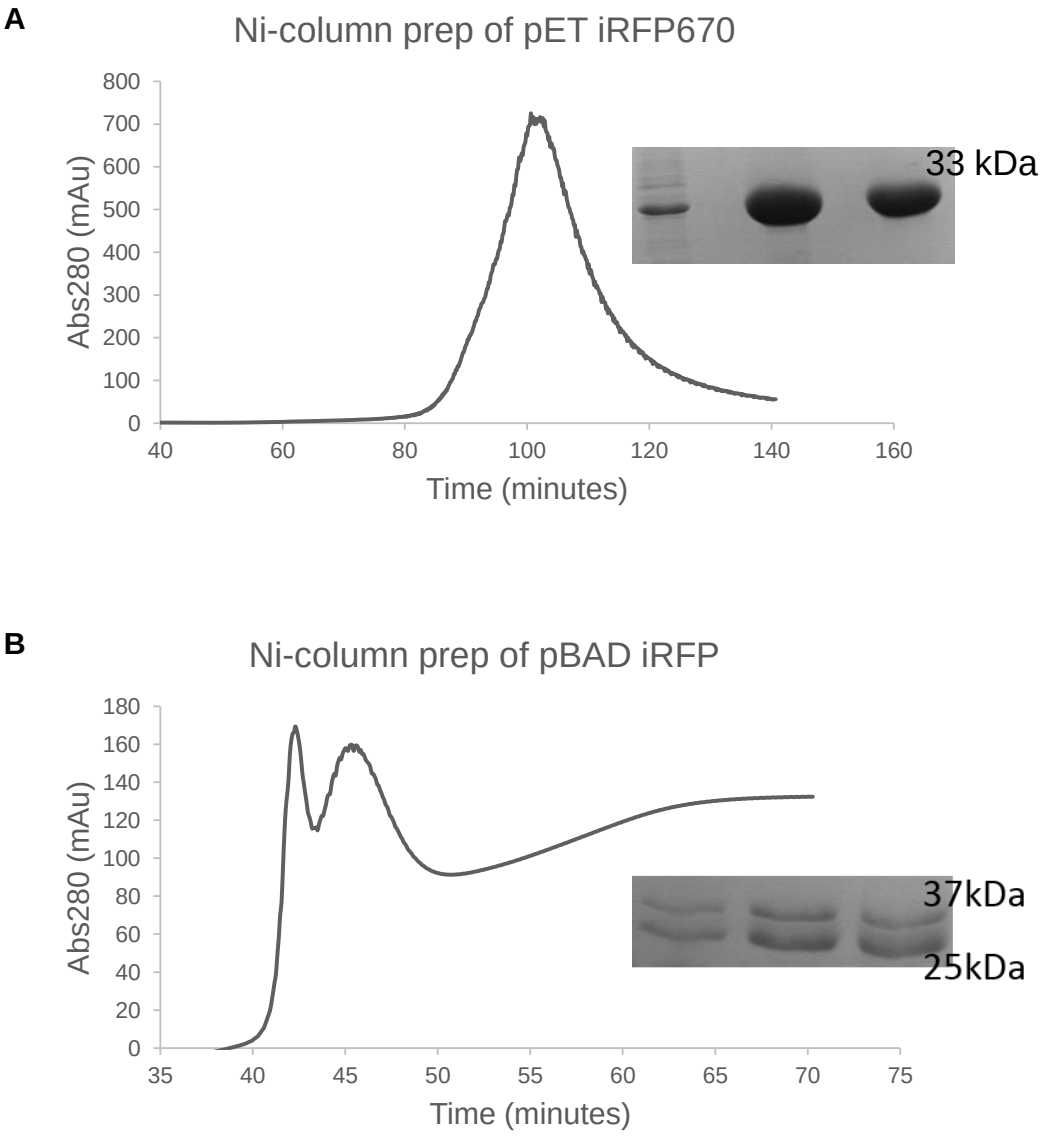


Figure 2. Final purification and crystallization of representative iRFP. A. Size exclusion chromatography was performed on (sp. abs: 2.2/1) fractions 14 and 15 on Ni column. B. SEC purified protein was concentrated and used to set up initial JCSG crystal screen. 0.1 M Na cacodylate pH 6.5, 40% MPD, 5% PEG 8000 at a concentration of 10 mg/ml. Appearance of small sharp needle/shard like crystals.

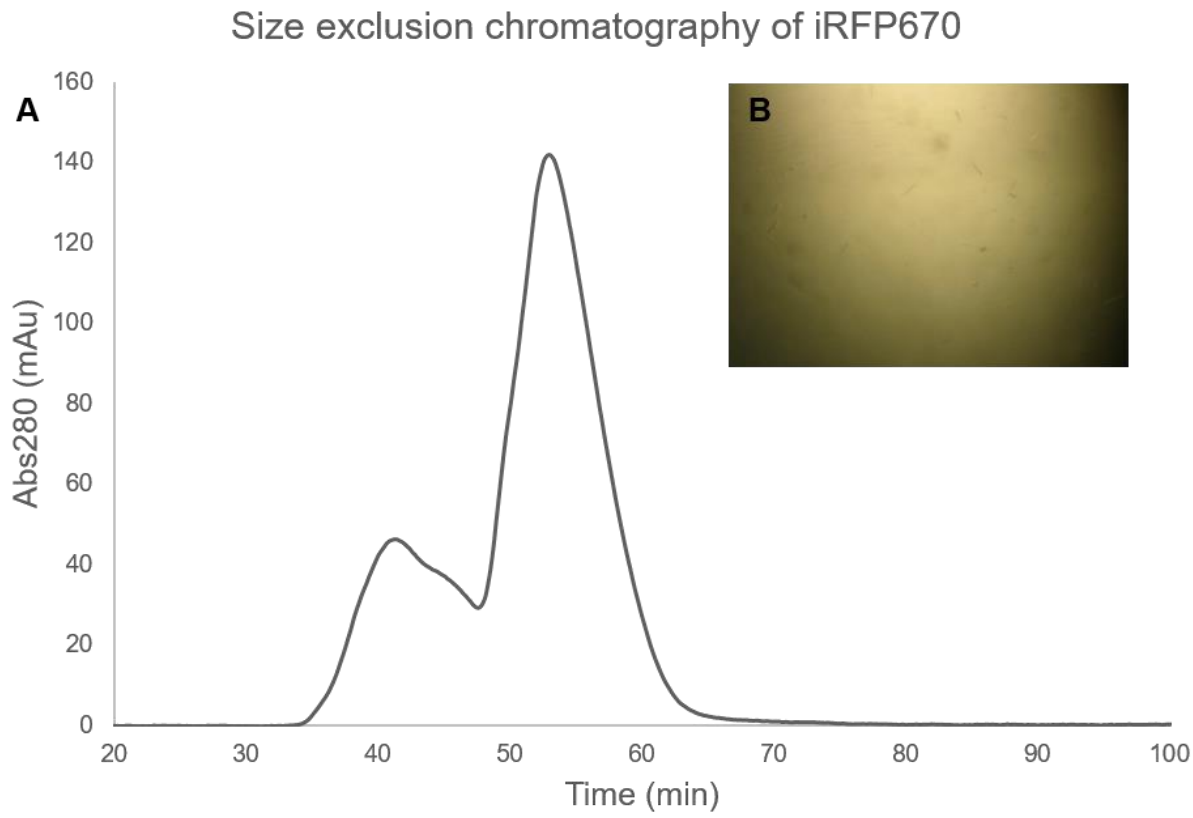


Figure 3. Mass Spectroscopy

Table 1: Purification chart of iRFPs. The best results we have obtained have been displayed

Protein sample	**Total amount purified (mg)**	**% polydispersity**	**% purity**	**Crystal conditions**	**status**
iRFP670	200	<10	97	700	+ hits
iRFP682	100	<10	97	200	N/A
iRFP702	100	<10	97	96	N/A
iRFP713	100	High	92	N/A	N/A
iRFP720	200	High	92	N/A	N/A

Table 2: Summary of crystal trials

Protein name and concentration	Buffer	Screen	Positive hit conditions
iRFP670 20 mg/ml	30 mM Tris pH 8.0	JCSG+	1 A2, 0.1 M tri-Na citrate pH 5.5, 20% PEG 3000 1C4, 0.1 M HEPES pH 7.5, 10% PEG 8000, 8% Ethylene glycol 1B5, 0.1 M Na cacodylate pH 6.5, 40% MPD, 5% PEG 8000 2B4, 0.1 M BICINE pH 9, 10% PEG 20000, 2 %v/v Dioxane 2C2, 0.2 M Magnesium chloride, 0.1 M HEPES pH 7.5, 30% PEG 400
iRFP670 10 mg/ml	30 mM Tris pH 8.0	Cryos	----

References

1. Jobsis, F. F. (1977) Noninvasive, infrared monitoring of cerebral and myocardial oxygen sufficiency and circulatory parameters. *Science*. **198**, 1264–1267
2. Wagner, J. R., Zhang, J., von Stetten, D., Günther, M., Murgida, D. H., Mroginski, M. A., Walker, J. M., Forest, K. T., Hildebrandt, P., and Vierstra, R. D. (2008) Mutational analysis of Deinococcus radiodurans bacteriophytochrome reveals key amino acids necessary for the photochromicity and proton exchange cycle of phytochromes. *J. Biol. Chem.* **283**, 12212–12226
3. Shu, X., Royant, A., Lin, M. Z., Aguilera, T. A., Lev-Ram, V., Steinbach, P. A., and Tsien, R. Y. (2009) Mammalian expression of infrared fluorescent proteins engineered from a bacterial phytochrome. *Science*. **324**, 804–807
4. Auldridge, M. E., Satyshur, K. a, Anstrom, D. M., and Forest, K. T. (2012) Structure-guided engineering enhances a phytochrome-based infrared fluorescent protein. *J. Biol. Chem.* **287**, 7000–7009
5. Filonov, G. S., Piatkevich, K. D., Ting, L.-M., Zhang, J., Kim, K., and Verkhusha, V. V (2011) Bright and stable near-infrared fluorescent protein for in vivo imaging. *Nat. Biotechnol.* **29**, 757–761
6. Shcherbakova, D. M., and Verkhusha, V. V. (2013) Near-infrared fluorescent proteins for multicolor in vivo imaging. *Nat. Methods*. **10**, 751–754
7. Bhattacharya, S., Auldridge, M. E., Lehtivuori, H., Ihalainen, J. A., and Forest, K. T. (2014) Origins of Fluorescence in Evolved Bacteriophytochromes. *J. Biol. Chem.* **289**, 32144–32152
8. Lehtivuori, H., Bhattacharya, S., Angenent-Mari, N. M., Satyshur, K. A., and Forest,

K. T. (2015) Removal of Chromophore-Proximal Polar Atoms Decreases Water Content and Increases Fluorescence in a Near Infrared Phytofluor. *Front. Mol. Biosci.* **2**, 65–69

9. Stepanenko, O. V., Verkhusha, V. V., Kuznetsova, I. M., Uversky, V. N., and Turoverov, K. K. (2008) Fluorescent proteins as biomarkers and biosensors: throwing color lights on molecular and cellular processes. *Curr. Protein Pept. Sci.* **9**, 338–369
10. Crameri, A., Whitehorn, E. A., Tate, E., and Stemmer, W. P. (1996) Improved green fluorescent protein by molecular evolution using DNA shuffling. *Nat Biotechnol*. **14**, 315–319
11. Yoshihara, S., Shimada, T., Matsuoka, D., Zikihara, K., Kohchi, T., and Tokutomi, S. (2006) Reconstitution of Blue−Green Reversible Photoconversion of a Cyanobacterial Photoreceptor, PixJ1, in Phycocyanobilin-Producing Escherichia coli. *Biochemistry*. **45**, 3775–3784
12. Tsien, R. (1998) The green fluorescent protein. *Annu. Rev. Biochem.* **67**, 509
13. Ando, R., Hama, H., Yamamoto-Hino, M., Mizuno, H., and Miyawaki, A. (2002) An optical marker based on the UV-induced green-to-red photoconversion of a fluorescent protein. *Proc. Natl. Acad. Sci.* **99**, 12651–12656
14. Ehrig, T., O'Kane, D. J., and Prendergast, F. G. (1995) Green-fluorescent protein mutants with altered fluorescence excitation spectra. *FEBS Lett*. **367**, 163–166
15. Reeder, P. J., Huang, Y.-M., Dordick, J. S., and Bystroff, C. (2010) A Rewired Green Fluorescent Protein: Folding and Function in a Nonsequential, Noncircular GFP Permutant. *Biochemistry*. **49**, 10773–10779

16. Hanson, G. T., McAnaney, T. B., Park, E. S., Rendell, M. E. P., Yarbrough, D. K., Chu, S., Xi, L., Boxer, S. G., Montrose, M. H., and Remington, S. J. (2002) Green Fluorescent Protein Variants as Ratiometric Dual Emission pH Sensors. 1. Structural Characterization and Preliminary Application. *Biochemistry*. **41**, 15477–15488

17. Davis, S. J., and Vierstra, R. D. (1998) Soluble, highly fluorescent variants of green fluorescent protein (GFP) for use in higher plants. *Plant Mol. Biol.* **36**, 521–528

18. Baulcombe, D. C., Chapman, S., and Cruz, S. (1995) Jellyfish green fluorescent protein as a reporter for virus infections. *Plant J.* **7**, 1045–1053

19. Peng, S.-Y., Yang, Y.-S., Chou, C.-J., Lin, K.-Y., and Wu, S.-C. (2015) Differentiation of Enhanced Green Fluorescent Protein-Labeled Mouse Amniotic Fluid-Derived Stem Cells into Cardiomyocyte-Like Beating Cells. *Acta Cardiol. Sin.* **31**, 209–214

20. Boas, D. A., Gaudette, T., Strangman, G., Cheng, X., Marota, J. J. A., and Mandeville, J. B. (2001) The Accuracy of Near Infrared Spectroscopy and Imaging during Focal Changes in Cerebral Hemodynamics. *Neuroimage*. **13**, 76–90

21. Mancini, D. M., Bolinger, L., Li, H., Kendrick, K., Chance, B., and Wilson, J. R. (1994) Validation of near-infrared spectroscopy in humans. *J. Appl. Physiol.* **77**, 2740–2747

22. Hilderbrand, S. A., and Weissleder, R. (2010) Near-infrared fluorescence: application to in vivo molecular imaging. *Curr. Opin. Chem. Biol.* **14**, 71–79

23. Luo, S., Zhang, E., Su, Y., Cheng, T., and Shi, C. (2011) A review of NIR dyes in cancer targeting and imaging. *Biomaterials*. **32**, 7127–7138

24. Ke, S., Wen, X., Gurfinkel, M., Charnsangavej, C., Wallace, S., Sevick-Muraca, E. M., and Li, C. (2003) Near-Infrared Optical Imaging of Epidermal Growth Factor Receptor in Breast Cancer Xenografts. *Cancer Res.* **63**, 7870–7875

25. Bugaj, J. E., Achilefu, S., Dorshow, R. B., and Rajagopalan, R. (2001) Novel fluorescent contrast agents for optical imaging of in vivo tumors based on a receptor-targeted dye-peptide conjugate platform. *J. Biomed. Opt.* **6**, 122–132

26. Achilefu, S., Dorshow, R. B., Bugaj, J. E., and Rajagopalan, R. (2000) Novel receptor-targeted fluorescent contrast agents for in vivo tumor imaging. *Invest. Radiol.* **35**, 479–485

27. Bongarzone, S., Staderini, M., and Bolognesi, M. L. (2014) Multitarget ligands and theranostics: sharpening the medicinal chemistry sword against prion diseases. *Future Med. Chem.* **6**, 1017–1029

28. Doerr, A. (2009) Fluorescent proteins: into the infrared. *Nat. Methods*. **6**, 482–483

29. Cashmore, A. R. (2003) Cryptochromes: Enabling plants and animals to determine circadian time. *Cell*. **114**, 537–543

30. Murphy, J. T., and Lagarias, J. C. (1997) The phytofluors: a new class of fluorescent protein probes. *Curr. Biol.* **7**, 870–876

31. Christie, J. M., Gawthorne, J., Young, G., Fraser, N. J., and Roe, A. J. (2012) LOV to BLUF: Flavoprotein Contributions to the Optogenetic Toolkit. *Mol. Plant*. **5**, 533–544

32. Pudasaini, A., El-Arab, K. K., and Zoltowski, B. D. (2015) LOV-based optogenetic devices: light-driven modules to impart photoregulated control of cellular signaling. *Front. Mol. Biosci.* **2**, 18–22

33. Möglich, A., Yang, X., Ayers, R. A., and Moffat, K. (2010) Structure and Function of Plant Photoreceptors. *Annu. Rev. Plant Biol.* **61**, 21–47

34. Rockwell, N. C., and Lagarias, J. C. (2010) A Brief History of Phytochromes. *ChemPhysChem*. **11**, 1172–1180

35. Auldridge, M. E., and Forest, K. T. (2011) Bacterial phytochromes: more than meets the light. *Crit. Rev. Biochem. Mol. Biol.* **46**, 67–88

36. Al-Sady, B., Ni, W., Kircher, S., Schaefer, E., and Quail, P. H. (2006) Photoactivated Phytochrome Induces Rapid PIF3 Phosphorylation Prior to Proteasome-Mediated Degradation. *Mol. Cell*. **23**, 439–446

37. Fankhauser, C. (2001) The Phytochromes, a Family of Red/Far-red Absorbing Photoreceptors. *J. Biol. Chem.* **276**, 11453–11456

38. Rizzini, L., Favory, J.-J., Cloix, C., Faggionato, D., O'Hara, A., Kaiserli, E., Baumeister, R., Schäfer, E., Nagy, F., Jenkins, G. I., and Ulm, R. (2011) Perception of UV-B by the Arabidopsis UVR8 protein. *Science*. **332**, 103–106

39. Quail, P. H. (2002) Phytochrome photosensory signalling networks. *Nat. Rev. Mol. Cell Biol.* **3**, 85–93

40. Kehoe, D. M., and Grossman, A. R. (1996) Similarity of a chromatic adaptation sensor to phytochrome and ethylene receptors. *Science*. **273**, 1409–1412

41. Terauchi, K., Montgomery, B. L., Grossman, A. R., Lagarias, J. C., and Kehoe, D. M. (2004) RcaE is a complementary chromatic adaptation photoreceptor required for green and red light responsiveness. *Mol. Microbiol.* **51**, 567–577

42. Davis, S. J., Vener, A. V, and Vierstra, R. D. (1999) Bacteriophytochromes: phytochrome-like photoreceptors from nonphotosynthetic eubacteria. *Science*.

286, 2517–2520

43. Bhoo, S. H., Davis, S. J., Walker, J., Karniol, B., and Vierstra, R. D. (2001) Bacteriophytochromes are photochromic histidine kinases using a biliverdin chromophore. *Nature*. **414**, 776–779
44. Butler, W. L., Norris, K. H., Siegelman, H. W., and Hendricks, S. B. (1959) Detection, assay, and preliminary purification of the pigment controlling photoresponsive development of plants. *Proc. Natl. Acad. Sci. U. S. A.* **45**, 1703–1708
45. Blaauw-Jansen, G. (1959) The influence of red and far red light on growth and phototropism of the avena seedling. *Acta Bot. Neerl.* **8**, 1–39
46. Wagner, J. R., Brunzelle, J. S., Forest, K. T., and Vierstra, R. D. (2005) A light-sensing knot revealed by the structure of the chromophore-binding domain of phytochrome. *Nature*. **438**, 325–331
47. Wagner, J. R., Zhang, J., Brunzelle, J. S., Vierstra, R. D., and Forest, K. T. (2007) High resolution structure of Deinococcus bacteriophytochrome yields new insights into phytochrome architecture and evolution. *J. Biol. Chem.* **282**, 12298–12309
48. Yang, X., Stojkovic, E. A., Kuk, J., and Moffat, K. (2007) Crystal structure of the chromophore binding domain of an unusual bacteriophytochrome, RpBphP3, reveals residues that modulate photoconversion. *Proc. Natl. Acad. Sci.* **104**, 12571–12576
49. Nagatani, A. (2010) Phytochrome: structural basis for its functions. *Curr. Opin. Plant Biol.* **13**, 565–570
50. Takala, H., Björling, A., Berntsson, O., Lehtivuori, H., Niebling, S., Hoernke, M.,

Kosheleva, I., Henning, R., Menzel, A., Ihalainen, J. A., and Westenhoff, S. (2014) Signal amplification and transduction in phytochrome photosensors. *Nature*. **509**, 245–248

51. Baker, A. W., and Forest, K. T. (2014) Structural biology: Action at a distance in a light receptor. *Nature*. **509**, 174–175
52. Burgie, E. S., Wang, T., Bussell, A. N., Walker, J. M., Li, H., and Vierstra, R. D. (2014) Crystallographic and Electron Microscopic Analyses of a Bacterial Phytochrome Reveal Local and Global Rearrangements During Photoconversion. *J. Biol. Chem.* **289**, 24573–24587
53. Yang, X., Kuk, J., and Moffat, K. (2009) Conformational differences between the Pfr and Pr states in Pseudomonas aeruginosa bacteriophytochrome. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 15639–15644
54. Bischoff, M., Hermann, G., Rentsch, S., and Strehlow, D. (2001) First Steps in the Phytochrome Phototransformation: A Comparative Femtosecond Study on the Forward (Pr → Pfr) and Back Reaction (Pfr → Pr). *Biochemistry*. **40**, 181–186
55. Andel, F., Hasson, K. C., Gai, F., Anfinrud, P. A., and Mathies, R. A. (1997) Femtosecond time-resolved spectroscopy of the primary photochemistry of phytochrome. *Biospectroscopy*. **3**, 421–433
56. Toh, K. C., Stojkovic, E. A., van Stokkum, I. H., Moffat, K., and Kennis, J. T. (2011) Fluorescence quantum yield and photochemistry of bacteriophytochrome constructs. *Phys Chem Chem Phys*. **13**, 11985–11997
57. Dasgupta, J., Frontiera, R. R., Taylor, K. C., Lagarias, J. C., and Mathies, R. A. (2009) Ultrafast excited-state isomerization in phytochrome revealed by

femtosecond stimulated Raman spectroscopy. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 1784–1789

58. Piatkevich, K. D., Subach, F. V, and Verkhusha, V. V. (2013) Engineering of bacterial phytochromes for near-infrared imaging, sensing, and light-control in mammals. *Chem. Soc. Rev.* **42**, 3441–3452

59. Yang, X., Stojković, E. A., Ozarowski, W. B., Kuk, J., Davydova, E., and Moffat, K. (2015) Light Signaling Mechanism of Two Tandem Bacteriophytochromes. *Structure*. **23**, 1179–1189

60. Rockwell, N. C., Su, Y.-S., and Lagarias, J. C. (2006) Phytochrome structure and signaling mechanisms. *Annu. Rev. Plant Biol.* **57**, 837–858

61. Kim, P. W., Freer, L. H., Rockwell, N. C., Martin, S. S., Lagarias, J. C., and Larsen, D. S. (2012) Femtosecond Photodynamics of the Red/Green Cyanobacteriochrome NpR6012g4 from *Nostoc punctiforme* . 2. Reverse Dynamics. *Biochemistry*. **51**, 619–630

62. Ulijasz, A. T., Cornilescu, G., Cornilescu, C. C., Zhang, J., Rivera, M., Markley, J. L., and Vierstra, R. D. (2010) Structural basis for the photoconversion of a phytochrome to the activated Pfr form. *Nature*. **463**, 250–254

63. Kneip, C., Hildebrandt, P., Schlamann, W., Braslavsky, S. E., Mark, F., and Schaffner, K. (1999) Protonation state and structural changes of the tetrapyrrole chromophore during the Pr/Pfr phototransformation of phytochrome: a resonance Raman spectroscopic study. *Biochemistry*. **38**, 15185–15192

64. Spillane, K. M., Dasgupta, J., and Mathies, R. A. (2012) Conformational Homogeneity and Excited-State Isomerization Dynamics of the Bilin Chromophore

in Phytochrome Cph1 from Resonance Raman Intensities. *Biophys. J.* **102**, 709–717

65. Toh, K. C., Stojkovic, E. A., van Stokkum, I. H. M., Moffat, K., and Kennis, J. T. M. (2010) Proton-transfer and hydrogen-bond interactions determine fluorescence quantum yield and photochemical efficiency of bacteriophytochrome. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 9170–9175
66. Cormack, B. P., Valdivia, R. H., and Falkow, S. (1996) FACS-optimized mutants of the green fluorescent protein (GFP). *Gene*. **173**, 33–38
67. Heim, R., and Tsien, R. Y. (1996) Engineering green fluorescent protein for improved brightness, longer wavelengths and fluorescence resonance energy transfer. *Curr Biol*. **6**, 178–182
68. Afar, B., Merrill, J., and Clark, E. A. (1991) Detection of lymphocyte subsets using three-color/single-laser flow cytometry and the fluorescent dye Peridinin chlorophyll-a protein. *J. Clin. Immunol.* **11**, 254–261
69. Nakajima, O., and Gray, C. H. (1967) Studies on haem alpha-methenyl oxygenase. Isomeric structure of formylbiliverdin, a possible precursor of biliverdin. *Biochem. J.* **104**, 20–22
70. Schäfer, F. P. (1973) Principles of dye laser operation. in *n Topics in Applied Physics Dye Lasers* (Schäfer, F. P. ed), pp. 1–89, Springer-Verlag, New York, 10.1007/3-540-51558-5_7
71. Filonov, G. S., Krumholz, A., Xia, J., Yao, J., Wang, L. V, and Verkhusha, V. V. (2012) Deep-tissue photoacoustic tomography of a genetically encoded near-infrared fluorescent probe. *Angew Chem Int Ed Engl*. **51**, 1448–1451

72. Bartz-Schmidt, K. U., Walter, P., Krott, R., Brunner, R., Esser, P., and Heimann, K. (1996) [Effects of fluorescein and indocyanine green angiography on subsequent dark adaptation and the electroretinogram]. *Klin Monbl Augenheilkd*. **208**, 224–228

73. Taylor, W. R. (2000) A deeply knotted protein structure and how it might fold. *Nature*. **406**, 916–919

74. Khatib, F., Weirauch, M. T., and Rohl, C. A. (2006) Rapid knot detection and application to protein structure prediction. *Bioinformatics*. **22**, e252–e259

75. Mallam, A. L., Morris, E. R., and Jackson, S. E. (2008) Exploring knotting mechanisms in protein folding. *Proc. Natl. Acad. Sci.* **105**, 18740–18745

76. Mallam, A. L., and Jackson, S. E. (2006) Probing Nature's Knots: The Folding Pathway of a Knotted Homodimeric Protein. *J. Mol. Biol.* **359**, 1420–1436

77. Mallam, A. L., Rogers, J. M., and Jackson, S. E. (2010) Experimental detection of knotted conformations in denatured proteins. *Proc Natl Acad Sci U S A*. **107**, 8189–8194

78. King, N. P., Yeates, E. O., and Yeates, T. O. (2007) Identification of Rare Slipknots in Proteins and Their Implications for Stability and Folding. **373**, 153–166

79. King, N. P., Jacobitz, A. W., Sawaya, M. R., Goldschmidt, L., and Yeates, T. O. (2010) Structure and folding of a designed knotted protein. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 20732–20737

80. Dill, K. A., and MacCallum, J. L. (2012) The protein-folding problem, 50 years on. **338**, 1042–1046

81. Andersson, F. I., Pina, D. G., Mallam, A. L., Blaser, G., and Jackson, S. E. (2009) Untangling the folding mechanism of the 52-knotted protein UCH-L3. **276**, 2625–

2635

82. Bornschlogl, T., Anstrom, D. M., Mey, E., Dzubiella, J., Rief, M., and Forest, K. T. (2009) Tightening the knot in phytochrome by single-molecule atomic force microscopy. *Biophys J*. **96**, 1508–1514

83. von Stetten, D., Seibeck, S., Michael, N., Scheerer, P., Mroginski, M. A., Murgida, D. H., Krauss, N., Heyn, M. P., Hildebrandt, P., Borucki, B., and Lamparter, T. (2007) Highly conserved residues Asp-197 and His-250 in Agp1 phytochrome control the proton affinity of the chromophore and Pfr formation. *J. Biol. Chem.* **282**, 2116–2123

84. Anders Borg, O., and Durbeej, B. (2008) Which factors determine the acidity of the phytochromobilin chromophore of plant phytochrome? *Phys. Chem. Chem. Phys.* **10**, 2528–2537

85. Lehtivuori, H., Bhattacharya, S., Angenent-Mari, N., Satyshur, K. A., and Forest, K. T. (2015) Removal of Chromophore-Proximal Polar Atoms Decreases Water Content and Increases Fluorescence in a Near Infrared Phytofluor. *Front. Mol. Biosci.* **2**, 65–69

86. Spillane, K. M., Dasgupta, J., Lagarias, J. C., and Mathies, R. A. (2009) Homogeneity of phytochrome Cph1 vibronic absorption revealed by resonance Raman intensity analysis. *J Am Chem Soc*. **131**, 13946–13948

87. Essen, L.-O., Mailliet, J., and Hughes, J. (2008) The structure of a complete phytochrome sensory module in the Pr ground state. *Proc. Natl. Acad. Sci.* **105**, 14709–14714

88. Yang, X., Kuk, J., and Moffat, K. (2008) Crystal structure of Pseudomonas

aeruginosa bacteriophytochrome: photoconversion and signal transduction. *Proc. Natl. Acad. Sci.* **105**, 14715–14720

89. von Stetten, D., Gunther, M., Scheerer, P., Murgida, D. H., Mroginski, M. A., Krauss, N., Lamparter, T., Zhang, J., Anstrom, D. M., Vierstra, R. D., Forest, K. T., and Hildebrandt, P. (2008) Chromophore heterogeneity and photoconversion in phytochrome crystals and solution studied by resonance Raman spectroscopy. *Angew Chem Int Ed Engl*. **47**, 4753–4755

90. Dasgupta, J., Frontiera, R. R., Taylor, K. C., Lagarias, J. C., and Mathies, R. A. (2009) Ultrafast excited-state isomerization in phytochrome revealed by femtosecond stimulated Raman spectroscopy. *Proc. Natl. Acad. Sci.* **106**, 1784–1789

91. Foerstendorf, H., Mummert, E., Sch??fer, E., Scheer, H., and Siebert, F. (1996) Fourier-transform infrared spectroscopy of phytochrome: Difference spectra of the intermediates of the photoreactions. *Biochemistry*. **35**, 10793–10799

92. Toh, K. C., van Stokkum, I. H., Hendriks, J., Alexandre, M. T., Arents, J. C., Perez, M. A., van Grondelle, R., Hellingwerf, K. J., and Kennis, J. T. (2008) On the signaling mechanism and the absence of photoreversibility in the AppA BLUF domain. *Biophys J*. **95**, 312–321

93. Fischer, A. J., and Lagarias, J. C. (2004) Harnessing phytochrome's glowing potential. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 17334–17339

94. Lehtivuori, H., Rissanen, I., Takala, H., Bamford, J., Tkachenko, N. V., and Ihalainen, J. A. (2013) Fluorescence Properties of the Chromophore-Binding Domain of Bacteriophytochrome from *Deinococcus radiodurans*. *J. Phys. Chem. B*.

117, 11049–11057

95. Otwinowski, Z., and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326

96. McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007) Phaser crystallographic software. *J Appl Crystallogr*. **40**, 658–674

97. Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F., and Vagin, A. A. (2011) REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. D. Biol. Crystallogr.* **67**, 355–367

98. Hahn, J., Strauss, H. M., Landgraf, F. T., Gimenèz, H. F., Lochnit, G., Schmieder, P., and Hughes, J. (2006) Probing protein-chromophore interactions in Cph1 phytochrome by mutagenesis. *FEBS J.* **273**, 1415–1429

99. Jamroz, M., Niemyska, W., Rawdon, E. J., Stasiak, A., Millett, K. C., Sułkowski, P., and Sulkowska, J. I. (2015) KnotProt: a database of proteins with knots and slipknots. *Nucleic Acids Res.* **43**, D306–D314

100. Lai, Y.-L., Yen, S.-C., Yu, S.-H., and Hwang, J.-K. (2007) pKNOT: the protein KNOT web server. *Nucleic Acids Res.* **35**, W420–W424

101. Millett, K. C., Rawdon, E. J., Stasiak, A., and Sułkowska, J. I. (2013) Identifying knots in proteins. *Biochem. Soc. Trans.* **41**, 533–537

102. Burgie, E. S., and Vierstra, R. D. (2014) Phytochromes: An Atomic Perspective on Photoactivation and Signaling. *Plant Cell Online*. **26**, 4568–4583

103. Rockwell, N. C., Martin, S. S., Gulevich, A. G., and Lagarias, J. C. (2012)

Phycoviolobilin Formation and Spectral Tuning in the DXCF Cyanobacteriochrome Subfamily. *Biochemistry*. **51**, 1449–1463

104. Sineshchekov, V., Mailliet, J., Psakis, G., Feilke, K., Kopycki, J., Zeidler, M., Essen, L.-O., and Hughes, J. (2014) Tyrosine 263 in cyanobacterial phytochrome Cph1 optimizes photochemistry at the prelumi-R→lumi-R step. *Photochem. Photobiol.* **90**, 786–795
105. Narikawa, R., Fukushima, Y., Ishizuka, T., Itoh, S., and Ikeuchi, M. (2008) A Novel Photoactive GAF Domain of Cyanobacteriochrome AnPixJ That Shows Reversible Green/Red Photoconversion. *J. Mol. Biol.* **380**, 844–855
106. Narikawa, R., Ishizuka, T., Muraki, N., Shiba, T., Kurisu, G., and Ikeuchi, M. (2013) Structures of cyanobacteriochromes from phototaxis regulators AnPixJ and TePixJ reveal general and specific photoconversion mechanism. *Proc. Natl. Acad. Sci.* **110**, 918–923
107. Narikawa, R., Nakajima, T., Aono, Y., Fushimi, K., Enomoto, G., Ni-Ni-Win, Itoh, S., Sato, M., Ikeuchi, M., Narikawa, R., Enomoto, G., Ni, W. N., Fushimi, K., and Ikeuchi, M. (2015) A biliverdin-binding cyanobacteriochrome from the chlorophyll d–bearing cyanobacterium Acaryochloris marina. *Sci. Rep.* **5**, 7950–7975
108. Zhang, J., Wu, X.-J., Wang, Z.-B., Chen, Y., Wang, X., Zhou, M., Scheer, H., and Zhao, K.-H. (2010) Fused-Gene Approach to Photoswitchable and Fluorescent Biliproteins. *Angew. Chemie Int. Ed.* **49**, 5456–5458
109. Rockwell, N. C., Martin, S. S., and Lagarias, J. C. (2012) Red/Green Cyanobacteriochromes: Sensors of Color and Power. *Biochemistry*. **51**, 9667–9677
110. Pal, D., and Chakrabarti, P. (1999) Cis peptide bonds in proteins: residues involved,

their conformations, interactions and locations. *J. Mol. Biol.* **294**, 271–288

111. Joseph, A. P., Srinivasan, N., and de Brevern, A. G. (2012) Cis-trans peptide variations in structurally similar proteins. *Amino Acids*. **43**, 1369–1381
112. MacArthur, M. W., and Thornton, J. M. (1991) Influence of proline residues on protein conformation. *J. Mol. Biol.* **218**, 397–412
113. Fischer, G., and Schmid, F. X. (1990) The mechanism of protein folding. Implications of in vitro refolding models for de novo protein folding and translocation in the cell. *Biochemistry*. **29**, 2205–2212
114. Hesterkamp, T., and Bukau, B. (1996) Identification of the prolyl isomerase domain of Escherichia coli trigger factor. *FEBS Lett.* **385**, 67–71
115. Kramer, G., Patzelt, H., Rauch, T., Kurz, T. A., Vorderwülbecke, S., Bukau, B., and Deuerling, E. (2004) Trigger Factor Peptidyl-prolyl cis/trans Isomerase Activity Is Not Essential for the Folding of Cytosolic Proteins in Escherichia coli. *J. Biol. Chem.* **279**, 14165–14170
116. Stoller, G., Rücknagel, K. P., Nierhaus, K. H., Schmid, F. X., Fischer, G., and Rahfeld, J. U. (1995) A ribosome-associated peptidyl-prolyl cis/trans isomerase identified as the trigger factor. *EMBO J.* **14**, 4939–4948
117. Kramer, G., Rauch, T., Rist, W., Vorderwülbecke, S., Patzelt, H., Schulze-Specking, A., Ban, N., Deuerling, E., and Bukau, B. (2002) L23 protein functions as a chaperone docking site on the ribosome. *Nature*. **419**, 171–174
118. Gupta, R., Lakshmipathy, S. K., Chang, H.-C., Etchells, S. A., and Hartl, F. U. (2010) Trigger factor lacking the PPIase domain can enhance the folding of eukaryotic multi-domain proteins in Escherichia coli. *FEBS Lett.* **584**, 3620–3624

119. Ferbitz, L., Maier, T., Patzelt, H., Bukau, B., Deuerling, E., and Ban, N. (2004) Trigger factor in complex with the ribosome forms a molecular cradle for nascent proteins. *Nature*. **431**, 590–596

120. Thomason, L. C., Costantino, N., Court, D. L., Thomason, L. C., Costantino, N., and Court, D. L. (2007) *E. coli* Genome Manipulation by P1 Transduction. in *Current Protocols in Molecular Biology*, pp. 1.17.1–1.17.8, John Wiley & Sons, Inc., Hoboken, NJ, USA, 10.1002/0471142727.mb0117s79

121. Burgie, E. S., Walker, J. M., Phillips, G. N., and Vierstra, R. D. (2013) A Photo-Labile Thioether Linkage to Phycoviolobilin Provides the Foundation for the Blue/Green Photocycles in DXCF-Cyanobacteriochromes. *Structure*. **21**, 88–97

122. Frey, U. H., Bachmann, H. S., Peters, J., and Siffert, W. (2008) PCR-amplification of GC-rich regions: “slowdown PCR.” *Nat. Protoc.* **3**, 1312–1317

123. Gibson, D. G., Young, L., Chuang, R.-Y., Venter, J. C., Hutchison, C. A., and Smith, H. O. (2009) Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat Meth*. **6**, 343–345

124. Schneider, C. A., Rasband, W. S., and Eliceiri, K. W. (2012) NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods*. **9**, 671–675

125. Rumyantsev, K. A., Shcherbakova, D. M., Zakharova, N. I., Emelyanov, A. V., Turoverov, K. K., and Verkhusha, V. V. (2015) Minimal domain of bacterial phytochrome required for chromophore binding and fluorescence. *Sci. Rep.* **5**, 18348–18352

126. Mallam, A. L. (2009) How does a knotted protein fold? *FEBS J.* **276**, 365–375

127. Sułkowska, J. I., Noel, J. K., Ramírez-Sarmiento, Esar A, Rawdon, E. J., Millett, K.

C., and Onuchic, J. N. (2013) Knotting pathways in proteins. *Biochem. Soc. Trans.* **41**, 523–527

128. Mallam, A. L., and Jackson, S. E. (2011) Knot formation in newly translated proteins is spontaneous and accelerated by chaperonins. *Nat. Chem. Biol.* **8**, 147–153
129. Yang, X., Stojković, E. A., Ozarowski, W. B., Kuk, J., Davydova, E., and Moffat, K. (2015) Light Signaling Mechanism of Two Tandem Bacteriophytochromes. *Structure*. **23**, 1179–1189
130. Yang, X., Ren, Z., Kuk, J., and Moffat, K. (2011) Temperature-scan cryocrystallography reveals reaction intermediates in bacteriophytochrome. *Nature*. **479**, 428–432
131. Yang, X., Stojkovic, E. A., Kuk, J., and Moffat, K. (2007) Crystal structure of the chromophore binding domain of an unusual bacteriophytochrome, RpBphP3, reveals residues that modulate photoconversion. *Proc Natl Acad Sci U S A*. **104**, 12571–12576
132. Burgie, E. S., Bussell, A. N., Walker, J. M., Dubiel, K., and Vierstra, R. D. (2014) Crystal structure of the photosensing module from a red/far-red light-absorbing plant phytochrome. *Proc. Natl. Acad. Sci.* **111**, 10179–10184
133. Yu, D., Gustafson, W. C., Han, C., Lafaye, C., Noirclerc-Savoye, M., Ge, W.-P., Thayer, D. A., Huang, H., Kornberg, T. B., Royant, A., Jan, L. Y., Jan, Y. N., Weiss, W. A., and Shu, X. (2014) An improved monomeric infrared fluorescent protein for neuronal and tumour brain imaging. *Nat. Commun.* **5**, 3626–3632
134. Giraud, E., Zappa, S., Vuillet, L., Adriano, J.-M., Hannibal, L., Fardoux, J., Berthomieu, C., Bouyer, P., Pignol, D., and Verméglio, A. (2005) A new type of

bacteriophytochrome acts in tandem with a classical bacteriophytochrome to control the antennae synthesis in Rhodopseudomonas palustris. *J. Biol. Chem.* **280**, 32389–32397

Appendix II. List of primers for chapter IV

Primer names	**Sequence**
knotlessRA-forward	5' CCGCCGCCGTGCCGCTCGATGGTGTCCTCAA 3'
knotlessRA-reverse	5' TGCGGGTTGAGGACACCATCGAGCGGCACGG 3'
P236G-forward	5’ CGTGCCGCTCGATGGAGTCCTCAACCCG 3’
P236G-reverse	5’ CGGGTTGAGGACTCCATCGAGCGGCACG 3’
P236-245-247G-Gibson 1	5'CATGTGCATGGGCGAGGTGGCGCGCAGCACGGCGCCGCCCAGACCGGTACCCGCATTCGTCTGCGGGTTGAG 3'
P236-245-247G-Gibson 2	5' GGCGAACTACTTACTCTAGCTTCCCG 3'
P236-245-247G-Gibson 3	5' CCACCTCGCCCATGCACATG 3'