Expression, purification and crystallization of phyB and characterization of putative PIF1 from *Sorghum bicolor*

Dissertation

zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

Am Fachbereich Biologie und Chemie der

Justus-Liebig-Universitat Giesen

vorgelegt von

Sintayehu Manaye Shenkutie

aus

Addis Ababa (Ethiopia)

Giessen 2017

Erstgutachter:

Prof. Ph.D. Jon Hughes

Fachbereich Biologie u. Chemie, Justus-Liebig-Universitat Giesen

Zweitgutachter:

Prof. Dr. Lars-Oliver Essen

Fachbereich Chemie, Philipps-Universitat Marburg

Prof. Dr. Guenter Lochnit

Institut des Fachbereichs Medizin der Justus-Liebig-Universität, Giessen

Prof. Dr. Gabriele Klug

Institut für Mikrobiologie und Molekularbiologie, Justus-Liebig-Universitat Giesen

Part of this work has been accepted for publication:

Francisco Velázquez Escobar, David Buhrke, Maria Fernandez Lopez, Sintayehu Manaye Shenkutie, Silke von Horsten, Lars-Oliver Essen, Jon Hughes, and Peter Hildebrandt.

Structural communication between the chromophore binding pocket and the N-terminal extension in plant phytochrome phyB

FEBS Letters: 2017

Contents

1 Introduction

Sunlight played a fundamental role in the origin of life and its subsequent evolution into highly complex and diverse living system. Light is, therefore, one of the most important environmental factors for living organisms. Consequently, in all kingdoms, most species are able to sense light through a set of photoreceptors. These sensory proteins enable living things to detect the quantity and quality of light and to transduce that physical signal into biochemical outputs which entrain their metabolism with the ambient light environment [1].

1.1 Prokaryotic phytochromes

Phytochromes are biliprotein photoreceptors ubiquitous in plants [2, 3], widespread in bacteria [4-7] and known in fungi [8, 9]. Linear tetrapyrroles (see Figure 1.1) such as biliverdin (BV), phycocyanobilin (PCB), phytochromobilin (PΦB), phycoviolobilin (PVB), phycoerythrobilin (PEB) and phycourobilin (PUB) are collectively known as bilins [1]. Bilins function as prosthetic groups for phytochrome photoreceptors [10-12]. Heme oxygenase catalyzes the oxidative cleavage of the heme ring to form biliverdin [13]. Biliverdin is subsequently converted to PCB or PΦB by ferredoxin-dependent bilin reductases (FDBR) [14]. The hallmarks of phytochrome photoreceptors [27] are bilin lyase activity [15], photoreversibility [2, 16] and signal transduction [7, 8].

The bilin lyase domain (BLD) has been defined as the minimal phytochrome structure that is capable of autocatalytic bilin assembly [15, 17]. All phytochromes show autocatalytic assembly with certain bilin cofactors [15]. However, the minimal structure that is necessary and sufficient for covalent attachment of bilin is different for different classes of phytochromes [15, 18]. The chromophore structure and the precise nature of the chromophore-protein linkage had been definitively established for cyanobacterial and plant phytochromes [12, 19].

Photoreversibility refers to a phenomenon in which physiological responses to weak red light could be nullified if the red treatment were followed immediately with far-red light. Phytochrome photoreversibility involves *cis-trans* (Z-E) isomerisation of the chromophore D-ring [19-23]. In red/far-red photoreversible phytochromes, Pr is generally the lowest energy ground state that absorbs exclusively in the red region (λ_{max} ~660 nm) whereas Pfr is a more or less thermally stable ground state that absorbs predominantly in the far-red region (λ_{max} ~720 nm) [2, 24]. In

prototypical phytochromes, Pr is more thermally stable than Pfr with dark reversion from Pfr to Pr (albeit often extremely slowly). Bathy-phytochromes are unusual in that, although they autoassemble as Pr, this rapidly converts into a Pfr-like state by an unknown thermal mechanism [6, 25, 26]. A difference spectrum can be obtained by subtracting the spectrum after saturating red irradiation from the spectrum after saturating far-red irradiation. The difference spectrum is a unique feature of phytochromes that was crucial in their initial detection [16] and subsequent quantification [27].

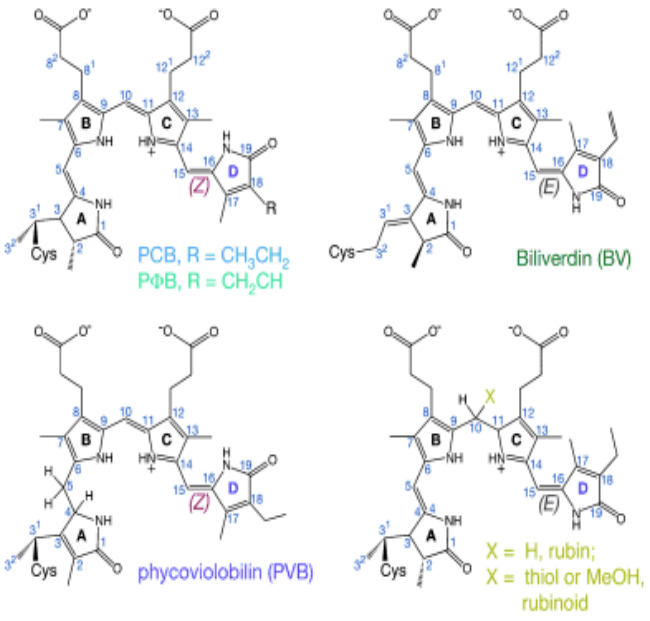


Figure 1.1. Bilin cofactor adducts of phytochromes and related photoreceptors. These bilins are shown in the (5Z)-syn, (10Z)-syn, (15Z)-anti configuration of the Pr state. Taken from [1].

As light sensitive signaling proteins, phytochromes are connected to specific signal transduction pathways [28, 29]. The cyanobacterial phytochrome (Cph1) from *Synechocystis* 6803 is a light-regulated histidine kinase that mediates reversible phosphorylation of a small response regulator [7]. The fungal phytochrome (FphA) from *Aspergillus nidulans* causes light-dependent phosphorylation of the MAP kinase SakA and its shuttling into nuclei [9]. Photosensory signalling by plant phytochromes involves light-induced, conformer-specific nuclear translocation and recognition of transcription factors [30, 31]. There is evidence for additional cytoplasmic signaling in plant phytochromes [32-34].

1.1.1 Domain organization

Canonical prokaryotic phytochromes, defined in Section 1.1.2, comprise an N-terminal photosensory module (PSM) and a C-terminal transmitter module (TM) [1, 17]. The PSM comprises PAS (Per/ARNT/Sim), GAF (cGMP phosphodiesterase/adenylate cyclase/FhlA) and PHY (phytochrome-specific) domains. The C-terminal TM comprises a dimerisation and histidine phosphoacceptor (DHp) domain and an ATPase catalytic domain (CA) (Figure 1.2) [17, 35]. Additional features of canonical plant phytochromes are discussed in Section 1.2.4.

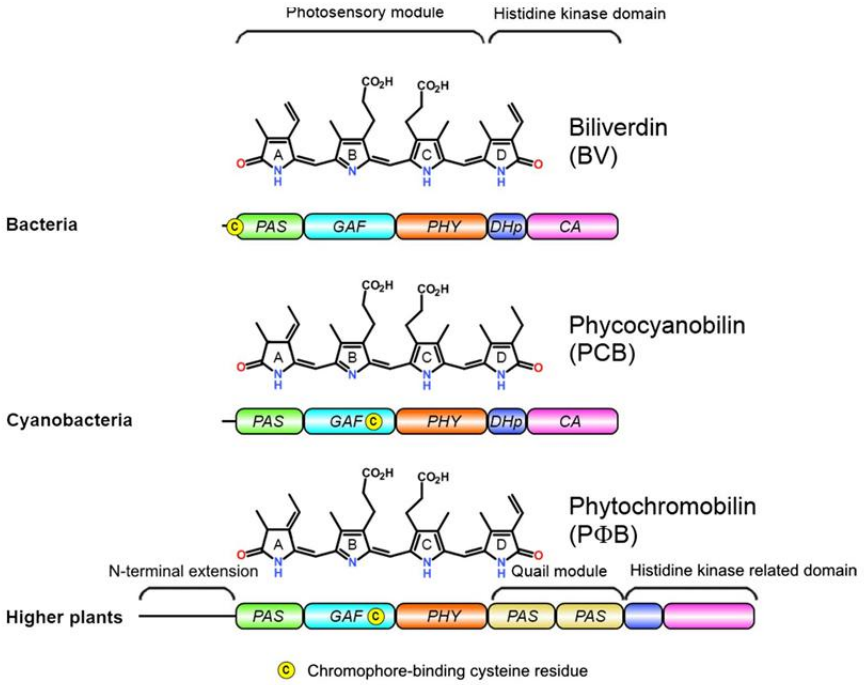


Figure 1.2. Schematic representation of domain architecture in different canonical phytochromes. The native bilin chromophore is shown on top of each phytochrome. Homologous domains have the same color. Adapted from [36].

The PAS, GAF and PHY domains are necessary and sufficient for light perception and photoconversion [17]. PAS domains, originally identified in algal phytochromes [37], have been found in proteins from all kingdoms of life and are often involved directly or indirectly in signal transduction [38]. In prokaryotes, PAS domains are found almost exclusively in sensors of two-component regulatory systems. In eukaryotes, three major classes of PAS proteins have been recognized: photoreceptors, transcription factors and voltage-sensitive ion channels [38]. Some PAS domains mediate signal transduction through protein-protein interaction and the PAS

domain may determine the specificity of interaction. Others bind small hydrophilic ligands with high affinity [39]. As PAS domains are quite diverse [40], they are not always recognizable by sequence comparison alone [41]. The 3D structure of the core PAS fold is, however, very conserved. It includes a five-stranded antiparallel β-sheet flanked by α-helices.

GAF domains are ubiquitous motifs present in cGMP-regulated cyclic nucleotide phosphodiesterases, certain adenylyl cyclases and the bacterial transcription factor FhlA [42]. In phytochromes, the GAF domain comprises most of the conserved bilin-binding pocket. The C3 side chain of the bilin A-ring is covalently attached to a conserved cysteine residue in phytochrome apoprotein. There is a correlation between the location of the bilin-binding cysteine residue and the type of the bilin. Phytochromes that utilize BV have a conserved bilin-binding cysteine N-terminal to the PAS domain [43] while phytochromes that utilize PCB or PΦB have a conserved cysteine in the GAF domain [10]. Bacteriophytochromes form the thioether bond between the $C3^2$ carbon atom of the BV A-ring and cysteine residue preceding the PAS domain [43]. On the other hand, Cph1 and plant phytochromes attach the $C3^1$ carbon atom of the bilin A-ring covalently to a cysteine residue in the GAF domain through a thioether bond [24, 35]. The bilin-bound GAF domain and the bilin-lyase activity are the distinguishing features of phytochrome photoreceptors. The 3D structure of the GAF domain showed that this domain shared a similar fold with the PAS domain [17, 35].

The PHY domain is unique to phytochromes, and hence its name [35]. Mutations affecting the PHY domain often lead to hypsochromic shifts and/or compromise photoconversion [24, 25, 45]. The PHY domain is required for the formation of a stable Pfr state [24] and efficient photoconversion [44]. Although the sequence similarity between PHY and GAF domains is weak [35], 3D structural studies (see Section 1.1.4) showed unequivocally that the PHY domain belongs to the GAF domain family [25]. The tongue-like protrusion from the PHY domain seals the chromophore-binding pocket and stabilizes the Pfr chromophore through direct contacts with the 'DIP' motif in the GAF domain [26, 45, 46].

1.1.2 Classification

Phytochromes can be classified into three classes on the basis of domain architecture of the photosensory module and the minimal bilin lyase domain [1]. The sensory modules of canonical phytochromes (plant phytochromes and Cph1-like cyanobacterial phytochromes) comprise a

PAS/GAF/PHY tridomain [4, 7]. A short sequence motif ('IQ' motif; I20-Q27 in Cph1) was found essential for the bilin lyase activity of canonical phytochromes [18, 47]. In Cph2-like phytochromes, the sensory modules comprise a GAF/PHY bidomain in which the PAS domain is missing [15, 45]. Multiple sequence alignments [15], biochemical and structural studies [45] support the conclusion that the PHY domains of canonical phytochromes are orthologous with the GAF2 domain of the Cph2s [17]. Cph2-like phytochromes are nevertheless red/far-red photoreversible. The sensory modules of cyanobacteriochromes (CBCRs) contain a solitary GAF that is able to bind the bilin chromophore and perform photochemistry autonomously [48]. Exclusively found in cyanobacteria, the CBCR sensors extend the photosensory range of the phytochrome superfamily to the entire visible light spectrum [49].

1.1.3 The 3D structure of the sensory module

With the discovery of prokaryotic phytochromes [4] it became possible to study the family with powerful biophysical methods such as X-ray crystallography [25, 26, 44, 50], nuclear magnetic resonance (NMR) [22, 51, 52], electron microscopy (EM) [53, 54], magic-angle spinning (MAS-NMR) [23, 55, 56] and small-angle X-ray scattering (SAXS) [46, 57]. The first phytochrome 3D structure (1ZTU) was obtained for the N-terminal 321 residues comprising the PAS-GAF bidomain of bacteriophytochrome from *Deinococcus radiodurans* assembled with BV (*Dr*CBD) in the Pr ground state. This groundbreaking structure showed that the polypeptide is knotted and confirmed that the chromophore is attached via a 2-carbon thioether linkage to C24 near the N-terminus. It also revealed key amino acids that form a solvent-shielded bilin-binding pocket. Moreover, the chromophore was in a periplanar ZZZssa configuration [44]. Subsequent mutagenesis led to a high resolution structure of *Dr*CBD. The 1.45 Å model for *Dr*CBD-Y307S (2O9C) showed all 318 residues of the *Dr*CBD with no gaps in the polypeptide chain [50].

Although *DrCBD* assembled with BV to generate a protein with a Pr-like spectrum, this construct was photochemically dead. The PHY domain is needed to stabilize the Pfr state [15]. Subsequently, the crystal structure of the complete, photochemically-competent sensory module of Cph1 from *Synechocystis* 6803 in the lowest energy Pr state was determined (2VEA) [25]. This structure showed a 1-carbon thioether link between ring-A of the PCB chromophore and the sulfur atom of C259, established the PHY domain as a member of the GAF domain family and identified the 'tongue' as a unique structural feature of the PHY domain. Like 1ZTU, the Cph1

structure showed that the chromophore adopts a periplanar ZZZssa geometry (with tilts of 9.8°, 1.4°, and 26.3° between rings A-B, B-C, and C-D, respectively) [25, 35].

The sensory module of Cph1 (shown in Figure 1.3) resembles an asymmetrical dumbbell, the N-terminal PAS domain and the central GAF domain comprising the large lobe (T4–S317) and the C-terminal PHY domain (Y327–E514) comprising the small lobe. The PAS fold (I26–D129) encompasses five-stranded antiparallel β-sheet flanked by helices, its concave front surface possibly representing a protein–protein interaction interface [25, 41, 44]. Also, α1 and α7 are collinear and might represent a docking site. Following the PAS domain is the GAF domain (N130–E323) comprising a five-stranded antiparallel β-sheet sandwiched between a three-helix bundle [25]. The GAF domain contains a large loop constituting the knot and a helix inserted in its core PAS fold. The PHY domain (T324–E514) structure was novel, revealing a PAS-like fold in which a central five-stranded antiparallel β-sheet has a single helix packed on one side and a helical bundle on the other. The PHY fold is very similar to that of the GAF domain but with an additional 'tongue'-like hairpin (P442–Q490). The tongue of the PHY domain seals the chromophore binding pocket and stabilizes the photoactivated far-red-absorbing state. The tandem-GAF dumbbell structure of the phytochrome sensory module is thereby remarkably similar to the regulatory domains of cyclic nucleotide (cNMP) phosphodiesterases and adenylyl cyclases [25, 35].

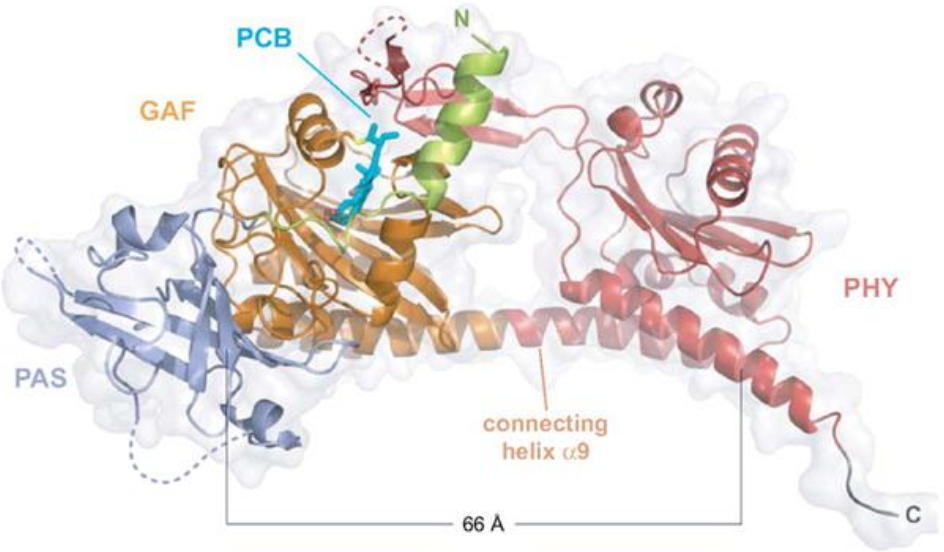


Figure 1.3. Ribbon diagram for the sensory module of Cph1 showing the N-terminal α-helix (green), PAS (blue), GAF (orange) and PHY (red) domains. The PCB chromophore (cyan) is covalently attached to C259. Disordered loop regions are indicated as dotted lines. Taken from [25].

The PAS–GAF and PHY lobes form an interconnected structural unit. In Cph1, a nearly linear 66 Å helix (F299–A345) connects the two lobes covalently, an additional connection being provided by the tongue of the PHY domain. The tongue extends back from the PHY lobe toward the GAF domain, making contact not only with the GAF surface but also with α1 (T4–L18) protruding through the knot (Figure 1.4). Whereas the earlier 1ZTU structure for *Dr*CBD showed solvent access to the chromophore [44], in Cph1 the tongue effectively closes the pocket, isolating the chromophore. In the complete sensory module, the bilin binding pocket comprises a tripartite shell containing the PAS domain, the GAF domain and the tongue of the PHY domain [25].

The chromophore interacts with the protein through a network of multiple, weak, non-covalent interactions. In Cph1 L18, L15 and I20 form part of the hydrophobic walling around rings A and B of the chromophore [25, 45, 46]. The conserved 'IQ' motif is important for proper folding and chromophore binding [18, 58]. Also, the thioether linkage is shielded from the solvent by Y458 and L469–P471 of the tongue. Similarly, the side chains of residues Y176, V186, Y203, P204, and Y263 form a conserved hydrophobic subpocket around ring D. The steric interaction between the ring D methyl carbon and Y263 hydroxyl group as well as that between the C-ring methyl and the D-ring nitrogen (N24) prevent a more coplanar conformation. F475 of the tongue closes the pocket, thereby shielding the chromophore from the solvent [25, 59].

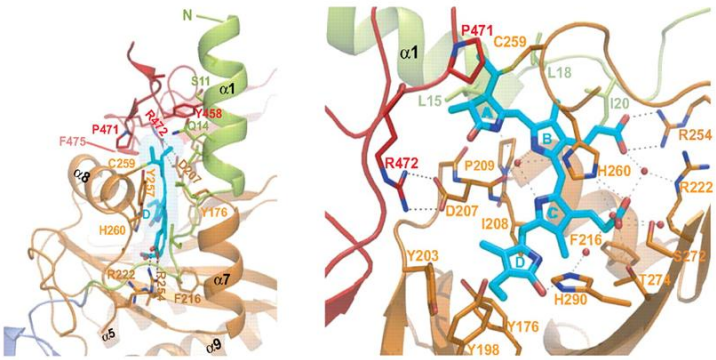


Figure 1.4. The chromophore-binding pocket in Cph1. The conformation of the PCB chromophore (cyan) within the PCB-binding site is shown to adopt a ZZZssa configuration in the left. Taken from [25].

In addition to hydrophobic and steric interactions, the chromophore is also anchored by hydrogen bonds and salt bridges. In Cph1, R222 is bridged through WAT10 both to propionate side chains of ring B and C and comes within 4 Å of the PAS–GAF domain interface. The C-ring propionate is H-

bonded to T274 (and through two waters to S272). Ring D forms a conserved H-bond through its carbonyl group with the imidazole moiety of H290. The loss of the ring D–H290 interaction could easily be transmitted to the protein surface by means of H291 [25, 45, 59]. P471 and R472 form a conserved PRxSF motif in the tongue. Whereas the side chain of P471 shields ring A of the chromophore, R472 acts as an arginine finger pointing into the PCB-binding cleft to form a salt bridge to the conserved acidic residue D207 in the GAF domain [25, 45, 59]. This interaction might be involved in signal transduction. R472 might couple changes upon ring D photoisomerization to the tongue conformation, because it H-bonds the main chain oxygen between G451 and G452 at the kink of the tongue [25, 45, 46]. The tongue is present in all canonical and Cph2-like phytochromes and includes several highly conserved residues, which contribute either to its interaction with the GAF domain (DIP motif) or to its unusual conformation along the GAF-PHY surfaces (WGG motif: W450–G452, E480) [25, 45]. Ring B propionate and R254 form a second salt bridge that might be involved in signal transduction on the other side of the chromophore.

The 1.95 Å resolution structure of the Y263F mutant of Cph1 (3ZQ5) resolved loops of the PAS (Q73-E80), GAF domain (R148-Q150) and tip of the tongue in the PHY domain (A460-K466). Whereas, its overall structure is similar to that of the wild type, the PHY domain is shifted by 17^{o} in relation to the larger lobe comprising the PAS-GAF bidomain. The tongue interacts with the GAF domain through the highly-conserved WGG and WxE motifs. The N-terminal extension adopts different main chain conformations in 3ZQ5 and 2VEA, whereby the native structure remains uncertain [25, 45, 59]. Cph1 represents a valuable evolutionary link between BphPs and plant phytochromes [25].

Another interesting Cph1 mutant with intense red fluorescence and severely impaired photochemistry was originally isolated by directed evolution. The identified mutation was substitution of a universally-conserved tyrosine residue in the GAF domain with a histidine (Y176H) [60]. Saturation mutagenesis established Y176 has been evolutionarily optimized to support photochemical isomerization better than any other residue [61]. Equivalent mutation in phyB from *Arabidopsis,* Y276H (commonly abbreviated YHB) showed a constitutively photomorphogenic phenotype [62]. This mutant was shown to interact with phytochrome interacting factor 6 (At.PIF6) independently of light [63]. Substitution of another conserved tyrosine residue in the GAF domain in *Arabidopsis* phyB with valine, Y303V (commonly

abbreviated YVB) was also shown to be constitutively active [61, 64]. The effect of these interesting mutations on the structure of the photosensory module remains to be elucidated.

1.1.4 Photochemical activation

Classically, two competing models have been proposed for photochemical activation of phytochromes on the basis of the bilin protonation state and the π electron conjugation system. The first model reaction was the sequence of oxidation, nucleophilic addition and tautomerization. It was based on reactivity studies of the A-dihydro-bilindione, which was taken as a model for the Pr chromophore [65]. The oxidation model explicitly predicted very specific functions of the apoprotein as both the oxidant and the nucleophile. Tryptophan and tyrosine residues were predicted to serve as the nucleophiles. The second model was a geometric isomerization of a double bond between rings A and B or C and D, first invoked on the basis of earlier studies on the Z/E isomerization of dipyrroles [19]. High-resolution crystallography [25, 26, 44, 45] and NMR spectroscopy [51] results unambiguously demonstrated that the Pr chromophore adopts ZZZssa geometry whereas the Pfr is ZZEssa. It seems, however, that besides the conformation, protonation of the bilin is an important factor in optimizing light absorption, although how protonation occurs is still uncertain. The conserved H260 in Cph1 probably acts as a proton source/sink coupling the C-ring propionic acid group by H-bonding through water-3 to the nitrogens of rings A–C.

1.1.5 Intramolecular signal transduction

The structure of Cph1 and other phytochromes indicate close packing around chromophore rings A–C. Such tight packing would rule out major conformational changes in that region of the chromophore without associated dramatic changes in the protein [23, 50]. Photochemical activation of the sensory module in the red-absorbing ground state (Pr) begins with a Z to E isomerisation (photoflip) of the chromophore D-ring. The resulting changes are transmitted to the protein surface by means of the evolutionarily conserved ‘tongue’ which is in direct contact with the chromophore. The refolding of the tongue further propagates into changes in the quaternary structure of phytochrome molecule connecting photon absorption to the cellular signaling network [25, 46, 59].

The first crystallographic information about the Pfr state came from studies of bathy-phytochrome from *Pseudomonas aeruginosa* (PaBphP-PCM) assembled with its chromophore BV in the dark-stable state (3C2W) [26]. A photoconversion mechanism from conformational differences between Cph1 Pr structure (2VEA) and PaBphP-PCM Pfr structure (3C2W) was deduced while maintaining the differences in the nature and binding site of the chromophore [35]. The configurations of the bilin chromophores are different along with tautomeric differences in several nearby tyrosine residues [25, 26]. Besides, the secondary structures of the N-terminus and the tongue are quite different. These differences implied putative structural changes occurring within and in the vicinity of the bilin-binding pocket upon photoconversion. To eliminate the difficulty in interpreting structural changes between two different proteins in two different conformational states unambiguously, the Pr and Pfr structures of the same molecule are required [35]. Currently, low-resolution X-ray structures are available for Q188L mutant of PaBphP-PCM in the predominantly Pr (3G6O) [66] and dark-stable Pfr-states (3C2W) [26] and the photosensory module of the bacteriophytochrome from *Deinococcus radiodurans* as Pr (4O0P) [54] and as R-illuminated (4O01) [46] and F479W mutated (5C5K) predominantly Pfr states [67].

Various models have been suggested for coupling structural changes of the chromophore with those of the protein. The flip-and-rotate model postulates Z to E photoisomerization of the chromophore (flip) is accompanied with sliding movement within the bilin-binding pocket (rotate). This model predicts increased hydrogen bonding interaction with the ring D carbonyl in the Pfr state and significant changes in the protein environment around the propionate carboxylate group of ring C [66]. The ‘‘toggle’’ model is just a different way of describing the same flip and rotate event [67, 68]. The Trp-swap model, on the other hand, describes the swap of the bulky tryptophan side chains of WGG and WxE motifs in the tongue-GAF interface to support photochromicity of Group I and II phytochromes [45]. The predictions made by the Trp-swap model were later confirmed by structural studies in bacteriophytochromes [46, 67].

Model "solution" structures for *Deinococcus* phytochrome were generated based on crystal structures from dark and illuminated crystals by molecular dynamics simulations [46] in order to reconcile the crystal structures with experimental data from small-angle X-ray scattering. On this basis a model was proposed for photoactivation whereby atomic rearrangements in the chromophore binding pocket, including displacement of D207 and Y263, cause the β-sheets in the tongue to unfold and nearby atoms to refold into an α-helix. The tongue is shortened as a result

and the distance between the GAF and PHY domain is reduced by 2.5 Å, as a consequence of which, the dimer opens up between the adjacent PHY domains by several nanometers, a movement that could be physically transduced to the C-terminal transmitter module and thereby its signalling activity [46]. A major limitation of this study, however, was the inability to resolve the conformation of the bilin and surrounding amino acids perhaps due to mixed Pr and Pfr states in illuminated-crystals.

1.2 Plant phytochromes

Light is the most important environmental factor in the life cycle of the green plant: from seed germination [2] through seedling development [69, 70], to flowering and senescence [71], thereby bringing about the most dramatic environmental effects on plant development known in nature [69]. Light is not only the source of energy for photosynthesis, but provides critical information about the environment which is used by plants to coordinate their growth and development [72]. Light affects chloroplast biogenesis [73], elongation of stem and hypocotyls [74], expansion of the leaf and floral induction. High light levels lead to the formation of reactive oxygen species which can cause severe cellular damage [73, 75]. To fine-tune cellular, physiological, morphological and developmental processes according to the quantity, quality, direction and duration of light, plants have evolved sophisticated photoreception mechanisms.

1.2.1 Photomorphogenesis

Higher plant seedlings are genetically-endowed with the ability to follow two different strategies of development, skoto- and photomorphogenesis, depending on the absence or presence of light, respectively [70, 76]. Light signals perceived by photoreceptors contribute to the decision to use either the skoto- or the photomorphogenic pathway. Phytochromes are the most extensively studied plant photoreceptors. They were first discovered by the R/FR photoreversibility of diverse physiological processes such as seed germination, stem elongation and photoperiodic control of flowering time [2]. The unique absorbance difference spectrum of phytochromes allowed early spectral detection in etiolated plants [16]. Functionally-important residues and sequence motifs in plant phytochromes were identified by molecular cloning of oat phytochrome [77] and biochemical isolation [78, 79] and characterization [80, 81] of the pigment *in vitro*. Physiological and biochemical studies were supplemented with forward genetics studies. The model plant

Arabidopsis played an important role in investigating the complex light signaling network in plants [72]. Four long-hypocotyl mutants (*hy1, hy2, hy3 and hy6*) were found to be associated with a deficiency of functional phytochrome, whereas a fifth locus, hy5, codes for a positive regulator of photomorphogenesis [82, 83]. *The Arabidopsis HY1* and *HY2* loci are required for phytochrome-chromophore biosynthesis, encoding heme oxygenase [13] and phytochromobilin synthase (ferredoxin-dependent biliverdin reductase) [14], respectively.

Plants have multiple phytochrome subtypes encoded by a small gene family [3]. Whereas the dicots appear to have evolved a five-membered phytochrome family (phyA to phyE), the monocots (such as *Sorghum*) have only three [84]. The *hy3* long hypocotyl mutant of *Arabidopsis* is deficient in phytochrome B [85]. Biochemical and physiological studies distinguished two phytochrome types differing in their stability as Pfr *in vivo*. The level of type I phytochrome decreases dramatically when dark-grown plants are transferred to light, whereas type II phytochromes persist at a similar level as in the dark [86]. phyA is a type I phytochrome and phyB to phyE are all type II phytochromes. Phytochrome A (*PHYA*, *HY8*) [87, 88] and phytochrome B (*PHYB*, *HY3*) [89] are the most extensively studied plant phytochromes. phyA and phyB have overlapping and distinct functions in *Arabidopsis* development [85]. phy B (and in some cases other type II phytochromes) mediates the classical R/FR photoreversible low fluence responses (LFRs) such as seed germination [90], stem elongation and photoperiodic control of flowering time. On the other hand, the very low fluence responses (VLFRs) of seed germination and LHC gene expression in *Arabidopsis* are induced by brief irradiation, are not R/FR reversible, and are mediated by phyA [90]. Whereas phyB plays the major role in continuous monochromatic red (Rc) light [89], phyA is the primary photoreceptor in continuous monochromatic far-red light [91]. Many photoresponses such as anthocyanin biosynthesis are most dramatic following prolonged exposures to high fluence rates of FR light. phyA is responsible for these high irradiance responses (HIRs) [92].

1.2.2 Phytochrome signaling in plants

Yeast two-hybrid screens identified potential phytochrome signaling partners. The first three included a basic helix-loop-helix transcription factor, phytochrome interacting factor 3 (PIF3) [30], a cytoplasmic protein PHYTOCHROME KINASE SUBSTRATE, PKS1 [93] and a nucleoside diphosphate kinase, NDPK2, which is localized to both cytoplasm and nucleus [94]. The bHLH

transcription factors (PIFs) and the bZIP transcription factors (HY5 and HYH) form antagonistic transcriptional modules that integrate light and reactive oxygen species (ROS) signaling [95, 96]. PIFs promote skotomorphogenesis whereas HY5 is a master regulator of photomorphogenesis. Phytochrome signaling regulates gene expression by initiating turnover of these key transcription factors. Genetic and molecular screens also identified many genes acting downstream of photoreceptors [70, 76, 82]. The purification of the major 'COP' complex revealed this protein complex plays a role in protein turnover [69]. CONSTITUTIVELY PHOTOMORPHOGENIC 1 (COP1), which is conserved in both higher plants and vertebrates, comprises RING finger, coiled-coil and WD40 domains. COP1 is a prototypical E3 ubiquitin ligase that mediates the ubiquitination and subsequent proteasomal degradation of HY5 and a myriad of other key regulators in other systems [97]. HY5 interacts with the WD40 repeat domain of COP1. The abundance of HY5 protein is inversely correlated with the nuclear abundance of COP1 [83].

Phytochrome molecules rapidly translocate from the cytoplasm into the nucleus following photoconversion to the Pfr form [98, 99] (see Fig. 1.5). Pfr physically interacts with nuclear transcription factors, PIFs [31], thereby transducing light signals to light-regulated genes. The G-box motif is the promoter element recognized by PIF3 [100]. Light induces the phosphorylation of PIFs [101-103], which is necessary for recruitment of the light-response BTB1 (LRB1) and BTB2 (LRB2) E3 ubiquitin ligases to the phyB-PIF3 complex. Mutually assured destruction (MAD) of the receptor and its immediate signaling partner by the proteasome couples signal transmission to signal attenuation [103].

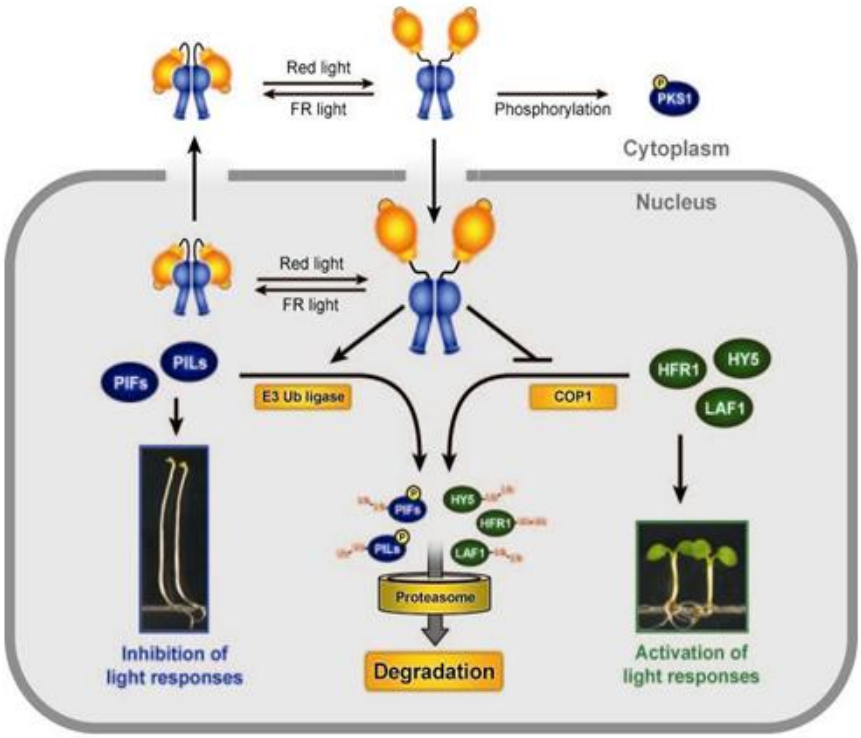


Figure 1.5. Phytochromes regulate light signaling through negative and positive transcription factors. PIFs/PILs inhibit light responses in the dark, whereas HY5/HYH and HFR1 are degraded by the nuclear-localized COP1. Light activated phytochromes translocate from the cytoplasm to the nucleus to promote the degradation of PIFs/PILs and inhibit COP1 by excluding it from the nucleus. Taken from [96].

1.2.3 Domain architecture of plant phytochromes

Plant phytochromes comprise a photosensory module (PSM) N-terminal of a PAS repeat domain (PRD) and a C-terminal histidine kinase-related domain (HKRD) [35]. The PRD is unique to plant phytochromes. It includes a cluster of amino acids named the Quail box [105]. Missense mutations in the Quail box impair sub-nuclear localization and downstream signaling [28]. The HKRD comprises a dimerisation domain lacking a histidine phosphoacceptor (D) and a catalytic domain (CA). The HKRD in phyB contains a putative split-nuclear localization signal [106] whose nature remains unknown. PSM carries the determinants for phytochrome specificity whereas PRD and HKRD are responsible for dimerisation and nuclear localization. PRD and HKRD are functionally interchangeable between isoforms [104].

Unlike Cph1 and other prokaryotic phytochromes, At.phyA, At.phyB and Sb.phyB contain highly divergent N-terminal extensions (NTE) of various lengths (68, 102 and 113 amino acids,

respectively) [105]. The NTE can be divided into a low-complexity region (LCR), a helix-forming motif (HFM) and an amino-terminal sequence (ATS). The HFM predicted *in silico* is likely to exhibit conformational change during photoconversion, explaining its state-dependent protease sensitivity [81, 107, 108]. This conformational change might be essential to stabilize Pfr [107]. The short, flexible ATS anchors the HFM into the PAS domain. Mutations of serine residues in the NTE of *Arabidopsis* phyA into alanines increased its biological activity [109]. Deletion of the NTE in *Arabidopsis* phyB decreased the efficiency of light signal perception [110]. Moreover, *Arabidopsis* response regulator 4 (ARR4) specifically binds the phyB NTE [111].

Phylogeny would appear to support kinase function for plant phytochromes; Cph1 and many other prokaryotic phytochromes function as light-regulated histidine kinases [7, 112]. PKS1 was identified as a substrate for light-regulated phytochrome kinase activity *in vitro*. *In vivo* experiments in *Arabidopsis* suggest that PKS1 is phosphorylated in a phytochrome-dependent manner and negatively regulates phytochrome signaling [93]. However, plant phytochromes signal from the N-terminal sensory module, the C-terminal half of the molecule being responsible for dimerisation and Pfr-dependent nuclear localization [113]. Isolated PAS-GAF bidomain of *Arabidopsis* phyB is fully functional *in vivo* when it is attached to domains that provide for dimerisation and nuclear localization [114]. The PHY domain is dispensable for phyB signal transduction but is required for stabilizing the Pfr form of phyB [113, 114]. The core machinery for plant phytochrome signal transduction is, therefore, the region containing the knotted PAS and GAF domains [113, 114]. The mechanism by which this core region transduces light signal to downstream components remains to be elucidated.

1.2.4 The structure of At.phyB

The 3.4 Å structure (4OUR) of the PSM from *Arabidopsis thaliana* phyB as Pr, the first crystal structure of a plant phytochrome, revealed strong structural homology (Figure 1.6) to its prokaryotic relatives, particularly Cph1. As expected, PΦB is attached by a thioether bond to C357 through its C3 atom. The C-ring propionate is parallel to the B-ring propionate and in contact with the adjacent R322. The D-ring is surrounded by H403 and three tyrosine residues (Y276, Y303, and Y361) forming a hydrophobic pocket. The ε-nitrogen of H403 forms an H-bond with the carbonyl of the D-ring, anchoring the D-ring in its Pr state. Two of the conserved tyrosines forming the hydrophobic pocket are probably those that have been known for many

years to be critical for the formation of Pfr in plant phytochromes [80]. The first could be Y276, located immediately below ring D [60, 61]. The second is likely Y303 located at the side [64]. Both Y276H (YHB) and Y303V (YVB) mutants have compromised photochemistry and constitutive photomorphogenic activity in *Arabidopsis* [61, 64]. Electron density for the PHY domain was less resolved, especially in the loop regions, probably due to domain wobbling within the crystal lattice [68].

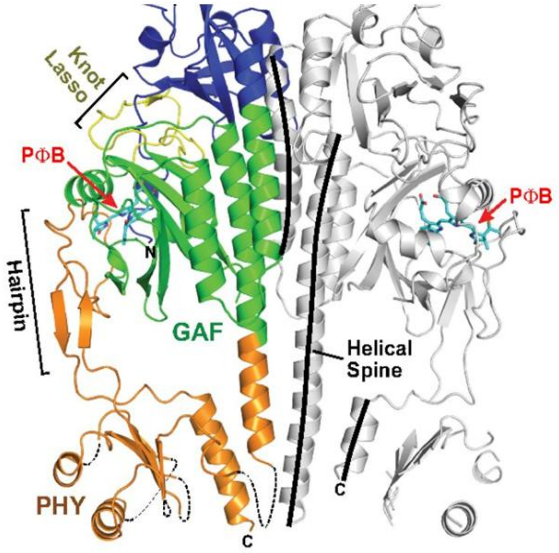


Figure 1.6. Ribbon diagram of the PSM dimer (residues 90–624) of At.phyB as Pr. Taken from [68].

Structural differences that might be important for plant specific signaling were also identified. Structural differences between phyB and Cph1 included the cap helix (Y104-R110), a new contact between the residual NTE (V96-A103) and the DIP (307-309) motif in the GAF domain, an elongated PAS-GAF connecting helix (L226-A247) and two extended loops: Q146–P154 (150s) and G379–S392 (380s). The sequence and presumably the structure of the cap helix is conserved among plant phytochromes but not Cph1. The cap helix forms a steric barrier for the A and B pyrrole rings with conserved residues Y104 and I108 directly leaning on the bilin. Accordingly, the Y104A substitutions of At.PSM accelerated dark reversion. The side chain of M159 near the 150s loop contributes to hydrophobic interactions with residues in the GAF domain lasso (VVQ: 334-336), perhaps stabilizing the phyB-PIF signaling interface [68, 115, 116]. The α-helix connecting the PAS and GAF domains (SIA helix, S227-A229) is distinctly longer than that of bacterial phytochromes by three rotations.

1.2.5 phyB from *Sorghum bicolor*

Sorghum is a tropical monocot of east African origin [84, 117]. It is one of the most important cereal crops and dietary staple of millions of people worldwide [118]. As a C4 plant, *Sorghum* has physiological, biochemical and morphological specialization to increase net-carbon assimilation at high temperatures [84]. *Sorghum* is drought-tolerant and can be grown in semi-arid conditions, where maize, wheat, and rice cannot be grown because of water scarcity [119]. *Sorghum* is a quantitative short day plant [117], floral induction being sensitive to both photoperiod and temperature [120].

In studies of floral induction in *Sorghum*, mutations at three loci involved in photoperiodic flowering behavior were identified [117, 120]. The maturity genes *Ma1*, *Ma2* and *Ma3* influence the night length which would promote or delay flowering [117]. *Ma3* shows abnormal phytochrome physiology [121], lacks the light-stable phytochrome that predominates in green tissue [122] and encodes phytochrome B [123].

1.3 Phytochrome interacting factors (PIFs)

At.PIF3 was originally discovered from a yeast two-hybrid screen using the C-terminal domain of phyB from *Arabidopsis thaliana* [30]. *poc1* (photocurrent-1), which exhibits enhanced responsiveness to red light, was identified in an independent genetic screen of T-DNA tagged *Arabidopsis* mutants defective in early light signaling. The T-DNA insertion in *poc1* was found to be located in the promoter region of *PIF3* [124]. There are at least seven PIFs in *Arabidopsis* that function in a partially-differential to a largely-overlapping manner to suppress photomorphogenesis [125]. PIFs are members of a sub-family of bHLH proteins that bind preferentially to photoactivated phytochromes.

PIF family members control light-regulated gene expression [126]. Dark-grown quadruple *pif1 pif3 pif4 pif5* mutants displayed constitutive photomorphogenic phenotypes both morphologically and at the gene expression level [127]. PIFs have an active phyB-binding (APB) domain and/or active phyA-binding (APA) domain located at the N-terminus that is necessary and sufficient for interaction with light-activated phytochromes [128]. Both phytochrome and a putative PIF transcription factor are present even in algae [129-131]. Thus, the molecular mechanisms for light

dependent transcriptional regulation mediated by the phy-PIF system might have been established early in land plant evolution [130, 132].

1.3.1 Domain organization of PIFs

PIFs, like other eukaryotic transcription factors, are assumed to be modular in structure containing an N-terminal transactivation domain (TAD) and a C-terminal DNA binding domain (DBD) (see Figure 1.7). The TAD domain contains an acidic region which facilitates transcriptional activation and/or dimerisation. The amino acids sequences of TAD domains are generally divergent, even among closely related species [133]. On the other hand, the bHLH domain is highly conserved and comprises approximately 60 amino acids with two functionally-distinct regions. The basic region contains 13-17 primarily basic amino acids and binds to DNA. The helix-loop-helix (HLH) region comprises two amphiphatic α-helices, mainly consisting of hydrophobic amino acids, which are connected by a loop of variable length. Proteins containing the HLH motif often form homo- or heterodimers with other HLH proteins, which is a prerequisite for DNA recognition and contributes to DNA-binding specificity [134].

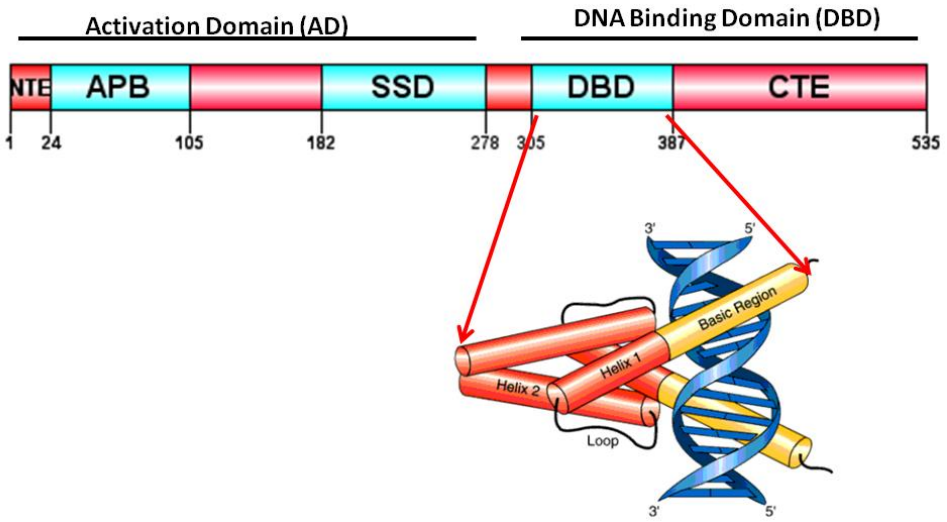


Figure 1.7. Hypothetical domain architecture of phytochrome interacting factors based on sequence similarity to mammalian transcription factors. The activation domain contains the N-terminal extension (NTE), activated phytochrome binding (APB) site and the signal sensing domain (SSD). The DNA binding domain contains the bHLH region and a long C-terminal extension. Built using DOG domain illustrator.

The bHLH domain represents an ancient family found in plants, fungi and animals, but not prokaryotes [133]. Phylogenetic analyses suggest that plant bHLH proteins are monophyletic and constitute 26 subfamilies [135]. 28 short, conserved motifs are present in the N-terminal regions of plant bHLH proteins, although none of these corresponds to a known domain in the PFAM database. Sub-group VII (a+b) contains proteins involved in light signaling. Motif-14 present in members of this group overlaps with the active phytochrome binding (APB) motif [128, 135].

1.3.2 The phy-PIF interaction interface

Different techniques have been employed to map the phy-PIF interaction interface. Yeast-reverse hybrid screen identified missense mutations in the PSM of At.phyB that are normal in signal perception but defective in the capacity to transduce light signals to At.PIF3 [115]. Genetic screen for *Arabidopsis* mutants showing reduced sensitivity to red light identified another set of missense mutations [116]. Critical mutations disrupting light-induced binding of At.phyB to At.PIF3 were mapped to the knot region in phyB [115, 116]. A novel molecular recognition motif, APB, necessary for targeting PIFs to photoactivated phytochrome was defined using *in vitro* co-immunoprecipitation assay [128]. To characterize and quantify the interaction between phyB and PIF3N (N-terminal region) further, isothermal titration calorimetry (ITC) has been employed [136]. To date, there is no published crystal structure of phytochrome in complex with its interacting partner.

1.3.3 *Sorghum* PIFs

Characterization of PIFs from only a single plant species such as *Arabidopsis* may be misleading, since at least some of the properties observed may be unique to the molecule from that species. This potential problem is complicated by the observations that PIFs exhibit multiple isoforms. A search for molecular features conserved among PIFs from evolutionarily-divergent plant species might help to understand the mechanism for signal transfer from light activated phytochromes to primary signaling partners (PIFs). PIFs in *Sorghum bicolor* remain to be characterized.

1.3.4 Aim of this study

The general aim of this study was to investigate the major structural differences between the photosensory modules of the Pr ground state of plant phytochrome and its constitutively-active mutant (YHB). *Sorghum* phyB was chosen primarily for crystallographic reasons as it might represent a particularly robust variant. The locus has also been the subject of extensive molecular genetic studies. Structural information on Sb.phyB might provide novel insights into phytochrome function not accessible from *Arabidopsis*. Crystallographic evidence from the photosensory module of YHB in complex with its PIF signaling partner might provide a model for the biochemical mechanism of signal transfer from photoactivated phyB to PIFs.

The specific objectives of this study were cloning, producing, purifying and crystallizing different Sb.phyB constructs, its constitutively-active mutant Y291H and its primary interacting partner Sb.PIF3. Multiple constructs were produced to explore the influence of various domains and sub-domains on photochemical and crystallization properties of Sb.phyB. In particular, the effects of the NTE and of the PAS and PHY domains in chromophore assembly, photochemistry and crystallization properties were examined. The utility of a novel approach involving co-expression of the YHB mutant with its putative signaling partner, Sb.PIF1, was explored. The effect of different chromophores on the properties of Sb.phyB was also addressed. Several biochemical and biophysical techniques were used to gain insight into the mechanism of photoconversion. This work represents a fundamental step in structure/functional studies of phyB and PIFs in monocotyledons.

2 Materials and methods

2.1 Materials

2.1.1 Equipment

2.1.1.1 Equipment for molecular biology

Scales	AJ50L (Mettler)
	PJ400 (Mettler)
Water purification	Ion separator GENO-sep (Gruenbeck)
systems	Ultrafiltration module (membraPure)
pH meter	inoLab pH level 1 (wtw)
PCR	T3000 thermocycler (Biometra)
Autoclave	VX-150 (Systec)
Clean bench	Laflow (Prettl)
Electroporation	Genezapper 450/2500 (IBI, Kodak) and Electroporation cuvettes (Biorad)
Temperature	Water bath 1002 (GFL)
control	Thermomixer comfort (Eppendorf)
Incubator	B 5042 (Heraeus), can be used as shaker
Shakers	TR-125 (Infors HT)
	Centromat HK and Centromat R (Sartorius)
	Innova 42/42R (Eppendorf)
Flasks	2L reusable baffled-glass culture flasks (PYREX) with broad bottoms and long necks to increase surface area/volume ratio and reduce splashing
Centrifuges	Biofuge 4 x 180 swinging bucket & 24 x 1.5/2 fixed angle rotors (Heraeus)
	5415 B with R7018 microcentrifuge fixed angle rotor (Eppendorf)
	5415 R with 24 x 1.5/2 ml standard rotor (Eppendorf)
	RC2-B, SLA-1500 and SA-800 fixed angle rotors (Sorvall)
	5810 R and A-4-81 swing-bucket rotor with 7 x 50 bucket (Eppendorf)
Agarose gel	Homemade chambers
UV-transilluminator	TCP-20-LM fluorescent lamp (Vilber Lourmat)
Digital camera	Coolpix 995 (Nikon)
Cold storage room	Eco Coils & Coolers (Luvata)
-80 ^{o}C freezer	Hera freezer (Heraeus)

2.1.1.2 Equipment for protein biochemistry

French pressure cell press	PP60KN (Watz Hydraulik)
Heating block	Driblock DB-3 (Techne)
Protein gel apparatus	Mini-PROTEAN Tetra Cell (Biorad)
FPLC	ÄKTA Purifier (GE Healthcare)
UV-Vis spectroscopy	8453 diode array detector spectrophotometer (Agilent) equipped with homemade sample holder for R/FR irradiation
	UV mini-1240 (Shimadzu)
	Quartz cuvettes 1 mm pathlength for R/FR irradiation (Hellma)
Protein concentrator	Stirred ultrafiltration cell (Millipore)
	Ultrafiltration membrane (membraPure)
Fluorescence spectroscopy	Fluoromax4 spectrofluorometer (HORIBA)
	Optical interference filters (Schott)
	Cut-off filters (Schott)
Illumination devices	LEDs (660 and 720 nm) & safelight (490 nm)(Roithner)
	Plexiglas Formmasse black 90053 0.2 mm filter (Rohm)
Circular dichroism	J-710 CD-Spectropolarimeter (Jasco)

2.1.1.3 Equipment for structural biology

Dialysis	Slide-A-Lyzer (Thermo Fisher)
High-throughput screening	Liquid Handling Sampling Robot (Zinsser Analytic)
	Honeybee 963 dispensing system (Genomic solutions)
Microscope and macroscope	B061 (Olympus), Z16 Apo A (Leica)
Camera	EC3 (Leica)
Liquid nitrogen Dewar	Foam Dewar (Molecular Dimensions)
	Stainless steel Dewar (Molecular Dimensions)
	Sample changer basket (Molecular Dimensions)
	Magnetic cryowand (Molecular Dimensions)
	Vial Tongs (45°) (Molecular Dimensions)
	Dry shipper (Molecular Dimensions)
Synchrotron radiation sources	Bessy II beamline MX 14.1

	European Synchrotron Radiation Facility beamline ID23-1
Crystal dehydration	HC1c Crystal Humidifier

2.1.2 Consumables

2.1.2.1 Consumable items

Column chromatography	Ni-NTA Superflow (Qiagen)
	Superdex 200 16/60 prep grade (GE)
	Superdex 200 26/60 prep grade (GE)
	TSK-SWxl (Tosoh Bioscience LLC)
Protein biochemistry	Disposable polystyrene cuvettes (Sarstedt)
	Filtropur S 0.2 µm syringe filter (Sarstedt)
	Ultrafiltration membrane discs (Millipore)
Protein crystal screening	24-well plates and micro-bridges (Sarstedt)
	18 mm and 22 mm cover slips (Sarstedt)
	Acupuncture needles (SEIRIN)
	Baysilone paste (GE, Bayer)
	Granada Crystallization Box (Triana Sci. and Tech.)
	CryoLoops (Hampton)
	EasyXtal 15-well plates (Qiagen)
	CrystalCap SPINE HT (Hampton Research)
	Magnetic CryoVial SPINE (Molecular Dimensions)

2.1.2.2 Reagents

All chemicals and reagents used were analytical grade purity. Chemicals were purchased from Applichem, Merck and Sigma-Aldrich. Kits from NeXtal suites (Qiagen), JBScreen Pentaerythritol (Jena Biosciences) and Morpheus (Molecular Dimensions) were used for initial protein crystallization screening.

2.1.2.3 Solutions

Bacterial growth media

SOC-medium	2 % (w/v) tryptone, 0.5 % (w/v) yeast extract, 2.5 mM KCl, 10 mM NaCl, 10 mM

	$MgCl_2$, 10 mM $MgSO_4$ and 20 mM glucose. Sterile filtered. Courtesy of Melanie Bingel.
LB medium	1 % (w/v) tryptone, 0.5 % (w/v) yeast extract and 1 % (w/v) NaCl to pH 7.5 with NaOH. Autoclaved + sterile antibiotic when the medium was cooled to ca. 50 ^{o}C.
2xYT broth	1.6 % (w/v) tryptone, 1 % (w/v) yeast extract and 0.5 % (w/v) NaCl to pH 7.0 with NaOH, autoclaved + sterile antibiotic when cooled to ca. 50 ^{o}C.
Super broth	3.2 % (w/v) tryptone, 2 % (w/v) yeast extract and 0.5 % (w/v) NaCl to pH 7.0 with NaOH, autoclaved + sterile antibiotic when cooled to ca. 50 ^{o}C.
Terrific broth	1.2 % (w/v) tryptone, 2.4 % (w/v) yeast extract and 0.04 % glycerol autoclaved + sterile filtered 17 mM KH_2PO_4, 72 mM K_2HPO_4 + sterile antibiotic when cooled to ca. 50 ^{o}C.
SOB medium	2 % Trypton, 0.5 % Yeast extract, 10 mM NaCl and 2.5 mM KCl. Autoclaved plus sterile filtered 10 mM $MgCl_2$ and 10 mM $MgSO_4$ + sterile antibiotic when cooled to ca. 50 ^{o}C.
LB agar	1.5 % agar in LB medium, autoclave, cool to 55 ^{o}C + sterile antibiotic and plate.
Antibiotic stock 1000x	100 mg/ml Ampicillin, 50 mg/mL Kanamycin, 50 mg/ml Spectinomycin and 10 mg/mL Tetracycline. Courtesy of Melanie Bingel.

Solutions for molecular biology

5x TBE buffer	500 mM Tris, 425 mM Boric acid, 50 mM EDTA pH 8.0. Courtesy of Melanie Bingel.
Agarose gel	0.8 – 1.2 % (w/v) Agarose in 0.5 % TBE buffer, melt in microwave and cast.
6x loading dye	0.03 % (w/v) Bromophenol blue, 0.03 % (w/v) Xylene blue, 60 % (v/v) Glycerol, 60 mM EDTA and 10 mM Tris pH 8.0.
DNA ladder	1 kb marker from NEB diluted in 6x loading dye.
EtBr staining	1 : 20,000 ethidium bromide stock solution (10 mg/ml) in 0.5 % TBE buffer.

Solutions for protein biochemistry

TES-β	50 mM Tris, 5 mM EDTA, 300 mM NaCl and 1 mM β-mercaptoethanol pH 7.8; include DNase I (1 μg/ml) and PMSF (2 mM) for use as extraction buffer.

	Filtered and degassed prior to use.
HES-β	50 mM HEPES, 5 mM EDTA, 300 mM NaCl and 1 mM β-mercaptoethanol pH 7.8. DNase I (1 µg/ml) and PMSF (2 mM) for use as extraction buffer. Filtered and degassed prior to use.
$TISI^{10}T^{1}$-β	PIF lysis buffer: 50 mM Tris, 1 mM IDA, 300 mM NaCl, 10 mM imidazole, 1 % triton-X 100 and 1 mM β-mercaptoethanol pH 7.8. Filtered and degassed prior to use.
AmS	Ammonium sulfate buffer: 50 mM Tris, 1 mM IDA, and 3.3 M $(NH_4)_2SO_4$ pH 7.8. Filtered and degassed prior to use. Stored in cold room.
$HISI^{10}$-β	Ni-NTA binding buffer: 50 mM HEPES, 1 mM IDA, 500 mM NaCl, 10 mM imidazole and 1 mM β-mercaptoethanol pH 7.8. Filtered and degassed prior to use.
$HISI^{50}$-β	Ni-NTA washing buffer: 50 mM HEPES, 1 mM IDA, 500 mM NaCl, 50 mM imidazole and 1 mM β-mercaptoethanol pH 7.8 filtered and degassed prior to use.
$HISI^{250}$-β	Ni-NTA elution buffer: 50 mM HEPES, 1 mM IDA, 500 mM NaCl, 250 mM imidazole and 1 mM β-mercaptoethanol pH 7.8. Filtered and degassed.
IEBB	Ion-exchange binding buffer: 20 mM Tris, 1 mM EDTA and 1 mM β-mercaptoethanol pH 7.8. Filtered and degassed.
IEEB	Ion-exchange elution buffer: 20 mM Tris, 1 mM EDTA, 1M NaCl and 1 mM β-mercaptoethanol pH 7.8. Filtered and degassed.
30 % AB solution	30 % (w/v) of 37.5 : 1 (w/w) (Acrylamide : Bisacrylamide).
APS	10 % (w/v) Ammonium persulphate (APS) in H_2O.
Stacking gel buffer (4x)	0.5 M Tris/HCl (pH 6.8) and 0.6 % (w/v) SDS.
Separating gel buffer (4x)	1.5 M Tris/HCl (pH 8.8) and 0.6 % (w/v) SDS.
6x SDS loading buffer	67 % (v/v) 4x stacking gel buffer, 30 % (v/v) glycerol, 1 % (w/v) SDS, 0.5 M DTT, 0.09 % (w/v) Bromphenol blue, 1 % (v/v) β-mercaptoethanol.
10 % separating gel	5 ml AB solution, 3.75 ml separating gel buffer (4x), 6.25 ml H_2O, 90 µl APS (10 %) and 15 µl TEMED.
4 % stacking gel	650 µl AB solution, 1.25 ml stacking gel buffer (4x), 3.1 ml H_2O, 25 µl APS (10 %) and 5 µl TEMED.
SDS-PAGE running buffer	TGS: 25 mM Tris, 192 mM Glycine and 0.1 % (w/v) SDS.

Bradford reagent	100 mg Coomassie Brilliant Blue G-250 + 50 ml 95 % Ethanol + 100 ml 85 % (w/v) H_3PO_4 + 500 ml H_2O to dissolve + 350 ml of H_20, filter before use.
Ellman's reagent	10 mM DTNB (5,5'-dithio-bis-(2-nitrobenzoic acid)) in 0.1 M Tris-HCl pH 7.5.
ZnAc stain	1 mM Zinc acetate solution.
Hot Coomassie stain	0.5 % (w/v) Coomassie Brilliant Blue R-250, 25 % (v/v) Isopropoanol and 10 % (v/v) Acetic acid.
Destaining solution	10 % Acetic acid.

Solutions for structural biology

TS/20	Crystallization buffer for dialysis containing 2.5 mM Tris and 15 mM NaCl pH 7.8 filtered and degassed.
HS/10-TCEP	Crystallization buffer for dialysis containing 5 mM HEPES, 50 mM NaCl and 0.3 mM TCEP pH 7.8 filtered and degassed.
Crystallization buffers	For optimization of crystal growth 0.1 M HEPES pH 7.5, 0.1 M CHES pH 9.5 and, 0.1 M CAPSO pH 9.5.
Major precipitants	12 % (v/v) PEG 400, 10 % (v/v) PEG 6,000, 10 % (v/v) PEG 8,000 and 1 M AmS.
Minor precipitants	0.2 M NaCl, 0.2 M $MgCl_2$ and 0.1 M Li_2SO_4.
Antifreeze solutions	Cryoprotectant agents used include 20 % (v/v) Glycerol, 20 % (v/v) Propylene glycol, 25 % (v/v) Ethylene glycol, 25 % (v/v) 2-methyl-2,4-pentane diol (MPD), 25 % (v/v) PEG 400.

2.1.3 Oligonucleotides and plasmids

2.1.3.1 Oligonucleotides

Table 1: The list of primers used for PCR

No.	Sequence	Description
J27	CCCGGTACCGAATTCATTAAAGAGGAGAAATTAACCATGGCTTCTGGTTCT	*EcoRI-Sb.NPGPFw*
J28	CGGGAATTCTTAAAGAGGAGAAATTAACCATGGctACCGAATCTGTTTCTAAAG	*EcoRI-RBS-start-Sb.HPGPFw*
J29	CGGGAATTCTTAAAGAGGAGAAATTAACCATGGctGCTGCTTACCTGTCTCGTA	*EcoRI-RBS-start-Sb.PGPFw*
J30	GTAGGGAAGCTTTTAGTGATGATGATGATGATGACCCAGCTGAACCTGACC	*HindIII-stop-6xHis-SbNPGPRev*
SS1	GTGGAAGCGGTGAACCATAAC	*Sb.PHYBY291HR1*
SS2	GGTTACGACCGTGTTATGGTTCACCGCTTCCACGAAGACGAACAC	*Sb.PHYBY291HF2*
SS3	GTGTTCGTCTTCGTGGAAGCGGTGAACCATAACACGGTCGTAACC	*Sb.PHYBY291HR2*
SS4	GTTATGGTTCACCGCTTCCAC	*Sb.PHYBY291HF1*
SS14	cggCATATGCATCATCATCATCACCAC	*NdeI-Sb.PIF1_APBFw*
SS15	aatCTCGAGTCACGGTGCTGCTGCTGC	*Xho1-stop-Sb.PIF1_APBRev*
SS16	cggCTCGAGTCATTCACTCGGACTTTCGCT	*Xho1-stop-Sb.PIF1_LEARev*
SS17	cggCTCGAGTCAGCCCATCCACATCATTTG	*Xho1-stop-Sb.PIF1_bHLHRev*
SS18	GCTAAAGAAATCAAATGGGAAGGTGCTAAACACCACCCG	*Sb.PHYB-G480E-F1*
SS19	CGGGTGGTGTTTAGCACCTTCCCATTTGATTTCTTTAGC	*Sb.PHYB-G480E-R1*
SS41	CGGCATATGAACCAGTTTGTGCCGGATTGGAGCAACATGGGC	*NdeI-Sb.PIF1CTHFw*
SS42	CGGCTCGAGTTAGTGGTGATGATGATGATGACCCGGACCACGGCTAGG	*XhoI -stop-6xhis-Sb.PIF1CTHR*
SS43	GTAGGGAAGCTTTTAGTGATGATGATGATGATGAGCCAGCTGCAGTTCCAT	*HindIII-stop-6xHis-Sb.PGRev*
SS48	ACCACACCTCTCCGCGTGCCATCCCGTTCCCGCTGCGTTA	*Sb.PGP-C425A-F*
SS49	TAACGCAGCGGGAACGGGATGGCACGCGGAGAGGTGTGGT	*Sb.PGP-C425A-R*
SS50	TCATGGACCTGGTTAAAGCCGACGGTGCTGCTCTGTACTA	*Sb.PGP-C492A-F*
SS51	TAGTACAGAGCAGCACCGTCGGCTTTAACCAGGTCCATGA	*Sb.PGP-C492A-R*
SS52	TTCTGGTGGTGACGACGCCGCGACCGGTCGTGGTGGTATC	*Sb.PGP-EQ399A-F*
SS53	GATACCACCACGACCGGTCGCGGCGTCGTCACCACCAGAA	*Sb.PGP-EQ399A-R*
SS56	ACCTCTCCGCGTGCCATCCCGTTCCCG	*Sb.PGP-C425A-F2*
SS57	CGGGAACGGGATGGCACGCGGAGAGGT	*Sb.PGP-C425A-R2*
SS58	GACCTGGTTAAAGCCGACGGTGCTGCT	*Sb.PGP-C492A-F2*
SS59	AGCAGCACCGTCGGCTTTAACCAGGTC	*Sb.PGP-C492A-R2*
SS60	GTGGTGACGACGCCGCGACCGGTCGTGG	*Sb.PGP-EQ399A-F2*
SS61	CCACGACCGGTCGCGGCGTCGTCACCAC	*Sb.PGP-EQ399A-R2*
SS62	TCTCCGCGTGCCATCCCGTTC	*Sb.PGP-C425A-F3*
SS63	GAACGGGATGGCACGCGGAGA	*Sb.PGP-C425A-R3*
SS64	TTAACCATGGctGCTCGTACCGAAGACCCGGCTCTGTCTA	*NcoI-Sb.GPFw*

2.1.3.2 Vectors

Table 2: Selection of plasmids for apophytochrome cloning and production

No.	Plasmid	Origin	Antibiotic	Promoter	N-tag	Map	Copy no
1	pPROLar.A	p15A	Kan	plac/para	None	Fig. 2.1	Low
2	pBAD-MCS	pMB1	Amp	pBAD	None	n/a	Low
3	pCDFDuet-1	CDF	Sp/St	T7	His.tag	Fig. 2.2.	Low
4	pET28a	ColE1	Kan	T7	None	n/a	High
5	pUC57	pMB1	Amp	library	None	n/a	High

Table 3: Selection of plasmids for chromophore production

No.	Plasmid	Origin	Antibiotic	Promoter	N-tag	Copy no
1	p183 (PφB)	ColE1	Amp	T5	n/a	Low
2	p171 (PCB)	ColE1	Amp	T5	n/a	Low
3	p45 (PCB)	p15A	Kan	plac/para	n/a	Low
4	pCL-PCB	p15A	Amp	plac/para	n/a	Low

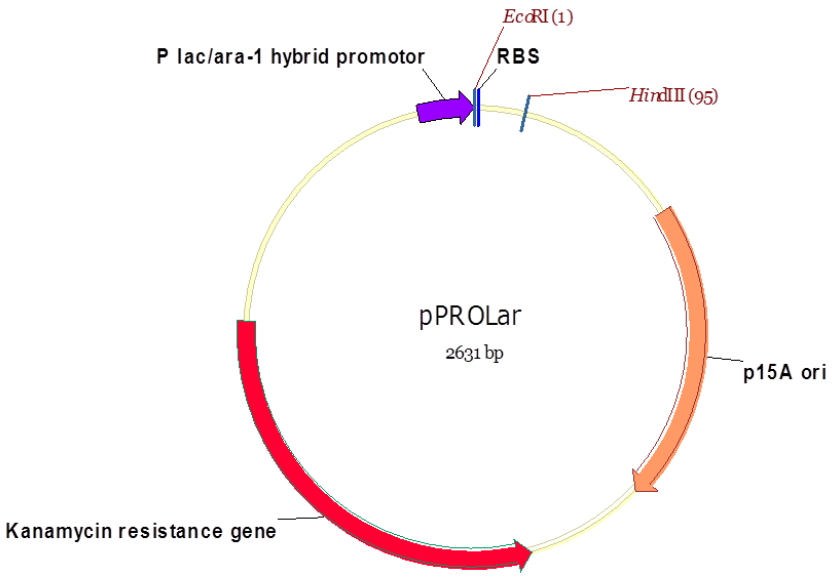


Figure 2.1. The pPROLar vector. pPROLar vector contains a hybrid promoter (lac/ara), origin of replication (p15A ori), antibiotic resistance gene (Kan^R) and multiple cloning site. EcoRI and HIndIII sites were used for ligation dependent cloning. The potential myc tag was not used.

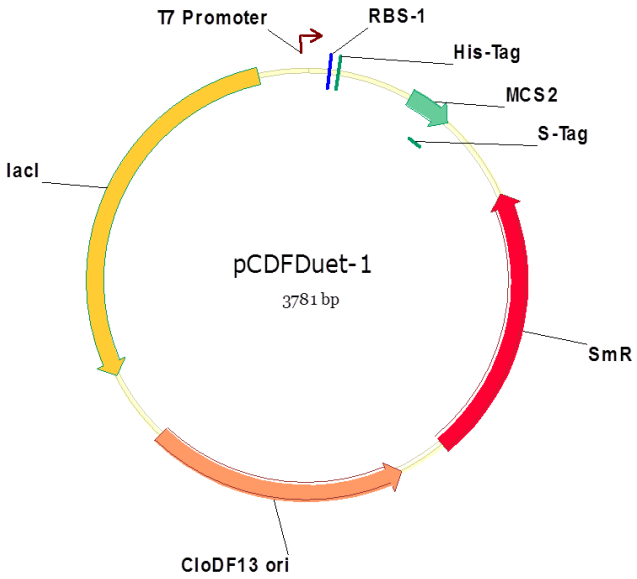


Figure 2.2. The pCDFDuet-1 vector. pCDFDuet vector contains a T7 promoter, origin of replication (CloDF13 ori), antibiotic resistance gene (Spectinomycin resistance) and two multiple cloning sites. MCS1 has an N-terminal His-tag and MCS2 has a C-terminal S-tag. MCS1 was used for ligation dependent cloning of phyB derived constructs. A C-terminal His-tag for purification was introduced by PCR.

2.1.4 Synthetic genes and bacterial genotypes

2.1.4.1 Synthetic genes

A codon-optimized DNA containing a ribosome binding site and the CDS encoding the first 655 amino acid residues (complete photosensory module) of *Sorghum PHYB* was synthesized by Eurofins Genomics (Germany). The synthetic 2kb gene was blunt-cloned into EcoRV site of pUC57. Similarly, the full *Sb.PIF1* gene (1669 bp) was synthesized and cloned into pUC57 by GenScript (NJ, USA).

2.1.4.2 *Escherichia coli* genotypes

For cloning and propagation of plasmids, library strains such as DH5α, TOP10F' and XL1-Blue were employed. For the production of recombinant proteins, expression strains such as BL21, $BL21_{PRO}$ and BL21(DE3) were used. $BL21_{PRO}$ constitutively expresses lac and Tet repressors for tight regulation of expression.

2.2 Methods

2.2.1 Molecular biology methods

2.2.1.1 Preparation of electrocompetent XL1-Blue cells by Tanja Gans

Electrocompetent cells were prepared by streaking glycerol stock XL1-Blue cells on fresh LB agar plate containing tetracycline and incubating it overnight at 37 °C. A single colony from this plate was used to inoculate LB + tetracycline (5 mg/ml) starter culture. The culture was incubated overnight at 37 °C while shaking at 140 rpm. 500 µl of this was inoculated into 250 ml TB containing tetracycline and further incubated at 18 °C until OD_{600} reached 0.4. The broth was then cooled on ice for 10 min. Subsequently, cells were harvested by centrifugation at 5,000 g and washed twice in ice-cold sterile water. DMSO was added with gentle swirling to a final concentration of 7 %. Aliquots of 100 µl were frozen in liquid nitrogen and stored at -80 °C.

2.2.1.2 Preparation of heatshock competent cells by Melanie Bingel

Heat competent DH5α cells were prepared by growing glycerol stock DH5α cells on fresh LB agar plate overnight at 37 °C. A single colony from this plate was used to inoculate 250 ml SOB in a 1L flask and incubated at 18 °C with vigorous shaking to OD_{600} ~ 0.4-0.6. The flask was then placed on ice for 10 min. Cells were harvested by centrifugation at 4,000 g and 4 °C for 20 min and resuspended gently in 80 ml ice-cold TB (or 0.1 M $CaCl_2$). Cells were once again pelleted at 4,000 g and 4 °C for 20 min. 20 ml ice-cold TB and and 1.4 ml DMSO (or 0.1 M $CaCl_2$ and 15 % glycerol) were added while swirling the culture gently for 10 minutes. Aliquots of 100 µl were immediately frozen in liquid nitrogen and stored at -80 °C.

2.2.1.3 Design of synthetic genes

Gene designer 2.0 is an open-access, easy-to-use bioinformatics software for *in silico* cloning, codon optimization, back translation and primer design developed by ATUM (Newark, California). *Sorghum PHYB* and *PIF1* 'native' DNA sequences were optimized with respect to codon adaptation index (CAI), GC content and mRNA-folding parameters using Gene designer 2.0. The sequences for the synthetic genes are given in Section 10.1.

2.2.1.4 Preparation of inserts

Forward and reverse primers containing endonuclease recognition sites (RE) for ligation-dependent cloning into pPROLar.A (Clontech) and pCDFDuet-1 (Novagen) were designed. PCR amplification of the synthetic genes was performed using Phusion high-fidelity DNA polymerase (NEB) according to manufacturer's instructions. Briefly, a 20 µl reaction mixture contained 5 ng of template DNA, 25 pmol of each primer, 5 nmol of dNTPs and 0.4 U of Phusion DNA polymerase. Annealing temperatures and extension times for the cycling parameters were adjusted based on the length of the primers and insert DNAs respectively. Reaction mixtures were first denatured for 30 s at 98 °C and then amplified for 35 cycles of 10 s at 98 °C, 30 s at 54/59 °C and 30 s at 72 °C and final extension of 10 min. at 72 °C using T3000 thermal cycler (Biometra). The amplified PCR products were analyzed by agarose gel electrophoresis.

Different fragments derived from the photosensory module of *Sorghum* phyB were named as follows: Sb.NPGP comprises residues 1 to 655, Sb.HPGP comprises residues 64 to 655 and Sb.PGP comprises residues 133 to 655. Likewise, different fragments derived from Sb.PIF1 were named as Sb.APB (residues 1 to 107), Sb.LEA (residues 1 to 303) and Sb.ALB (residues 1 to 387).

Table 4: Constructs, primers restriction sites and molar extinction coefficients.

Gene	Constructs	Primer pairs	Restriction sites	Encoded residues	ε_{280nm} (M^{-1} cm^{-1})
Sb.PHYB	*NPGP*	J27 & J30	EcoRI & HindIII	1 to 655	58,330
	HPGP	J28 & J30	EcoRI & HindIII	64 to 655	58,330
	PGP	J29 &J30	EcoRI &HindIII	113 to 655	55,350
	GP	SS64 & J30	NcoI & *HindIII*	267 to 655	46,870
	NPG	J27 & SS43	EcoRI & HindIII	1 to 458	25,900
	PG	J29 & SS43	EcoRI & HindIII	113 to 458	22,920
Sb.PIF1	*PIF1*		NdeI & XhoI	1 to 535	36,900
	APB	SS14 & SS5	NdeI & XhoI	1 to 107	17,990
	LEA	SS14& SS16	NdeI & XhoI	1 to 303	19,480
	ALB	SS14 & SS17	NdeI & XhoI	1 to 387	26,470

2.2.1.5 Agarose gel electrophoresis

A horizontal agarose gel-electrophoresis was used for separation and extraction of target DNAs. Samples and molecular weight standards in 6x loading dye were applied to 1 - 2 % agarose gel in TBE-running buffer. After running the gel at 200 V for 20 - 30 min, DNA bands were stained in ethidium bromide bath for 10 minutes. Gels were documented using a UV-B transilluminator (TCP-20-LM, Vilber Lourmat) equipped with a CCD-camera (Nikon).

2.2.1.6 Ligation-dependent cloning

Various PCR produced-inserts for *Sb.PHYB (NPGP*, *HPGP*, *PGP,* and *PG)* and *Sb.PIF1* (*APB, LEA, ALB, & PIF1*) as well as miniprep purified-vectors (pPROLar and pCDFDuet-1) were digested with restriction enzymes (NEB) for 1 h at 37 °C. EcoRI and HindIII sites were used for cloning into pPROLar vector whereas NcoI and HindIII sites were used for cloning into pCDFDuet vector. Digested vectors and PCR products were gel-purified using MinElute Gel extraction kit (Qiagen) according to the manufacturer's instructions. Gel-purified-linearized vectors were de-phosphorylated by antarctic phosphatase (NEB). Generally, 3-30 fmol of vector ends were ligated with a 3-fold molar excess of insert DNAs. The ligation mixtures containing digested-inserts, digested-vectors, T4-ligase and ligase buffer were incubated at 16 °C overnight. T4-ligase was heat deactivated prior to transformation.

2.2.1.7 Transformation of *E. coli* cells by heat-shock

Frozen aliquots of heat-shock competent cells were first thawed on ice for 5-10 minutes, mixed with 20 ng of total DNA (plasmid DNA or ligation mixture) and incubated on ice for 30 min. The cells were then heat-shocked at 42 °C for 45 s and then cold-shocked on ice for 5 min. The cells were incubated at 37 °C for 1 hr in 800 µl of SOC medium. Finally, cells were pour-plated on agar medium containing antibiotic for selection and incubated overnight at 37 °C.

2.2.1.8 Transformation of *E. coli* cells by electroporation

Frozen aliquots of electrocompetent cells were first thawed on ice for 5-10 minutes. Thawed-cells were mixed with 10 ng of total DNA (plasmid DNA or ligation mixture). The mixture was carefully transferred into pre-cooled cuvettes (2 mm width, Biorad). Cells were electroporated at 1,550 V, 21 µF and 400 Ω using Genezapper 450/2500 (IBI, Kodak)). Cells were transferred from the

cuvettes into 2 ml tubes in 800 µl of LB-Mg with a Pasteur pipette and incubated at 37 °C for 1 hr (TR-125, Infors HT). Finally, cells were pour-plated on agar medium containing antibiotic for selection and incubated overnight at 37 °C (B 5042, Heraeus).

2.2.1.9 Identification of positive clones

The presence of the target-DNA in the plasmid was quickly screened directly from *E. coli* colonies by colony PCR using homemade Taq polymerase (Courtesy of Melanie Bingel). The presence of the desired-insert was also independently tested by digestion of the plasmid with restriction enzymes that excise the insert, followed by separation by agarose gel electrophoresis. The reading frame of the construct was verified by test expression and Sanger's sequencing (GATC Biotech). The vector maps of the positive clones are given in Section 10.2.4.

2.2.1.10 Site-directed mutagenesis

Site-directed mutants of *Sb.PHYB* were generated from the wild-type plasmid by using either overlapping (QuikChange, Agilent) or back-to-back PCR-based methods. Primer pairs carrying the mismatch of interest were used to amplify the complete plasmid linearly (QuikChange) or exponentially (back-to-back). Primers SS1 to SS4 were used to generate the constitutively active Y291H, SS18 and SS19 to generate the Pfr stable G480E mutant, and SS52 and SS53 to generate the surface entropy reduction (E399A/Q400A) mutants. PCR products were then DpnI (NEB) digested for the removal of methylated parental plasmid DNA and self-ligated prior to bacterial transformation.

2.2.1.11 Transformation for production of recombinant proteins

The pPROLar-*NPGP*, pPROLar-*HPGP* and pPROLar-*PGP* plasmids were cotransformed into heat-competent $BL21_{PRO}$ cells with the p183 plasmid expressing the *Synechocystis* sp. PCC 6803 *HO1* heme oxygenase and *Arabidopsis thaliana* PΦB synthase (*HY2)* enzymes for production of recombinant holophytochrome. Similarly, *Sb.PHYB* inserts cloned into pCDFDuet (pCDFDuet–*NPGP*, pCDFDuet–*PGP*, and pCDFDuet–*PG*) were cotransformed into heat-shock-competent BL21(DE3) cells with p183 plasmid to perform *in vivo* assembly. Colonies from freshly transformed *E. coli* cells were selected on LB agar supplemented with glycerol and glucose for carbon catabolite repression and antibiotics (Amp and Spec) for selection. Various inserts of *Sb.PIF1*

cloned into pCDFDuet (pCDFDuet-*APB*, pCDFDuet-*LEA,* pCDFDuet-*ALB* and pCDFDuet-*PIF1*) were transformed into BL21(DE3) cells and directly screened on spectinomycin plates. For test expression, 100 ml LB media in 500 ml baffled flasks inoculated with freshly transformed *E. coli* cells were induced overnight at 37 °C after addition of 0.2 % arabinose and 1 mM IPTG at OD_{600} of 0.5 (UV-mini orange) shaking at 140 rpm (Centromat R, Sartorius).

2.2.1.12 Optimization

The Box-Behnken independent quadratic design in which the treatment combinations are at the midpoints and at the centre of the process space was used for optimization of culture conditions. For three factors, the design offers the advantage in requiring fewer runs than complete factorial design [137]. Culture conditions such as induction OD_{600}, inducer concentration and temperature for soluble phytochrome production were optimized accordingly. Induction OD_{600} was tested at 0.5, 0.75 and 1 (UV-mini orange), arabinose concentration 0.2 %, 0.4 % and 0.6 %, and temperature at 18 ^{o}C (-1), 28 ^{o}C (0) and 37 ^{o}C (+1). Expression levels were also compared between different growth media such as LB, TB and SB.

2.2.1.13 Starter culture

Once the optimal clone and culture conditions were established, the corresponding clone was used to prepare glycerol stock and to inoculate an LB pre-culture containing the desired antibiotics. Starter cultures were grown by inoculating colonies from freshly transformed *E. coli* cells into 2 ml LB medium in 10 ml culture tubes at 37 ^{o}C shaking at 140 rpm. The optical density of the starter culture was carefully monitored to avoid overgrowth. The starter culture was allowed to grow at 37 °C with vigorous shaking at 140 rpm (TR-125, Infors HT) until OD_{600} (UV-mini orange) was ~0.6.

2.2.1.14 Production of *Sorghum* phytochrome in SB medium

400 ml autoclaved SB main culture supplemented with the appropriate antibiotics (Amp and Kan or Spec) in 2L baffled flask with high surface area/volume ratio was inoculated with 400 µl of the starter culture. The main culture was incubated at 37 °C shaking at 140 rpm (TR-125, Infors HT) until an OD_{600} (UV-mini orange) of 0.8 was achieved. The flask was then cooled in an ice-water bath for 0.5 hr and holoprotein production was induced by adding 0.2 % arabinose and 1 mM

IPTG for pPROLar or 1 mM IPTG alone for pCDFDuet clones. The ice-cooled culture was then grown for 16 h at optimal temperature in darkness (B 5042, Heraeus) or (Innova 42/42R, Eppendorf) shaking at 140 and 120 rpm respectively. The optimal temperature was found to be protein specific. 24 °C, 28 °C, and 30 °C were optimal for NPGP, PGP and PG respectively. Protein producing cells were harvested by centrifugation (RC2-B, Sorvall) at 6,000 g and 4 °C for 30 min. Cells were frozen and stored at -80 °C in freezer (Hera freezer, Heraeus) until the next step.

2.2.2 Methods in protein biochemistry

2.2.2.1 Extraction and clarification of phytochrome from *E. coli* cells

To test phytochrome production, induced cultures were transferred into conical centrifuge tubes after measuring the OD_{600} at harvest. The cells were harvested by centrifugation at 4,000 rpm for 20 minutes (Biofuge, Heraeus). The harvested cells were either used directly or stored in -80 °C freezer. The amount of TES-β for extraction was calculated by dividing the product of OD_{600} at harvest and the volume of the main culture (ml) by 20. Cell density at harvest was measured by five-fold dilution of the cell culture in SB medium. For example, 4 ml of TES-β was used for 100 ml of pPROLar-*PGP* inoculated culture harvested at OD_{600} of 0.8. The cells were thoroughly resuspended by pipetting up and down with 5 ml pipette and cracked using a 10 ml pre-cooled French pressure cell with two passages at maximum force of 10,000 N (PP60KN, Watz Hydraulik).

The supernatant was clarified by centrifugation in Oakridge tubes at 4 °C and 12,000 g for 30 min (RC2-B, Sorvall). The clarified supernatants were transferred into clean conical centrifuge tubes and 0.6 ml of each sample was taken for absorbance measurement. Absorbance spectra of the clarified supernatants were taken after FR and R irradiation, respectively in a quartz cuvette on Agilent 8453 UV-Vis spectrophotometer. The spectrophotometer in Hughes lab is equipped with homemade cuvette holder and R/FR light sources which allow fast and reproducible irradiation of multiple samples.

For phytochrome purification, frozen cells were thawed at 4 °C overnight and resuspended in cold extraction buffer. The amount of extraction buffer was obtained by dividing the product of OD_{600} at harvest & culture volume (ml) by 20. Resuspended cells were cracked by using a 50 ml pre-cooled French pressure cell with two passages with a force of 21,000 N, and cell debris was removed by centrifugation at 4 °C and 15,000 g for 30 minutes (RC2-B, Sorvall). Ice-chilled AmS buffer was then added to the supernatant to bring final concentration to 60 % saturation (3 : 2

ratio V_{Ams} : V_{sample}) and stored on ice for 0.5 hr. The resulting slurry was transferred into pre-cooled Oakridge tubes and centrifuged at 4 °C and 5,000 g for 30 minutes. The supernatant was decanted carefully and the pellet was spun briefly to remove the remaining liquid. The dark-green pellet was resuspended in binding buffer ($TISI^{10}$-β, half the volume of extraction buffer) and centrifuged at 4 °C and 15,000 g for 20 minutes. The clarified supernatant was used for Ni-NTA affinity chromatography.

2.2.2.2 Purification of phytochrome by Ni-NTA affinity chromatography

Original protocol

The original protocol used Tris-based buffers (TES-β and $TISI^{10}$-β) for extraction and affinity purification. The AmS precipitated and resuspened supernatant was loaded to a 5-10 ml Ni-NTA column (Qiagen) previously equilibrated with 5 column volumes (CV) of binding buffer. The flow-through was allowed to exit the column very slowly by the action of gravity. The column was washed with 3CV of binding buffer to remove unbound protein and then with 3CV of the washing buffer to remove weakly bound proteins. Finally, the bound phytochrome was eluted by 3CV of elution buffer and collected in 2 ml fractions. Sb.phyB constructs typically eluted at a concentration of 100 mM imidazole. The collected fractions were analyzed by UV-Vis spectroscopy and SDS-PAGE. Fractions displaying green color and high A_{660}/A_{280} specific absorbance ratio (SAR) were pooled and the protein yield was determined by UV-Vis spectroscopy.

Improved protocol

The improved purification protocol was based on HEPES buffer for cell lysis, Ni-NTA and SEC. The modified extraction buffer was 50 mM Na HEPES pH 7.8, 5 mM EDTA, 300 mM NaCl and 1 mM β-ME. Likewise, the modified binding buffer for Ni-NTA was 50 mM HEPES pH 7.8, 500 mM NaCl, 1 mM IDA, 10 mM imidazole and 1 mM β-ME. Bound proteins were washed with 5CV of the binding buffer followed by 3CV of 50 mM imidazole in binding buffer. Finally, the sample was eluted with 5CV of 250 mM imidazole in the binding buffer.

2.2.2.3 Buffer exchange

AmS precipitation, desalting columns and dialysis were used for buffer exchanging phytochrome samples.

2.2.2.4 Size-exclusion chromatography (SEC)

A Superdex 200 26/60 prep grade column (V_t = 318 ml, V_0 = 105 ml) connected to an Äkta Purifier (GE Healthcare) and equilibrated with TES-β or HES-β was used for preparative SEC at room temperature. Highly-concentrated phytochrome samples (5 to 10 mg ml^{-1}) were irradiated with saturating FR-light in a 250 µl HPLC glass syringe prior to loading. SEC was carried out in safelight (ca. 490 nm) to avoid photoconversion using TES-β as running buffer. The eluate was monitored at 280 nm (predominantly tryptophan absorption) and at 660 and 700 nm where only phytochromes absorb. Molecular weight standards used to calibrate the column include thyroglobulin (669 kDa), ferritin (440 kDa), catalase (232 kDa), aldolase (158 kDa), albumin (67 kDa), ovalbumin (43 kDa), chymotrypsinogen A (25 kDa) and RNAse A (14 kDa) (LMW calibration kit, GE Healthcare). Eluates were collected in 1 ml fractions and analyzed by UV-Vis spectroscopy and SDS-PAGE. Fractions displaying high SARs were pooled and stored at 4 °C.

2.2.2.5 Ion-exchange chromatography (IEC)

A Mono Q 5/50 GL high performance column (V_t = 1 ml, Q= +) connected to Äkta Purifier and equilibrated with ion-exchange binding buffer was used at room temperature. After the column was appropriately equilibrated (~20CV) and A_{280} and the conductivity baseline stabilized, 2 ml concentrated (5 to 10 mg ml^{-1}) phytochrome was injected. Unbound material was washed with 10CV binding buffer and bound sample was eluted with a continuous gradient of ion-exchange elution buffer (20CV). Mono Q eluates were collected in 0.4 ml fractions and subsequently analyzed by UV-Vis spectroscopy and SDS-PAGE.

2.2.2.6 Extraction and purification of Sb.PIF1

Frozen cell pellets were lysed by French press in 50 mM Tris pH 7.8, 300 mM NaCl, 5 mM imidazole, 1 mM β-mercaptoethanol, 1 % triton X-100, 2 mM phenylmethylsulfonyl fluoride (PMSF), and 1 µg/ml DNase I. Triton X-100 at 1 % was added to solubilize proteins from the cell debris. DNase I was used for degradation of nucleic acids dissociated from DNA-protein complex

in the presence of high salt concentration. β–mercaptoethanol was included to maintain reducing environment.

2.2.2.7 SDS-PAGE

Laemmlli denaturing sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) was used to analyze recombinant proteins [138]. Protein samples (V_t=50 µl, C_{max}= 1 mg ml^{-1}) were mixed with 10 µl of 6x SDS sample loading buffer and incubated for 10 minutes at 95 °C. High range PM-III (AppliChem) was used as molecular mass standard. Boiled samples were spun briefly at 10,000 rpm for 30 s and the supernatant (10-15 µl) loaded on to 0.75 mm 10 - 15 % SDS-PAGE gels. Proteins were separated at 180 V in TGS running buffer until the dye front reached the end of the gel (40 to 60 minutes).

2.2.2.8 Protein staining

After separation, proteins were detected by staining with zinc acetate and/or Coomassie Brilliant Blue R-250. For zinc-induced fluorescence, the gel was incubated in 1 mM zinc acetate for 20 minutes, washed with water and photographed under UV light to detect covalently-attached bilin [139]. For hot Coomassie staining, the gel was warmed in a microwave (600 W, 30 s) in hot Coomassie solution and allowed to cool while shaking. The staining solution was discarded and the gel was rinsed with water and warmed in destaining solution. Destaining was repeated three times until the background was completely removed.

Densitometric estimation of protein bands from Coomassie-stained gels was carried out using ImageJ (NIH, Bethesda). The picture mode of high resolution scans for SDS-PAGE gels was changed to ‘Grayscale’. A lane was selected as a region of interest (ROI) using rectangular selection tool and marked for analysis using ‘Gels’ and ‘Select first lane’ under the “Analyze” menu. The different bands of the selected lane were measured using the ‘Plot Lanes’ function and the ‘Wand’ tracing tool after setting a horizontal baseline. Estimated densities for the different bands were analyzed by exporting the data into spreadsheet.

2.2.2.9 Protein concentration

The volume of the sample after SEC was reduced by different membrane ultrafiltration methods. A pressure-based stirred ultrafiltration cell (Amicon) was used for large sample volumes (>5 ml)

whereas Vivaspins were used for small sample volumes (<1 ml). Membranes (MWCO 30 KDa) were rinsed extensively with milli-Q water and buffer in advance of concentration. Monodispersed phytochrome preparations were concentrated to 15 mg ml^{-1} (0.21 mM) by a combination of stirred-ultrafiltration membrane and Vivaspin.

2.2.2.10 Protein quantification

For determination of total protein concentration by the Bradford assay, 20 µL of appropriately-diluted protein sample was added to 1 ml of Bradford reagent [140]. At the same time, multiple standard lysozyme samples (10 - 100 µg) were mixed with the reagent. The color of the mixture was allowed to develop for 10 minutes at room temperature and absorbance of the solution was measured at 595 nm. A calibration curve, prepared by plotting lysozyme concentration versus absorbance at 595 nm, was used to estimate the concentration of phytochrome.

For determination of total protein concentration from $A_{280\ nm}$, calculated ε and $\varepsilon_{0.1\%}$ values were used. Many sources calculate absorbance ($A_{280\ nm}$, $\varepsilon_{0.1\%}$) values for 0.1 % (= 0.1 g/100 ml) protein solutions measured in a 1 cm path length. These values can be understood as attenuation coefficients having units of $(mg/ml)^{-1}\ cm^{-1}$ instead of $M^{-1}\ cm^{-1}$. Consequently, when these values are applied as extinction coefficients, the unit for protein concentration, c, is mg/ml. Molar concentrations for normalization in units $M^{-1}\ cm^{-1}$ can be obtained by multiplying the calculated $\varepsilon_{0.1\%}$ by the molar mass of the protein.

2.2.2.11 Quantitative determination of -SH groups using Ellman's reagent

Ellman's reagent (5,5'-dithio-bis-(2-nitrobenzoic acid)), also known as DTNB, reacts with a free sulfhydryl group and the efficiency of labeling is sensitive to the accessibility of the reactive group at the protein surface [141]. Phytochrome samples were diluted before this assay. DTNB stock solutions were prepared in 0.1 M Tris-HCl pH 7.5. Cysteine was used as a standard to build a calibration curve 10 to 500 µM. DTNB working solution (950 µl) was mixed with 50 µl test sample or standards. The color of the mixture was allowed to develop for 5 minutes at room temperature and absorbance of the solution was measured at 412 nm. A calibration curve, prepared by plotting standard cysteine concentrations versus absorbance at 412 nm, was used to estimate the number of free surface exposed-cysteines in phytochrome samples.

2.2.2.12 Limited proteolysis

Sequencing grade trypsin (Roche) that specifically cleaves peptide bonds C-terminally at lysine and arginine residues was used. Purified phytochrome samples were irradiated with FR-light and subsequently kept in total darkness. Samples were digested with increasing protease concentrations at 18 °C for 2 hours in darkness. Proteolysis was stopped by addition of SDS sample loading buffer.

2.2.3 Methods in structural biology

2.2.3.1 Crystallization screening and optimization

For setting up initial screening, commercially available precipitants were used at two protein concentrations (ca. 8 and ca. 15 mg ml^{-1}). A Honeybee 963 bench-top robotic system (Genomic Solutions) was employed to mix protein solutions with NeXtal suite screening kits (Qiagen). A drop size of 400 nl was formed by mixing 200 nl protein solution with 200 nl screening solution using the sitting-drop vapor-diffusion method in 96-well MRC plates (Jena Bioscience). The same screening solution (80 µl) was added to the reservoir. The entire procedure was carried out in blue-green safelight (ca. 490 nm) where plates were irradiated with far-red (720 nm) light after pipetting and stored at 18 ^{o}C in darkness (Initial screening was kindly conducted by Ralf Poeschke).

Experimental drops were observed under safelight (490 nm) daily for the first week and weekly thereafter. During high-throughput initial screening, crystallization conditions for Sb.NPGP, Sb.PGP and Sb.PG were identified. These conditions were systematically optimized in 24-well plates (Sarstedt) with 2 µl drop size (1 µl protein + 1 µl reservoir solution) using hanging-drop vapor diffusion method with 500 µl crystallization solution in the reservoir.

2.2.3.2 Harvesting and cryoprotection

Different concentrations of cryoprotecting agents such as glycerol, ethylene glycol, 2-methyl-2,4-pentanediol (MPD) and low-molecular-weight PEGs were added to the mother liquor to make cryosolutions. The resulting solutions were tested for their ability to vitrify by flash-freezing 3 µl of the solutions within 10 µl tips in liquid nitrogen. Acupuncture needles were used to remove ‘proteinaceous-skin’ and to free NPGP crystals from the surface during harvesting. Crystals were

cryoprotected by brief transfer to the crystallization solution supplemented with either 20 % glycerol or 25 % PEG 400 and then flash-frozen in liquid nitrogen.

2.2.3.3 Alternative crystallization methods

Counter-diffusion using Granada crystallization boxes and sitting drop using micro-bridges were carried out to grow NPGP crystals without the formation of the proteinaceous skin. Granada crystallization box [142] was used for growing protein crystals inside capillaries with gelled and un-gelled precipitating agent by the counter-diffusion technique. The protein solution was introduced by capillarity. The crystallization cocktail was filled into the box and was allowed to diffuse by the action of gravity. A micro-bridge with a smooth concave depression at the center was inserted into the reservoir of 24-well plates to perform sitting drop vapor diffusion experiments with 4 µl drops.

2.2.3.4 Post-crystallization treatments (PXT)

Post-crystallization treatments such as annealing, dehydration, cross-linking and reduction [143, 144] were used to improve the diffraction quality of NPGP and PGP crystals. Crystal annealing was done by warming the frozen crystal to room temperature for 10-20 seconds followed by flash-freezing. Since phytochromes are photosensitive, annealing was carried out in near darkness in the experimental hutch. Phytochrome crystals were taken in their plates to the BESSY II synchrotron facility for dehydration experiments. The HC1c dehydration device installed on BL14.3 produced an open stream of air at room temperature with the relative humidity (RH) adjustable between 45 % and 99.7 % [145]. The stream was dispensed by a nozzle similar to cryostream devices. Crystals were placed in the air stream mounted on mesh loops. For cross-linking experiments, glutaraldehyde (0.1 %) was added directly to the reservoir or micro-bridge and allowed to equilibrate. Similarly, β-mercaptoethanol (1.43 mM) was added to the reservoir after the protein drop was set up to grow big protein crystals upon reduction.

2.2.3.5 Carboxymethylation of cysteines

Phytochrome samples purified by the improved Ni-NTA protocol and SEC were AmS precipitated, the pellets resuspended in 2 ml of 100 mM ammonium bicarbonate (pH 8.0) and protein concentration determined by Bradford assay. The sample was then incubated at 4 ^{o}C overnight

with five-fold molar excess of DTT for complete reduction. The sample was kept in darkness (wrapped in aluminum foil) and freshly prepared iodoacetamide (IAA) was added to final concentration of 20 mM. The solution was stirred for 1 hour at 4 ^{o}C in darkness and DTT added to final concentration of 20 mM to quench excess IAA. The protein solution was buffer exchanged prior to crystallization screening.

2.2.3.6 Surface entropy reduction (SER)

Surface entropy reduction is a strategy of mutating surface K and E or Q residues [146]. Two surface exposed high entropy residues in Sb.PGP and Sb.NPGP (E399 and Q400) were identified by automated homology modeling (ITASSER) [147]. These residues were mutated into alanines by site-directed mutagenesis. The resulting surface-engineered protein was subjected to crystallization screening (AG Essen, Marburg).

2.2.3.7 Truncation

Protein engineering strategies to improve diffraction quality include N-terminal and C-terminal truncations as well as internal loop deletion [148]. In the recently-solved crystal structure of At.phyB (4OUR), electron density for the PHY domain was poorly resolved. Based on this, a PG construct lacking the PHY domain was produced, purified and subjected to crystallization screening (AG Essen, Marburg).

2.2.3.8 Additive screening

Additives were screened using the sitting drop vapor-diffusion method in 96-well format to evaluate the effect of small molecules on the diffraction-quality of the obtained crystals. HR2-138 additive screening kit (Hampton) includes small molecules such as amino acids, dissociating agents, linkers and polyamines, polymers, chelating agents, detergent and non-detergent amphiphiles, among others. All crystallization conditions obtained during high-throughput initial screening for all purified constructs (NPGP, PGP-WT, PGP-SER and PG) were optimized by additive screening.

2.2.3.9 Diffraction experiments

Diffraction experiments were conducted at beamlines 14.3 at BESSY II, EMBL Hamburg (courtesy of Ankan Banerjee) and ID14-3, ESRF Grenoble (courtesy of Sophie Franz and Marian Vogt). Complete data set were collected for PGP (Courtesy of Sophie Franz at ESRF) and PG (BESSY II).

2.2.4 Sequence analysis

A comprehensive resource for protein sequence (UniProt) and a database of manually-curated disordered protein regions (Disprot) were used for sequence retrieval. PONDR-FIT was employed to predict intrinsically-disordered regions from protein sequences. Atomic coordinates for various structures were obtained from the Protein Structure Databank (PDB). Multiple sequence alignments were generated with ClustalW. STRING was used to identify putative phytochrome-interacting proteins. Conserved protein domains in a set of PIFs from different plants were detected with MEME and SMART (EMBL). Principal component analysis (PCA) using MATLAB was utilized to classify monocot and dicot sequences based on dipeptide percentage composition of NTE in plant phytochromes.

3 Results

3.1 The photosensory module of Sb.phyB

The 3D structures of plant phytochromes provide valuable insights into how fundamental photochemical and biochemical processes are conducted at the atomic level in this group of photoreceptors [149]. Since plant phytochromes signal from the N-terminal photosensory module, determining the structure of this module from phyB in complex with PIF3 would be a cornerstone of plant photobiology [113]. *MA3*, which encodes phytochrome B, regulates photoperiodic sensitivity of flowering in *Sorghum* [123]. Structural-function studies of phyB and PIF from *Sorghum bicolor* would shed light on the evolutionary divergence of the phy-PIF system between dicots and monocots.

3.1.1 Molecular cloning and expression

The full-length *Sorghum* phytochrome B apoprotein is a large soluble protein with a molecular mass of 129 kDa comprising 1178 amino acids [123]. The PΦB A ring is expected to be covalently linked to the apoprotein at C372 to form the holoprotein (Fig. 3.1). The domain organization of Sb.phyB and the sequence of the NTE were analyzed before molecular cloning.

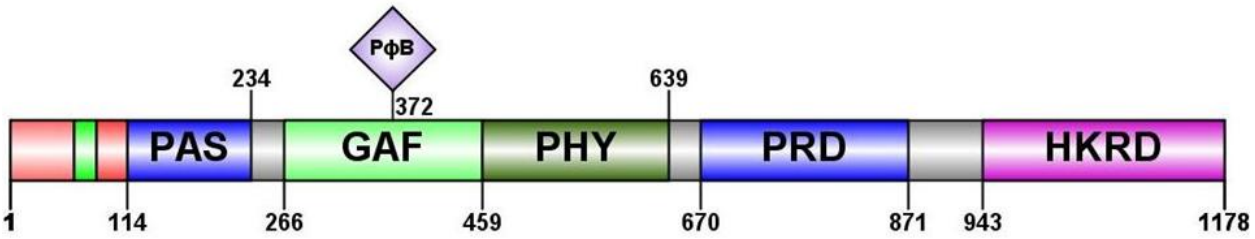


Figure 3.1. Domain organization of *Sorghum* phyB. The first 113 residues belong to the N-terminal extension. NTE, PAS, GAF and PHY domains comprise the photosensory module followed by the PAS repeat domain (PRD) and histidine kinase related domain (HKRD).

3.1.1.1 Sequence analysis of the N-terminal extension (NTE)

A dataset comprising the NTEs of phytochromes from mono- and dicotyledonous plants was compiled (Section 10.2.1) based on published diversity of canonical phytochromes in green plants [130]. The various sequence features of the NTE of plant phytochromes were compared with those of prokaryotic phytochromes. A heat-map was generated to reflect the general percentage

dipeptide compositional bias of the NTE. Sequence analysis of the NTE in this dataset revealed the following features. The NTE is unique to plant phytochromes and it is highly variable in sequence compared to all other phytochrome regions. The NTE is devoid of tryptophan and cysteine residues. GG, GS and SQ were among the most abundant dipeptides in the NTE of plant phytochromes (Figure 3.2). Generally, monocots have longer NTE than dicots which in turn have longer NTE than lower plants. The NTE of monocots has different compositional bias from that of dicots. The NTE is more heterogeneous in dicots than monocots. All monocots clustered into the same quadrant by PCA analysis (Figure 3.3). Both A and B-type phytochromes contain a low-complexity region containing tandem array of glycine and/or histidine and/or asparagine residues.

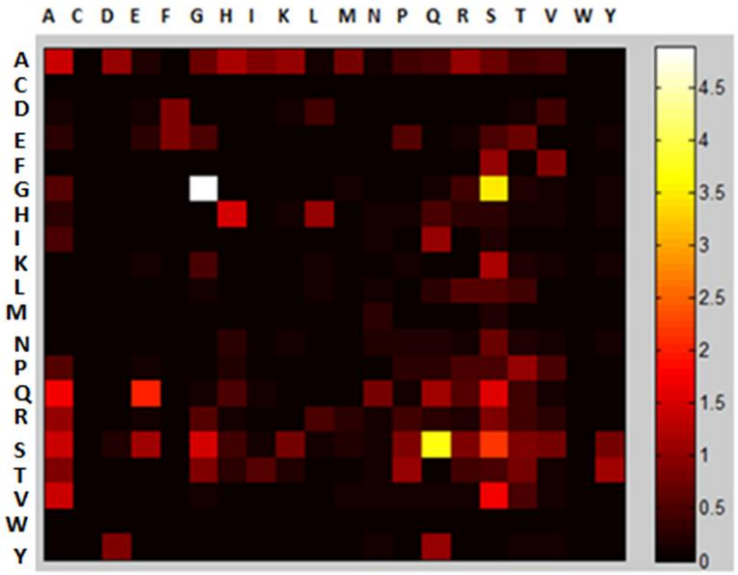


Figure 3.2. The percentage dipeptide composition of NTE in plant phytochromes. The relative composition of four hundred dipeptides in the NTE of plant phytochromes was calculated and weighted into heat map.

The NTE in Sb.phyB is rich in charged and polar amino acids (pI 10.2) but devoid of hydrophobic residues (GRAVY -0.75). It can be divided into a low complexity region (LCR), a helix forming motif (HFM) and an amino-terminal sequence (ATS). LCR in Sb.phyB (Fig. 3.4) contains serine-, histidine- and glycine-rich segments (S-, H- and G-segments). Secondary structure prediction for the NTE of Sb.phyB postulates LCR and ATS to be intrinsically disordered, interspaced by a small helical region (HFM). While many proteins need to adopt a well-defined structure to carry out their

function, intrinsically disordered regions are not able autonomously to fold stably but can nevertheless be functional [151, 152].

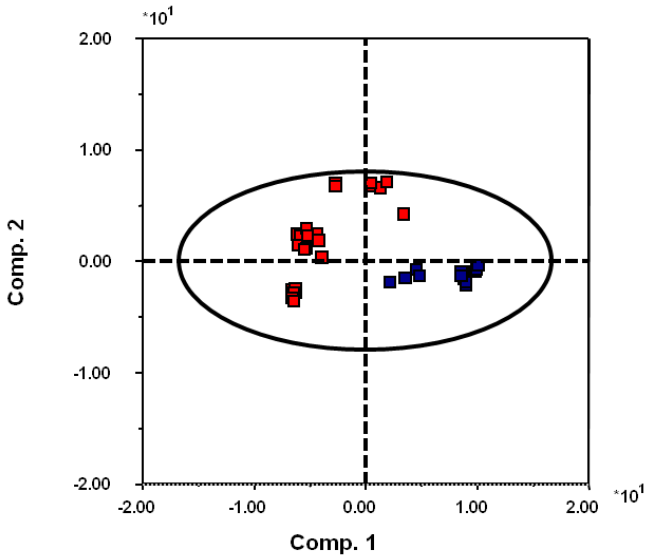


Figure 3.3. Graphical representation of the compositional bias in the NTE of plant phytochromes. Correlated dipeptide compositions of the NTE were transformed by linear combination into a small number of uncorrelated variables called principal components. The first two principal components were used to categorize monocots and dictots. Monocots are labeled blue and dicots are labeled red.

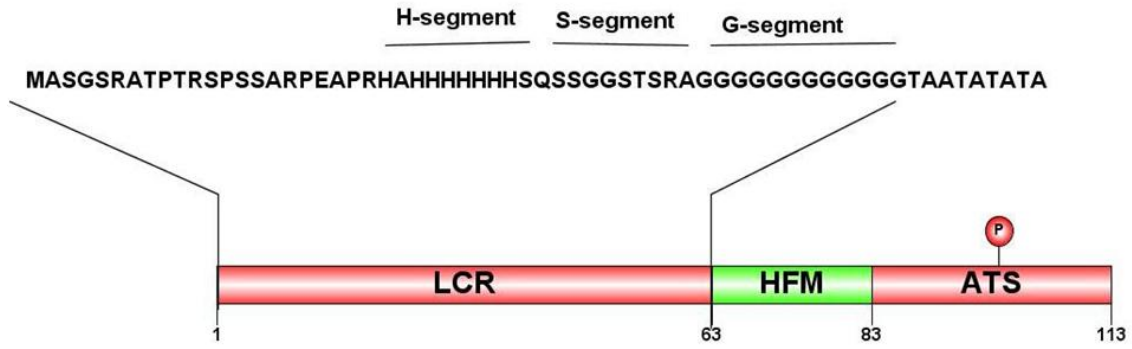


Figure 3.4. Schematic representation of the NTE of Sb.phyB. The positions of the different segments in the low-complexity region (LCR) are mapped onto the sequence. HFM is putative helix-forming motif where as ATS is the amino-terminal sequence. Phosphorylation site is based on similarity to At.phyB.

3.1.1.2 Custom designed synthetic gene for Sb.phyB

The Gene Designer 2.0 software package was used to optimize the coding sequence of *Sb.PHYB* (*AF182394.1*) for heterologous expression in *E. coli* [153]. A synthetic gene encoding the complete photosensory module (1-655) was designed while considering the following parameters. The correlation between the codon bias of *AF182394.1* and its expression host *E. coli* was maximized as measured by codon adaptation index (CAI) [154]. Other considerations included modifying translation initiation regions [155], avoidance of mRNA secondary structures [154] and the presence or absence of specific restriction sites. The final CAI for the synthetic gene was made 0.85. The optimized CDS given in Section 10.1 was synthesized by Eurofins Genomics (Germany).

3.1.1.3 Construct design

In designing the constructs, a truncated NTE was included, namely the predicted helix-forming motif (HFM). Three constructs, NPGP (NTE-PAS-GAF-PHY), HPGP (HFM-PAS-GAF-PHY) and PGP (PAS-GAF-PHY), were designed *in silico* (see Table 4 and Fig. 3.5). Smaller constructs are presented in Sections 3.2 & 3.3.

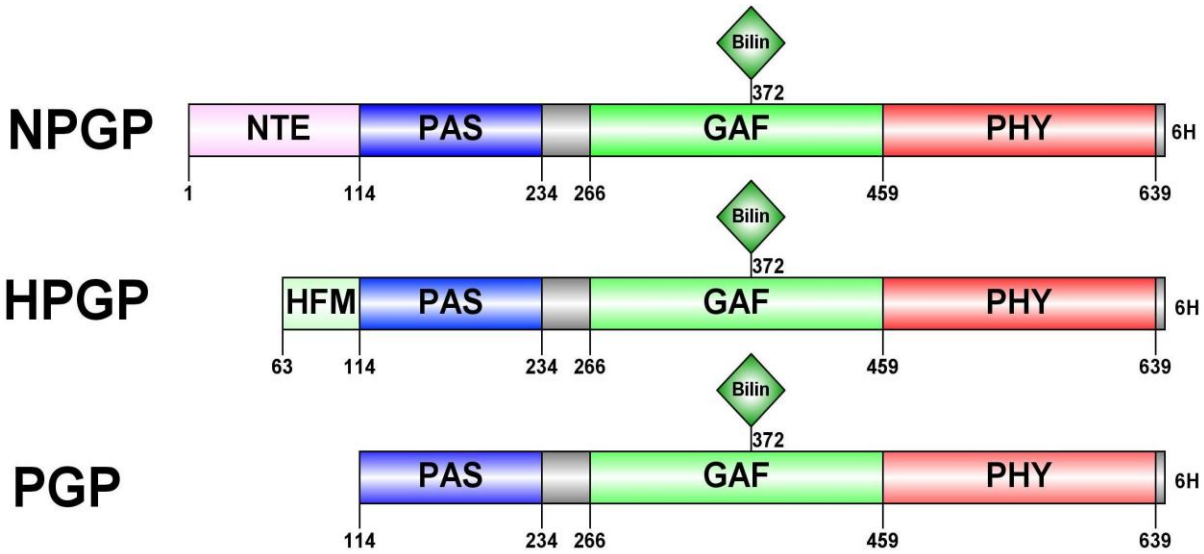


Figure 3.5. Multi-construct approach for the photosensory module Sb.phyB. The three constructs, NPGP, HPGP and PGP, contained different segments of the NTE. In all constructs, the C-terminal PAS repeat domain (PRD) and transmitter like module (TLM) were deleted and a $(His)_6$ tag added.

3.1.1.4 Cloning *Sb.PHYB* inserts into pPROLar

DNA fragments for insertion into pPROLar vector were generated by PCR amplification of the synthetic gene with the error-checking Phusion polymerase (NEB). Gene-specific primers containing restriction sites and an 18-base 3' addition encoding a His_6-tag for affinity purification were used to amplify *in silico* designed fragments (Fig. 3.5) from the template synthetic gene. The insert PCRs yielded the expected 2, 1.8 and 1.7 kb products after agarose gel electrophoresis (Figure 3.6) corresponding to *Sb.NPGP*, *Sb.HPGP* and *Sb.PGP* respectively.

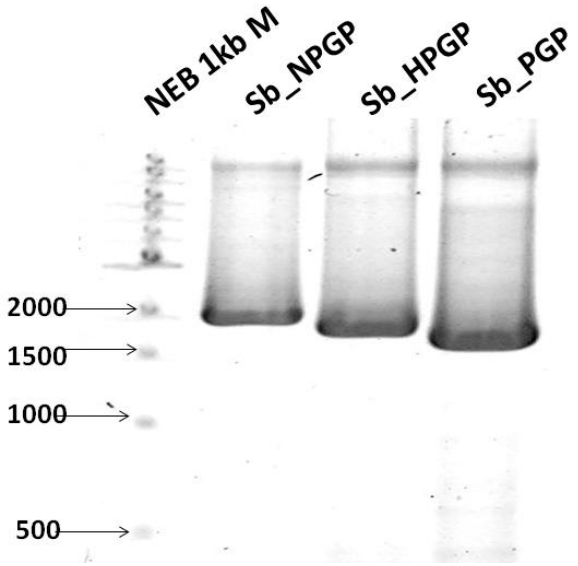


Figure 3.6. Preparation of inserts by PCR to create expression clones. PCR products of different sizes (*NPGP, HPGP and PGP)* were separated on 1 % agarose gel and stained with ethidium bromide. Molecular weight markers in bp (M).

The linearized and gel purified pPROLar vector and digested PCR products were ligated and transformed into XL1-Blue cells by heat-shock. *Sb.PGP* ligations were rather inefficient yielding 50 to 100 clones per µl ligate. Ligation was much more inefficient for *Sb.NPGP* with about ten-fold fewer colonies in *Sb.NPGP* than *Sb.PGP*. Sixteen colonies from each ligation were picked and screened by colony PCR. Positive colonies were independently screened by restriction mapping of miniprep DNA. Three positive colonies were identified for each construct and the integrity of the reading frame for each construct was validated by sequencing. The sequenced insert had the same DNA sequence as the one produced by *in silico* cloning (Figure 3.7).

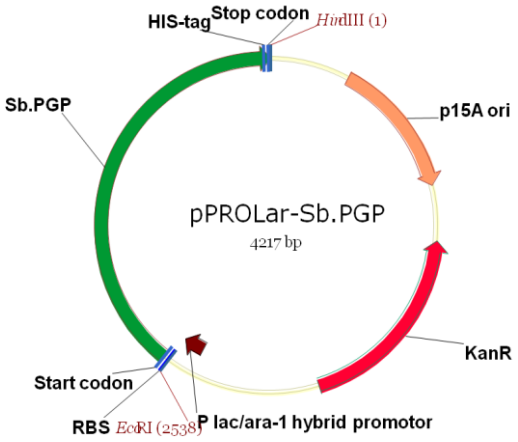


Figure 3.7. The pPROLar-*Sb.PGP* plasmid. *Sb.PGP* was inserted at EcoRI and HindIII sites. The ribosome binding site (RBS) was introduced by PCR primers. PGP has C-terminal His-tag for purification followed by stop codon.

3.1.1.5 *In vivo* assembly of holophytochrome

In the past, attempts to express plant phytochromes in *E. coli* met with limited success. Apophytochromes require the correct covalent and autocatalytic attachment of the chromophore to form a functional, photoactive holophytochrome [156, 157]. Hemeoxygenase (*HO1*) from *Synechocystis* and *HY2* from *Arabidopsis thaliana* were co-expressed in *E. coli* in the Hughes lab by F.T Landgraf (p183) to overproduce PΦB. This plasmid was cotransformed into $BL21_{PRO}$ together with each of the three Sb.phyB plasmid-constructs: pPROLar-*NPGP*, pPROLar-*HPGP* and pPROLar-*PGP*. ApoPGP can autocatalytically assemble with PΦB *in vivo* to produce holoPGP (Figure 3.8).

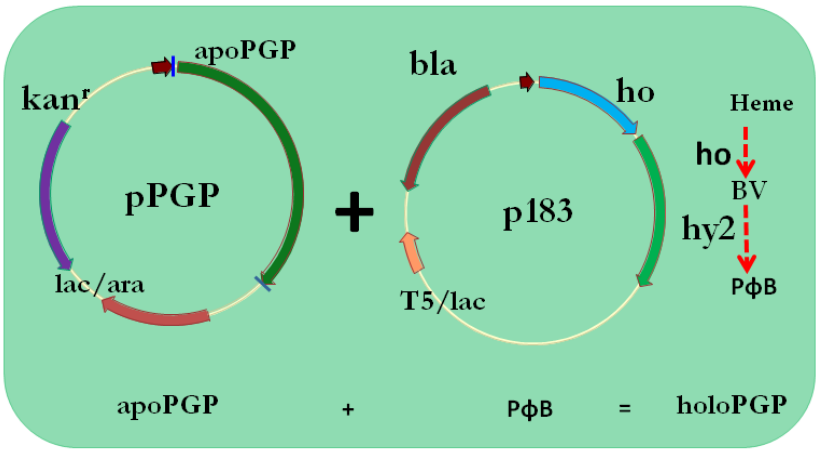


Figure 3.8. Schematic representation of holophytochrome biosynthesis in *E. coli*. Phytochrome apoprotein (apoPGP) is synthesized from pPGP plasmid whereas heme oxygenase (ho) and hy2 which converts endogenous heme to BV and BV to PΦB respectively are produced from p183 plasmid. apoPGP and PΦB autocatalytically assemble into holoPGP.

3.1.1.6 Expression screening from pPROLar

The different pPROLar Sb.phyB constructs (pPROLar-*NPGP*, pPROLar-*HPGP* and pPROLar-*PGP*) were cotransformed with p183 into heat-shock competent $BL21_{PRO}$ cells for production of holophytochromes by *in vivo* assembly with PφB. Small-scale preliminary expression and extraction experiments were carried out for three independent colonies from each construct. The spectra after FR and R irradiation were collected and the levels of soluble phytochrome in the clarified supernatants were compared from the absorbance difference spectra (FR - R). The production of soluble holophytochrome was further checked by comparing induced and uninduced samples on SDS-PAGE followed by zinc-induced fluorescence and hot Coomassie staining. Based on lack of absorbance difference spectrum and zinc-induced fluorescence after SDS-PAGE from all the supernatants, it was concluded that phytochrome production from all constructs in LB medium was unsuccessful. Green color observed for pPROLar-*NPGP*-PφB cells derived from bilin production is necessary but not sufficient for holophytochrome production.

At this point, different bacterial strains ($BL21_{PRO}$ and BL21) under different growth media (LB, TB and SB) and different induction temperatures (30 ^{o}C and 37 ^{o}C) were compared for the production of the three constructs (pPROLar-*PGP*, pROLar-*HPGP* and pPROLar-*NPGP*). However, only Sb.PGP was produced robustly from pPROLar clones (see Fig. 3.9). PGP was produced at 30 and 37 ^{o}C in SB medium from $BL21_{PRO}$ by adding 0.2 % arabinose and 1 mM IPTG overnight. The production level was too low to detect by Coomassie staining, but the specific and sensitive zinc-induced fluorescence after SDS-PAGE revealed that holoPGP was assembled. HPGP and NPGP found to be more challenging as the production level was barely detectable (see Figure 3.9).

The yield of photoreversible holoPGP obtained after lysing 100 ml of induced-culture harvested at OD_{600} of 0.8 in 4 ml TES-β estimated by Δ(ΔA) was 40 mAU (Figure 3.10). Hence it still required additional optimization efforts. The Box-Behnken design (BBD) was used for optimization of culture conditions such as growth temperature, inducer concentration and induction OD_{600}.

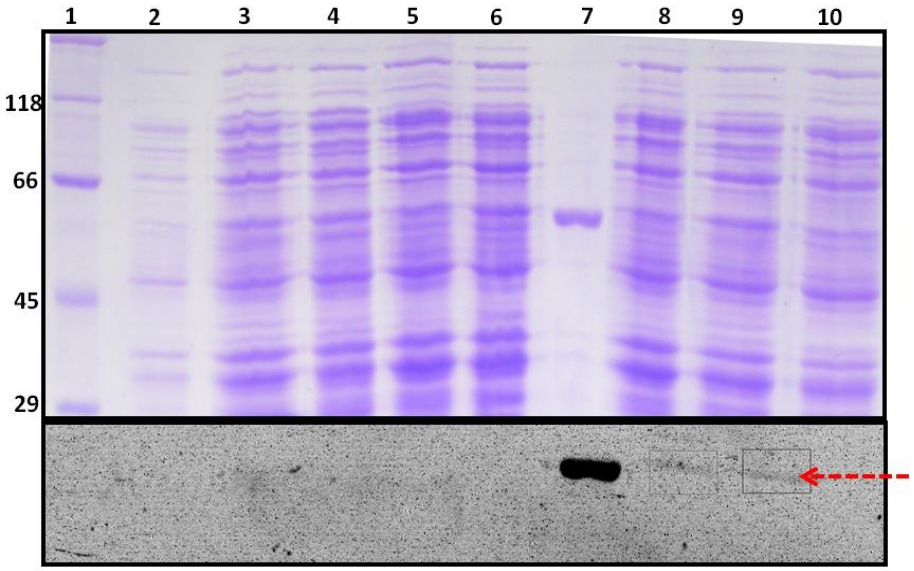


Figure 3.9. SDS-PAGE gel showing test expression of pPROLar clones in $BL21_{PRO}$. Coomassie-stained proteins (above) and zinc-induced fluorescence from bilins (bottom). Molecular weight markers in kDa (1), NPGP and PGP uninduced controls (2 & 10), NPGP, HPGP and PGP induced at 30 ^{o}C (3, 5 & 8), at 37 ^{o}C (4, 6 & 9) and a positive control Cph1Δ2 (7). The red arrow indicates the position of assembled Sb.PGP.

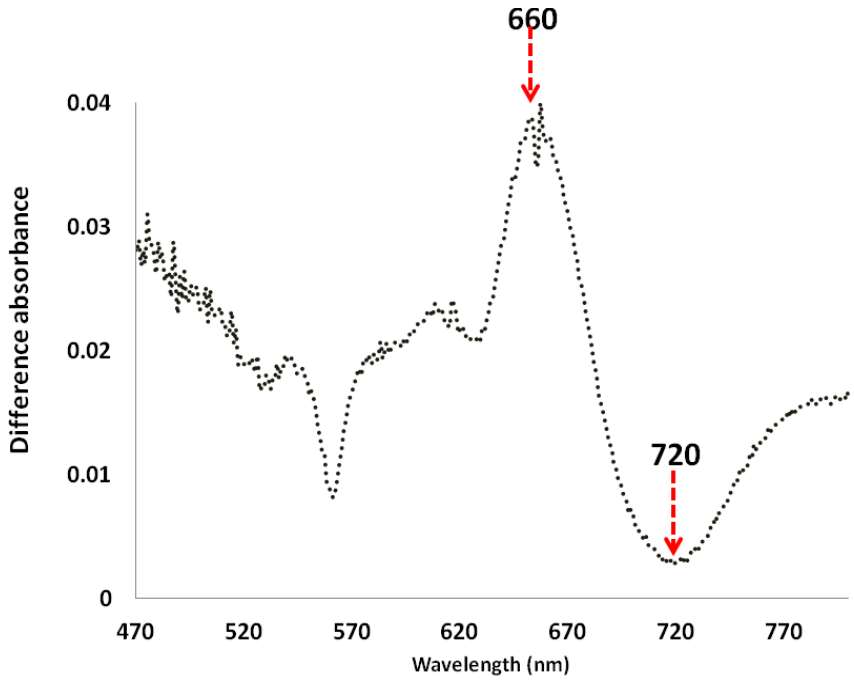


Figure 3.10. Difference spectrum of the soluble fraction obtained after *in vivo* assembly of apoPGP from pPROLar clone and PΦB from p183. Red arrows indicate the maximum and minimum of the difference spectrum.

3.1.1.7 Box-Behnken design for optimization

In vivo assembly was expected to help not only the proper folding of recombinant phytochrome but also quantitative optimization of the amount of phytochrome produced in different culture conditions from the absorbance difference spectra. Single factor optimization by maintaining all other factors at certain constant levels does not consider non-linear factor interaction [158]. The Box-Behnken independent-quadratic design for statistical optimization can be applied in situations in which the experimental variability is sufficiently small to obtain the necessary precision with only few replicates. Culture conditions used for Box-Behnken optimization here included temperature (18 ^{o}C, 37 ^{o}C), cell density (OD_{600} 0.5, 1) and arabinose concentration (0.2 %, 0.6 %). Experiments at the central point (28 ^{o}C, 0.75 & 0.4 %) were performed in triplicate to determine the standard deviation. 1 mM IPTG was used for all conditions. A total of 15 expression tests (Section 10.3) were carried out for pPROLar-*PGP* in 100 ml SB medium.

The amount of lysis buffer for normalization of total cell protein for each condition in Box-Behnken design was calculated by dividing the product of OD_{600} at harvest and culture volume by 20. OD_{600} at harvest and the volumes of lysis buffer are given in Section 10.3. The levels of recombinant holoPGP for each condition were quantitatively estimated from the difference spectra in Figure 3.11 and are presented in Table 5. The production of soluble holophytochrome was independently checked by SDS-PAGE analysis of small aliquots (10 µL) of total cell protein from each condition. The signal to noise ratio was very low for detection by Coomassie stain but comparison of the Coomassie stained bands clearly showed the total cell protein was well balanced across the Box-Behnken conditions. Moreover, holoPGP specific bands were revealed by zinc-induced fluorescence (Fig. 3.12 and 3.13). The optimal condition at BB4, which gave the highest amount of photoreversible Sb.PGP, was temperature at 28 ^{o}C, induction OD of 1 and 0.6 % arabinose concentration. The same conclusion was made from qualitative analysis of zinc-induced fluorescence after SDS-PAGE (Figure 3.12).

Table 5: Box-Behnken optimization of culture conditions for pPROLar-*PGP*

	Induction OD	Arabinose (%)	Temperature (°C)	Δ (ΔA) (mOD)
BB1	0.5	0.2	28	26.2
BB2	1	0.2	28	34.9
BB3	0.5	0.6	28	25.5
BB4	1	0.6	28	71.3
BB5	0.5	0.4	18	15.0
BB6	1	0.4	18	12.3
BB7	0.5	0.4	37	9.2
BB8	1	0.4	37	26.4
BB9	0.75	0.2	18	45.3
BB10	0.75	0.6	18	35.4
BB11	0.75	0.2	37	22.9
BB12	0.75	0.6	37	22.3
BB13	0.75	0.4	28	23.3
BB14	0.75	0.4	28	23.7
BB15	0.75	0.4	28	32.7

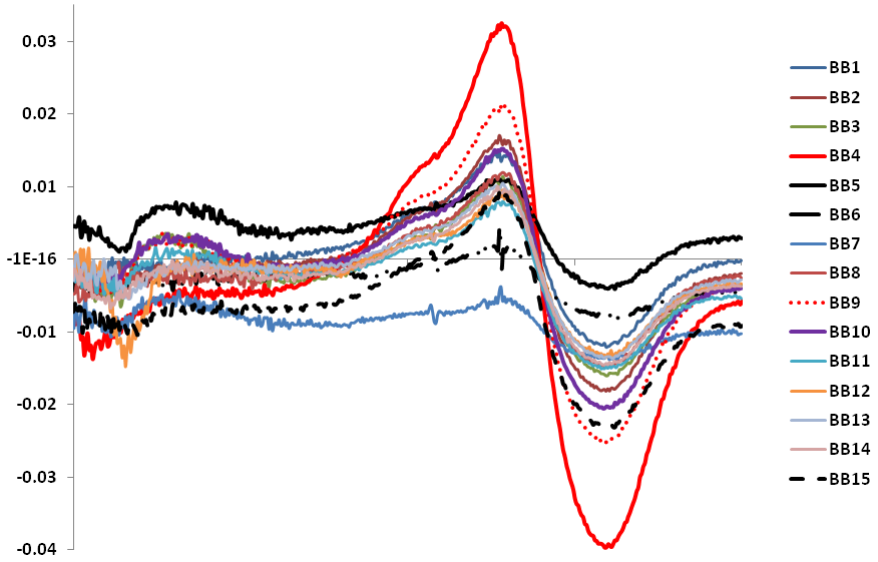


Figure 3.11. Difference spectra of the soluble fractions obtained after induction of PGP holoprotein from pPROLar-*PGP* in combination with p183 as a function of temperature, inducer concentration and induction OD for Box-Behnken optimization.

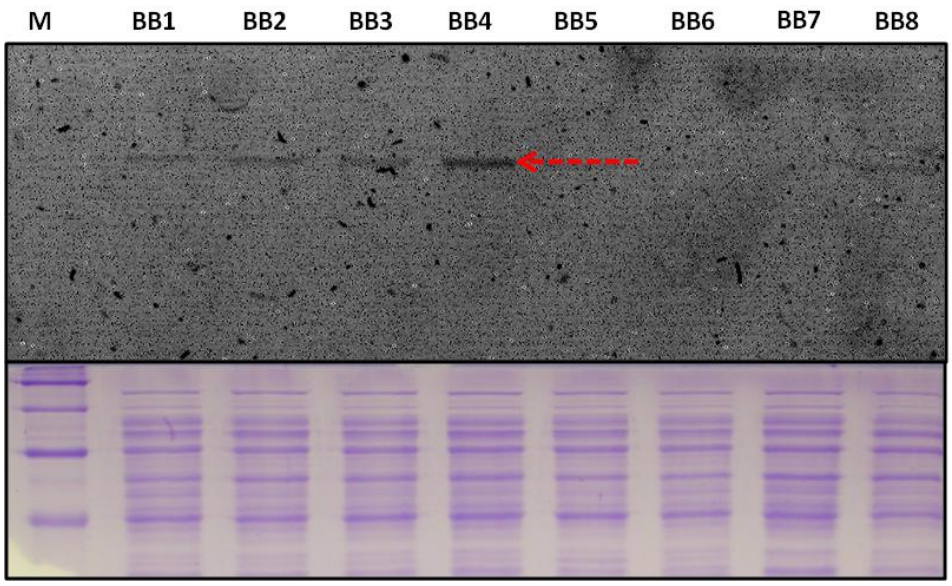


Figure 3.12. SDS-PAGE gel showing zinc-induced fluorescence from bilins (above) and Coomassie-stained proteins (below) of the soluble fractions obtained after induction of PGP holoprotein from pPROLar-*PGP* in combination with p183 as a function of temperature, inducer concentration and induction OD for Box-Behnken optimization. The red arrow indicates of assembled Sb.PGP for the optimal condition.

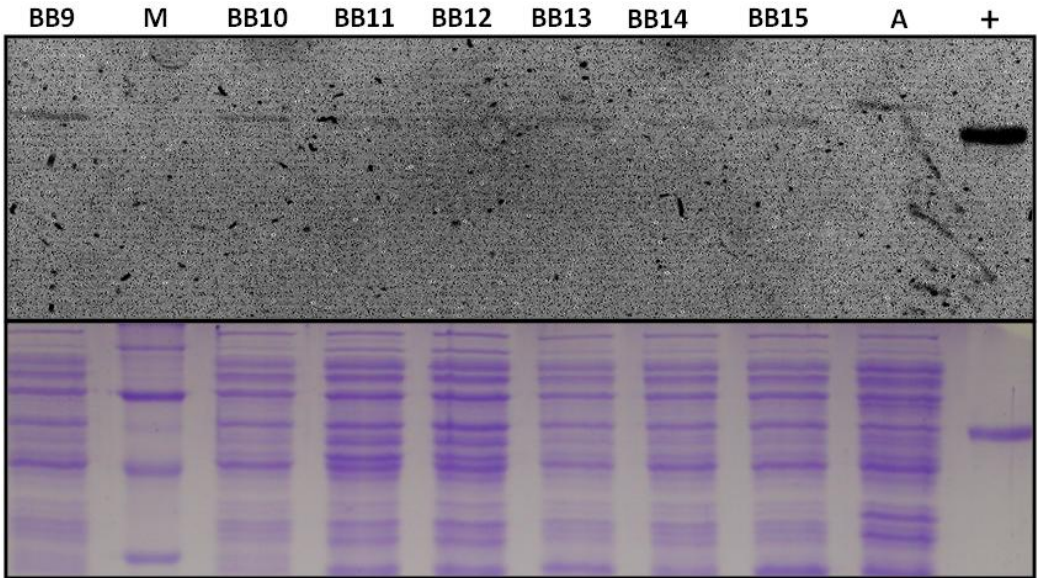


Figure 3.13. SDS-PAGE gel showing zinc-induced fluorescence from bilins (above) and Coomassie-stained proteins (below) of the soluble fraction obtained after induction of PGP holoprotein from pPROLar-*PGP* in combination with p183 as a function of temperature, inducer concentration and induction OD for Box-Behnken optimization. Molecular markers (M), NPGP from pCDFDuet (A) and CphΔ2 (+).

3.1.1.8 Cloning *Sb.PHYB* inserts into pCDFDuet

Despite extensive work to optimize NPGP and HPGP production from pPROLar, it proved difficult to establish a procedure giving consistently high yields. An alternative vector, pCDFDuet, was tested in the hope of improving the yield. All the three inserts were recloned with NcoI and HindIII into MCS1 of pCDFDuet. One of the resulting plasmids, pCDFDuet-*NPGP*, is illustrated in Figure 3.14.

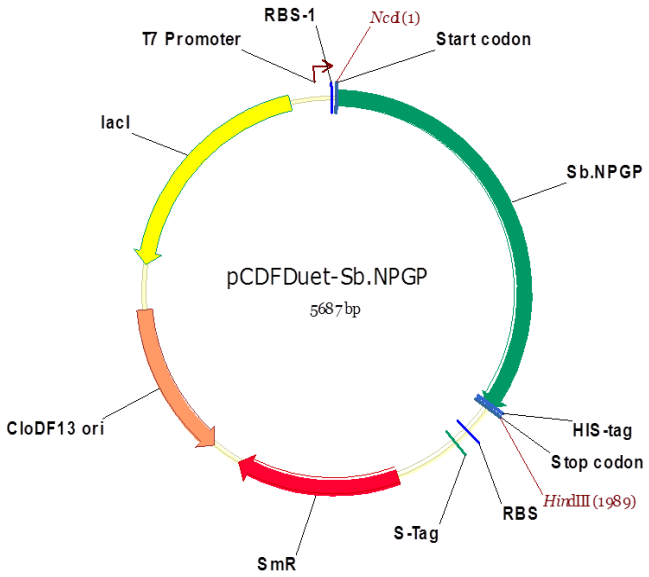


Figure 3.14. The pCDFDuet-*Sb.NPGP* plasmid. *Sb.NPGP* was inserted at NcoI and HindIII sites of MCS1 with primer-introduced C-terminal His_6-tag for purification.

pCDFDuet-*NPGP,* pCDFDuet-*HPGP* & pCDFDuet-*PGP* were cotransformed with p183 into heat-shock competent BL21(DE3) cells for *in vivo* assembly with PϕB. Expression and cell lysis were performed as described in the materials and methods section. OD_{600} at harvest for pCDFDuet-*NPGP* and pCDFDuet-*PGP* were determined to be 1.52 and 1.56 respectively. Induced cultures were lysed in 8 ml of TES-β. Holophytochrome production was assayed by measuring the difference absorption spectra following cell lysis and clarification. Both NPGP and PGP were produced as holoproteins in much higher concentration in pCDFDuet (Figure 3.15) than from pPROLar (Figure 3.10). The production level and solubility profiles of these proteins were further optimized. The production of HPGP was barely detectable from pCDFDuet-*HPGP* construct.

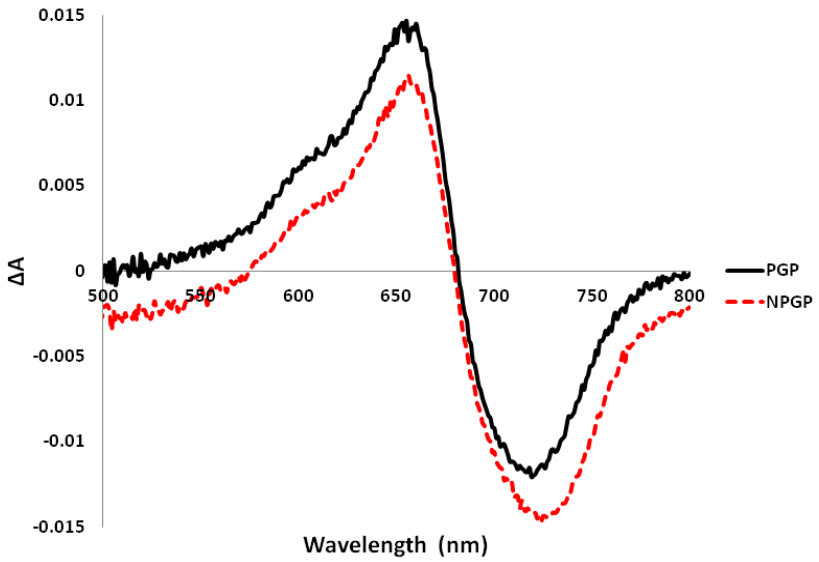


Figure 3.15. Difference spectra of the soluble fractions obtained after induction of PGP and NPGP holophytochromes from the first multi-cloning site (MCS1) of pCDFDuet in BL21(DE3) cells.

3.1.1.9 Production of NPGP-PϕB and PGP-PϕB from pCDFDuet

Here, optimized production protocols were developed for NPGP and PGP from pCDFDuet constructs in conjunction with p183 in *E. coli* grown in SB medium. A maximum yield (8 mg/L) of photoreversible phytochrome was obtained for the PGP fragment. The same procedure supported the production of NPGP at ca. 10-fold lower levels. *E. coli* cultures for recombinant phytochrome production were always cultivated in baffled-shake flasks. Production of recombinant phytochromes was reproducible in B 5042 (Heraeus) and Innova 42/42R (Eppendorf) shakers at 140 and 120 rpm, respectively. Other shakers tend to heat up during overnight incubation, making it difficult to control induction temperature. An important factor to achieve reproducible production of NPGP from pCDFDuet-*NPGP* construct was aeration (i.e. surface area /culture volume) in association with the geometry of culture flasks. Adequate aeration was achieved by using a maximum of 400 ml culture volume in broad-bottom and long-neck 2L baffled flasks. PGP production was stronger than that of the other constructs. The maximum PGP production was attained using SB medium and the following conditions: temperature of 28 ^{o}C, cell density at induction equivalent to OD_{600} 0.8, and inducer concentration of 1 mM IPTG. Optimal NPGP production was attained using SB medium supplemented with glucose and glycerol and the following optimized conditions: temperature of 24 ^{o}C, cell density at induction equivalent to OD_{600} 0.8, and inducer concentration of 1 mM IPTG.

3.1.2 Extraction and purification

An optimized purification protocol was developed using holophytochrome from the pPROLar clones. Cells were harvested, lysed and the soluble fraction was taken for further processing. A three-phase purification strategy was designed and optimized for recovery and purity. This strategy involved three consecutive steps: ammonium sulfate (AmS) precipitation, Ni-NTA affinity purification and size-exclusion chromatography. Ni-NTA affinity purification was optimized with respect to the nature of the buffer, the nature of additives and salt concentration. SEC was used as a polishing step to remove contaminants and establish monodispersion.

3.1.2.1 Cell lysis

The compatibility of different components of the lysis buffer with each other and with the subsequent purification step was carefully examined. TES-β, which has been routinely used in Hughes lab, was adopted as a lysis buffer because it has high buffering capacity and high ionic strength to maintain protein stability and to minimize interaction with column matrices, contains EDTA as metal chlelator thus inhibiting metal-dependent proteases and β-mercaptoethanol to maintain a reducing environment. The amount of extraction buffer was obtained by dividing the product of OD_{600} at harvest and total culture volume by 20. Frozen *E. coli* cells were thawed at 4 °C overnight and resuspended in cold lysis buffer (TES-β). The cells were cracked by using a 50 ml pre-cooled French pressure cell with two passages at the maximum force of 21,000 N (PP60KN, Watz Hydraulik). The extract was clarified by centrifugation at 4 °C and 15,000 g for 30 minutes (RC2-B, Sorvall).

3.1.2.2 AmS precipitation

Two AmS precipitation steps were performed. Firstly, saturated AmS buffer (3.3 M, 100 %) was added into the supernatant to 25 % saturation to precipitate several contaminant proteins. Following incubation on ice-water bath for 20 minutes, these were removed by centrifugation at 4 °C and 4,000 rpm for 20 min. (Biofuge, Heraeus). Secondly, phytochrome was precipitated by adding AmS to 60 % saturation, leaving contaminants in the supernatant. Following 30 minutes on ice-water bath, the sample was pelleted at 4 °C and 4,000 rpm for 30 min.

3.1.2.3 Ni-NTA affinity chromatography

All Sb.phyB derived constructs were cloned with primer-derived C-terminal oligohistidine tags to facilitate protein purification via Ni-NTA affinity chromatography. Test purification was conducted by re-suspending equal amounts of AmS-precipitated PGP samples in equal volumes of different binding buffers and loading 10 ml of the clarified supernatant onto appropriately-equilibrated 0.5 ml Ni-NTA columns. Buffers tested included 50 mM sodium phosphate pH 8.0, 50 mM Tris/HCl pH 7.5 and 50 mM sodium HEPES pH 7.8. As a starting point, 300 mM NaCl was included in all buffers. After the supernatant was loaded on the Ni-NTA column, it was successively purged with the binding buffer and washing buffer including an intermediate concentration of imidazole. The washing step was optimized with respect to the concentration of imidazole as well as volume. The bound phytochrome was finally eluted at high imidazole concentration. To monitor and compare phytochrome yield and purity following Ni-NTA affinity chromatography, recombinant Sb.phyB fractions were analyzed on SDS-PAGE followed by staining with hot Coomassie.

Tris buffer was selected as there was no significant difference among the tested buffers in terms of purity and yield of the purified phytochrome. The optimal concentration of imidazole in the binding buffer to reduce non-specific binding to the Ni-NTA column was experimentally determined to be 10 mM. The effect of ionic strength on purity and recovery of phytochrome was explored by comparing different concentrations of NaCl (0 and 500 mM) in 50 mM Tris pH 7.8 containing 1 mM IDA, 1 mM β-ME and 10 mM imidazole (see Figures 3.16 & 3.17). High salt concentration was clearly necessary for effective purification.

At first, Ni-NTA affinity-purification protocol was carried out by gravity flow using 50 mM Tris pH 7.8, 300 mM NaCl, 1 mM IDA, 1 mM β-ME and 10 mM imidazole as binding buffer. This protocol was named as the 'original' purification protocol to distinguish it from the HEPES based 'improved' purification protocol to be encountered later. After sample loading, proteins were washed with three column volumes of binding buffer and three column volumes of binding buffer supplemented with 40 mM imidazole. Finally, bound phytochromes were eluted with three column volumes of the binding buffer containing 250 mM imidazole. After Ni-NTA, the major protein fractions (Figure 3.18) were pooled, AmS-precipitated and resuspended in TES-β for size-exclusion chromatography.

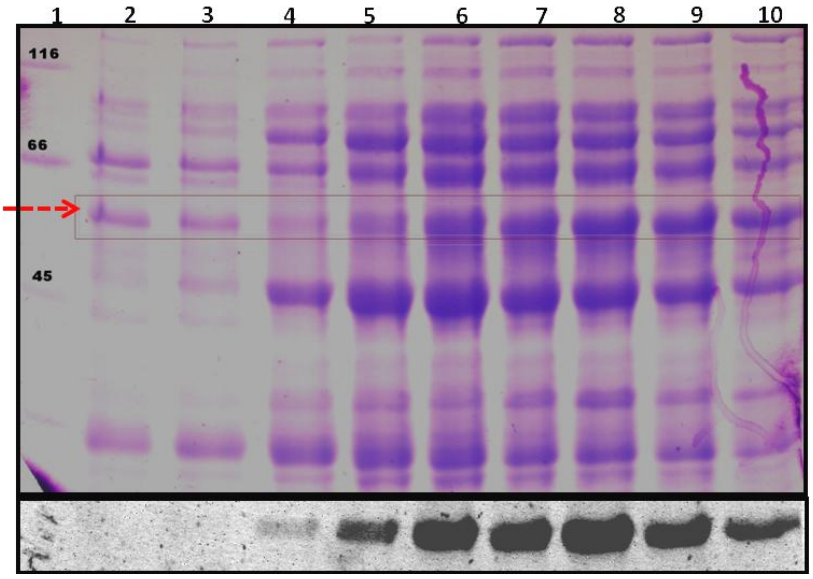


Figure 3.16. Optimization of purification. SDS-PAGE analysis of test purification for PGP produced as 58 kDa fusion protein detected by Coomassie stain (top) and zinc-induced fluorescence (bottom) in 50 mM Tris pH 7.8, 1 mM IDA and 1 mM β-ME and 10 mM imidazole without salt. Markers in kDa (1), flow-through (2), washing fractions (3-5) and elution fractions (6-10). The red arrow indicates bands for PGP.

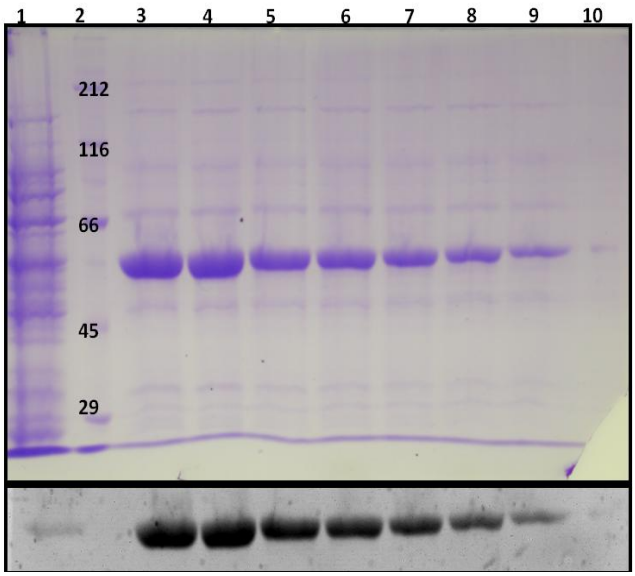


Figure 3.17. Optimization of purification. SDS-PAGE analysis of test purification of PGP as detected by Coomassie stain (top) and zinc-induced fluorescence (bottom) in 50 mM Tris pH 7.8, 500 mM NaCl, 1 mM IDA, 1 mM β-ME and 10 mM imidazole. Input (1), markers in kDa (2) and different elution fractions (3-10).

Figure 3.18. Blue-green protein solution from the original purification protocol. The blue-green coloration of *in vivo* assembled PGP in different fractions from the Ni-NTA column allows the elution process to be monitored by eye.

3.1.2.4 Size-exclusion chromatography (SEC)

Ni-NTA purified NPGP and PGP were polished by SEC in TES-β at room temperature using a Superdex 200 26/60 preparative grade column and 1 ml/min flow rate. SEC also allowed the degree of aggregation of the sample to be estimated. No phytochrome was eluted in the void volume, indicating negligible aggregation (Figures 3.19 and 3.21). The major elution peak for PGP was at 190.5 ± 3 ml obtained from multiple runs under identical conditions, equivalent to a globular protein with a molecular mass of 83 kDa (Fig. 3.19). The discrepancy in relation to its actual molecular mass of 60 kDa is perhaps due to its elongated shape. PGP purified by the original purification protocol showed a main peak accompanied by a small higher molecular weight shoulder at 183.6 ± 3 ml, perhaps a dimer.

Densitometeric estimation of Coomassie-stained bands on SDS-PAGE gel (described in Section 2.2.2.6) after SEC showed the final PGP preparation was about 85 % pure (Figure 3.20). Contaminating proteins were still evident after Ni-NTA and SEC. Analytical SEC (Figure 3.21) confirmed aggregation for purified PGP in TES-β was negligible. PGP was eluted as a 58 kDa major peak on TSK-SWxl Tosoh analytical column at flow rate of 0.3 ml/min. Besides, a small 99 kDa phytochrome peak was also observed. This peak might represent a dimer of the sensory module in solution. Purified NPGP and PGP samples were characterized by various spectroscopic techniques (Section 3.1.3) and concentrated to nearly 10 mg/ml and buffer exchanged for crystallization screening (Section 3.1.4). No crystal growth was observed from crystallization screening from this purification protocol, however (Fig. 3.41).

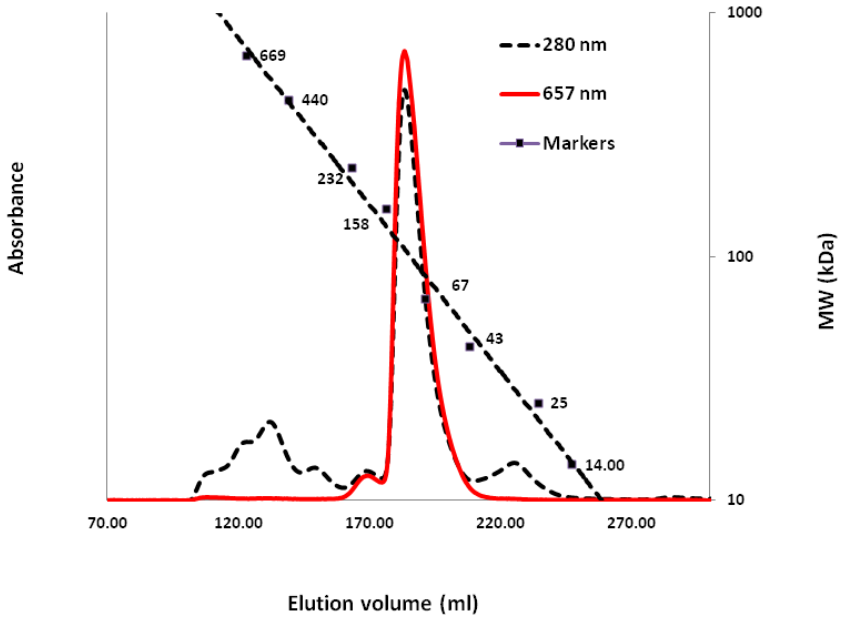


Figure 3.19. SEC elution profile of PGP purified by the original purification protocol in TES-β after saturating FR-irradiation in Superdex 200 26/60 preparative column at 1 ml/min flow rate monitored at 280 nm and 657 nm. The secondary axis represents the calibration curve with molecular weight standards: thyroglobulin (669), ferritin (440), catalase (232), aldolase (158), albumin (67), ovalbumin (43), chymotrypsinogen A (25) and RNAse A (14) in kDa.

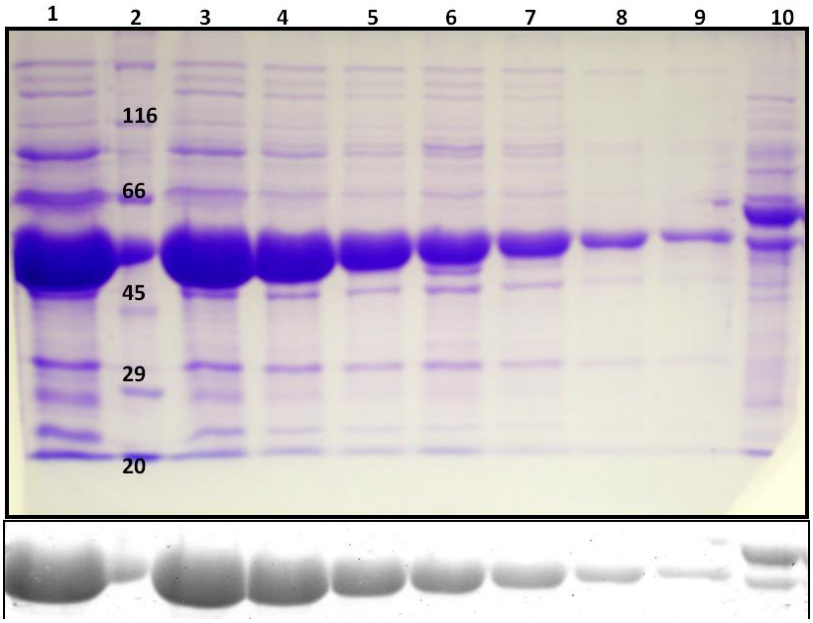


Figure 3.20. SDS-PAGE of PGP from the original Ni-NTA affinity purification protocol followed by preparative SEC as detected by Coomassie stain (top) and zinc-induced fluorescence (bottom). SEC fractions (1, 3-9), markers in kDa (2) and NPGP (10).

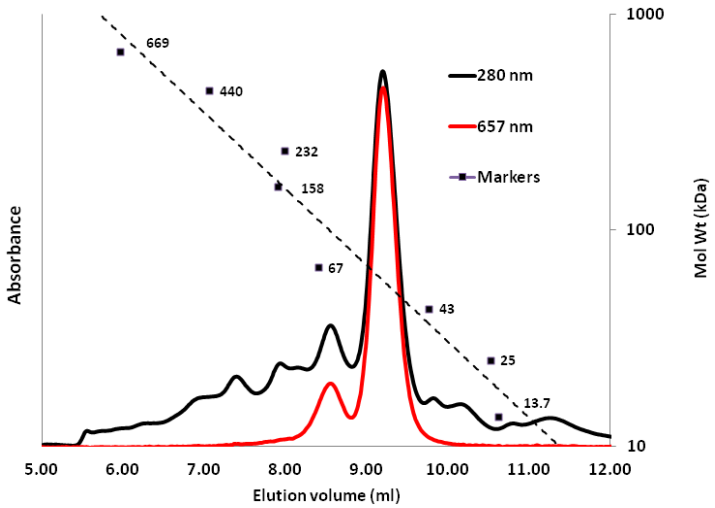


Figure 3.21. Analytical SEC elution profile of PGP purified the original purification protocol in TES-β after saturating FR-irradiation in TSK SWxl-Tosoh analytical column at 0.3 ml/min flow rate monitored at 280 nm and 657 nm. The secondary axis represents the calibration curve with molecular weight standards: thyroglobulin (669), ferritin (440), catalase (232), aldolase (158), albumin (67), ovalbumin (43), chymotrypsinogen A (25) and RNAse A (13.7) in kDa.

3.1.2.5 Storage and stability of purified Sb.phyB

The homogenous Sb.PGP preparation in Tris buffer (Figure 3.21) showed aggregation during storage at 4 °C (see Figure 3.22), an effect perhaps related to the poor results from crystallization screens using this material (Figure 3.41). Various parameters were investigated to improve the stability of the material. Indeed, the use of HEPES buffer largely eliminated the problem, thus the entire purification protocol was modified to HEPES-based buffer containing high salt.

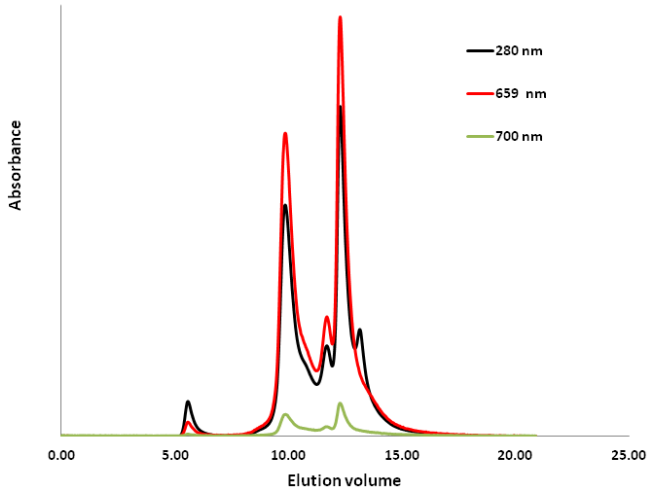


Figure 3.22. Analytical SEC elution profile of PGP from the original purification protocol stored for ca. 4 weeks in TES-β after FR-irradiation (TSK SWxl-Tosoh column at 0.3 ml/min flow rate monitored at 280 nm, 659 nm and 700 nm).

3.1.2.6 Improved purification protocol

An improved HEPES-based purification protocol was developed leading to recovery of ca. 8 mg of purified PGP per litre of culture volume with an SAR of 1.2 – 1.4 (> 95 % purity). For the improved purification protocol, about 10 ltr of *E. coli* culture was harvested. The yield and purity of purification fractions was monitored by measuring SAR and visualizing Coomassie-stained bands after SDS-PAGE. Following extraction and purification, the soluble fraction yielded around 0.1 g of purified photoreversible PGP (Figure 3.23), better than 90 % purity after Ni-NTA (Figure 3.24), and near perfect monodispersion (Figure 3.25) and better than 95 % purity after SEC (Figure 3.26). The poorly-resolved shoulder observed during preparative SEC using the original procedure was significantly reduced in the improved purification protocol (Figure 3.25). No protein degradation or aggregation was observed after storing the purified holophytochrome at 4 °C for ca. 4 weeks. On analytical SEC (Figure 3.27), PGP purified by the improved purification protocol was monodispersed but eluted aberrantly, perhaps due to interaction with the matrix despite the high salt concentration (Figure 3.27).

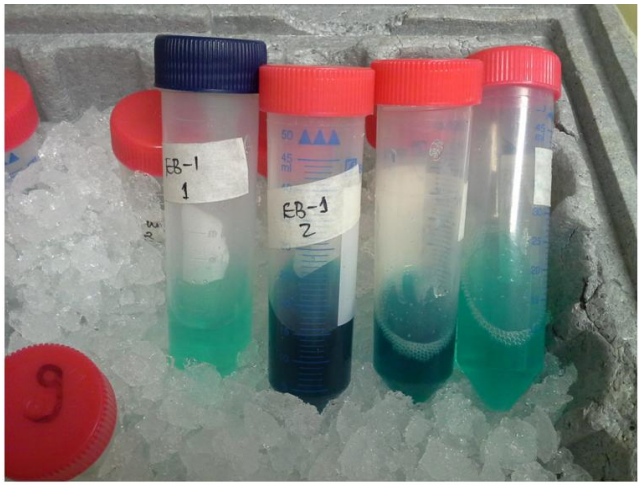


Figure 3.23. Sb.phyB preparations by the improved purification protocol. Nearly 0.1 gram of *In vivo* assembled PGP was purified by the improved protocol.

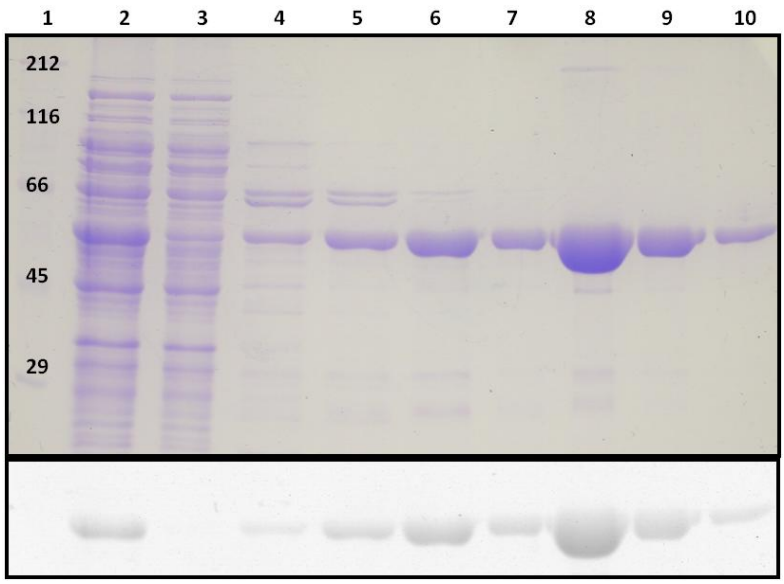


Figure 3.24. Improved PGP purification protocol. SDS-PAGE analysis of large-scale Ni-NTA purification of PGP detected by Coomassie stain (top) and zinc-induced fluorescence (bottom) in 50 mM HEPES pH 7.8, 500 mM NaCl, 1 mM IDA, 1 mM β-ME and 10 mM imidazole. Markers in kDa (1), input (2), flow-through (3) washing fractions (4-5) and elution fractions (6-10).

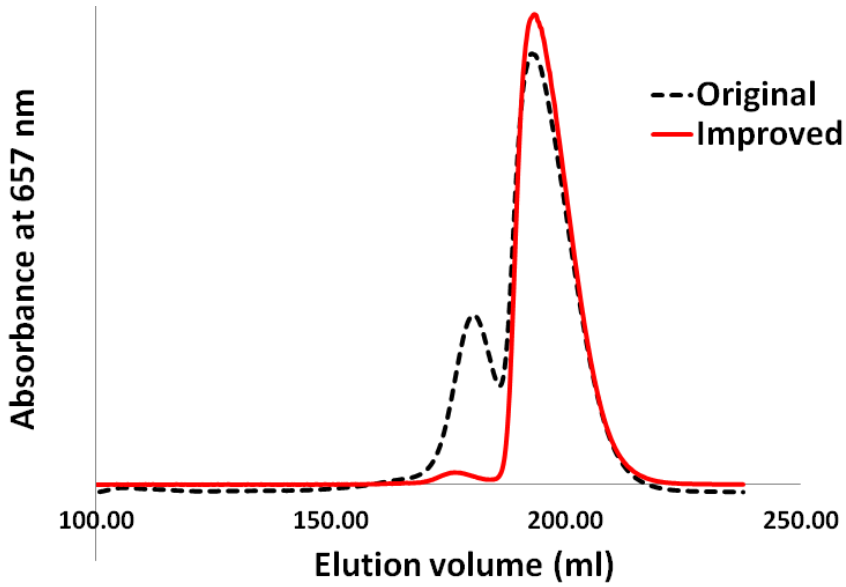


Figure 3.25. Comparison of SEC elution profiles of PGP purified by the original and the improved purification protocols using Superdex 200 26/60 preparative column at 1 ml/min flow rate monitored at 657 nm. The elution volumes were reproducible to 0.5 ml allowing the raw absorption data to be plotted in the same horizontal axis.

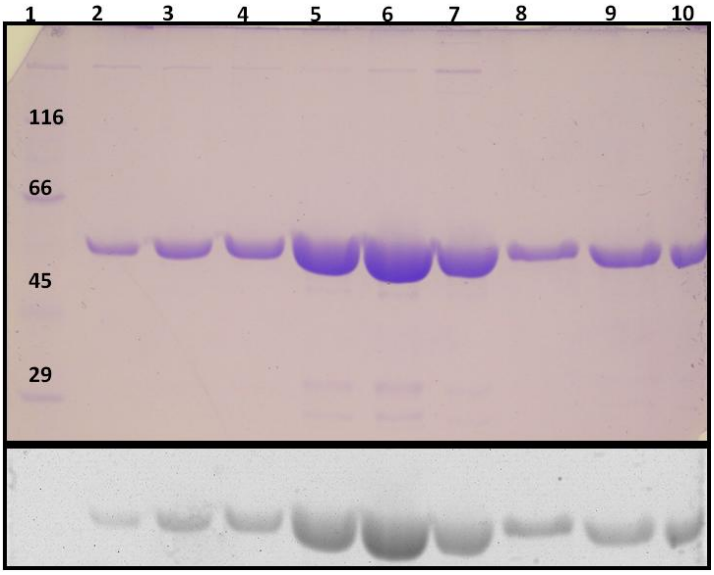


Figure 3.26. SDS-PAGE analysis of PGP purified by the improved Ni-NTA purification protocol after SEC detected by Coomassie stain (top) and zinc-induced fluorescence (bottom). Markers in kDa (1) and SEC fractions C1, C7, C11, C14, D12, D10, E9, E1 and D1 (2-10).

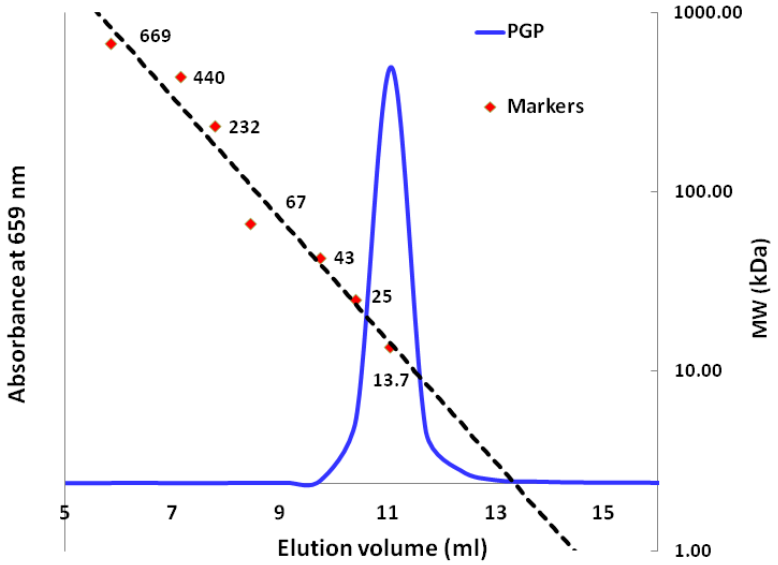


Figure 3.27. SEC elution profile of PGP after FR-irradiation purified by the improved purification protocol in HES-β using TSK Tosoh analytical column at 0.3 ml/min flow rate monitored at 659 nm. The secondary axis represents the calibration curve with molecular weight standards: thyroglobulin (669), ferritin (440), catalase (232), aldolase (158), albumin (67), ovalbumin (43), chymotrypsinogen A (25) and RNAse A (13.7) in kDa.

3.1.3 Characterization

In vivo assembled Sb.NPGP and Sb.PGP overproduced in *E. coli* were purified to homogeneity by a combination of Ni-NTA and SEC (Section 3.1.2). These purified holophytochromes were characterized by using the following spectroscopic and biophysical techniques.

3.1.3.1 UV-Vis absorption spectroscopy

For interpretation of spectroscopic data from Sb.phyB derived preparations, limitation originating from accurate determination of protein concentration and establishing protein purity were taken into consideration. All wild-type phytochromes preparations including NPGP, PGP (Section 3.1), PG (Section 3.2) and their Y291H mutants (Section 3.4) were assumed to be properly folded and 95 % homogeneous (Figure 3.52) as estimated by densitometry. Spectroscopic data obtained from purified phytochrome preparations were normalized to molar concentrations using calculated molar extinction coefficients (Table 4). Protein concentrations determined from the difference spectra, $A_{280\ nm}$ and Bradford assay were quite reproducible (~5 % error). There was no evidence of contaminating peak during additional purification of SEC purified proteins by IEC.

To gain insight into the structural and functional effects of the NTE in Sb.phyB, the spectroscopic and photochemical properties NPGP were compared with that of PGP. Removal of the NTE caused 6 nm and 10 nm hypsochromic shift from the Pr and Pfr absorption bands respectively (Figure 3.28). The Pfr peak intensity was notably lower in PGP relative to that of NPGP (Figure 3.29). Moreover, PGP showed a nearly 10-fold increase in the rate of dark reversion of Pfr to Pr compared to NPGP, with $t_{1/2}$ of 3.5 and 35 minutes respectively (Figure 3.30). Two-term exponential model was used to fit the experimental dark-reversion data suggesting different sub-states with different rates of dark reversion occurring simultaneously. The total half-life was determined from the sum averaged product of partial half-lives. The observed increased accessibility of the surface exposed cysteines to Ellman's reagent upon loss of the NTE (Figure 3.46) suggested a molecular shielding function for the NTE.

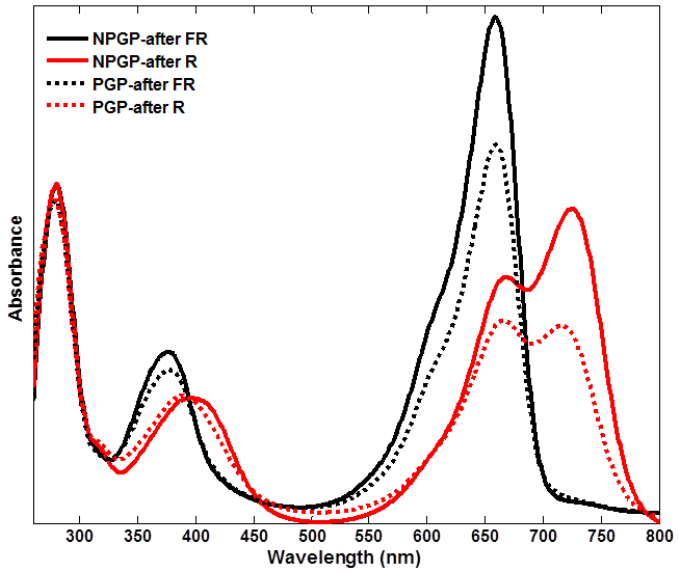


Figure 3.28. Absorbance spectra of NPGP and PGP after saturating R/FR irradiation normalized for molar protein concentration. Calculated ε at 280 nm used for normalization are given in Table 4. Constructs were ca. 95 % pure as judged by SDS-PAGE shown in Figure 3.52.

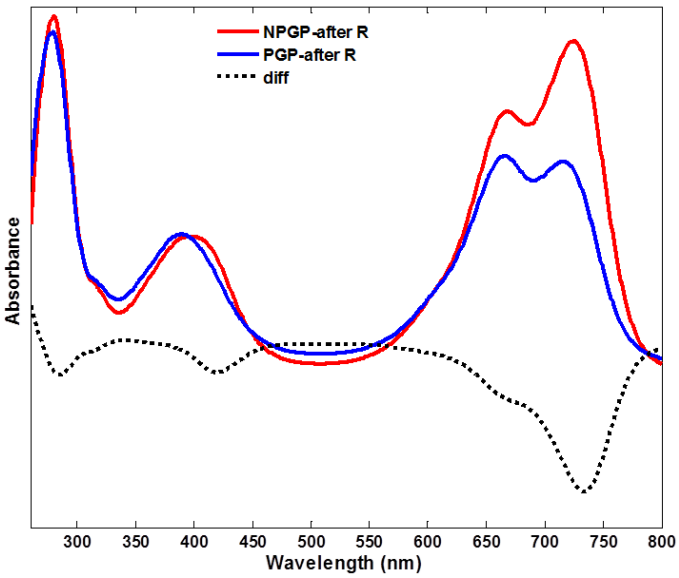


Figure 3.29. Absorbance spectra of NPGP and PGP normalized for molar protein concentration after saturating red light irradiation. The difference spectrum (diff) was obtained by subtracting the spectrum of NPGP from that of PGP.

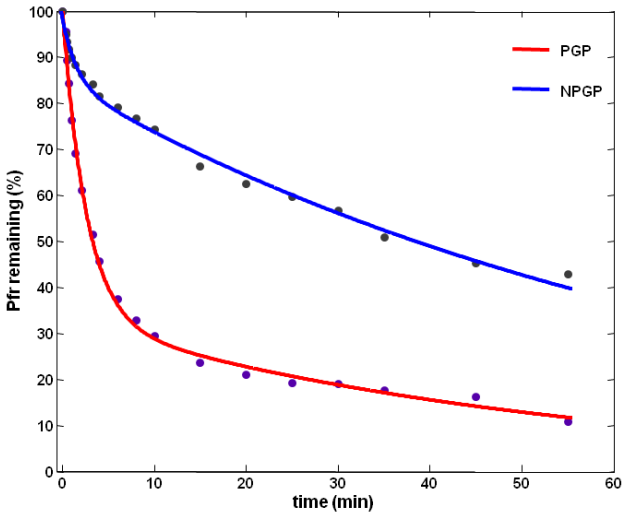


Figure 3.30. The Pfr to Pr dark reversion kinetics measured for PGP and NPGP in darkness at 25 °C in HES-β buffer.

To investigate the effect of chromophore replacement on the photochemical properties of Sb.phyB, NPGP was assembled *in vivo* with the native chromophore PϕB and PCB. As expected, NPGP red band exhibited 13 nm hypsochromic-shift when assembled with PCB compared to with PϕB (Figure 3.31). The R/FR difference spectrum for PCB-conjugated NPGP indicated maxima at 655 and 708 nm and an isosbestic point at 677 nm (Figure 3.32). The most remarkable difference due to chromophore replacement was observed during dark reversion studies (Figure 3.33) in which the half life for Pfr of NPGP with PCB was much longer than for PϕB (75 vs. 15 min.).

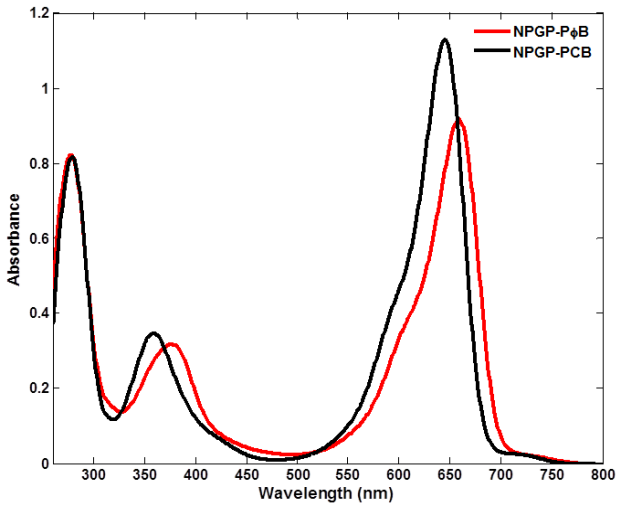


Figure 3.31. Absorbance spectra of NPGP assembled with PCB and PϕB normalized for molar concentration after saturating FR light irradiation in HES-β.

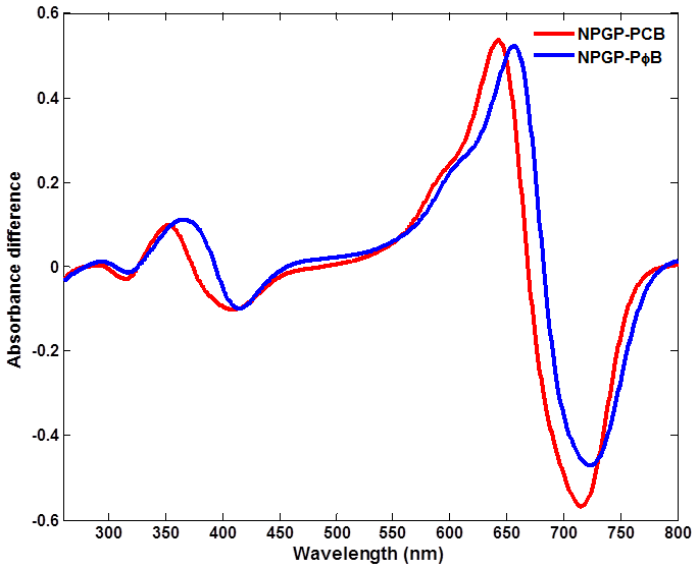


Figure 3.32. Phytochrome difference absorption spectra for NPGP assembled with PCB and PϕB normalized for molar concentration in HES-β.

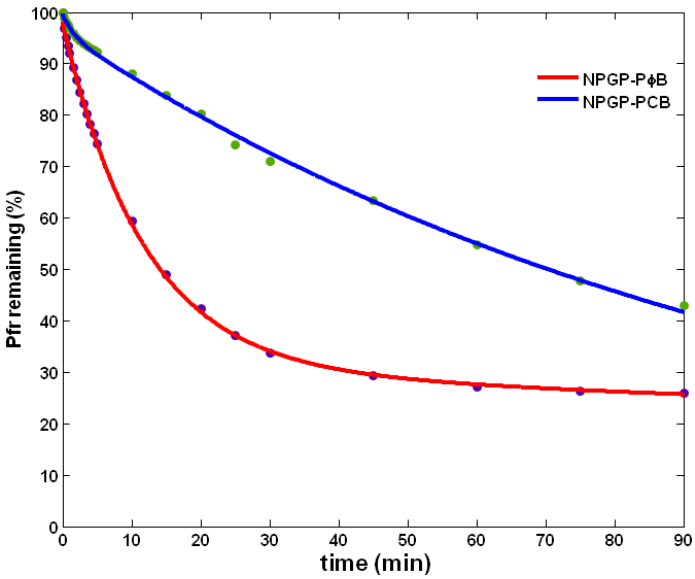


Figure 3.33. The Pfr to Pr dark reversion kinetics measured for NPGP assembled with PCB and PϕB at 30 °C in HES-β.

3.1.3.2 CD spectroscopy

The changes in secondary structures of NPGP and PGP after saturating R/FR irradiation were also characterized by far-UV circular dichroism (CD) spectroscopy. Marginally broad shifts in the negative CD bands near 222 nm and 208 nm that narrow down in either direction along the

wavelength-axis were observed upon photoconversion of both NPGP and PGP, the observed changes being more pronounced for NPGP (Fig. 3.34 A) than PGP (Fig. 3.34 B). For NPGP upon Pr to Pfr photoconversion, Pr showed an ~1 % larger ellipticity at 210 nm than Pfr. The ellipticity of Pr was lower than that of Pfr between 215 and 235 nm, at 220 nm by ~2 % (Figure 3.34 A). These spectral changes reflected an increase of α-helical structures during the Pr to Pfr conversion. Because the CD spectrum of the Pfr state was recorded with a sample containing the mixture of Pr and Pfr, the increase in the α-helical content of the Pfr state is reduced by at least 10 %.

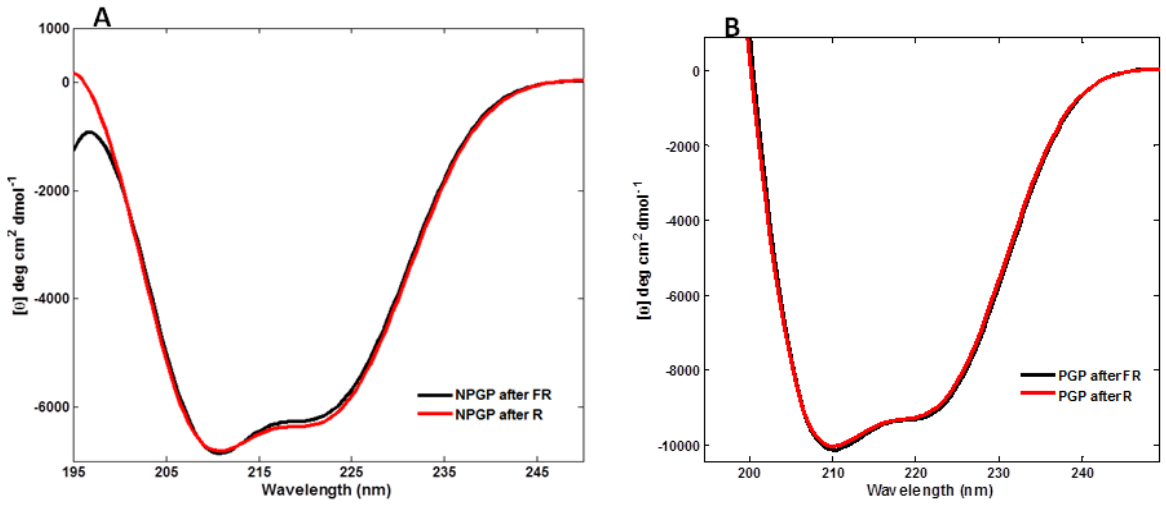


Figure 3.34. The CD spectra of NPGP (A) and PGP (B) after R/FR irradiation in 10 mM K-phosphate, pH 7.8, containing 50 mM ammonium sulfate. Spectra were recorded in 0.5 mm pathlength cell at 25 ^{o}C.

The CD spectrum of NPGP was also directly compared with that of PGP after far-red light treatment. The mean residue ellipticity [θ] was used to normalize for protein concentration ($A_{280\ nm}$) and to reflect the fact that the peptide bond is the active species. The CD spectra in Fig. 3.35 revealed unexpectedly large ellipticity difference between PGP and NPGP in the Pr state. At 210 nm, PGP showed ~45 % larger negative ellipticity than NPGP. The observed drop in CD signal for NPGP (45 %) was so large that it might not be merely due to impurities and misestimating protein concentrations. The CD spectrum for NPGP assuming a completely disordered NTE was calculated by using the ratio of the polypeptide length of PGP (550) to NPGP (661). The calculated CD spectrum suggested that a completely unstructured 111 residues in the NTE of NPGP would produce ~20 % smaller ellipticity than PGP. Taken together the calculated and observed CD spectra implied small structural changes, presumably unfolding, outside the NTE. This was independently tested by comparing the temperature-induced unfolding of NPGP and PGP (Figs.

3.36 A & B) using CD data collected as a function of temperature. Contrary to intuition, removal of the NTE caused ~5 ^{o}C rise in melting temperature (T_m) (Fig. 3.37). Increased thermal stability of PGP compared to NPGP suggested the NTE might destabilize the rest of the molecule.

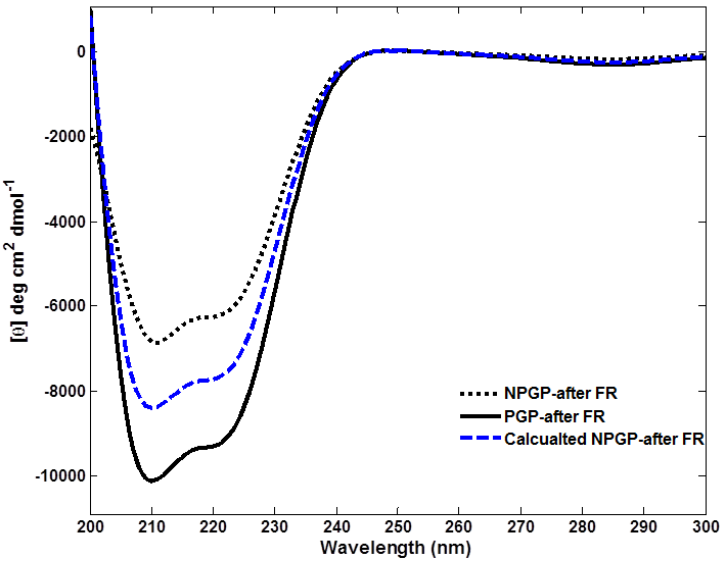


Figure 3.35. The near and far-UV CD spectra of NPGP and PGP after saturating FR-light irradiation in 10 mM K-phosphate, pH 7.8, containing 50 mM ammonium sulfate. Spectra were recorded in 0.5 mm pathlength cell at 25 ^{o}C. Constructs were ca. 95 % pure as judged by SDS-PAGE shown in Figure 3.52.

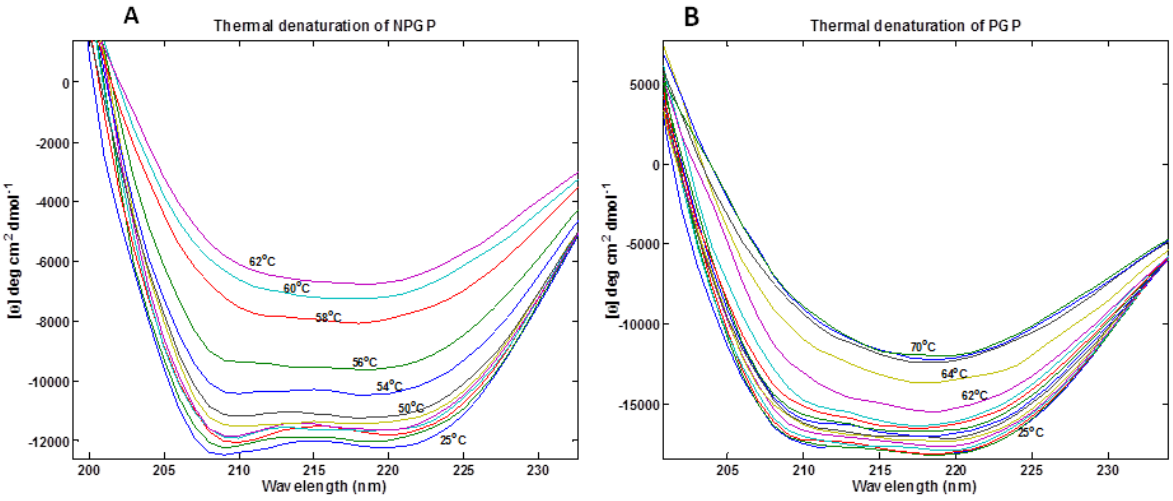


Figure 3.36. Changes in CD spectra of NPGP (A) and PGP (B) after FR-irradiation collected as a function of temperature, ramp rate of 0.25 ^{o}C/minute and 5 minutes of equilibration time at the measurement temperature.

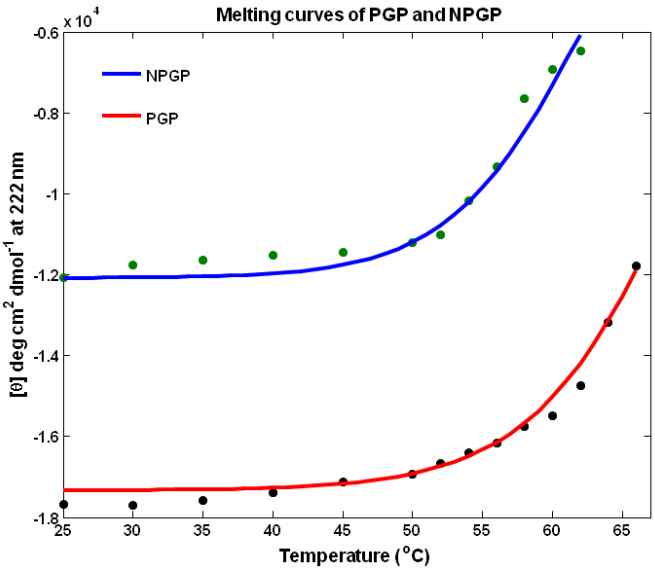


Figure 3.37. Changes in ellipticity of NPGP and PGP at 222 nm plotted as a function of temperature (^{o}C). The samples were held for 5 minutes at the measuring temperature and ramp rate was 0.25 ^{o}C/minute.

The replacement of PϕB with PCB affected the absorbance properties of NPGP (Fig. 3.31). To characterize secondary structural changes associated with chromophore replacement, CD spectra of NPGP assembled with PCB and PϕB were measured after far-red irradiation. NPGP-PCB and NPGP-PϕB showed very small differences in ellipticity between 210 nm and 220 nm. Big effects were observed below 200 nm, but absorption by buffer components and oxygen compromise interpretation of CD spectra in this region. The spectra were essentially identical between 225 nm and 250 nm suggesting minimal conformational differences (Fig. 3.38). On the other hand, PCB replacement in PGP was accompanied by significantly larger loss of negative ellipticity (~45 % at 210 nm) than PCB replacement in NPGP (Fig. 3.39). A major limitation of CD experiments was accurate determination of protein concentration and establishing protein purity. These conclusions should be independently verified by other robust methods.

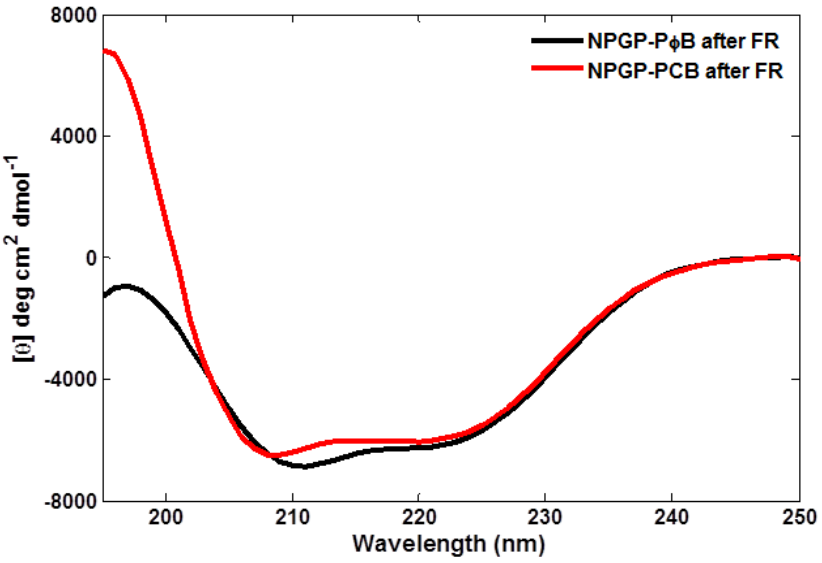


Figure 3.38. The CD spectra for NPGP assembled with PCB and PϕB after saturating FR-light irradiation in 10 mM K-phosphate, pH 7.8 containing 50 mM ammonium sulfate. Spectra were recorded in 0.5 mm pathlength cell at 25 ^{o}C.

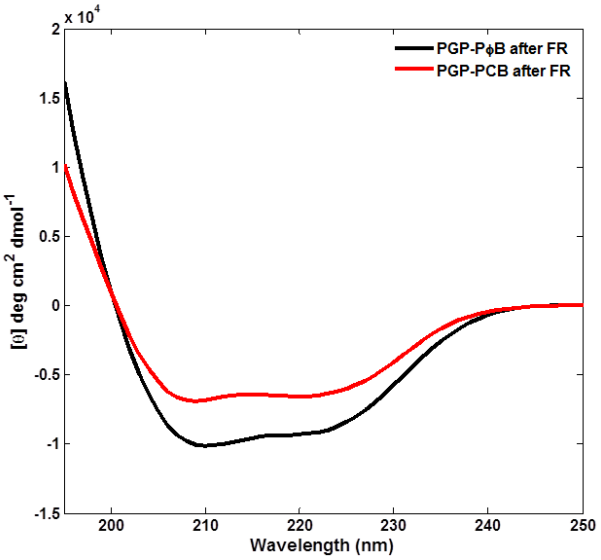


Figure 3.39. The CD spectra for PGP assembled with PCB and PϕB after saturating FR-light irradiation in 10 mM K-phosphate, pH 7.8 containing 50 mM ammonium sulfate. Spectra were recorded in 0.5 mm pathlength cell at 25 ^{o}C

3.1.3.3 Resonance Raman spectroscopy

The structure of the bilin chromophore and its interactions with amino acid residues in the binding pocket can be characterized by Resonance Raman spectroscopy using three phytochrome

specific spectral regions. These regions first described for oat phyA include: region-I (750 - 850 cm^{-1}), region-II (1520 - 1680 cm^{-1}) and region-III (1712 - 1729 cm^{-1}). Region-I is characterized by less intense bands than the other regions originating from hydrogen-out-of-plane (HOOP) modes. The most prominent band in this region is near 804 cm^{-1}, a HOOP mode of CD methine bridge. Isotopic labeling and temperature dependent studies of phytochromes in region-I identified two bands (804 and 820 cm^{-1}) separated by 16 cm^{-1} implying conformational heterogeneity in the Pfr state. The low frequency HOOP mode (804 cm^{-1}) corresponds to Pfr-I and the high frequency HOOP mode (820 cm^{-1}) corresponds to Pfr-II. The ratio of the intensities of these bands can be taken as a measure of relative populations of Pfr sub-states. Region-II is dominated by three peaks in the Pr state originating from N-H in plane (N-H in) bending of rings B and C (ca. 1575 cm^{-1}), the C=C stretching modes of the AB methine bridge (1646 - 1650 cm^{-1}) and the C=C stretching modes of the CD methine bridge (1626 - 1637 cm^{-1}). A region between 1600 and 1620 cm^{-1} due to RR-active C=C stretching modes of the CD methine bridge at lower frequency can also be used to study conformational heterogeneity in the Pfr state. Region-III contains bands originating from carbonyl stretching modes of ring D (ca. 1712 cm^{-1}) and ring A (ca. 1729 cm^{-1}).

Resonance Raman spectroscopy was used to characterize chromophore structures in NPGP and PGP. Spectra were collected by Dr. Francisco Escobar and analyzed by Professor Peter Hildebrandt (TU Berlin). No evidence was obtained for Pr heterogeneity in NPGP and PGP. On the other hand, removal of the NTE after R-light irradiation showed down shift of high frequency HOOP component in region-I accompanied by reversal of intensity ratios (Section 10.5). The torsion of the C-D ring methine bridge and the surrounding hydrogen bonding network in Pfr state was also altered upon removal of the NTE. These changes were accompanied by a shift of the conformational equilibria between Pfr-I and Pfr-II. The result was consistent with increasing Pfr-I to Pfr-II ratio. A detailed comparative description of these and other RR effects in Sb.phyB assembled with either PCB or PϕB has been published (Escobar et al., 2017; see Section 10.5).

3.1.3.4 Hydrogen/Deuterium exchange

Differential hydrogen/deuterium exchange (HDX) coupled with mass spectrometry was used to characterize conformational change of NPGP following R/FR irradiation. Fresh NPGP sample was purified by Ni-NTA and SEC. SEC purified sample was concentrated to 25 mg/ml by pressure cell and Vivaspin (Millipore). The concentrated NPGP sample was dialyzed overnight in HDX buffer. 20

ml HDX buffers (5 mM HEPES pH 7.8, 50 mM NaCl and 0.3 mM TCEP) were prepared by dilution of stock solutions in H_2O and D_2O. The final pH was adjusted using either HCl or DCl. A working (66 µM) NPGP solution was prepared by diluting the stock in HDX buffer. Two aliquots of 800 µL phytochrome working solutions were prepared to compare Pr and Pfr. State 1 (equivalent to Pr) was irradiated by far-red light in 250 µL HPLC syringe for one minute and thereafter kept in darkness whereas State 2 (Pr + Pfr mixture) was irradiated by red light. State 2 was periodically refreshed by 660 nm laser light. The five time points for exchange reactions were 0 s, 15 s, 60 s, 120 s and 300 s. A heat map was generated to represent differential relative uptake between State 1 and State 2 recorded as a function of time (Figure 3.40). (This data set was collected and analyzed by Dr. Andreas Mielcarek, AG Essen).

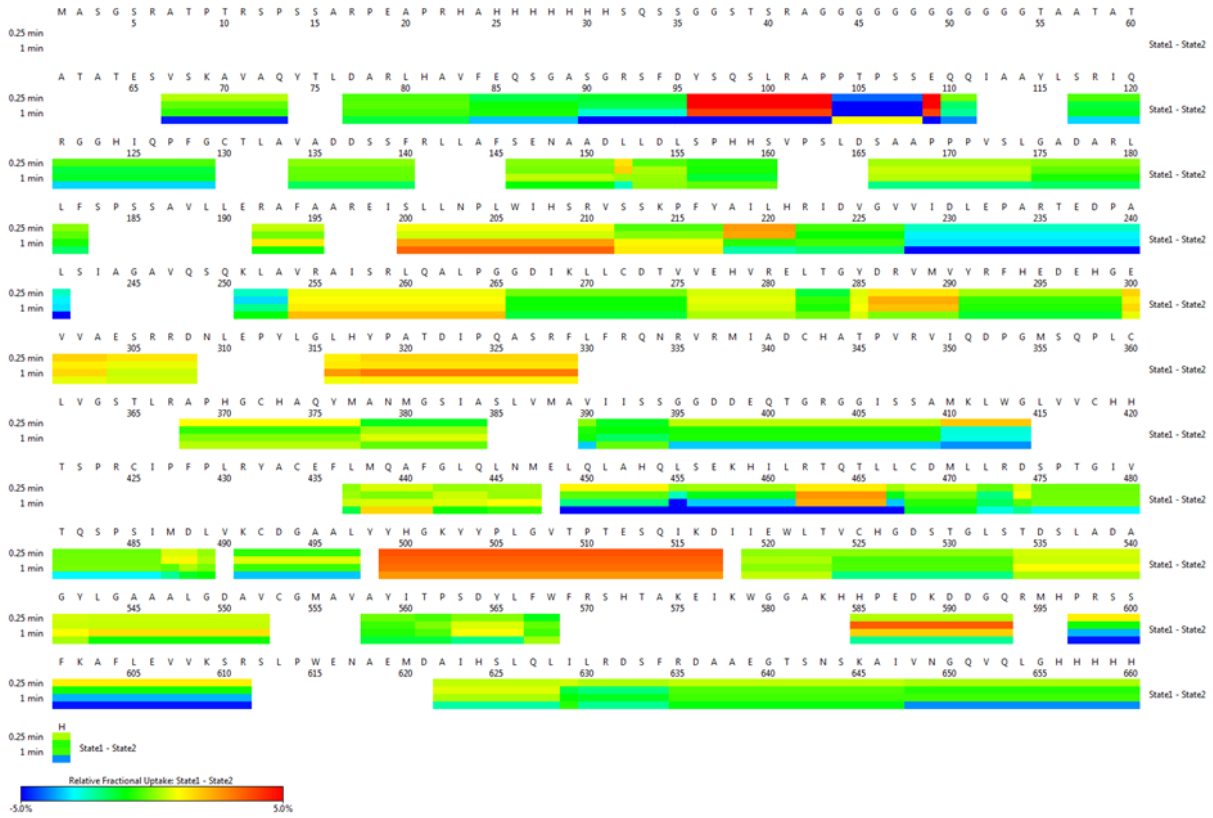


Figure 3.40. Relative fractional deuterium uptake for Sb.NPGP after saturating R/FR irradiation. The horizontal axes of the colored-bars denote residue numbers, colors represent differential relative uptake and the vertical axes represent different time points. The map was generated by the difference exchange rate (State1-State2) as a function of time for selected HDX detected peptides by Dr. Mielcarek.

Peptide identification was incomplete. Peptides that were not identified include the first 66 residues comprising the LCR region of NTE, the lasso in the GAF domain (F331-L366), the first 8 residues of the connecting helix (F428-F436) and ~20 residues containing the WGG motif in the tongue of the PHY domain. On the other hand, high differential relative uptakes for the two states were observed in the NTE (Y96-Q110), part of the helical spine (L455-L467) and part of the tongue of the PHY domain (H596-R611). Some regions of the NTE (L76-D95), the PAS knot (L116-G129) and the bilin-binding pocket (A368-A409) did not show substantial difference in relative uptake between the two states, suggesting minimal conformational change (Figure 3.40).

3.1.4 Crystallization and diffraction experiments

The purified and characterized NPGP and PGP preparations were subjected to crystallization screening to obtain diffraction-quality crystals for structural studies. Phytochromes undergo conformational changes upon exposure to different wavelengths of light. In order to maximize the relative probability of one conformational state over another, phytochrome crystallization procedures should ideally be performed in darkness or very dim safelight (~ 510 nm) [46, 159]. To this end, the crystallization robot in MarXtal crystallization facility at Philipps University of Marburg (AG Essen) was housed in a room that can be made completely dark with movable black shades. The shades were organized and installed by Ralf Poeschke.

3.1.4.1 Initial crystallization screening

PGP purified by the original purification protocol was concentrated to 10 mg/ml in 30 KDa MWCO centrifugal membrane-filters (Millipore) at 10,000 rpm and 4 ^{o}C for 10 minutes (5415 R, Eppendorf) and exchanged into crystallization buffer (2.5 mM Tris pH 7.8 and 15 mM NaCl) by dialysis tubing (MWCO 30 KDa). The dialyzed sample was clarified by centrifugation at 23,300 rpm for 10 minutes (Biofuge, Heraeus) and submitted to MarXtal crystallization facility at Philipps University of Marburg (Ralf Poeschke, AG Essen). Crystallization trials were performed in 96-well sitting-drop vapor-diffusion plates using a Honeybee 963 robot (Genomic solutions). Each trial contained 200 nl of PGP solution mixed with 200 nl of crystallization reagents equilibrating against 80 µl of reservoir. Eight crystallization kits from Qiagen were used. Crystallization was setup in green safelight and plates were incubated at 20 °C in darkness. Experimental drops were periodically examined but no lead condition was identified from this purification batch (see Figure 3.41).

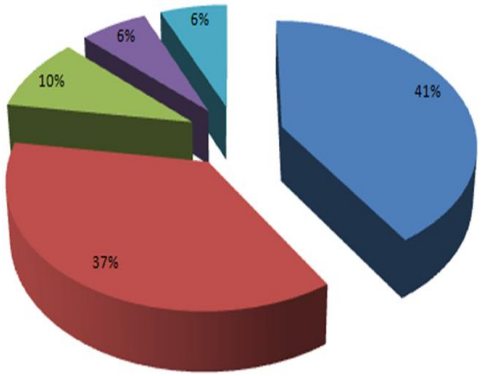


Figure 3.41. Outcomes of initial crystallization screening of PGP prepared by the original purification protocol at 10 mg/ml. Experimental drops were qualitatively examined and scored after 7 days in darkness in terms of heavy precipitate (41 %), precipitate (37 %), light precipitate (10 %), phase separation (6 %) and clear drops (6 %).

As no crystal was grown from the original purification protocol, various improvements were introduced (see section 3.1.2.6). The improved purification protocol resulted in enhanced purity and homogeneity as measured by analytical SEC (see Fig. 3.27). PGP prepared accordingly was concentrated to 20 mg/ml and again subjected to crystallization screening with the eight Qiagen kits (Ralf Poeschke, AG Essen). After equilibration for 7 days at 20 ^{o}C in darkness, several conditions buffered at high pH showed promising hits. The identified lead conditions were 0.1 M CAPSO pH 9.5, 0.2 M NaCl and 12 % PEG 4,000; 0.1 M Tris-HCl pH 8.5, 0.2 M $MgCl_2$ and 30 % PEG 4000; 0.1 M HEPES pH 7.5, 0.2 M $MgCl_2$, and 30 % PEG 4000; and 0.1 M HEPES pH 7.5 and 1.26 M AmS. Some of these crystals were observed under the microscope in white light (B061, Olympus) revealing that dark-green, rod-shaped phytochrome crystals had formed. NPGP purified by the improved purification protocol was also subjected to crystallization screening at 10 mg/ml with the Qiagen kits (Ralf Poeschke, AG Essen). This trial produced rod-shaped blue-green NPGP crystals in 0.1 M CAPSO pH 9.5, 0.1 M $LiSO_4$, 0.1 M NaCl and 12 % PEG4000 (Figure 3.42). No other crystallization condition was identified for NPGP.

3.1.4.2 Optimization of lead conditions

The conditions identified during initial screening were manually optimized with respect to protein concentration, precipitant concentration, pH and additives. A simple grid screening method was applied with precipitant concentration and pH as two major factors to obtain diffraction-quality crystals. To facilitate harvesting, optimization was carried out in hanging and sitting drops using the method of vapor diffusion in 24-well plates (Sarstedt) with 500 µl of reservoir solution. In

sitting drops, 2 µl of concentrated protein solution was mixed with 2 µl of reservoir solution while in hanging drops 1 µl of each solution was mixed. Protein concentration was varied between 20 and 30 mg/ml. Additional grid screening at 10 ^{o}C was also included to slow down nucleation rate relative to the standard 18 ^{o}C.

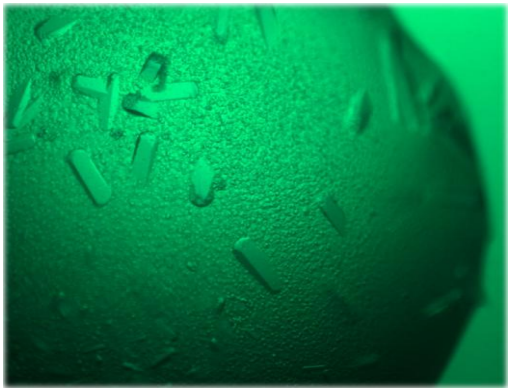


Figure 3.42. Photograph of NPGP crystals grown in 0.1 M CAPSO pH 9.5, 0.1 M $LiSO_4$, 0.1 M NaCl and 12 % PEG 4,000 after incubation for 7 days at 20 ^{o}C under green light. NPGP crystals were formed at the skin of denatured protein.

The optimal protein concentration for reproducible crystal growth at 18 ^{o}C was determined to be 24 mg/ml for both NPGP and PGP. PGP crystals appeared immediately after setting up crystallization plates and already reached maximum size overnight. Glycerol was added to the crystallization cocktail to slow down nucleation and allow the formation of big crystals. In the presence of 1.5 % glycerol, PGP crystals appeared after 3 days at 18 °C in darkness and continued to grow for a week (Figure 3.43).

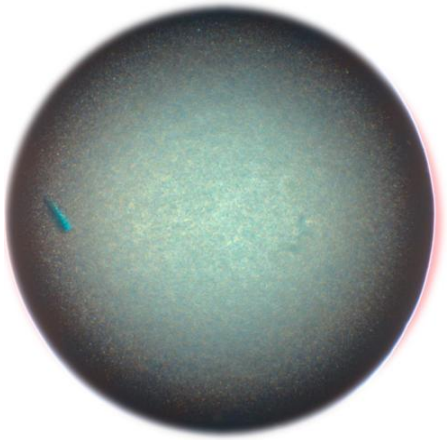


Figure 3.43. Photograph of a single crystal of PGP grown at 18 ^{o}C in 0.1 M CAPSO pH 9.3, 0.2 M NaCl, 1.5 % glycerol and 7.5 % PEG 8000 after optimization.

Crystals of NPGP appeared overnight and resisted all attempts to slow down nucleation. NPGP crystals were deeply rooted on a skin of denatured proteins at the air-water interface of the experimental drop (Fig. 3.42). These crystals required gentle surgery using acupuncture needles to remove the skin. Many separated crystals were harvested together by the fishing net effect of the covering skin. Because of the difficulty associated with harvesting NPGP crystals, alternative crystallization methods such as Granada crystallization boxes for counter-diffusion experiments were tested. NPGP crystals grew by using both sitting and hanging drop vapor diffusion but not counter-diffusion methods. Both PGP and NPGP failed to crystallize at 10 ^{o}C. The optimized robust crystallization condition for both NPGP and PGP rod-shaped green crystals was 0.1 M CAPSO pH 9.3, 0.2 M NaCl, 1.5 % glycerol and 7.5 % PEG 8,000.

3.1.4.3 Harvesting and cryoprotection

After optimizing crystal growth conditions, many types of cryoprotectants were tested by simple vitrification test. For the optimized crystallization cocktails, 20 % (v/v) glycerol, 20 % (v/v) PEG 400 and 17.5 % (v/v) propylene glycol were found to be suitable when mixed with the crystal and the reservoir solution (v/v). On the other hand, both NPGP and PGP crystals shrank in 20 % ethylene glycol. PGP crystals were picked into cryoloops, soaked in the selected cryoprotectant solutions and then flash-frozen in liquid nitrogen under safelight. A maximum of 10 crystals could be fished in one day for NPGP out of several optimization plates because of the skin-surgery problem. Crystals of NPGP harvested after painstaking surgeries were frozen in liquid nitrogen using selected cryoprotectant solutions prior to diffraction experiments.

3.1.4.4 X-ray diffraction experiments

The diffraction qualities of the harvested crystals were tested at HZB BESSY II on the fixed-wavelength 14.3 beamline equipped with a MAR-dtb goniometer and a MX225 CCD detector [145]. Test images were collected at $\varphi = 0^{o}$ and $\varphi = 90^{o}$ and 5 s exposure time at 100 K. PGP crystals showed diffraction to 6 Å resolution at best, often with strong anisotropy, although some showed Bragg peaks at 4 Å. Only few PGP crystals did not diffract at all. In a separate experiment, a complete data set usable to 6 Å was recorded for PGP at ESRF (Courtesy of Sophie Franz, AG Essen). On the other hand, most of the NPGP crystals examined did not diffract at all with a few poorly-diffracting exceptions.

3.1.4.5 Preliminary crystallographic analysis

X-ray diffraction data from PGP crystals were processed with XDS using the graphical user interface XDSAPP2 [160]. The diffraction data for PGP exhibited severe anisotropy. Diffraction extended to 6 Å in one plane but only to 15 Å along the perpendicular plane (Figure 3.44). Complete X-ray diffraction data set was collected to a maximum resolution of 5.4 Å for PGP (see Section 10.4.4). PGP crystals belonged to space group $P4_122$, with unit-cell parameters a = b = 122.54, c = 310.19 Å, α = β = γ = 90°.

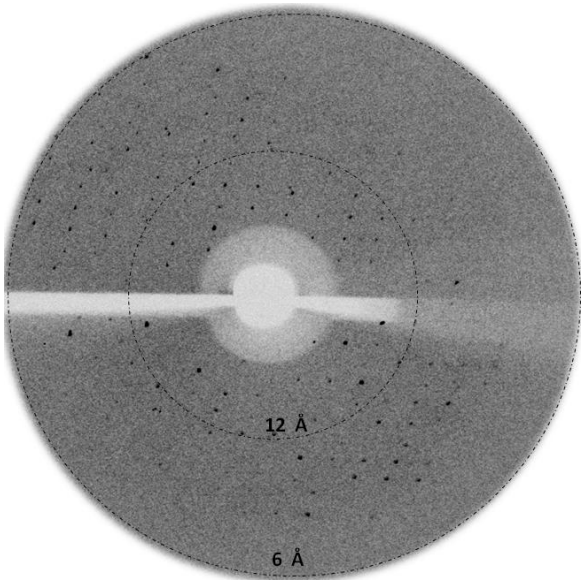


Figure 3.44. Diffraction image of PGP crystal collected at HZB BESSY II BL14.3 using X-rays of wavelength 0.898 Å.

3.1.4.6 Post-crystallization treatment (PXT)

Post-crystallization methods such as cryo-annealing, controlled dehydration and chemical cross-linking were tested to improve the quality of diffraction from PGP and NPGP crystals.

Cryo-annealing

Automated cryo-annealing of frozen crystals was performed by using the remotely controlled cryo-shutter available on BL14.3 at Bessy II for short time intervals (10 to 20 s) [162]. Cryo-annealing performed in darkness had no significant effect on the diffraction quality of PGP crystals.

Controlled dehydration

Flash-freezing protocols can affect the diffraction quality of protein crystals. The problem can be distinguished from the native crystal quality by room-temperature diffraction measurements [144]. 24-well plates containing PGP crystals were taken to HZB Bessy II for dehydration experiments. Crystals were picked in green safelight at the beamline. The native crystal quality was determined by initial assessment of images collected at room temperature under controlled humidity using the automated HC1c device installed on BL14.3. To determine the phase transition point, PGP crystals were dehydrated at reduced relative humidity (RH) [145]. PGP crystals diffracted to 6 Å at 95 % relative humidity at room temperature in darkness. The effect of flash-freezing on diffraction quality using the selected cryoprotecting solutions was therefore negligible. Diffraction was lost completely for all the crystals tested over an RH range of 90 % to 80 %, presumably due to lattice breakdown. Diffraction was not recovered by returning the RH to the starting point, suggesting lack of phase transition.

Chemical cross-linking

Gluteraldehyde cross-linking was used to harden PGP crystals. Cross-linking was carried out by equilibrating hanging droplets holding PGP crystals over reservoirs containing precipitant solution and a final concentration of 0.1 % gluteraldehyde for 10-20 minutes. Crystals became extremely dehydrated and shrank upon extended incubation (>1 hr). Chemical cross-linking performed in darkness at room temperature had no significant effect on the diffraction quality of PGP crystals.

Additive screening

Small molecules included in the Hampton high throughput additive screening kit (HR2-138) were systematically introduced into the mother liquors during robotic crystallization screening of the lead conditions identified for NPGP and PGP crystallization (Ralf Poeschke, AG Essen). Several additives (including 10 mM barium chloride dehydrate, 10 mM strontium chloride hexahydrate, 10 mM cesium chloride, 3 % w/v 6-aminohexanoic acid and 3 % w/v 1,5-diaminopentane dihydrochloride) were identified that supported crystal growth for both NPGP (Fig. 3.45) and PGP. The effect of these additives on the diffraction quality of these crystals remains to be measured.

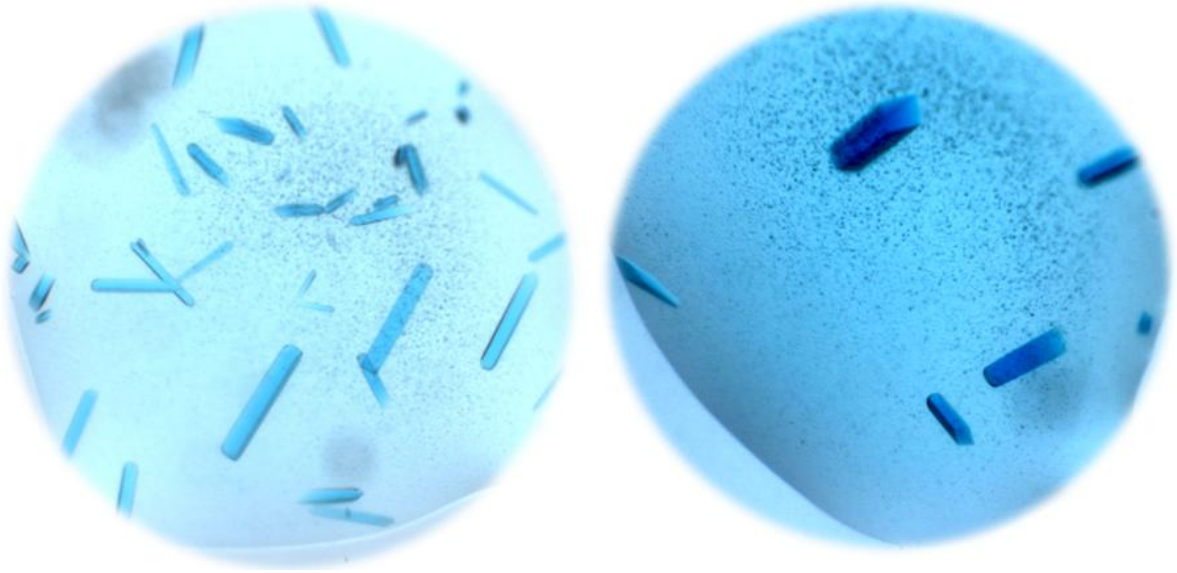


Figure 3.45. Rod-shaped crystals of NPGP grown in 0.1 M CAPSO pH 9.3, 0.2 M NaCl, 1.5 % glycerol, 0.03 M Glycyl-glycyl-glycine and 7.5 % PEG 8000.

3.1.4.7 Protein modification for crystallization

Protein modification techniques such as chemical modifications of surface exposed residues and chromophore replacement were employed to improve the diffraction qualities of *Sorghum* NPGP and PGP crystals.

Modification of cysteine

The numbers of surface-exposed sulfhydryl (-SH) groups in Sb.NPGP and Sb.PGP were quantitatively determined by using Ellman's reagent. Free-cysteine was used as a standard to build a calibration curve (see Fig. 3.46). The obtained data suggested that NPGP carries only one surface-exposed cysteine residue whereas PGP carries three, implying a shielding effect for the NTE. These thiol groups were blocked by iodoacetamide (IAA) to produce S-carboxyamidomethylcysteine. Carboxymethylated PGP and NPGP samples were purified by SEC and subjected to crystallization screening. Diffraction patterns of carboxymethylated crystals were recorded at ESRF Grenoble (Courtesy of Marian Vogt, AG Essen). Chemical modification of cysteines in NPGP and PGP preparations using the improved protocol had no significant effect on the diffraction quality of both PGP and NPGP crystals, however.

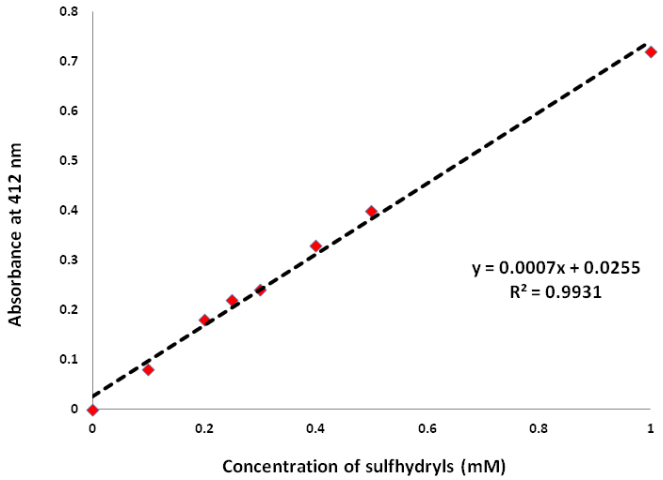


Figure 3.46. Calibration curve for quantitative determination of free sulfhydryls in constructs derived from the photosensory module of *Sorghum* phyB using Ellman's method.

Chromophore replacement

Both PGP and NPGP were assembled *in vivo* with phycocyanobilin (PCB) instead of the native PϕB chromophore. Both PCB-conjugated phytochromes crystallized under similar conditions to those successful for PGP-PϕB and NPGP-PϕB, implying minimal structural differences at the surface of the protein. PGP-PCB (see Figure 3.47) crystallized under the conditions essentially used for the equivalent PϕB adducts. The diffraction quality of PCB-conjugated crystals has not been measured yet. Consequently, the effect of chromophore replacement on the diffraction quality of Sb.PGP and Sb.NPGP crystals remains to be determined.

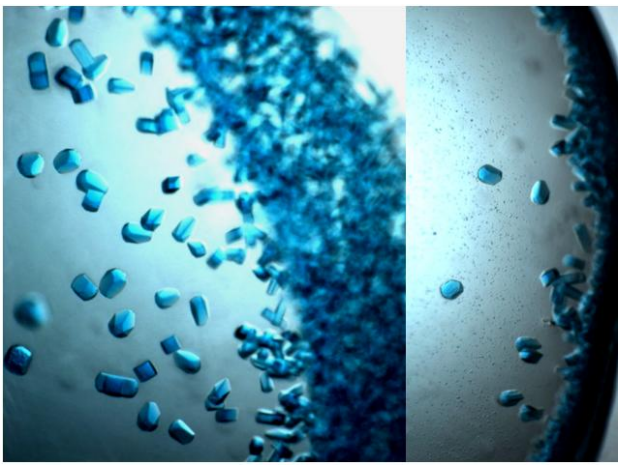


Figure 3.47. Photograph of crystals of the PGP-PCB adducts grown in 0.1 M CAPSO pH 9.3, 0.2 M NaCl, 1.5 % glycerol and 7.5 % PEG 8,000 for chromophore replacement.

3.2 The PAS-GAF bidomain of Sb.phyB

3.2.1 Molecular cloning

In addition to chemical modifications, systematic N- and C-terminal deletions have been reported to promote crystallization of recombinant proteins [163]. In the available crystal structure of At.phyB (4OUR), electron density for the PHY domain was poorly resolved, suggesting domain wobbling [68]. Deletion of the entire PHY domain to avoid domain wobbling would produce PHY-less *Sorghum* phyB that might produce diffraction quality crystals. Such a construct was generated by PCR using newly designed primer SS43.

By combining SS1 and SS2 forward primers and SS43 reverse primer, two new inserts were PCR amplified and cloned into pCDFDuet. The map of sequence verified construct, designated PG, is given in Figure 3.48. The sequence of the second construct containing the N-terminal extension but not PHY domain, designated NPG, is given in Section 10.2. pCDFDuet-*PG and* pCDFDuet-*NPG* were cotransformed with PΦB synthesizing plasmid into BL21(DE3) cells for *in vivo* assembly. NPG was produced but the yield was comparably lower than that of NPGP, hence no large scale purification effort was exerted. On the other hand PG was produced in fairly good yield under similar conditions of induction OD and inducer concentration previously used for PGP but at a different optimal temperature (30 ^{o}C).

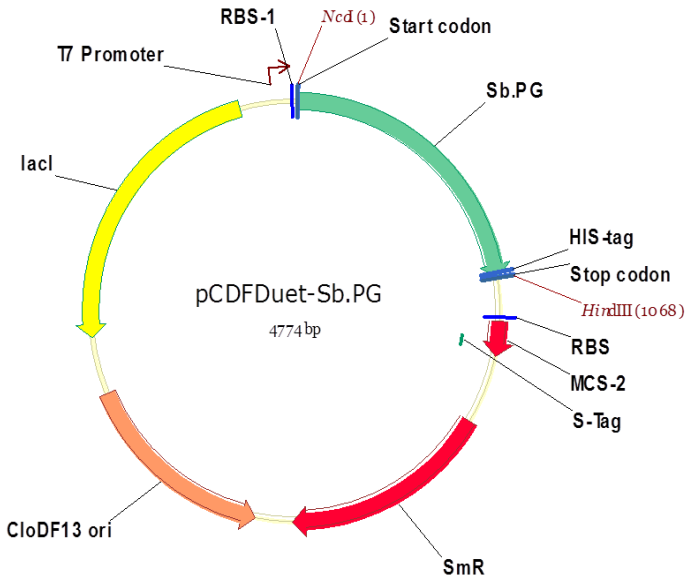


Figure 3.48. The pCDFDuet-*Sb.PG* plasmid for holophytochrome production in combination with p183 (PΦB synthesizing plasmid). The PHY-less construct has C-terminal His-tag for purification by Ni-NTA.

3.2.2 Purification

High yield of recombinant PG enabled rapid purification by the original (Figure 3.49) and improved (Figures 3.50 and 3.51) purification protocols. PG was assembled as holoprotein *in vivo* (Fig. 3.50). Thus, the PHY domain is dispensable for covalent bilin binding. PG binding to the Ni-NTA column was less efficient in the original purification protocol. Consequently, this preparation showed higher contamination (Figure 3.49) than the improved protocol (Figure 3.50). Endogenous contaminant proteins prevalent after Ni-NTA by the improved purification protocol (Figure 3.50) were successfully removed by subsequent SEC (Figure 3.51). SEC cleanup for PG was particularly effective perhaps due to its relatively small size compared to the other constructs.

NPGP and PGP were purified to homogeneity by the improved purification protocol as described in Section 3.1. This purification protocol was effectively applied to PG. Apart from potential degradation due to multiple rounds of freezing and thawing of phytochrome preparations, there was no evidence of proteolysis when samples were kept at 4 ^{o}C for extended periods of time. Altogether, the improved purification protocol was successfully utilized to purify three Sb.phyB derived fragments: NPGP, PGP and PG to 95 % homogeneity (Figure 3.52) suitable for spectroscopic and crystallographic studies.

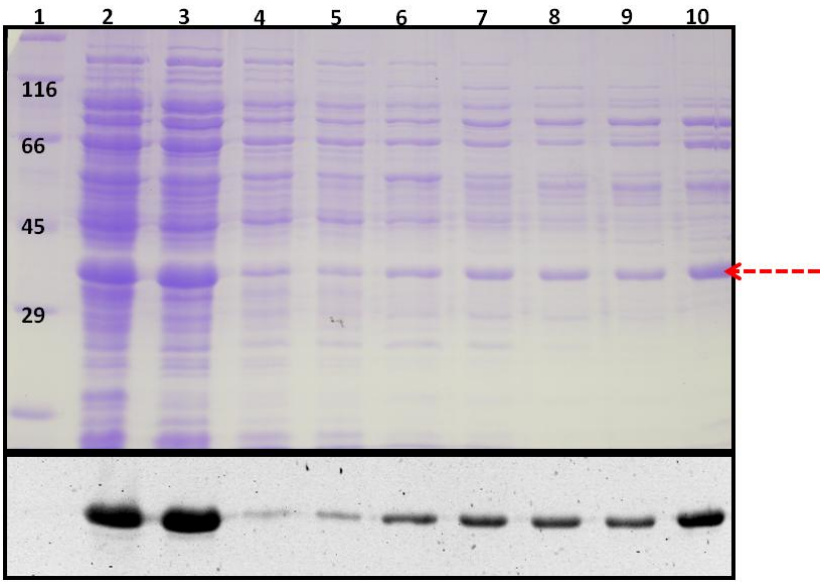


Figure 3.49. SDS-PAGE analysis of PG purification by the original Ni-NTA affinity chromatography protocol detected by Coomassie stain (top) and zinc-induced fluorescence (bottom). Molecular markers in kDa (1) soluble fraction (2), input (3), flow-through (4), washing fractions (5 to 9), elution fraction (10).

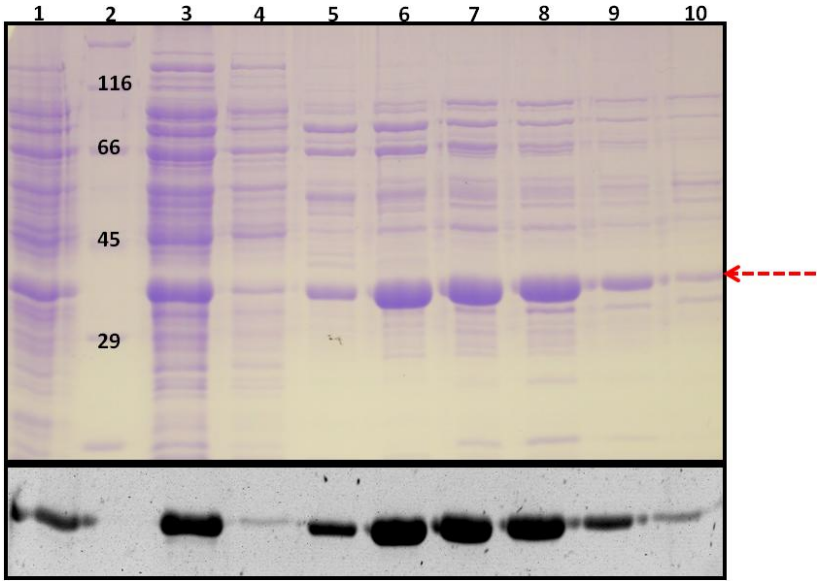


Figure 3.50. SDS-PAGE analysis of purification of PG by the improved Ni-NTA affinity chromatography detected by Coomassie stain (top) and zinc-induced fluorescence (bottom). Soluble fraction (1), molecular markers in kDa (2), input (3), flow-through (4) and elution fractions (5-10).

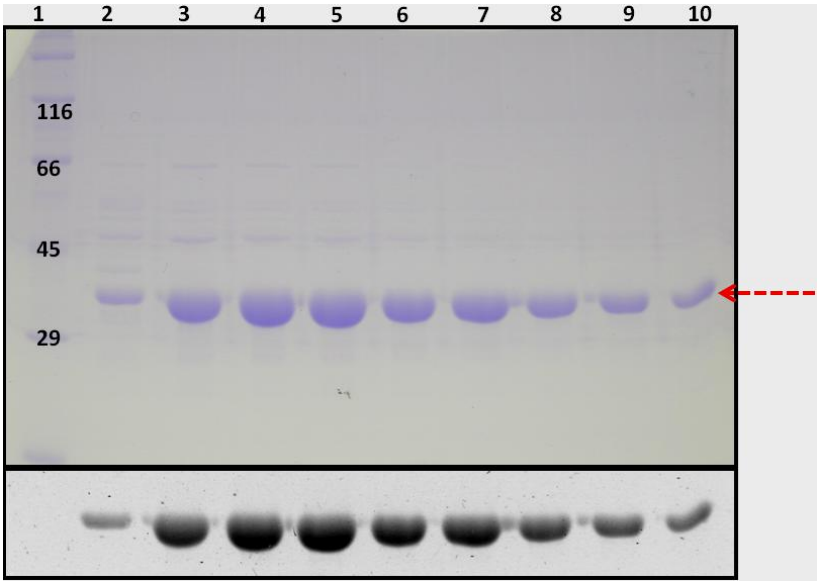


Figure 3.51. SDS-PAGE after preparative SEC of PG purified by the improved purification protocol detected by Coomassie stain (top) and zinc-induced fluorescence (bottom). Molecular markers in kDa (1), SEC fractions C6, C9, C11, C13, C15, D15, D14, D13 and D12 (2-10).

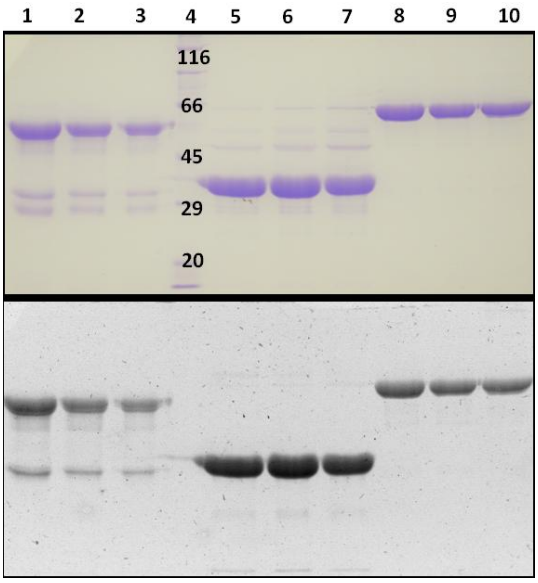


Figure 3.52. SDS-PAGE of phyB derived samples prepared by the improved purification protocol. Coomassie stain (top) and zinc-induced fluorescence (bottom). PGP (1-3), molecular markers in kDa (4), PG (5-7) and NPGP (8-10). Each sample was loaded into three replicates.

3.2.3 Characterization

The assumptions made in Section 3.1.3.1 and the limitation originating from accurate determination of protein concentration and establishing protein purity were taken into consideration to interpret spectroscopic data from PG and other Sb.phyB derived preparations.

3.2.3.1 UV-Vis spectroscopy

UV-Vis absorbance spectra for PG were measured after saturating R/FR irradiation. Negative absorbance for PG around 500 nm, originating from fluorescence, observed in the raw data showed apparently high absorbance after baseline correction (Figure 3.53). As the PHY domain is thought to stabilize the Pfr state through direct contacts with the D-ring pyrrole [45], it was not surprising that PG did not show the UV-Vis spectrum typical of the Pfr + Pr mixture formed by phytochromes following R irradiation. PGP and PG constructs from Sb.phyB were compared to dissect the spectroscopic effect the PHY domain in the Pr state. Removal of PHY domain from PGP led to 7 nm (Figure 3.54) hypsochromic shift of the red band with presumably strong hypochromicity. Tyrosine and tryptophan residues, mainly responsible for protein absorbance of UV light, are predominantly localized in the PHY domain resulting very low ε for PG. The observed hypochromicity (>100 %) was, however, extremely conspicuous. This pronounced shift could,

therefore, not entirely be due to underestimation of PG protein concentration. The oscillator ratio, the ratio of the peak absorbance value of the red band divided by that of the Soret band, is a photochemical property unbiased by protein contamination. The oscillator ratio for PG (=1.6) was significantly lower than that of PGP (=2.9) [61]. The drop in oscillator ratio upon removal of the PHY domain cannot be entirely due to protein contamination.

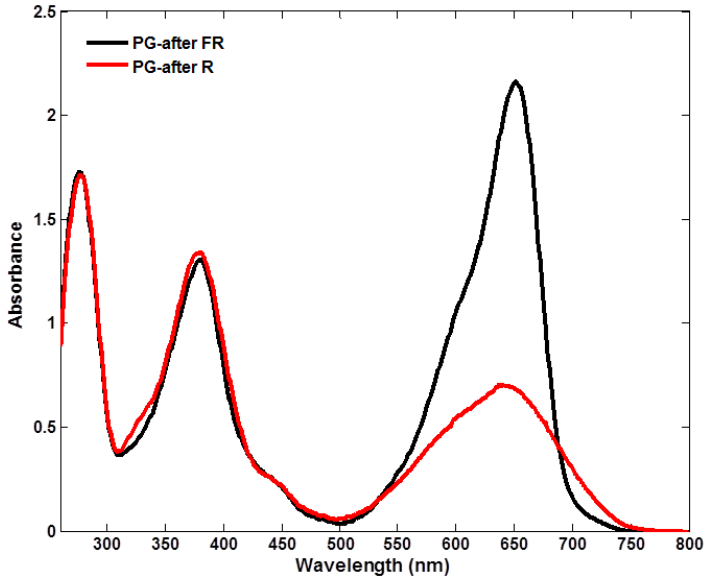


Figure 3.53. Absorption spectra of PG after saturating R/FR irradiation.

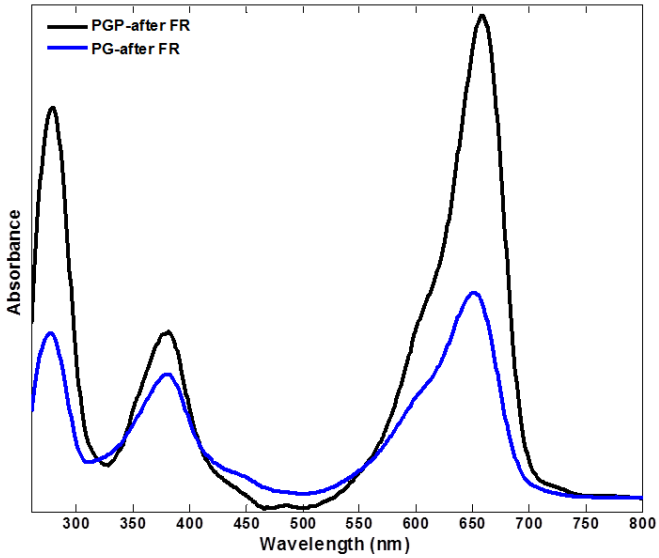


Figure 3.54. Absorbance spectra of PGP and PG after saturating FR-light irradiation normalized for molar concentration in HES-β. Calculated ε at 280 nm used for normalization are given in Table 4. Both constructs were more than 95 % pure as judged by SDS-PAGE in Fig. 3.52.

3.2.3.2 CD spectroscopy

Conformational changes in PG were also characterized by CD spectroscopy following R/FR irradiation (Figure 3.55). Generally, PG after R-light irradiation showed slightly lower CD signal at all wavelengths than PG after FR-light irradiation. The very large (~45 %) loss of CD signals for the Pr state upon removal of the PHY domain from PGP and NPGP needs to be further investigated.

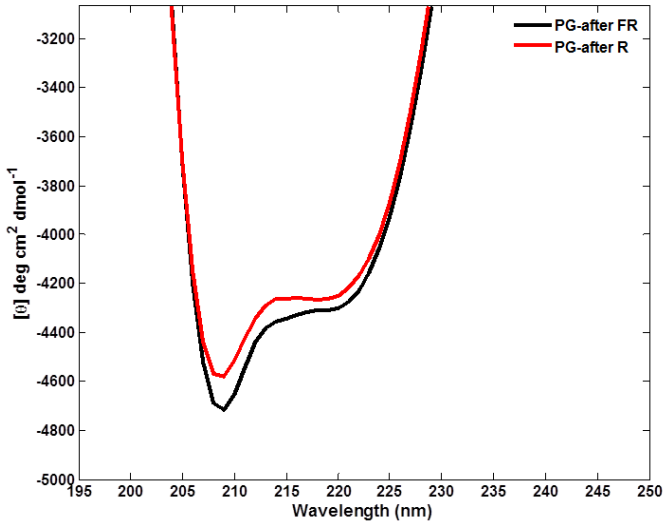


Figure 3.55. The CD spectra of the PG after R/FR irradiation in 10 mM K-phosphate, pH 7.8 and 50 mM ammonium sulfate. Spectra were recorded in 0.5 mm path length cell at 25 ^{o}C.

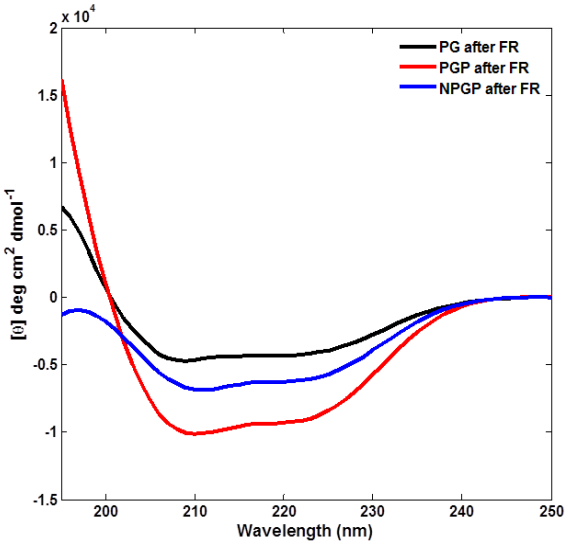


Figure 3.56. The CD spectra of PG, PGP & NPGP after saturating FR irradiation in 10 mM K-phosphate buffer pH 7.8 containing 50 mM ammonium sulfate. Constructs were ca. 95 % pure as judged by SDS-PAGE shown in Figure 3.52.

3.2.3.3 SEC

In SEC, PG eluted at 87 ml corresponding to a 29 kDa globular protein. Large deviation between the experimental molecular weights of Pr of PG from SEC calibration curve (29 ± 5) and theoretical molecular weight calculated from the amino acid sequence (39 kDa) suggests interaction with the gel matrix despite high salt concentration or deviation from the spherical form in the absence of the PHY domain (Fig. 3.57).

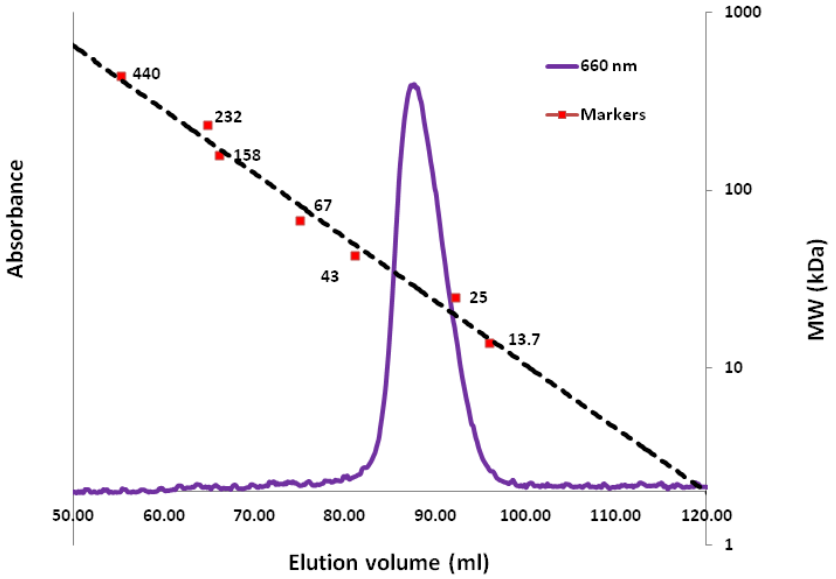


Figure 3.57. SEC elution profile of PG after FR-light irradiation in HES-β using Superdex 200 16/60 preparative column at 1 ml/min flow rate monitored at 660 nm. The secondary axis represents the calibration curve with molecular weight standards.

3.2.4 Crystallization and diffraction experiments

3.2.4.1 Initial crystallization screening

Initial crystallization screening for PG was performed at 10 mg/ml concentration against 10 different crystallization kits from Qiagen (Ralf Poeschke, AG Essen). Screens were setup in green safelight and stored at 18 ^{o}C in the dark. Blue-green needle-like PG crystals appeared after two weeks in 0.05 M Tris-HCl pH 8.5, 0.5 M NaCl and 10 % PEG 4000 (Figure 3.58).

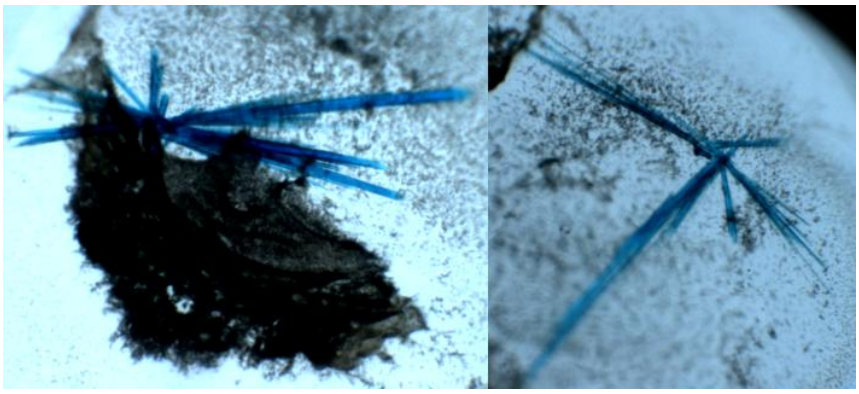


Figure 3.58. Photograph showing a cluster of needle-like crystals of PG construct grown in 0.05 M Tris-HCl pH 8.5, 0.5 M NaCl and 10 % PEG 4000.

3.2.4.2 Optimization of lead condition

PG crystals were optimized manually with respect to protein concentration, precipitant concentration and pH. Clusters of dark-green needle shaped crystals appeared after two weeks in 0.05 M Tris-HCl pH 8.5, 0.1 M NaCl and 7.5 % PEG 6,000. PG crystals formed as stacks of thin needles making them very difficult for single crystal harvesting. Preliminary measurements showed diffraction to near-atomic resolution (3.5 Å) (Figure 3.59). Removing the PHY domain thus leads to significantly improved diffraction, further optimization is necessary, however.

3.2.4.3 Preliminary crystallographic analysis

A diffraction data set was collected for PG crystals using the fixed-energy 13.8 keV BL14.3 beamline operating at Bessy II [162]. The data set was 99.9 % complete with Bragg peaks traceable from 3.2 Å to 47.2 Å. The crystals belonged to the trigonal space group P321 with unit cell dimensions a = b = 134 Å, c = 47 Å, $\alpha = \beta = 90^{o}$ $\gamma = 120^{o}$. Determination of phase from the acquired dataset by molecular replacement was not possible. Diffraction from a single crystal produces evenly-spaced reflections (e.g. Fig. 3.44). The weakly exposed diffraction image in Fig. 3.59 shows anomalous pattern. The magnified view shows unevenly-spaced, circular reflections very close to each other. This might be due to superimposed diffraction patterns originating from more than one crystal. PG crystals were formed as a stack of thin needles that were very difficult to separate (Fig. 3.58). The unusual reflections observed in Fig 3.59 might however be caused by radiation-induced and/or other physical damage to the crystals during harvesting. Irradiation of protein crystals with X-ray radiations could result in fading of diffraction spots resulting in weak

intensity. In any case, the stack of thin, intertwined crystals (Fig. 3.58) was not suitable for single crystal diffraction data collection.

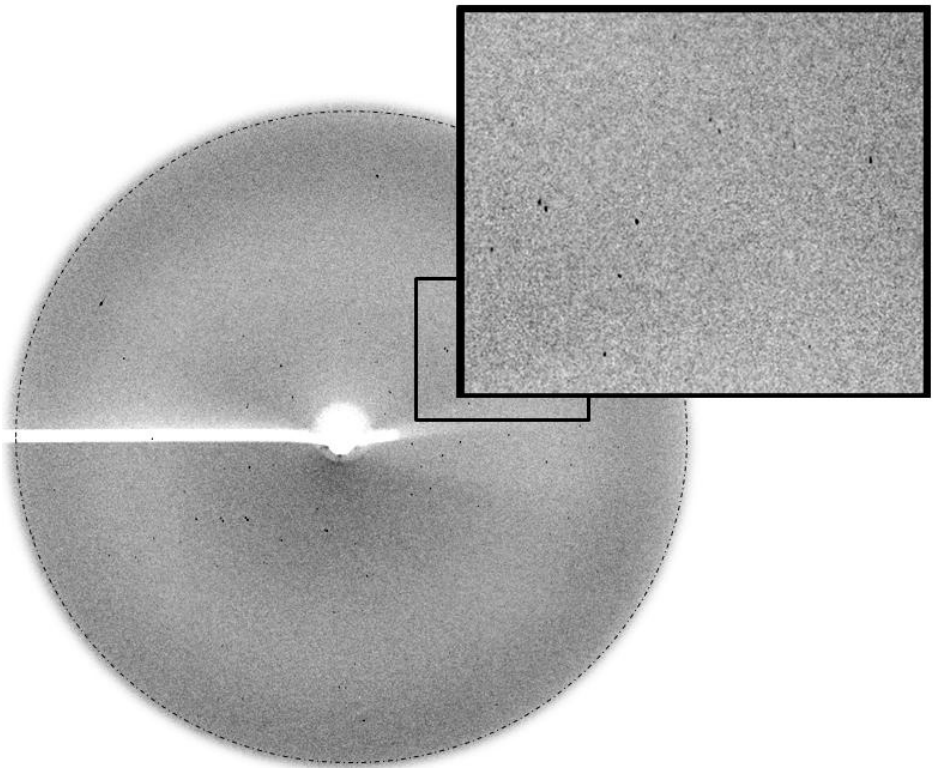


Figure 3.59. Diffraction image of the needle-like PG crystals collected at HZB BESSY II BL14.3 using X-rays of wavelength 0.898 Å. Resolution ring was roughly 3.2 Å. Magnified view of the area delimited by the square is also given.

3.2.4.4 Rescuing strategies

Several techniques are available to improve crystal morphology and diffraction quality. One useful method to improve crystal morphology is additive screening. To this end, the HR2-138 high-throughput additive screening kit (Hampton) was used in a screen with the crystallization condition for needle-shaped PG crystals (0.05 M Tris-HCl pH 8.5, 0.5 M NaCl and 10 % PEG 4000). A few additives were discovered that supported the growth of less-stacked needles (Fig. 3.60) than the parent condition (Fig. 3.58). The diffraction qualities of these crystals remain to be tested. Once suitable morphology is obtained by seeding or additive screening the problem of crystal damage due to physical stress during harvesting or radiation damage can be substantially reduced.

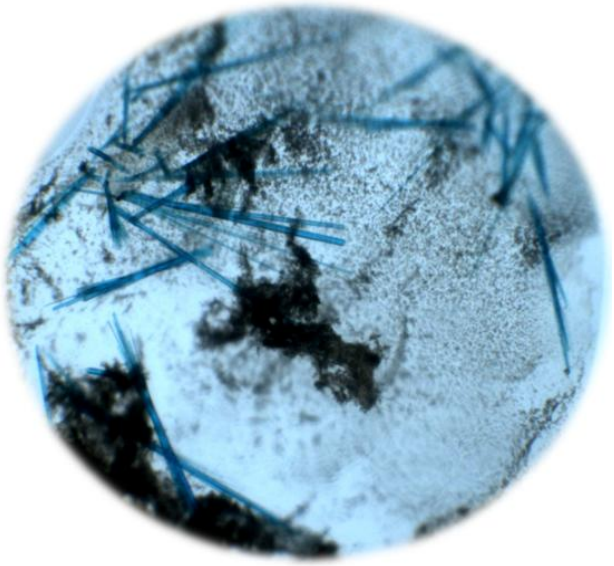


Figure 3.60. Needle-like crystals of the PG construct grown in 0.05 M Tris-HCl pH 8.5, 0.5 M NaCl, 0. 5 % w/v n-Octyl-β-D-glucoside and 10 % PEG 4000. Detergents promoted PG crystallization during additive screening.

3.3 The GAF-PHY bidomain of Sb.phyB

In Cph2, the GAF1 (P3) domain is necessary and sufficient for covalent PCB attachment [15]. Structural alignment of the P3 domains in Cph1 and Cph2 shown in Figure 3.61 was used to construct a GAF-PHY bidomain that can be applied to Sb.phyB. The photosensory module of Cph2 was used because it is the only structure known to utilize the GAF and PHY domains to form a holophytochrome. GAF domain in Cph1 has an extra long loop that extends from the chromophore binding pocket to the surface of the protein. On the other hand, Cph2 has a short β-strand that seals the chromophore binding pocket; a feature absent in Cph1.

Based on the structural alignment in Fig. 3.61 and domain architecture of Sb.phyB, a 44 kDa PAS-less deletion mutant, GP (residues 267 to 655), was cloned into pCDFDuet and co-expressed in *E. coli* with the bilin synthesizing plasmid. Recombinant GP showed higher protein production than PG when run adjacent to uninduced controls. GP partitioned into both soluble and insoluble fractions after cell lysis and centrifugation (Figure 3.62). However, GP failed to autoassemble, showing neither zinc-induced fluorescence nor photochromic activity (Figure 3.62). All attempts to produce a chromoprotein by *in vitro* assembly with PCB also failed. Therefore, it was concluded that complete removal of the PAS domain prevents proper folding in *Sorghum* phyB. Because the

major aim of this work was to identify Sb.phyB variants for crystallization, no attempt was made to characterize this protein further.

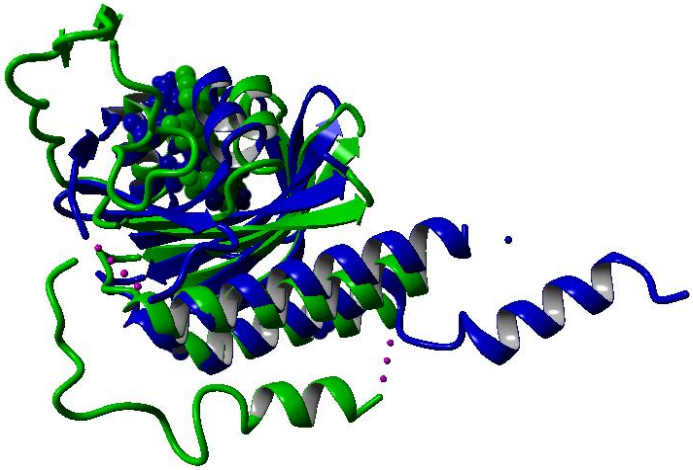


Figure 3.61. Structural alignment of the GAF domains from Cph1 (green, 2VEA) and Cph2 (blue, 4BWI). Apart from the PAS domain, Cph1 also has additional extended loop in the GAF domain that is absent in Cph2.

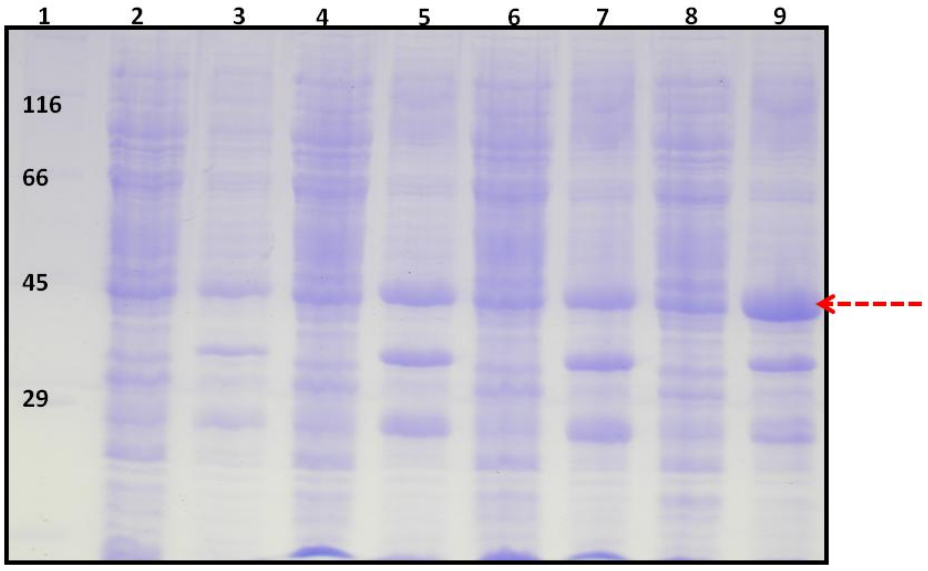


Figure 3.62. SDS-PAGE for test solubility of GP from pCDFDuet construct in combination with PϕB synthesizing plasmid. Molecular markers in kDa (1), soluble fractions (2, 4, 6 & 8), insoluble fractions (3, 5, 7 & 9). No zinc-induced fluorescent band was detected at the desired molecular weight (red arrow).

3.4 The SER mutant of Sb.phyB

3.4.1 Homology modeling and site-directed mutagenesis

Surface entropy reduction (SER) involves mutating high-conformational entropy surface residues such as K, E or Q into A to enhance crystallization of target proteins. SER mutations have been particularly successful when introduced into solvent-exposed loops [146]. Two residues suitable for SER strategy were selected based on automated assessment (SERp sever) [164] and homology modeling. Sb.PGP was sequence-threaded with proteins of known structure that share homology within the photosensory module. Residues in PGP (114–655) were modeled on the crystal structure of Cph1 (2VEA) using the I-TASSER server, which uses threading, *ab initio* modeling, and structural refinement. The missing sequences in 4OUR and its poor resolution made it less preferable template structure than 2VEA for the I-TASSER server (Figure 3.63).

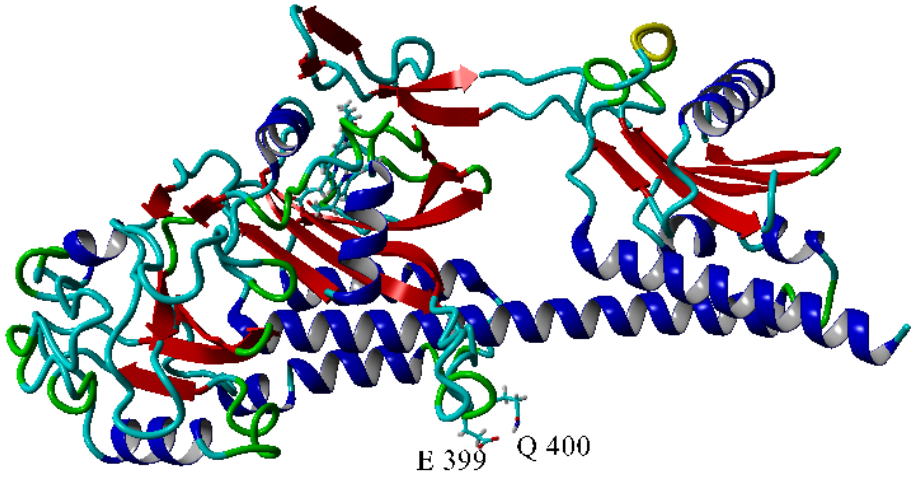


Figure 3.63. Homology model of the Pr state of Sb.PGP based on Cph1Δ2 structure (2VEA). The ribbon diagram of the model is shown using different colors for different secondary structural elements of PGP. The labeled residues were mutated into 'A' for Surface Entropy Reduction.

The identified putative SER target sites (Fig. 3.63) were residues 399 and 400 (EQ). The corresponding positions in At.phyB were structurally disordered as indicated by missing electron density [68]. These residues were engineering into a combination of two A residues by PCR-based *in vitro* site-directed mutagenesis using SS52 and SS53 primers. Site-directed mutants were created through the Quikchange method (Stratagene) from the wild-type recombinant clones in pCDFDuet plasmid. Positive clones were identified by colony PCR and the desired substitutions were verified by test expression and sequencing.

3.4.2 Extraction and purification

Surface entropy reduction (SER) mutants were produced in the hope of improving the diffraction quality of the wild-type PGP crystals. Surface engineered (ΔA399 - ΔA400) mutant protein, labeled PGP-SER1, was purified by the improved Ni-NTA affinity purification protocol (Fig. 3.64) and SEC (Fig. 3.65) into a homogeneous (Fig. 3.66) and photoreversible preparation (Fig. 3.67).

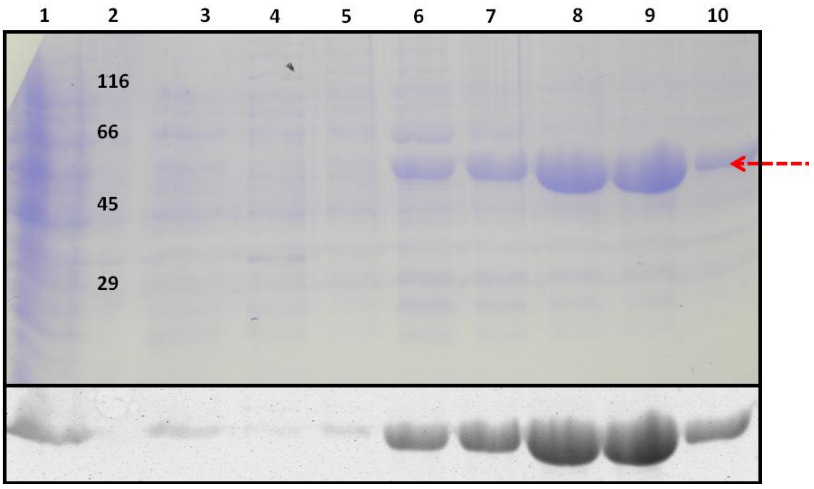


Figure 3.64. SDS-PAGE analysis of PGP-SER1 after the improved Ni-NTA affinity chromatography. Coomassie stain (top) and zinc-induced fluorescence (bottom). Input (1), molecular markers in kDa (2), flow-through (3), washing fractions (4, 5) and elution fractions (6-10). The red arrow indicates the position of assembled PGP-SER1.

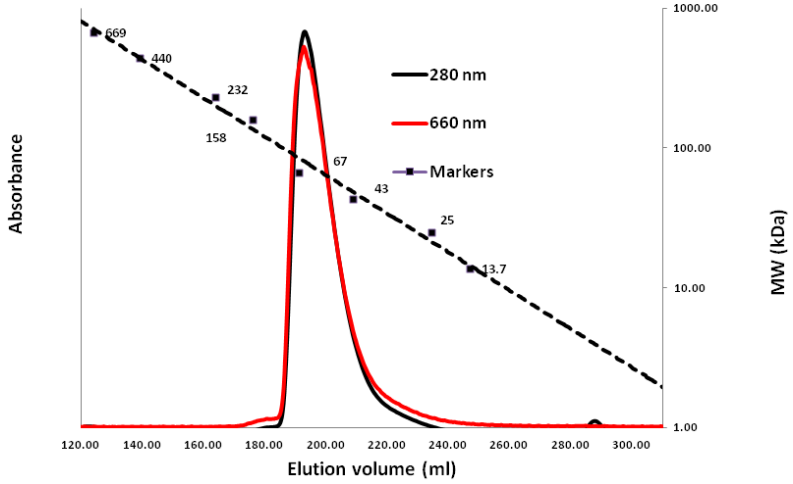


Figure 3.65. SEC profile of PGP-SER1 after saturating FR-light irradiation in HES-β using Superdex 200 26/60 preparative column at 1 ml/min flow rate monitored at 280 nm and 660 nm. The secondary axis represents the calibration curve with molecular weight standards: thyroglobulin (669), ferritin (440), catalase (232), aldolase (158), albumin (67), ovalbumin (43), chymotrypsinogen A (25) and RNAse A (13.7) in kDa.

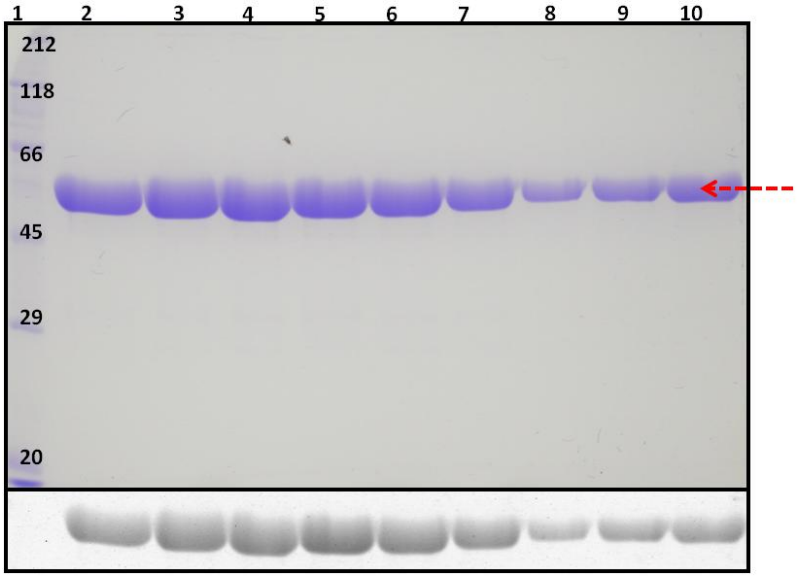


Figure 3.66. SDS-PAGE analysis of PGP-SER1 after the improved Ni-NTA protocol and SEC. Coomassie stain (top) and zinc-induced fluorescence (bottom). Molecular markers in kDa (1), SEC fractions C16, D13, D10, D7, D4, D1, E9, E6 and E3 (2-10). The red arrow indicates the position of assembled PGP-SER1.

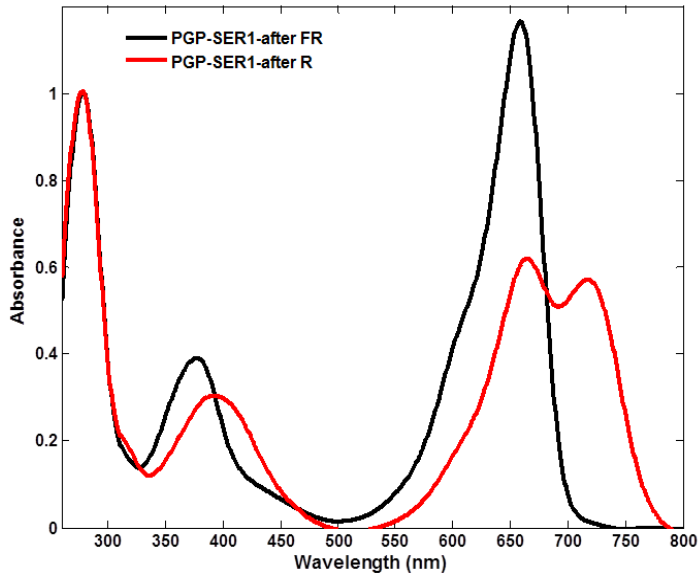


Figure 3.67. Absorption spectra of PGP-SER1 after saturating R/FR irradiation.

3.4.3 Crystallization screening

Because this particular mutant was generated to enhance crystallization, its spectroscopic and photochemical properties were not studied in any detail (see Fig. 3.67 for UV-Vis spectra). Initial

crystallization screening for PGP-SER1 was carried out at two different protein concentrations (24 mg/ml and 12 mg/ml) against 15 crystallization screening kits from Qiagen, Jena Bioscience and Molecular dimensions (Ralf Poeschke, AG Essen). Several lead conditions were identified that showed crystal growth after a week, several of which had been successful for wild-type PGP. However, a few novel crystallization conditions were also discovered, namely 0.1 M imidazole pH 8.0, 2.5 M NaCl and 0.05 M Tris HCl pH 8.5, 0.5 M NaCl and 10 % PEG 4,000 as a precipitating agent. The diffraction qualities of these crystals remain to be investigated.

3.4.4 Additive screening

To manipulate sample-sample and sample-solvent interactions which likely improve crystals or alter sample solubility, the HR2-138 additive screening kit (Hampton) was used. Several additives were identified that grew bar-shaped crystals for PGP-SER1 (Fig. 3.68). These include 10 mM calcium chloride dehydrate, 10 mM magnesium chloride hexahydrate, 10 mM phenol, 10 mM spermine tetrahydrochloride and 0.5 % polyvinylpyrrolidone K15.

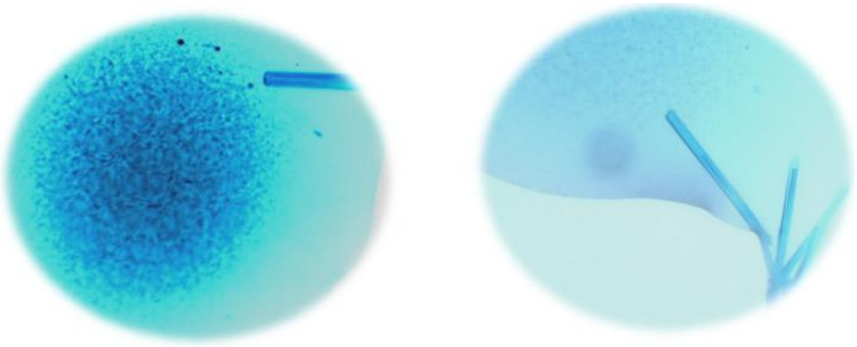


Figure 3.68. Rod-shaped crystals of surface-engineered mutants of PGP (PGP-SER1) after additive screening.

3.5 The Y291H mutant of Sb.phyB

3.5.1 Site-directed mutagenesis

The poorly photoactive and fluorescent Y276H allele of PHYB or YHB from *Arabidopsis thaliana* was shown to exhibit light-independent activation of gene expression and COP development [62]. For characterization and crystallization of this very interesting mutation from Sb.phyB, various mutant constructs (Y291H) were generated by *in vitro* site-directed mutagenesis. These constructs were designated as Y291HNPGP, Y291HPGP and Y291HPG. All mutations were verified by DNA sequencing.

3.5.2 Extraction and purification

The expression and purification protocols for the Y291H mutants of NPGP, PGP and PG were essentially the same as that of the corresponding wild-type proteins. For example, both the original (Figures 3.69 and 3.70) and improved (Figures 3.71 and 3.72) purification protocols are shown as applied to Y291HNPGP.

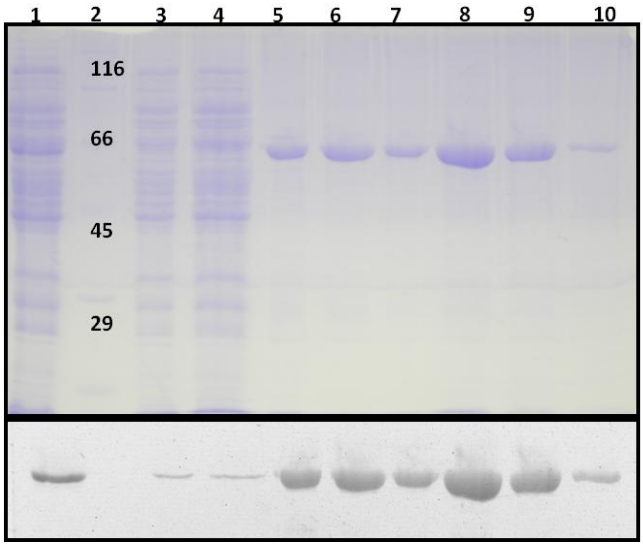


Figure 3.69. SDS-PAGE for purification of Y291HNPGP by the original Ni-NTA chromatography detected by Coomassie stain (top) and zinc-induced fluorescence (bottom). Input (1), Markers in kDa (2), flow-through (3), wash (4) and elution fractions (5-10).

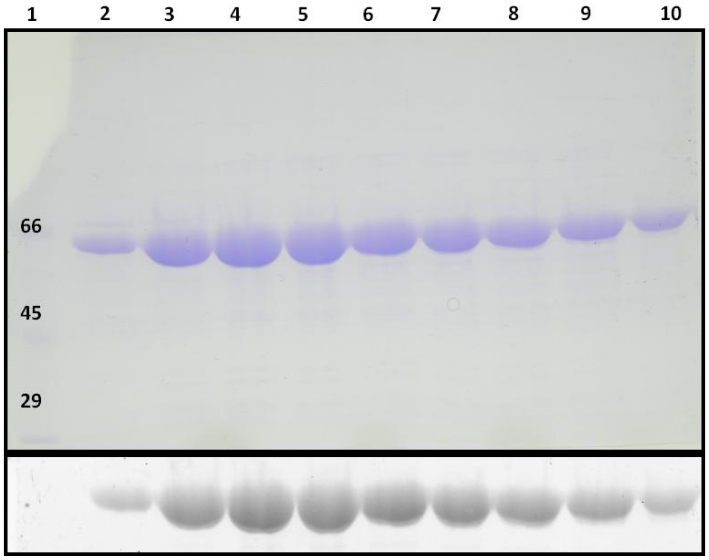


Figure 3.70. SDS-PAGE after preparative SEC for Y291HNPGP purified by the original purification protocol. Coomassie stain (top) and zinc-induced fluorescence (bottom). Markers in kDa (1), SEC fractions: C1, C6, C9, C12, C15, D15, D12, D9 and D6 (2-10).

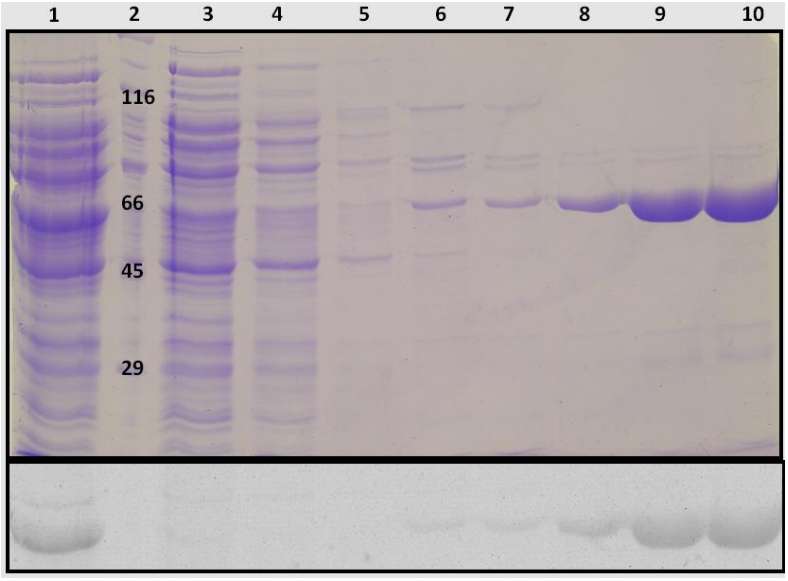


Figure 3.71. SDS-PAGE for Y291HNPGP using the improved Ni-NTA affinity chromatography detected by Coomassie-stain (top) and zinc-induced fluorescence (bottom). Input (1), Markers in kDa (2), flow-through (3), wash (4-7) and elution fractions (8-10).

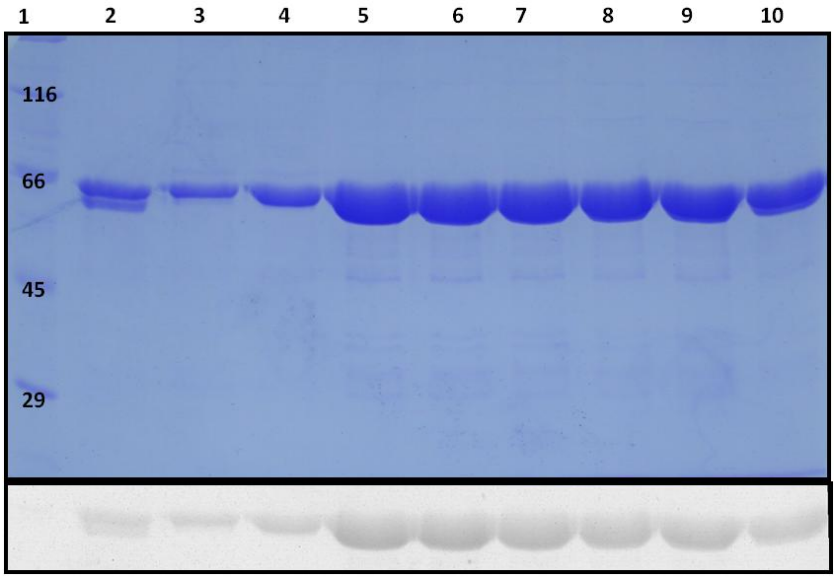


Figure 3.72. SDS-PAGE analysis of Y291HNPGP after the improved Ni-NTA affinity purification and preparative SEC as detected by Coomassie-stain (top) and zinc-induced fluorescence (bottom). Markers in kDa (1), SEC fractions D7, C4, C5, C7, C9, C11, C13, C15 and C16 (2-10).

3.5.3 Characterization

To study the effects of Y291H mutation on the structure of the photosensory module, spectroscopic, photochemical and SEC properties of purified Y291H mutants of Sb.phyB

(Y291HNPGP, Y291HPGP and Y291HPG) were compared to each other and to the corresponding wild-type phytochromes.

3.5.3.1 UV-Vis spectroscopy

The UV-Vis spectra of Y291HNPGP and Y291HPG were recorded after saturating R/FR irradiation. As expected, both Y291HNPGP and Y291HPG exhibited severely-compromised photoconversion (Figs. 3.73 & 3.75) and reduced difference spectra (Figs. 3.74 & 3.76) compared to their WT counterparts.

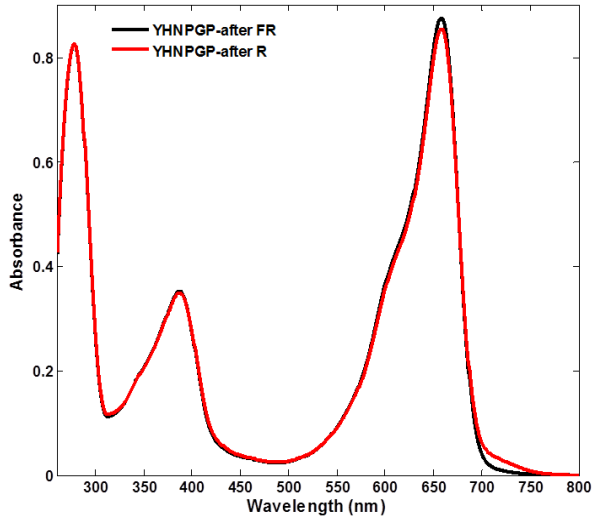


Figure 3.73. Absorbance spectra of YHNPGP (Y291H) after saturating R/FR irradiation.

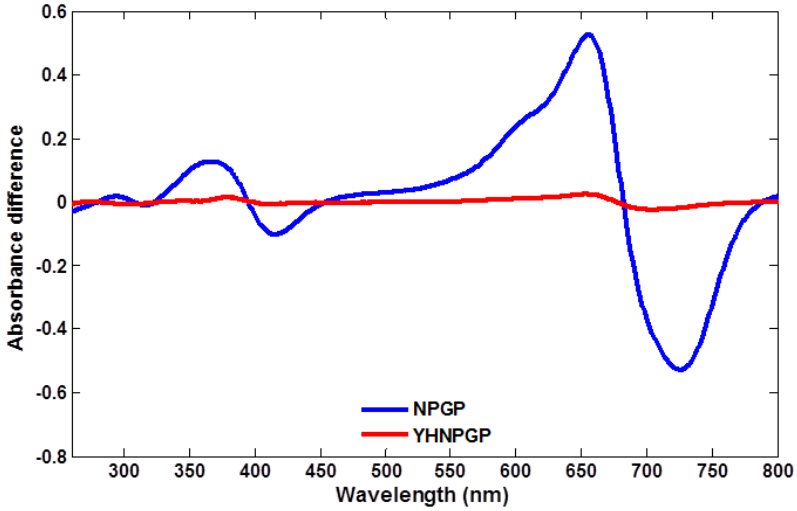


Figure 3.74. Difference absorption spectra for NPGP (WT) and YHNPGP (Y291H) normalized for molar concentration.

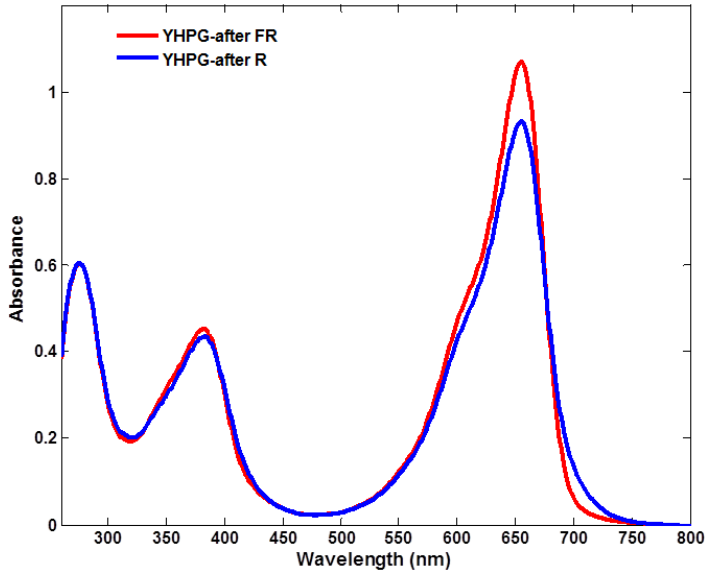


Figure 3.75. Absorbance spectra of YHPG (291H) after saturating R/FR irradiation.

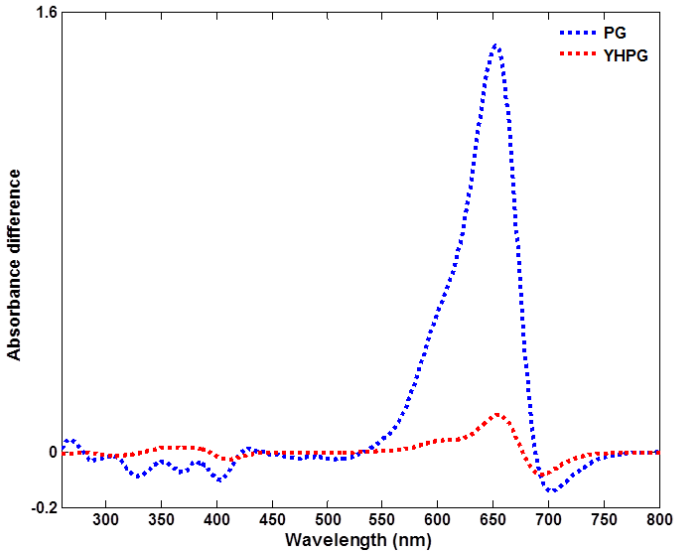


Figure 3.76. Difference absorption spectra of PG (WT) compared to YHPG (Y291H) normalized for molar concentration.

Comparison of the spectra of NPGP and Y291HNPGP after FR irradiation revealed a bathochromic shift near 388 nm accompanied by hypochromicity near 659 nm (~27 %) without any detectable hyperchromicity near 720 nm (Fig. 3.77). Bathochromic shift near the blue, large hypochromicity in red coupled with hyperchromicity in the far-red is a characteristic feature of photoconversion in wild-type phytochromes. Y291HNPGP after FR-irradiation, therefore, faithfully imitated the

photochemical properties of NPGP after R-irradiation in the blue and red region but not in the FR region. Spectroscopic evidence for hyperchromicity in the FR-region for Y291H mutant compared to the wild-type was observed for Y291HPGP (Figure 3.78). Uniquely, Y291HPGP showed a clear shoulder between 700 and 750 nm even after saturating FR-light irradiation implying conformational heterogeneity.

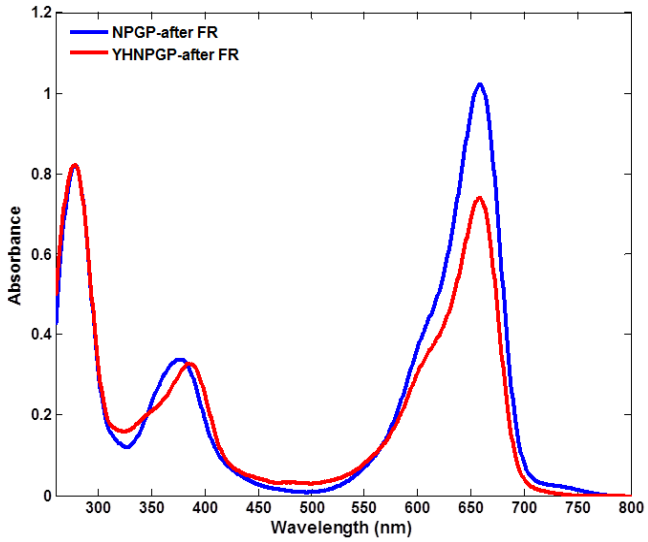


Figure 3.77. Absorbance spectra of NPGP (WT) and YHNPGP (Y291H) after saturating FR-light irradiation in HES-β normalized for molar concentration.

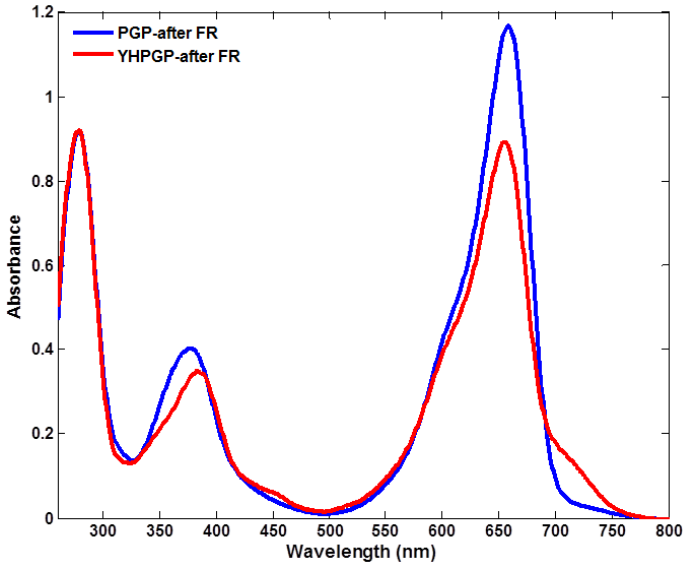


Figure 3.78. Absorbance spectra of PGP (WT) and YHPGP (Y291H) after saturating FR-irradiation normalized for molar concentration.

Reduced oscillator ratio is thought to indicate a more cyclic bilin conformation [60, 61]. The degree of coupling between conformational change in the protein and conformational change in the chromophore can be demonstrated by a plot of the difference absorption against oscillator ratio. There was a strong correlation between increasing $\Delta(\Delta A)$ and increasing oscillator ratio for wild-type constructs, but this was not apparent for the Y291H mutants (Figure 3.79).

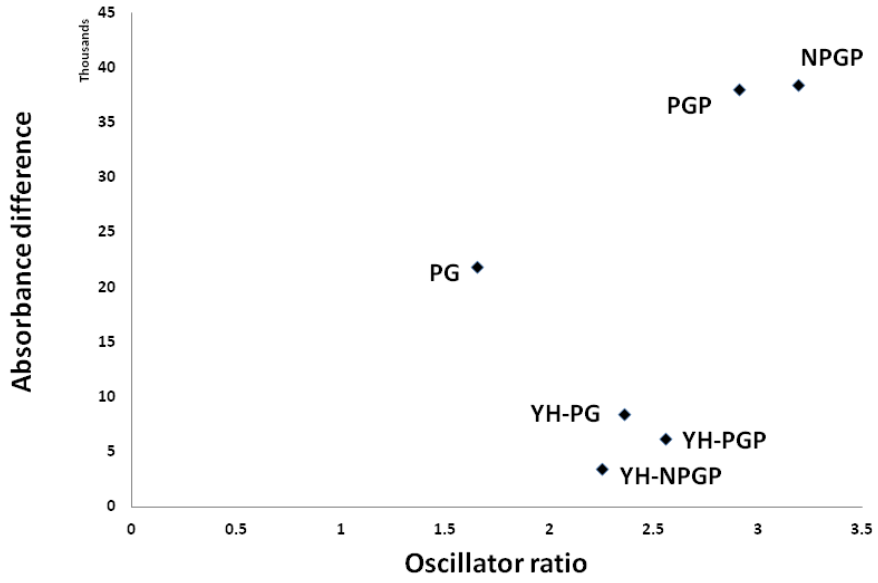


Figure 3.79. Plot of normalized Δ(ΔA) vs. oscillator ratio for the wild-type and Y291H substitution mutant proteins normalized for molar concentration. Calculated ε at 280 nm used for normalization are given in Table 4. All constructs were more than 95 % pure as judged by SDS-PAGE in Figures 3.52 & 3.70.

3.5.3.2 Fluorescence spectroscopy

Fluorescence spectroscopy at room temperature was used to characterize Y291H mutant phytochromes. The characteristics of the absorbance, excitation and emission spectra of Y291HNPGP were compared with that of NPGP. The fluorescence excitation and emission spectra of Y291HNPGP (Figure 3.80 B) showed maxima near 655 nm and 680 nm, respectively. Several differences in the spectra were observed. The excitation peak of NPGP was significantly broader (~10 nm) than its absorbance peak (Fig. 3.80 A) whereas they were similar in Y291HNPGP (Fig. 3.80 B). Besides, the observed Stokes shift for NPGP (~28 nm) was slightly larger than that of Y291HNPGP (~24 nm). The fluorescence characteristics of PGP compared to Y291HPGP (Fig. 3.81) and PG compared to Y291HPG (Fig. 3.82) were very similar to that of NPGP compared to Y291HNPGP.

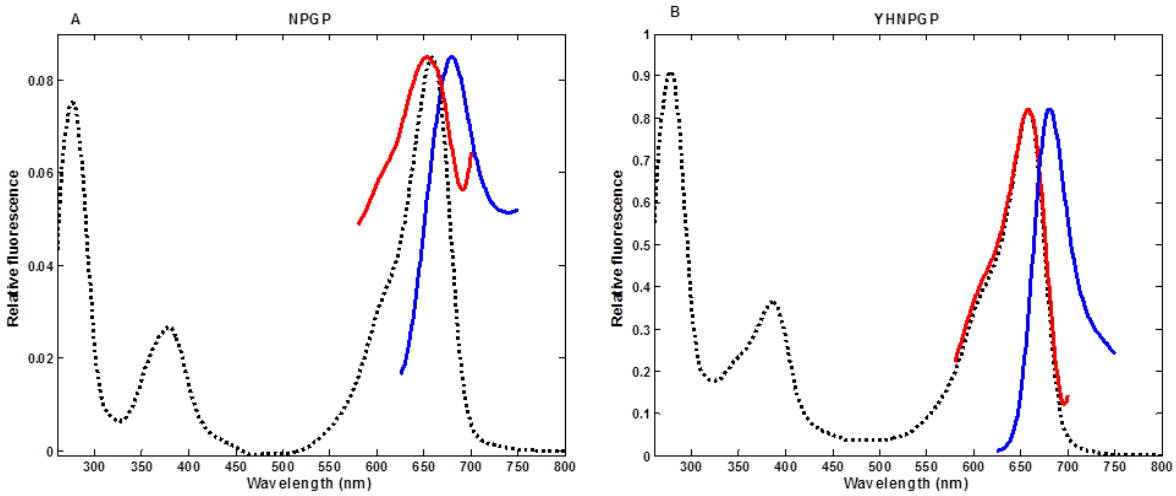


Figure 3.80. Fluorescence spectra of NPGP (A) and YHNPGP (Y291H) (B) at room temperature. Excitation and emission spectra (shown in red and blue, respectively) were normalized to their maxima. The absorbance spectrum (represented as dashed lines) was normalized to the maximum in the red region).

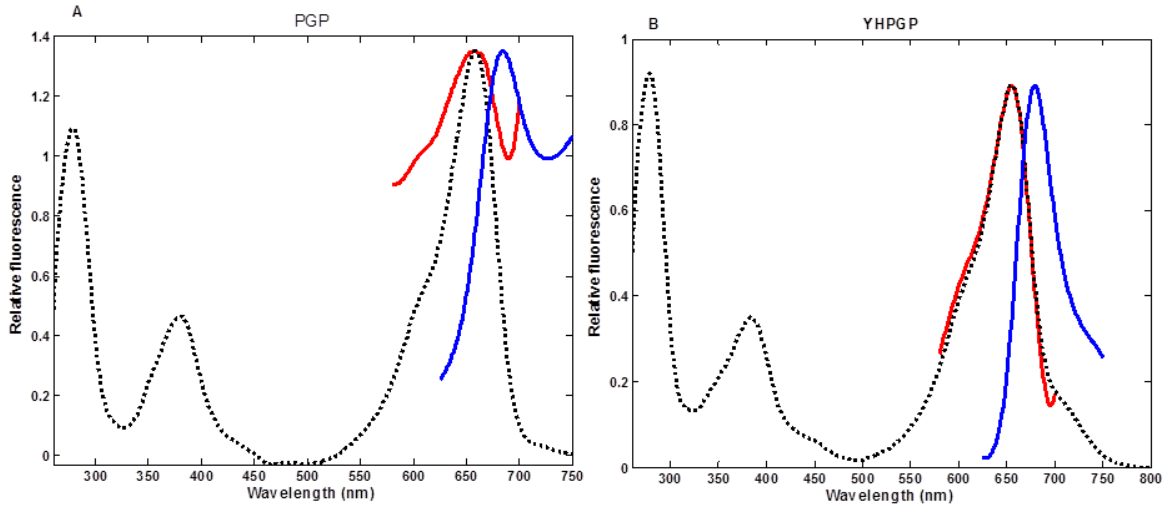


Figure 3.81. Fluorescence spectra of PGP (A) and YHPGP (Y291H) (B) at room temperature. Excitation and emission spectra (shown in red and blue, respectively) were normalized to their maxima. The absorbance spectrum (represented as dashed lines) was normalized to the maximum in the red region).

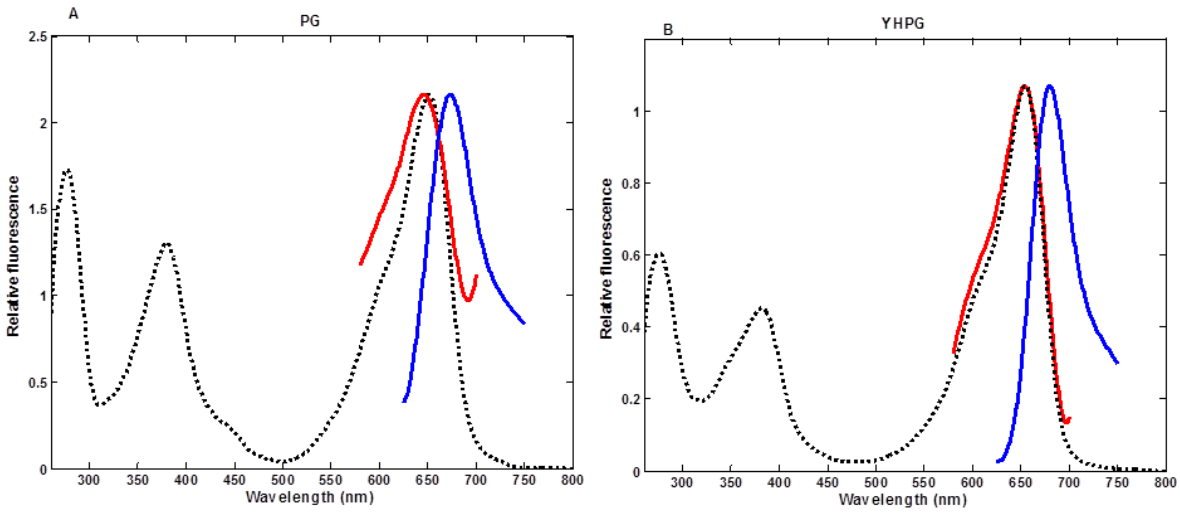


Figure 3.82. Fluorescence spectra of PG (A) and YHPG (Y291H) (B) at room temperature. Excitation and emission spectra (shown in red and blue, respectively) were normalized to their maxima. The absorbance spectrum (represented as dashed lines) was normalized to the maximum in the red region).

A key parameter for comparing fluorescent properties of different proteins is fluorescence quantum yield. As determination of the fluorescence quantum yield is beyond the scope of the present work, the emission and excitation spectra of wild-type and mutant phytochromes were compared after correction and normalization for molar concentrations. Furthermore, dilute samples were used such that photoreversibility was negligible and fluorescence measurements were within the range of linearity.

Wild-type PG exhibited pronounced fluorescence excitation compared to PGP and NPGP (Fig. 3.83). This construct showed the highest normalized fluorescence (~2 fold) in the Pr form compared to both PGP and NPGP (Fig. 3.83). On the other hand, all Y291H mutant proteins exhibited enhanced fluorescence (Figure 3.84). The Y291H mutant derived from the PG construct appeared to have reduced normalized fluorescence compared with Y291HPGP and Y291HNPGP mutants. This suggested that more energy is lost through futile modes of de-excitation upon removal of the PHY domain. The differences between the normalized fluorescence intensities in Y291H mutants were not however very large. Consequently, the observed lower normalized fluorescence for Y291HPG than the other constructs could be caused by protein contamination and misestimated protein concentrations. Therefore, the effect of the PHY domain on the ability of Y291H mutation to enhance fluorescence needs to be further investigated using more precise methods.

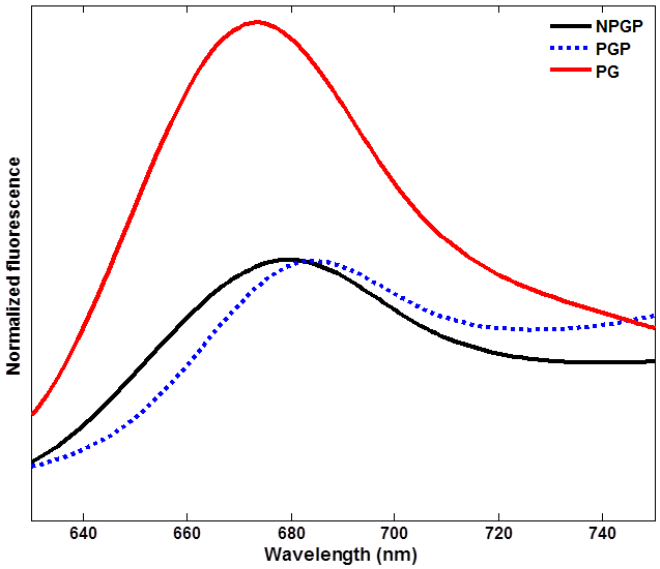


Figure 3.83. Concentration-normalized fluorescence emission spectra for NPGP, PGP and PG. Calculated ε at 280 nm used for normalization are given in Table 4. Constructs were more than 95 % pure judged by SDS-PAGE Figure 3.52.

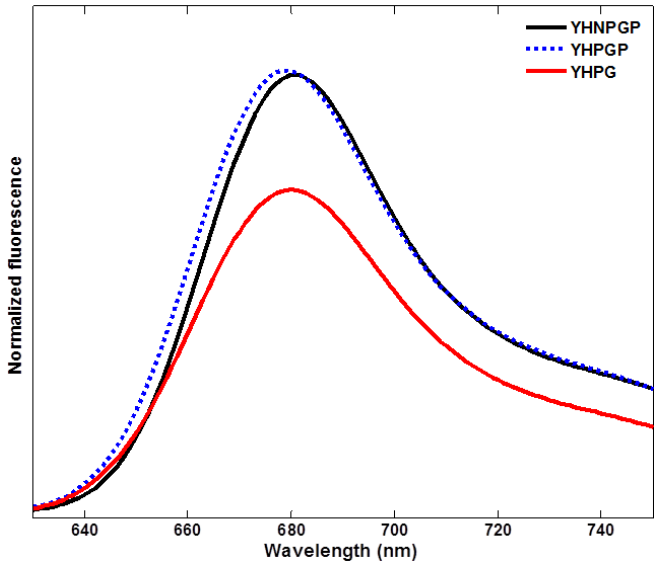


Figure 3.84. Concentration-normalized fluorescence emission spectra for YHNPGP (Y291H), YHPGP (Y291H) and YHPG (Y291H). Calculated ε at 280 nm used for normalization are given in Table 4.

3.5.3.3 CD spectroscopy

CD spectroscopy was employed to characterize the effect of Y291H mutation on secondary structure of the photosensory module of Sb.phyB. The CD spectra of PGP and Y291HPGP were compared after saturating FR-light irradiation. The far-UV CD spectrum of Y291HPGP showed

large negative ellipticity near 222 and 218 nm. The features of this spectrum were typical of globular proteins containing α-helices and β-sheets (Figure 3.85) implying properly folded protein. CD spectra also showed small differences in secondary structures between PGP and Y291HPGP after FR-light irradiation. At first glance, the two spectra appear quite similar for all structural elements with nearly identical amplitudes for the two extrema suggesting very reproducible signals. At 210 nm, PGP showed an ~2 % larger negative ellipticity than YHPGP. Between 215 nm and 245 nm the negative ellipticity in Y291HPGP was lower, at 220 nm by ~5 % (Figure 3.85) than that of PGP. These spectral changes implied higher α-helical content in Y291HPGP than PGP after FR irradiation. Therefore, Y291HPGP after FR-irradiation (Figure 3.85) faithfully imitated the circular dichroic properties of PGP after R-irradiation.

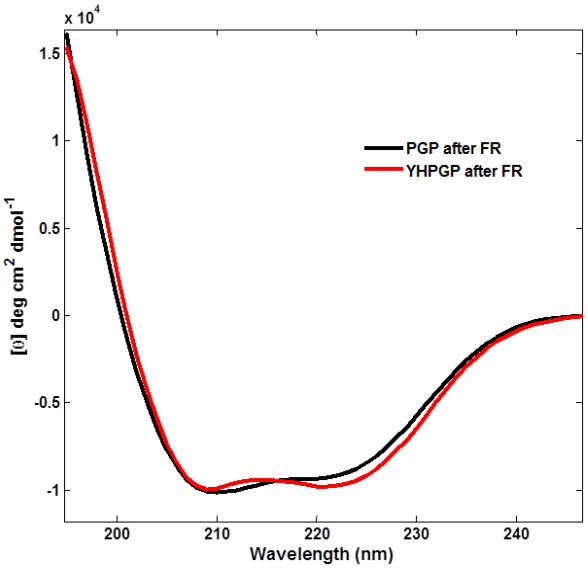


Figure 3.85. The CD spectra of PGP (WT) and YHPGP (Y291H) after FR-light irradiation in 10 mM K-phosphate, pH 7.8 containing 50 mM ammonium sulfate. Spectra were recorded in 0.5 mm pathlength cell at 25 ^{o}C.

CD data collected as a function of temperature (Figure 3.86) were used to determine the effect of the Y291H mutation on thermal stability of NPGP. Phytochrome samples were heated at the rate of 0.25 ^{o}C per minute and held for 5 minutes of equilibration time before measurement. Between 25 ^{o}C and 50 ^{o}C, whereas the CD spectra for NPGP were quite closely spaced (Fig. 3.36 A), Y291HNPGP showed larger effects (Fig. 3.86), implying larger conformational rearrangement. Y291HNPGP exhibited sharp transition at ~55 °C (Fig. 3.86) compared to the sharp transition at 59

°C for NPGP (Fig. 3.36 A). The slope of the transition for Y291HNPGP was steeper than that of NPGP (Figure 3.87).

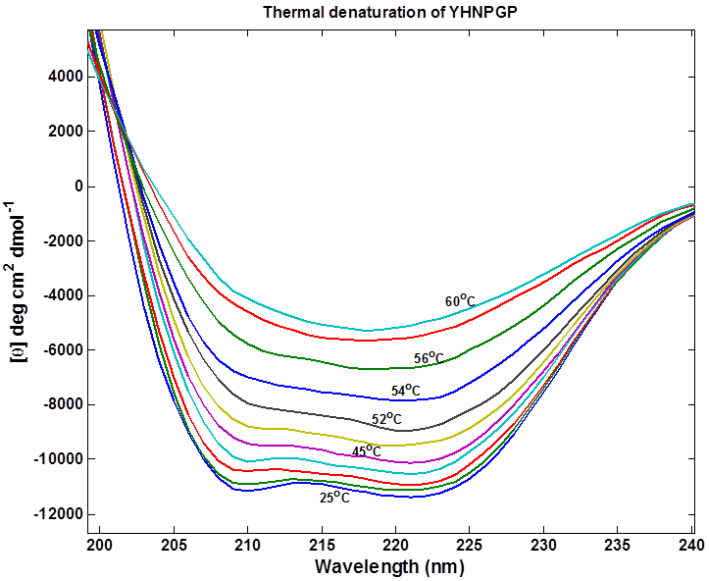


Figure 3.86. Changes in CD spectra of YHNPGP (Y291H) after FR-irradiation collected as a function of temperature, ramp rate of 0.25 °C/minute and 5 minutes of equilibration time before measurement.

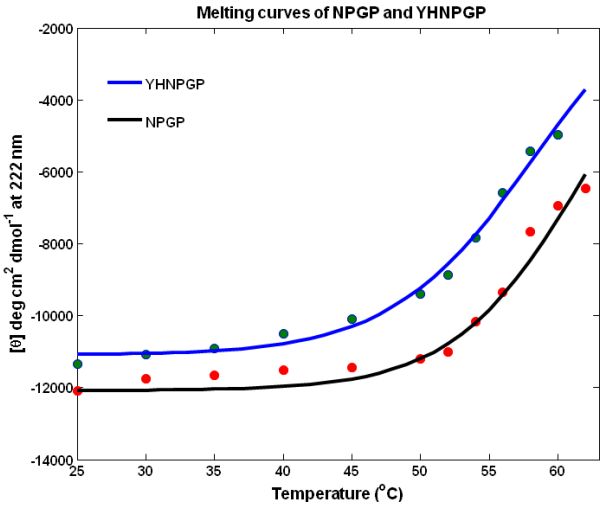


Figure 3.87. Changes in ellipticity of NPGP (WT) and YHNPGP (Y291H) at 222 nm plotted as a function of temperature (°C). The samples were held for 5 minutes at the measuring temperature and ramp rate was 0.25 °C/minute.

3.5.3.4 SEC

In SEC Y291HNPGP eluted earlier than NPGP by 3 ml and Y291HNPGP peak was broader than NPGP (Fig. 3.88). Repeated runs on this column showed elution volumes were reproducible with ± 3 ml uncertainty; hence this small shift might not be significant. The unusually broad, asymmetrical peak for Y291HNPGP on SEC suggests heterogeneity.

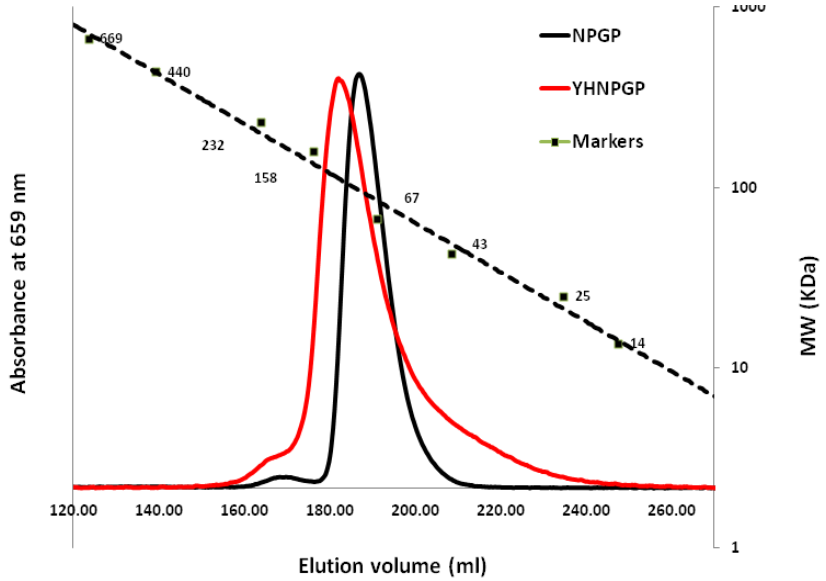


Figure 3.88. SEC elution profiles of YHNPGP (Y291H) and NPGP after FR-irradiation in HES-β using Superdex 200 26/60 preparative column at 1 ml/min flow rate monitored at 659 nm. The secondary axis represents the calibration curve with molecular weight standards: thyroglobulin (669), ferritin (440), catalase (232), aldolase (158), albumin (67), ovalbumin (43), chymotrypsinogen A (25) and RNAse A (14) in kDa.

3.5.3.5 Limited proteolysis

Putative conformational heterogeneity in Y291HNPGP was further examined by comparing limited proteolysis of Y291HNPGP and NPGP after saturating FR-light irradiation in the dark. Proteolysis experiments firmly showed a significant difference in cleavage patterns between Y291HNPGP and PGP (Fig. 3.89). From comparison of the number of proteolytic fragments, it can be seen that Y291HNPGP was more susceptible to cleavage than NPGP. CD data ruled out backbone disorder in Y291H mutants (Fig. 3.85). Pronounced proteolytic susceptibility apparent for the Y291HNPGP mutant (Figure 3.89) might thus be due to conformational heterogeneity. Taken together, broad band on SEC, susceptibility to proteolysis and thermal denaturation CD data implied that the YHB mutant phytochromes exhibited conformational heterogeneity.

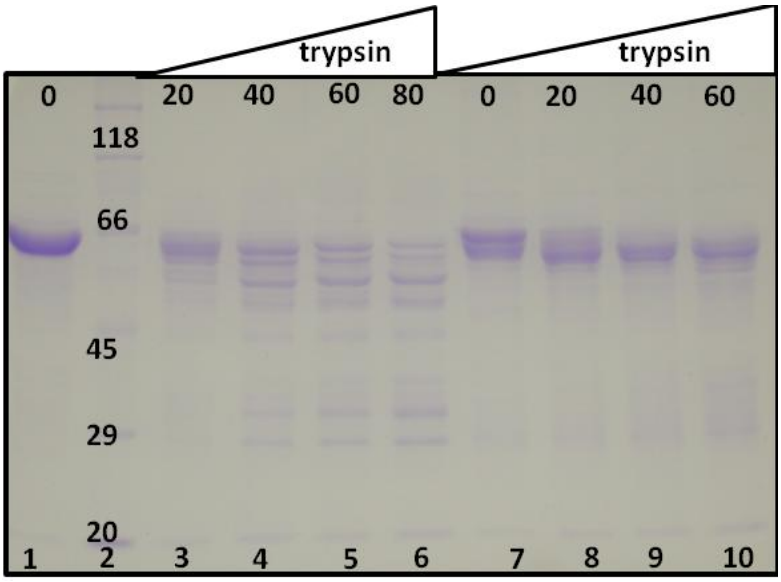


Figure 3.89. Comparison of protease susceptibility of Y291HNPGP (Lane 1, 3-6), 2 (markers in kDa) and NPGP (Lane 7-10) after saturating FR-light irradiation in 50 µL reaction containing 0.5 mg/ml substrates and different concentrations of trypsin in units ng labeled at the top. The labels at the bottom are lanes. 7 µL of the reaction mixture was loaded.

3.5.4 Initial crystallization screening

Purified Y291HNPGP, Y291HPGP and Y291HPG proteins were subjected to crystallization screening at 30 mg/ml and 15 mg/ml with 16 commercially-available crystallization kits from Qiagen, Jena Bioscience and Molecular Dimensions (Ralf Poeschke, AG Essen). The plates were incubated at 18 °C in darkness, and were regularly inspected. Crystallization was also attempted under conditions which had been successfully used for the wild-type proteins but none was observed. Many trials showed heavy precipitate indicating poor protein stability. Despite these extensive efforts, no lead crystallization condition was identified implying this particular mutation might not be suitable for crystallization studies.

3.6 Phytochrome interacting factors

Phytochromes regulate changes in gene expression in part by physically interacting with nuclear-localized phytochrome-interacting factors (PIFs) [165]. Because these interactions are determined by the structure of the phy-PIF binding interface as well as subsequent conformational changes, biophysical studies of these interactions are important for better understanding of the biochemical mechanism of signal transfer from photoactivated phytochromes.

3.6.1 PIFs in *Sorghum*

No PIF from *Sorghum bicolor* has been molecularly characterized. Moreover, the orthology of the group could not be established with certainty using phylogenetic methods [166]. Nevertheless, the two sequence signatures of At.PIFs were used to identify a set of putative Sb.PIFs from bHLH transcription factors in *Sorghum*. PIFs contain a signature bHLH DNA binding region [30, 100] and an active phyB (APB) binding region (see Fig. 3.90) [135].

3.6.1.1 Target selection

Using the seven well-characterized PIFs from *Arabidopsis thaliana* as querying probes, BLAST searches were performed for homologous predicted proteins in *Sorghum*. Six putative PIF sequences were identified containing the distinguishing bHLH and APB regions (see Section 10.2.4 for sequences). One of the identified bHLH transcription factors (Sb09g003090) had no conserved APB site. Upon closer examination, the corresponding cDNA showed an additional upstream start codon which would provide for production of the missing APB. BLAST searches using PIF3 from *Arabidopsis* as a query identified Sb.PIF1 from *Sorghum* EST database of GenBank. It failed to hit Sb.PIF3 (XP_002459338.1) perhaps due to the misannotated APB region. The amino acids sequences outside the core bHLH region of the putative Sb.PIFs are very divergent. Due to relatively high sequence similarity between At.PIF3 and one of the putative PIFs in *Sorghum* other than the misannotated Sb.PIF3, the next similar PIF to At.PIF3, Sb.PIF1, was selected for further characterization.

Multiple sequence alignment was performed by taking the first 110 amino acids of the APB-containing bHLH transcription factors from representative monocots and dicots (Figure 3.90). This strategy was employed to assign a simple 'orthology' for the different putative PIFs in *Sorghum* (Section 10.2.4). The predicted *Sorghum* PIF family contains at least 6 members. These members were numbered according to the heterologous nomenclature of *Arabidopsis* PIFs. Putative PIFs in *Sorghum* include Sb.PIF1 (XP_002466767.1), Sb.PIF3 (XP_002459338.1), Sb.PIF4 (KXG32305.1), Sb.PIF5 (XP_002466350.1), Sb.PIF6 (XP_002439246.1) and Sb.PIF7 (XP_002464941.1). Based on the highly-conserved bHLH regions (Figure 3.90) from diverse groups of representative PIFs given in Section 10.2.3, a putative phylogenetic tree was constructed using the www.phylogeny.fr web server in advanced mode (Figure 3.91).

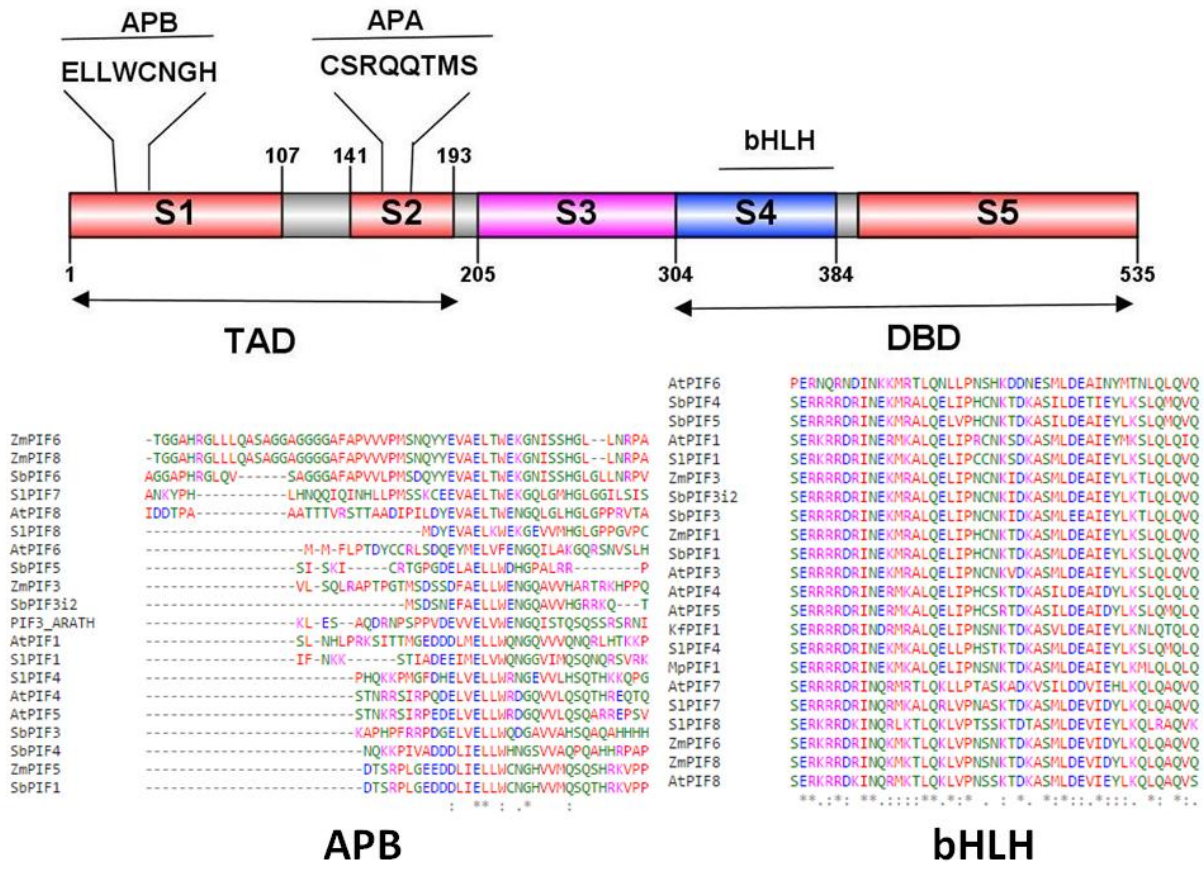


Figure 3.90. Domain structure and sequence alignment of the PIF family. The full-length Sb.PIF1 comprises a putative transactivation domain (TAD) and a putative DNA binding domain (DBD) connected by a long segment S3 (functional annotations are based on *Arabidopsis* PIF3). Sequence alignments of the conserved S1 (APB) and S4 (bHLH) motifs in the TAD and DBD domains respectively are shown. Homologous regions from *Arabidopsis* (At), Tomato (Sl), *Sorghum* (Sb), Maize (Zm), *Marchantia* (Mp) and *Klebsormidium* (Kf) were used for the alignments. The predicted PIF family in *Sorghum* contains at least 6 members containing sequence homologies to the APB motif of *Arabidopsis.*

3.6.1.2 Domain organization of Sb.PIF1

Secondary structure prediction, functional annotations in *Arabidopsis* PIFs and disorder prediction were used to construct preliminary domain architecture for Sb.PIF1. Sb.PIF1 polypeptide was divided into five segments designated S1 to S5 from N to C-terminus (Figs. 3.90). APB [128] and APA [167] are located on S1 and S2 respectively. S4 contains the bHLH region which is predicted to bind DNA [100]. S3 and S5 have no known functions. Sb.PIF1 is probably an intrinsically disordered protein with 74 % of its residues predicted to have no secondary structure (Figure 3.92).

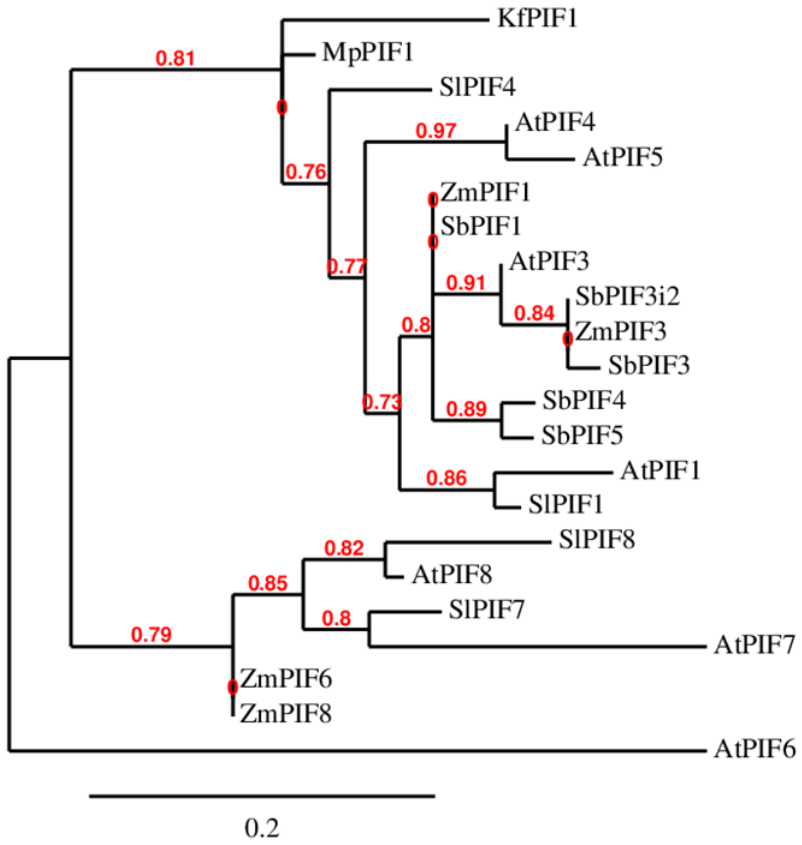


Figure 3.91. Phylogenetic tree of PIF homologs in different plant species. PIF homologs were identified by protein BLAST. Bootstrap values are given at the branches. Scale bar represents protein distance as substitutions per site. *At* (*Arabidopsis thaliana*), *Zm* (*Zea mays*), *Sb* (*Sorghum bicolor*), *Sl* (*Solanum lycopedium*), *Mp* (*Marchantia polymorpha*) and *Kf* (*Klebsormidium flacidum*). Accession numbers together with sequence information are given in Section 10.2.3.

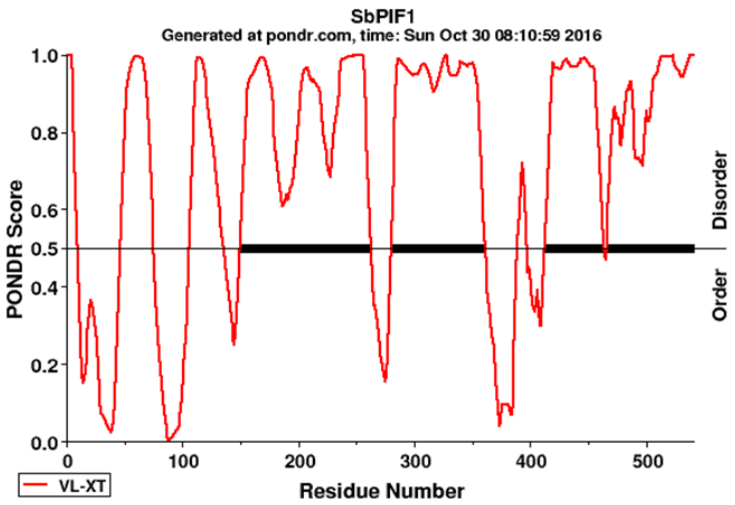


Figure 3.92. Intrinsically disordered regions in Sb.PIF1 predicted by the PONDR-FIT VL-XT algorithm.

3.6.1.3 Construct design for *Sb.PIF1*

Based on the preliminary domain architecture, four different inserts were designed *in silico* for PCR from the synthetic *Sb.PIF1* gene. The smallest construct encoded only the first 107 (APB, N107) residues in S1. A medium construct (LEA, N303) and a third construct containing the bHLH region (ALB, N387) were also included for comparison with the full-length Sb.PIF1 (Figs 3.93 & 3.94). The synthetic gene encoding the full-length Sb.PIF1 and other truncated constructs were first cloned into MCS2 of pCDFDuet with N-terminal 6x-His tag (Figure 3.95). The 6x-His-tag was then switched to the C-terminus of the full-length Sb.PIF1 to improve purification.

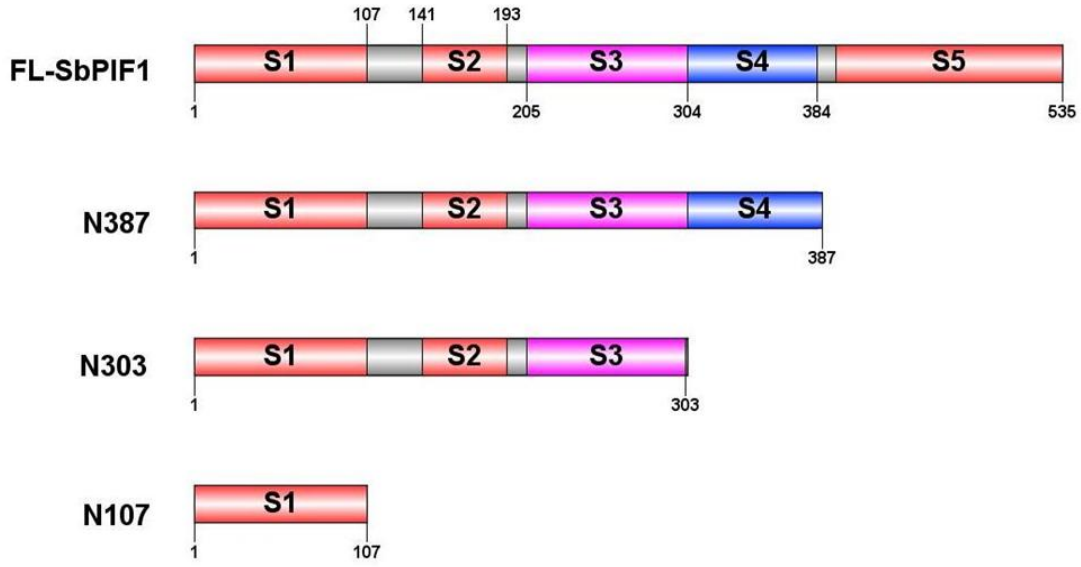


Figure 3.93. Multiple constructs designed for recombinant expression of Sb.PIF1. These constructs were designated by the number of residues in the N-terminal end. N107 (APB) is the minimal domain. N303 (LEA) and N387 (ALB) are without and with the bHLH region respectively. Functional evidences are from *Arabidopsis* by similarity.

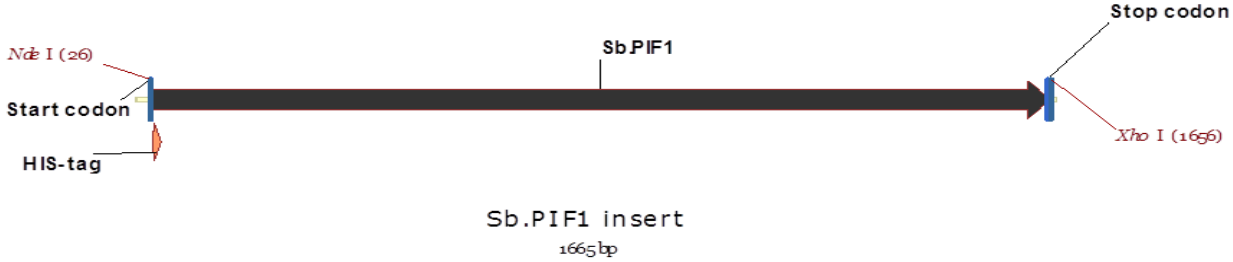


Figure 3.94. Schematics representation of *Sorghum PIF1*. The synthetic gene has a start codon, N-terminal His-tag coding sequence and a stop codon. NdeI and XhoI sites were used for cloning into MCS2 of pCDFDuet vector.

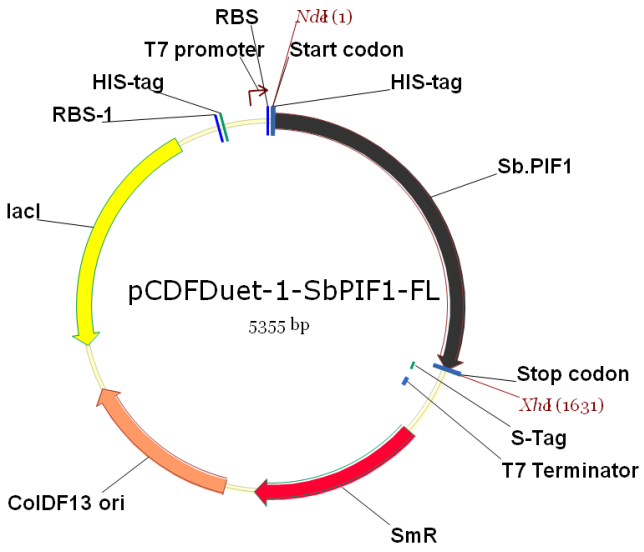
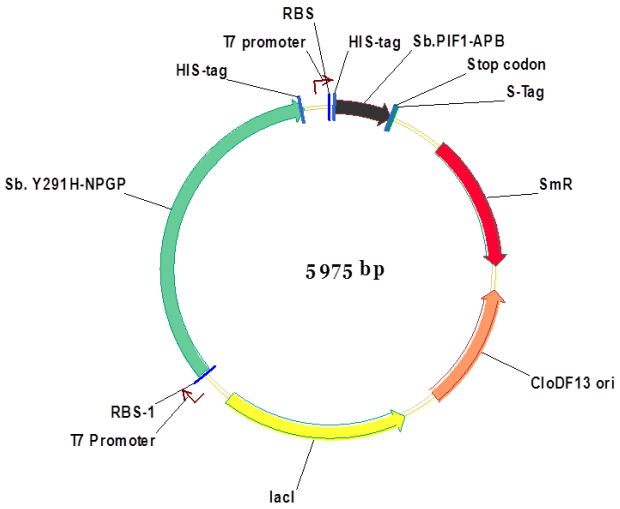


Figure 3.95. The pCDFDuet-*Sb.PIF1* plasmid for the production of full-length PIF1 from *Sorghum bicolor* with an N-terminal His-tag for purification.

3.6.1.4 Cloning *Sb.PIF1* for co-expression

The pCDFDuet plasmid has two multiple cloning sites, allowing simultaneous expression of two proteins. A dominant GAF domain tyrosine mutation (Y276H) of *Arabidopsis* phyB was shown to interact with At.PIF6 independently of light [63]. For co-expression of the analogous Sb.phyB mutant (Y291H) and its putative signaling partner (Sb.PIF1), the recombinant clone containing the Y291H mutant generated in Section 3.5 was linearized at MCS2 with NdeI and XhoI and the different PCR-amplified inserts derived from *Sb.PIF1* (*N107, N303, N387* and *Sb.PIF1*) were cloned accordingly. Thereby various clones were generated containing the presumably constitutively-active phytochrome mutants (*Y291HNPGP* or *Y291HPGP*) in MCS1 and different segments of *Sb.PIF1* (APB, LEA, ALB) in MCS2 for production of the Y291HNPGP-PIF complex. The integrity of the reading frames for all *Sb.PIF1* inserts were verified by sequencing prior to transformation into expression host. The clones with double inserts were labeled as pCDFDuet-*Y291HNPGP-N107Sb.PIF1* (Fig. 3.96), *Y291HPGP-N107-Sb.PIF1* and so on.



pCDFDuet-Sb.phyB-YHNPGP-Sb.PIF1-APB

Figure 3.96. The pCDFDuet-*Sb.Y291HNPGP-Sb.PIF1-APB* plasmid for holophytochrome and APB production in combination with the PΦB synthesizing plasmid. *Y291HNPGP* was inserted into MCS1 and *APB* of *Sb.PIF1* was inserted into MCS2 of pCDFDuet for co-expression.

3.6.1.5 Expression of *Sb.PIF1*

Verified recombinant clones were transformed into heat-shock-competent BL21(DE3) cells. Expressions of the different *Sb.PIF1* constructs were compared in LB and SB media at 37 ^{o}C with 1 mM IPTG. None of the constructs could be expressed in LB medium. On the other hand, the full-length Sb.PIF1 and N304 constructs were produced in SB medium (Figs. 3.97 & 3.98). No production of N107 could be established. The full-length Sb.PIF1 accounted for more than 85 % of the total protein when produced as insoluble inclusion body under these conditions (see Figure 3.97). The expected molecular weight of Sb.PIF1 is 56 kDa. It showed anomalous migration on SDS-PAGE near 64 kDa, a property of intrinsically disordered proteins.

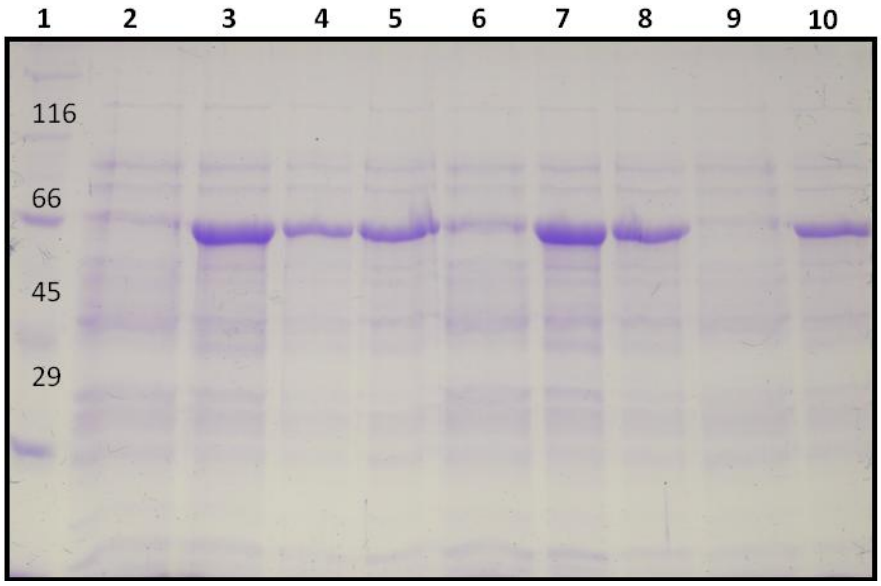


Figure 3.97. SDS-PAGE gel showing Coomassie-stained proteins for test expression of the full-length Sb.PIF1 from pCDFDuet-*Sb.PIF1* construct. Six different colonies were induced with 1 mM IPTG at 37 ^{o}C in SB medium and the level of expression was compared with uninduced negative controls. Molecular weight markers in kDa (1), uninduced negative controls (2, 6, 9) and induced clones (3-5, 7, 8 & 10).

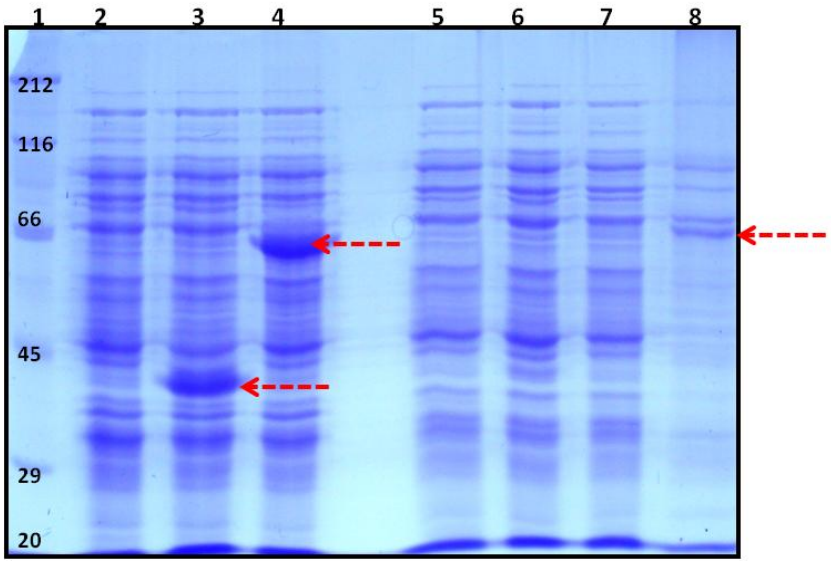


Figure 3.98. SDS-PAGE gel showing Coomassie-stained proteins for test expression and solubility of different constructs of Sb.PIF1 from MCS2 of pCDFDuet. Molecular markers in kDa (1), total cell protein from N107 (2), N304 (3), full-length Sb.PIF1 (4) and soluble fractions of N107 (5), N304 (6) and full-length Sb.PIF1 (7). A partially purified Sb.PIF1 was used as a positive control (8). Red arrows indicate recombinant proteins.

3.6.1.6 Co-expression of YHB-PIF complex

The pCDFDuet co-expression system was used to produce Sb.Y291HNPGP and Sb.PIF1 as complex in *E. coli*. Both Y291H mutant of Sb.phyB holoproteins and PIF1 were successfully produced within the same *E. coli* host (Figure 3.99). However, only limited improvement in the yield of soluble Sb.PIF1 was observed upon co-expression compared to expression of Sb.PIF1 alone. One unexpected observation during co-expression experiments was the presence of two separate fluorescent bands when SDS-PAGE gels were examined under ultraviolet light in the presence of zinc acetate (Figure 3.99). Fluorescence from PGP was accompanied by a second band corresponding to Sb.PIF1. The latter was, however, much weaker, despite the stronger Coomassie-stained band.

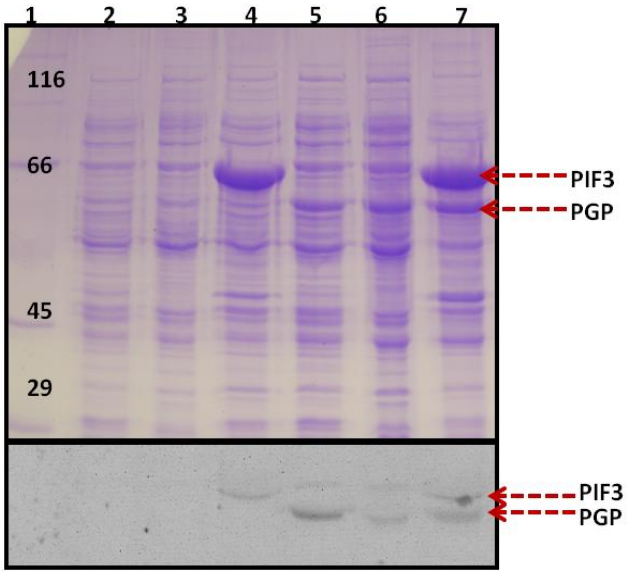


Figure 3.99. SDS-PAGE gel showing co-expressed genes as detected by Coomassie stain (top) and zinc-induced fluorescence (bottom). pCDFDuet harboring the genes encoding Y291H mutant of Sb.phyB in MCS1 and Sb.PIF1 in MCS2 was co-expressed with p183 harboring the genes for PϕB biosynthesis. Molecular weight markers in kDa (1), soluble fractions from two independent transformations (2 & 3) and insoluble fraction (4) for Y291HNPGP-Sb.PIF1 co-expression. Likewise, soluble (5 & 6) and insoluble fractions (7) were compared for Y291HPGP-Sb.PIF1 co-expression. Red arrows indicate recombinant proteins.

3.6.1.7 Optimized production protocol for Sb.PIF1 in *E. coli*

Soluble Sb.PIF1 could be expressed with reasonable yield (1.5 mg l^{-1}) from BL21(DE3) cells starting from overnight cultures. Cells were grown in rich medium such as SB until induction OD reached 0.5, whereupon cells were cooled on ice for 0.5 h and induced with 0.3 mM IPTG at 18 ^{o}C.

3.6.2 Purification of *Sorghum* PIF1

After induction of expression, a reasonable soluble yield was established both for solo and co-expression protocols. Protein solubility after centrifugation of the lysate was used as a predictor for properly folded recombinant protein production. As described in the following sections, this assumption can be misleading. The minimal fragment of Sb.PIF1 (N107) encompassing the putative APB region could not be expressed successfully.

3.6.2.1 Cell lysis

Frozen cell pellets were lysed as described in the materials and methods section (Section 2.2.2.6). After the crude lysate was incubated on ice for 30 minutes and the supernatant was clarified by centrifugation, potassium ferricyanide (PFC) was added to final concentration of 5 mM. The intensity of zinc-induced fluorescence was high under oxidizing conditions (Fig. 3.100). PFC was added to maintain strongly oxidizing condition necessary for the observed posttranslational modification. Unorthodox mixture of β-ME and PFC used in the binding buffer represented a compromise solution which significantly improved the purity of Sb.PIF1 while maintaining its modification by oxidation to yet unknown ligand (Figure 3.101). The crude *E. coli* lysate was clarified by filtration with 0.22 µm pore filter membranes. AmS precipitation was avoided due to incompatibility of AmS with triton X-100.

3.6.2.2 Purification of Sb.PIF1 by Ni-NTA

Sb.PIF1 showed very poor recovery when the lysis buffer, with 0.1 % triton X-100, was used as binding buffer for Ni-NTA. On the other hand, binding was improved when the concentration of triton X-100 in the binding buffer was increased to 1 %, suggesting poor binding caused by soluble aggregation.

3.6.2.3 Optimized purification protocol for Sb.PIF1

The clarified-lysate was purified by Ni-NTA column equilibrated with the lysis buffer. The bound protein was washed with two steps of 15 and 25 mM imidazole, two column volumes each. The bound PIF1 was then eluted in 100 mM imidzaole. The purification protocol showed multiple contaminant proteins when purification was done under oxidizing condition (Figure 3.100). A hybrid purification protocol was developed in which extraction was done under reducing condition but purification was done under oxidizing condition (Figure 3.101). When the purified full-length Sb.PIF1 fractions eluted from the Ni-NTA column, the protein solution was amber-colored (Figure 3.102). The pooled fractions from the hybrid protocol were dialyzed into TES-β for SEC.

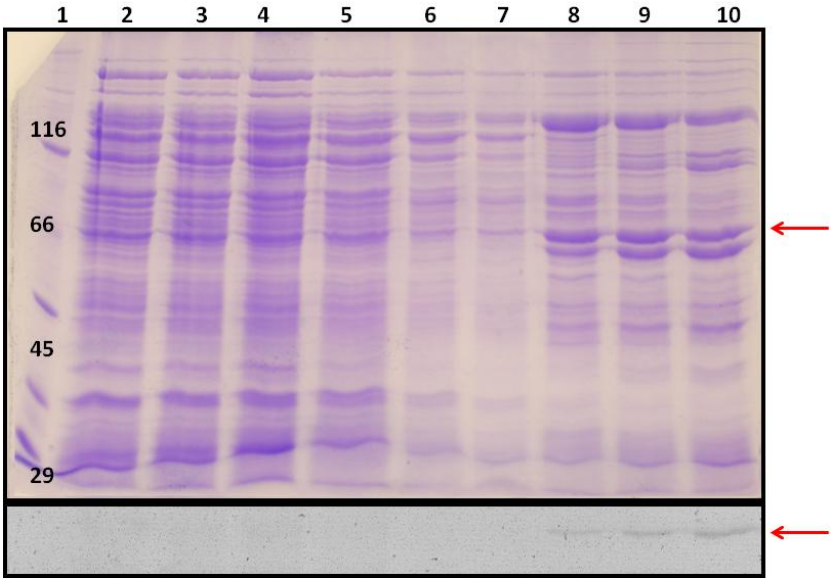


Figure 3.100. Extraction and purification of Sb.PIF1 under oxidizing condition. Coomassie stain (top) and zinc-induced fluorescence (bottom). Markers in kDa (1), input (2), flow-through (3-4), washing fractions (5-8) and elution fractions (9-10). Red arrows indicate the band presumably due to post translational modification.

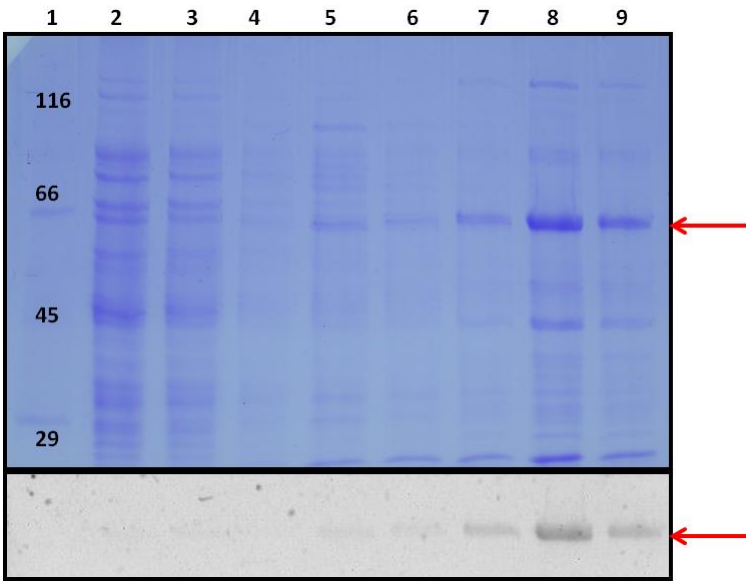


Figure 3.101. Extraction of Sb.PIF1 under reducing condition and purification under oxidizing condition. Coomassie stain (top) and zinc-induced fluorescence (bottom). Markers in kDa (1), flow-through (2), washing fractions (4-6) and elution fractions (7-9). Red arrows indicate Sb.PIF1 with the putative cofactor.

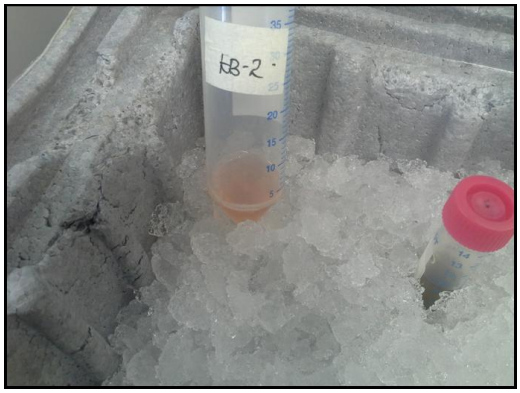


Figure 3.102. Amber-colored elution fraction of Sb.PIF1 purified by the hybrid purification protocol due to hitherto uncharacterized cofactor or contamination.

3.6.2.4 SEC of Sb.PIF1

Affinity purified Sb.PIF1 was applied to a Superdex 200 preparative SEC column equilibrated in TES-β. Most of the sample eluted in the void volume, suggesting soluble protein aggregation (Figure 3.103). Solubility of Sb.PIF1 was not an all-or-none phenomenon; there was very small but detectable peak of Sb.PIF1 near 67 kDa, presumably a monomer. Interestingly the copurified blue-absorbing cofactor remained associated with the Sb.PIF1 during SEC (Figure 3.103).

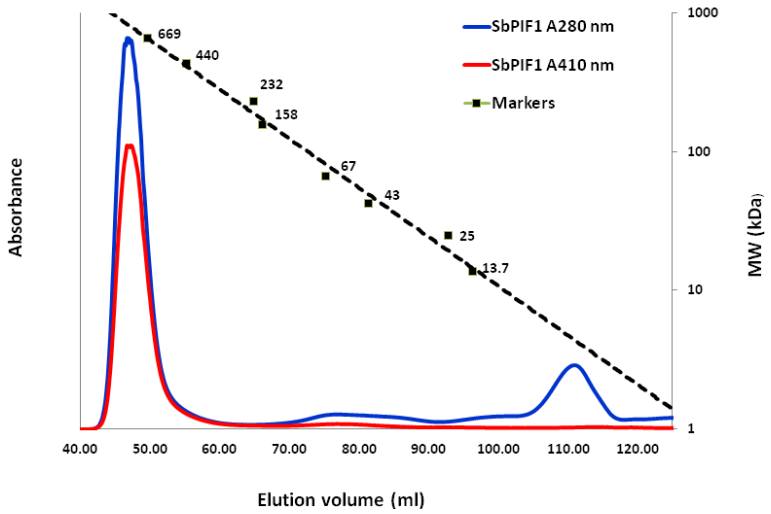


Figure 3.103. SEC elution profile of Sb.PIF1in TES-β using Superdex 200 16/60 preparative column at 1 ml/min flow rate monitored at 410 nm and 280 nm. The secondary axis represents the calibration curve with molecular weight standards: thyroglobulin (669), ferritin (440), catalase (232), aldolase (158), albumin (67), ovalbumin (43), chymotrypsinogen A (25) and RNAse A (13.7) in kDa.

3.6.2.5 Characterization of Sb.PIF1

The absorption spectrum of the SEC-purified Sb.PIF1 showed a peak at ~413 nm (Figure 3.104) corresponding to the amber coloration after Ni-NTA previously observed (Fig. 3.102) and the unexpected zinc-induced fluorescence after SDS-PAGE (Fig. 3.99). Such a peak is typical of absorption by heme proteins.

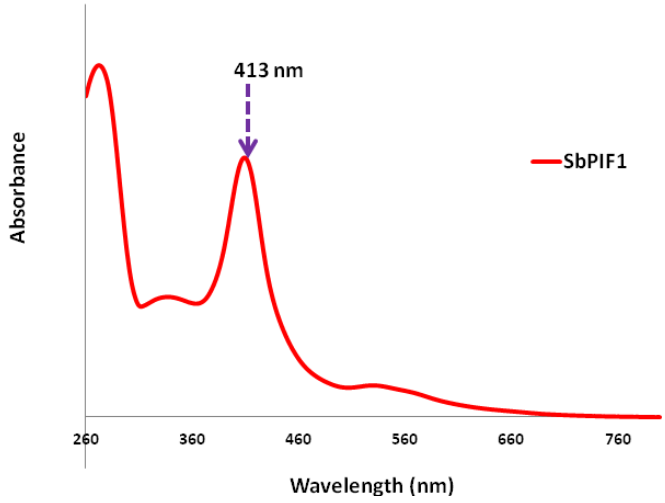


Figure 3.104. UV-VIS absorption spectrum of Sb.PIF1 after SEC measured in TES-β at 25 ^{o}C.

MALDI-TOF-MS analysis of tryptic fragments of Sb.PIF1 detected numerous peptides from Sb.PIF1 but none corresponding to the APB site (Figure 3.105). This might result from fragments with a very different m/z ratio from that predicted for the apoprotein, perhaps due to a covalently-linked prosthetic group. Clearly, however, other interpretations are at least as likely (data collected and analyzed by Professor Guenter Lochnit).

Figure 3.105. MALDI-TOF detected fragments of unmodified Sb.PIF1 (A) and modified Sb.PIF1 (B). Prepared apoprotein under reducing condition and holoprotein under hybrid condition were separated by SDS-PAGE. Coomassie stained bands were cut for MS analysis. The modified band showing zinc-induced fluorescent was compared with unmodified band with no zinc-induced fluorescence.

Secondary structure predictions of Sb.PIF1 gave five helical regions, a couple of β-strands and predominantly disordered segments. CD was used to estimate the content of secondary structure from recombinant Sb.PIF1 preparations (Figure 3.106). From this spectrum the K2D2 secondary structure server [168] predicted 11 % α-helix, 34 % β-strand and 55 % unfolded region. On the other hand, attempt to express and purify the smallest construct which was very interesting for co-crystallization trial was unsuccessful.

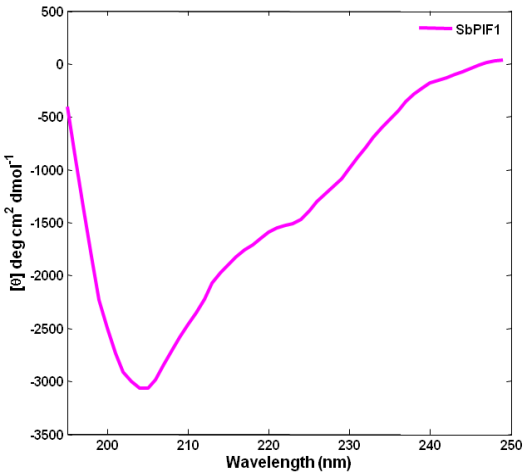


Figure 3.106. The CD spectrum of Sb.PIF1 purified by Ni-NTA and SEC in 10 mM K-phosphate, pH 7.8 containing 50 mM ammonium sulfate. Spectra were recorded in 0.5 mm path length cell at 25 ^{o}C.

SEC might give indications of the quaternary structure and aggregation behavior of proteins and protein complexes unless they are highly asymmetrical or interact directly with the column matrix [169]. A mixture containing molar excess of Sb.PIF1 over Y291HNPGP was incubated overnight at 4 ^{o}C and applied to SEC column. No evidence of interaction was obtained by comparing these data (Figs. 3.107-3.109) with Figures 3.88 (Sb.Y291HNPGP) and 3.103 (Sb.PIF1) for the separate proteins.

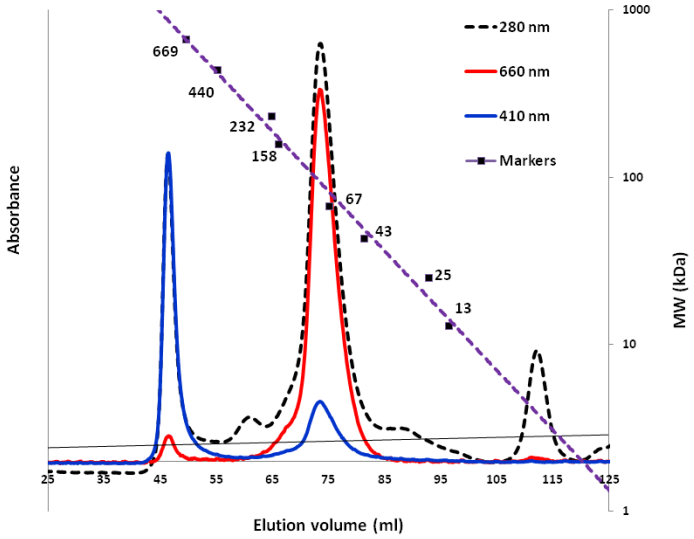


Figure 3.107. SEC elution profile of Sb.PIF1 and Y291HNPGP mixture inTES-β using Superdex 200 16/60 preparative column at 1 ml/min flow rate monitored at 280, 410 and 660 nm. The secondary axis represents the calibration curve with molecular weight standards (kDa).

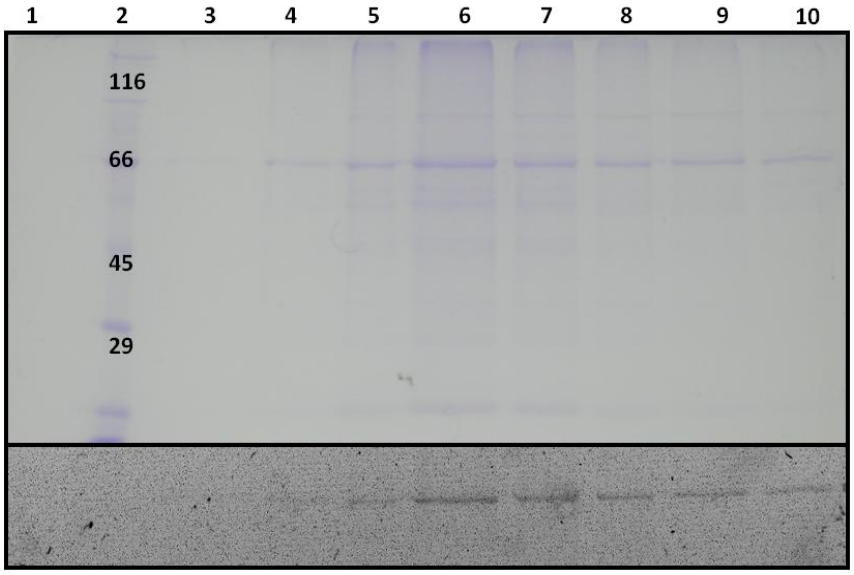


Figure 3.108. SDS-PAGE of the first peak of SEC containing resolved Sb.PIF1 in TES-β from the mixture of Y291HNPGP and Sb.PIF1. A1 (1), Markers in kDa (2), A2 to A9 (3-10) detected by Coomassie stain (top) and zinc-induced fluorescence (bottom).

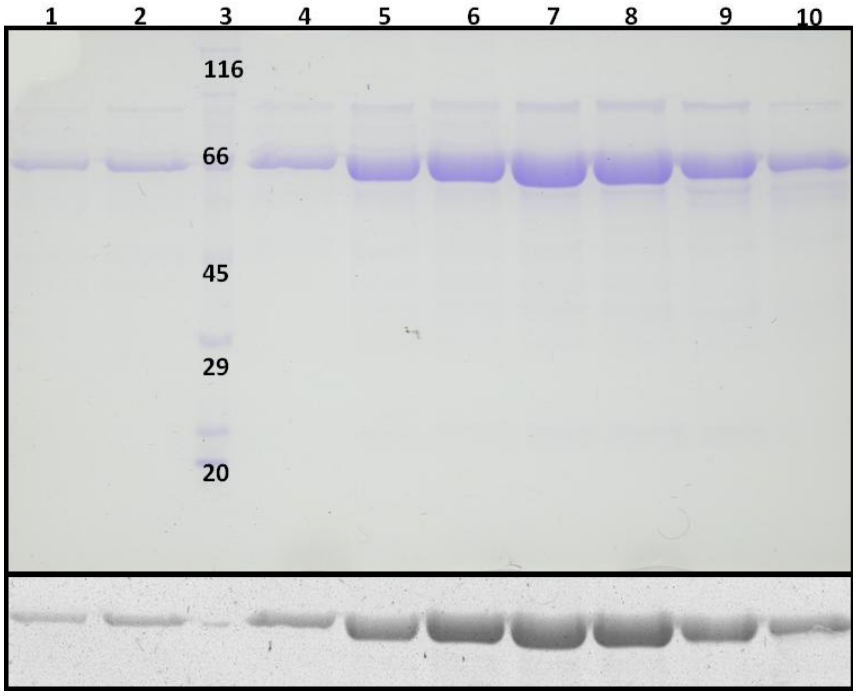


Figure 3.109. SDS-PAGE of the second peak from SEC in TES-β containing resolved Y291HNPGP from the mixture of Y291HNPGP and Sb.PIF1. All fractions come in non-void peak B7 (1), B5 (2), markers in kDa (3), B3 (4), B1 (5) and C1 to C6 (6-10) as detected by Coomassie stain (top) and zinc-induced fluorescence (bottom).

4 Discussion

4.1 Prokaryotic to plant phytochromes

The discovery and development of prokaryotic phytochromes as models for plant phytochromes lead to tremendous progress in understanding the primary phytochrome photochemistry and light-dependent conformational changes. Phytochrome photochemistry involves a Z-to-E isomerization about the C15=C16 double bond of the bilin chromophore [21]. Isomerisation of the chromophore is accompanied by conformational changes in the photosensory module. The Trp-swap model is a simple and elegant model that extends the ligand switching phenomena observed in hemeproteins to conformational changes in phytochrome photoreceptors [45]. The remarkable predictions made by Trp-swap model were later confirmed by other ground breaking works in prokaryotic phytochromes [46, 67]. Conformational changes during photoconversion involve the tongue region of the PHY domain and deformation of the helical spine. However, the utility of Trp-swap model for plant phytochromes needs to be carefully examined.

The major differences between the photosensory modules of plant phytochromes and that of cyanobacterial phytochromes include the presence of Ser-rich NTE, the use of PφB in place of PCB, and long flexible loops. The conserved WxE motif in the tongue of the PHY domain of cyanobacterial phytochromes is substituted for FxE motif in plant phytochromes. While such aromatic substitutions might have very little effect on protein structure, the effect on conformational dynamics and relative stability of Pr and Pfr should be examined. Neither the PHY domain nor the helical spine is indispensable for light signaling in plant phytochromes [114]. Further research is required to understand conformational changes at the PAS-GAF bidomain that make the 'docking site' accessible for PIF binding in conformer specific manner.

4.2 Dicot to monocot phytochromes

Arabidopsis thaliana has been the major plant model system to study light signaling networks in plants [72]. The 3.4 Å crystal structure of the truncated photosensory module of phyB from *Arabidopsis thaliana* showed strong structural homology to cyanobacterial phytochromes. However, the Ser-rich NTE was truncated and the PHY domain, the PφB chromophore as well as the long flexible loops unique to plant phytochromes were barely resolved. There is currently no

structural information on phyB from monocotyledon plants. Nearly 150 million years of evolutionary divergence between dicots and monocots limits direct transfer of knowledge [170].

4.2.1 Production and purification of Sb.phyB constructs

Protein crystallography to investigate phytochrome function in *Sorghum* supplements structural studies of phyB from *Arabidopsis thaliana*. The use of multiple constructs might increase the probability of generating diffraction quality crystals. To this end, multiple constructs derived from the photosensory module of Sb.phyB were heterologously produced in *E. coli* as a polyhistidine tagged holoproteins assembled *in vivo* with either phytochromobilin (PΦB) or phycocyanobilin (PCB). Low complexity regions in proteins are often missing in electron density maps from X-ray crystallography experiments. PGP, without the flexible NTE, was designed to promote crystallization. Sometimes missing density may correspond to wobbly, ordered domains that are disordered by occupancy of different positions rather than to low sequence complexity. PG was designed to reduce the presumed domain wobbling in the PHY domain [68].

The pPROLar (Figure 2.1) and pCDFDuet (Figure 2.2) vectors were used for cloning and expression of Sb.phyB derived fragments. The product regioselectivity of *HY2* from dicotyledon and monocotyledon plants was assumed to be similar if not identical. Consequently, p183 plasmid which uses hemeoxygenase (*HO*) from *Synechocystis* and *HY2* from *Arabidopsis thaliana* was used to produce the native PΦB chromophore. p171 plasmid which uses hemeoxygenase (*HO1*) and PcyA from *Synechocystis* was also used to synthesize PCB for chromophore replacement. Colony-PCR screening is very fast method, but holophytochrome production would be efficient if co-expression screening is used from the start. The NTE seems to hamper production of recombinant NPGP in *E. coli,* as the truncated PGP was produced quite efficiently from the pPROLar vector, while NPGP was not (Figure 3.9). However, although NPGP was not produced from pPROLar, it was produced effectively from the pCDFDuet vector (Figure 3.15). The production of Sb.phyB constructs was not high enough to reveal fat bands after Coomassie staining when induced samples were run in SDS-PAGE (Figs 3.12 & 3.13). Nevertheless, working on phytochromes has analytical advantage due to the specific and sensitive spectrophotometric difference assay (Figure 3.10) developed by Butler et al. (1959) [16] and rapid detection by zinc-induced fluorescence after SDS-PAGE (Figs. 3.12 & 3.13) developed by Berkelman et al. (1986) [139]. The use of *in vivo* assembly strategy was valuable to benefit from these sensitive and specific analytical methods.

Sb.NPGP was successfully produced and purified from codon-optimized synthetic gene and pCDFDuet plasmid in *E. coli* strains bioengineered to produce the native phytochromobilin (PΦB) chromophore. Production and purification protocols were developed for three recombinant His-tagged constructs: NPGP, PGP and PG. These major constructs were effectively produced in *E. coli* grown in SB medium (Sections 3.1 & 3.2). Major improvements resulted from the use of the native chromophore, fresh transformation, catabolite repression and improved aeration. *In vivo* assembly with the native chromophore might improve the structural integrity and stability of the chromophore binding pocket. Recombinant phytochrome production decreased with aging probably due to copy number related issues. Catabolite repression was used to control basal expression from T7 promoter. The effect of the geometry of the baffled-flasks on recombinant phytochrome production is often overlooked. In shake flasks, transfer of heat is influenced by the geometry of the rotating bulk liquid (the flask inner wall). On the other hand, oxygen transfer is influenced by the surface area exposed to the surrounding air. Modifications of shake flasks by the introduction of baffles provide sufficient aeration. A very important factor to achieve reproducible production of NPGP was aeration (i.e. mass exchange surface area/culture volume). Aeration was improved by increasing the surface area/volume ratio of the culture in baffled flasks.

Cell lysis and purification protocols for NPGP and PGP were systematically optimized. Increasing salt concentration reduced non-specific binding to Ni-NTA column and improved the purity of PGP (Figure 3.17). Exploration of different buffering agents and optimization of various buffer additives led to the production of more than 95 % pure PGP as judged by SDS-PAGE (Figure 3.26). In SEC, both NPGP and PGP holoproteins purified by the original purification protocol eluted with one major peak, accompanied by a small high molecular-weight shoulder peak (Figs. 3.21 & 3.25). The stability and homogeneity of Sb.PGP was found to be far better in HEPES than in Tris buffer (Figures 3.22 & 3.27). The shoulder peak observed during preparative SEC of Sb.PGP by the original purification protocol was removed in the improved HEPES-based purification protocol (Figure 3.25). Both NPGP and PGP holoproteins purified by the improved purification protocol were eluted with a single peak, implying the photosensory module exist as monomer in solution (Figs. 3.25 & 3.27). The difference in mobility on analytical SEC between PGP purified by the original and improved purification protocols suggested that different purification protocols might enrich different phytochrome conformations (Figs. 3.21 & 3.26). Dimerisation and conformational changes that alter the global shape of the protein might be potential reasons for these mobility

differences. Alternatively, different interactions with the gel matrix might also affect mobility on SEC. In any case, SEC reveals mobility differences between Pr forms under different purification conditions implying potential heterogeneity in Pr state. However, the cause and chemical nature of this heterogeneity was not clearly defined in the present study.

NPGP, PGP and PG were successfully purified to homogeneity by a combination of improved affinity chromatography protocol and SEC (Fig. 3.52). The improved purification protocol helped to increase the yield and homogeneity of holophytochromes facilitating crystallization. However, it is possible the microheterogenity observed in the original purification protocol (Figure 3.22) might reappear during prolonged crystallization setup for phytochromes prepared by the improved protocol. Structural variations in phytochrome preparations due to known heterogeneities in Pr or unknown non-enzymatic modifications *in vitro* might lead to the observed microheterogeneity. Microheterogeneity in phytochromes is usually undetected when only photochemical properties are monitored during purification and is also difficult to detect when SDS-PAGE is used to determine purity. The sample conditions from purification to crystallization need to be carefully reassessed to identify what might lead to sample heterogeneity in the original purification protocol.

4.2.2 Characterization of Sb.phyB constructs

Sb.phyB derived wild-type constructs, the SER mutant and Y291H mutant phytochromes were successfully purified to homogeneity by a combination of improved affinity chromatography protocol and SEC (Fig. 3.52). There is no means for directly quantifying the purity of a protein sample and determination of protein concentration is subjected to misestimation. Moreover, an apparently pure preparation that is correctly estimated may harbor undetectable protease contaminant affecting its structural and functional integrity. There was no evidence for proteolytic cleavage for all major Sb.phyB derived constructs. The spectroscopic, photochemical and biophysical properties these constructs were compared with each other and with the constitutively active (Y291H) mutants taking contamination and concentration misestimation into considerations.

4.2.2.1 The photosensory module of Sb.phyB

The first 113 amino acid residues in NPGP constitute the NTE (Figure 3.4). It is 11 amino acids longer than the NTE in At.phyB [68]. To gain insight into the functional effects of the NTE in Sb.phyB, photochemical properties NPGP were compared with that of PGP. Deletion of the NTE caused 6 and 10 nm hypsochromic shift from the Pr and Pfr absorption bands respectively (Figure 3.28). NTE of Sb.phyB slowed down the rate of dark reversion of Pfr to Pr (Figure 3.30). The NTE was found to be important for the integrity and higher Pfr peak intensity of NPGP than PGP (Figure 3.29). Deletion of the NTE increased the melting temperature of Sb.PGP by ~5 ^{o}C compared to Sb.NPGP (Figure 3.37). The CD spectra of NPGP and PGP of Sb.phyB implied structural role of NTE in reorganization of the PAS-GAF-PHY domains into a less packed (Figure 3.35), more flexible and less thermally stable structure (Figure 3.36).

In HDX measurements, the NTE of Sb.phyB exhibited fast exchange similar to the NTE of At.phyB [172]. HDX experiments on Sb.NPGP showed not only protected regions of the NTE through potential interactions with PGP but also distinct regions showing light dependent variable exchange rate (Figure 3.40). Light dependent changes were particularly high for the region Y96-E109 (Figure 4.1); a very similar region to recently identified peptide in At.phyB [172]. Furthermore, part of the long helical spine in Sb.NPGP (L499 to L467) showed pronounced light dependent exchange rates (Figure 4.2). These findings were consistent with HDX results reported for At.phyB [172] and the signal transduction and amplification model for bacteriophytochromes [46, 67].

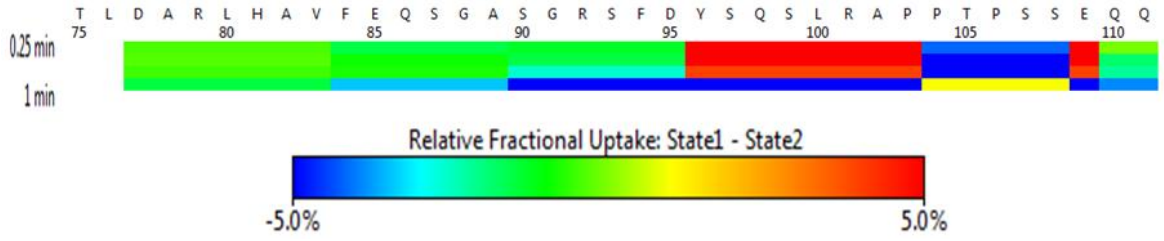


Figure 4.1. Deuterium uptake for the NTE of Sb.NPGP. A region (F94-E109) showed light-dependent changes whereas another region (R79-S90) showed light-independent reduced uptake. State 1 is after saturating FR-light irradiation and state 2 is after saturating R-light irradiation. The map is generated by the difference exchange rate for different peptides in the NTE as a function of time.

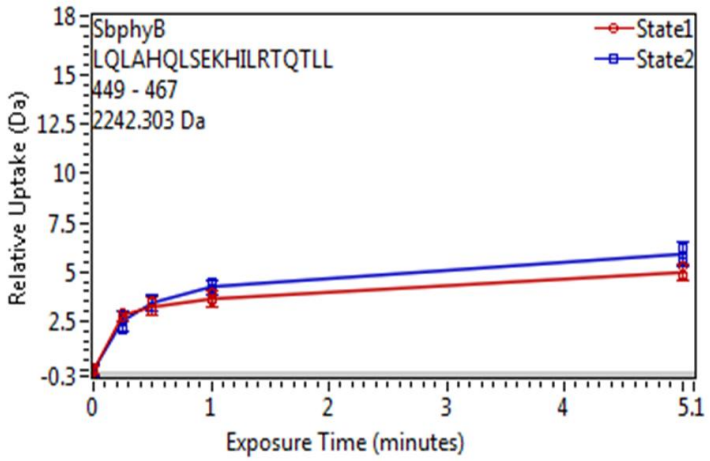


Figure 4.2. Relative deuterium uptake as a function of time for a region (L449-L467) in the connecting helix of Sb.NPGP. State 1 is after saturating FR-light irradiation and state 2 is after saturating red-light irradiation.

Based on differential H/D exchange rate between At.NPGP and At.PGP assembled with PCB, von Horsten et al. (2016) postulated multiple interaction sites for the NTE and PGP designated as slow exchanging loops (s-loops). According to the proposed packing model, the observed slow exchange for presumably exposed loops in PGP might be caused by interaction of these loops with the NTE. Particularly, NTE-s-loop1 interaction might block the binding of At.PIF3 in the Pr state [172]. HDX measurement for NPGP of Sb.phyB showed rather slow exchange for f-loop3 and much reduced uptake for some of the regions in the PHY domain that were not defined in At.phyB structure (Figure 3.40).

Based on CD measurements and surface cysteine labeling experiment, a shielding model is suggested to be appropriate to emphasize the effect of NTE on solvent accessibility of the chromophore, overall surface charge of the protein and intrinsic conformational heterogeneity of plant phytochromes. Several lines of evidences support a shielding function for the NTE. Burgie et al. (2014) showed the cap helix at the base of NTE interacts with the chromophore in Pr through hydrogen bond and van der Waals interactions involving Y104. This interaction shields the chromophore from solvent exposure [68]. HDX measurements for Sb.NPGP identified two regions of the NTE, R79-A89 and P103-S108, exhibiting significantly reduced uptake (Figure 4.1). The number of surface exposed cysteines accessible to Ellman's reagent increased from 1 to 3 upon removal of the NTE in PGP (Figure 3.46). Taken together, these experiments suggest shielding effect of the NTE on the rest of the photosensory module.

Plant phytochromes use PφB which interacts with the plant specific conserved Y104 residue in the cap helix [68]. Cph1, a Pr-active histidine-kinase, uses PCB [7]. The substituents on C-18 of PφB and PCB have different space filling and electronic properties. The vinyl group of PφB is relatively smaller and more polar than the ethyl group of PCB [173, 174]. Thus, the chemical nature of the chromophore and its local microenvironment appeared to be slightly changed while the bilin binding cysteine has been preserved over billions of years of evolution [60]. The different Sb.phyB derived holoproteins showed photoreversibility with both PCB and PφB chromophores (Fig. 3.32). Dark reversion kinetics of NPGP assembled with PφB was much faster than that with PCB *in vitro* (Figure 3.33). Hence, the true physicochemical properties of plant phytochromes assembled with PCB remain in doubt [172, 175]. The possible emergence of the NTE, hy2 and PIFs during light signaling networks in land plants needs to be examined.

Deletion of the NTE in At.phyB caused a large increase in the fluence rate of continuous red light (Rc) required for maximal seedling responsiveness [110]. This result indicated the NTE might have improved the quantum efficiency of photoconversion and the efficiency of light signal transduction through increased conformational flexibility for protein-protein interaction in land plants. Besides, additional level of regulation can be mediated by post-translational modification. Linear motifs in the NTE might be involved in weak and transient interactions with the chromophore and PGP forming hydrogen bonding and electrostatic interactions [176].

A prototypical plant phytochrome gene contains six exons and five introns. The photosensory module corresponds roughly to the major 5' exon. This pattern suggested the NTE, PAS, GAF and PHY domains might have arisen through the duplication of pre-existing exons [77, 171]. There is very weak sequence similarity between the NTE and PHY domain of Sb.phyB. The observed 23 % identity and 31 % similarity (Figure 4.3) suggested structural and functional similarities between the NTE and the PHY domain are not implausible. A unique structural feature of the PHY domain is the long unusually flexible tongue having periodically conserved aromatic residues (Y559, Y569 & W579) and a conserved motif (PRxSF) that seals the chromophore binding pocket [25]. The NTE seems to share this unusual flexible feature of the tongue having three periodically conserved aromatic amino acids (Y74, F84 & F94). The core PAS fold, however, might have been evolutionarily lost due to indels.

Figure 4.3. Pairwise sequence alignment tool was used to identify regions of similarity that may indicate functional, structural and/or evolutionary relationships between the NTE and PHY domain of Sb.phyB.

Canonical phytochromes in *Zygnematales* algae are very closely related to land plant phytochromes [130]. Examination of domain organization in the canonical phytochrome from *Zygnematales* green alga *Gonatozygon kinahani* (Figure 4.4) suggested features of a fusion protein. This putative fusion protein comprises an N-terminal 260 residues long protein of unknown function resembling the NTE of land plants and a C-terminal full-length phytochrome. This pattern suggested the NTE might have arisen as an N-terminal fusion protein to ancestral phytochrome photoreceptors.

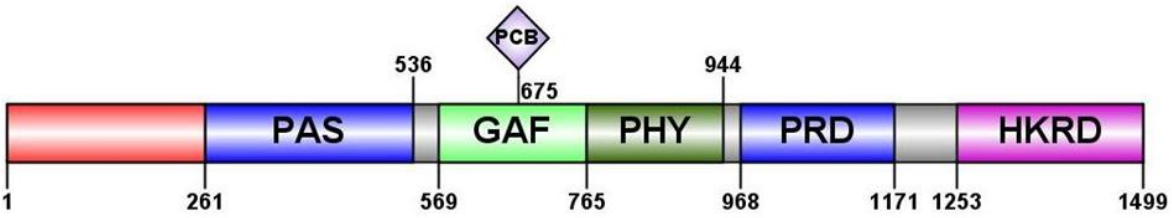


Figure 4.4. Unusual organization of *Gonatozygon* phytochrome. The N-terminal protein of unknown function resembling NTE of plant phytochromes seems to be fused to canonical phytochrome. The PAS domain in PSM is larger than that of plant phytochromes. The PSM of this phytochrome is followed by the PAS repeat domain (PRD) and histidine kinase related domain (HKRD).

4.2.2.2 The PAS-GAF bidomain

The exclusion of solvent water from the chromophore binding pocket is taught to be a key factor in ensuring efficient photochemistry [180]. Consequently, a functional PHY domain is absolutely required for the formation of the typical spectrum for Pr + Pfr (Figure 3.53). Two slow-exchanging loops of Sb.phyB exclusively localized in PG, s-loop1 (L116-G129) and s-loop2 (S200-Y217), showed state-dependent exchange rates in HDX measurements (Fig. 3.40). CD measurement in PG after R/FR irradiation revealed increased negative ellipticity upon photoconversion implying order-disorder transition (Figure 3.55). Light induced changes in CD signals for PG (Figure 3.55) were clearly different from PGP (Figure 3.34 B) implying conformational changes localized in the PHY domain for the latter. Taken together, these results indicated s-loop1, s-loop2, the tongue of the PHY domain [45] and the helical spine are fundamental elements of signal transduction and amplification in plant phytochromes [46].

4.2.2.3 The GAF-PHY bidomain

In canonical phytochromes, a nearly 30-residues insertion in the sequence of a prototypical GAF domain forms a lasso in which the N-terminus of the PAS domain passes to create a figure-eight polypeptide knot [25, 44]. The structural integrity of this knot is maintained by a combination of hydrogen bonding, hydrophobic and van der Walls interactions [44]. A PAS-less deletion mutant of Sb.phyB, designated GP, was cloned into pCDFDuet vector as alternative variant for crystallization. Test expression studies of the GP bidomain showed fat bands on SDS-PAGE when run adjacent to uninduced samples. However, Sb.GP failed to form a chromoprotein upon complete removal of the PAS domain (Figure 3.62) possibly due to misfolding and aggregation. Similarly, in attempts to generate a novel fluorescent protein label, although removal of the PAS domain from a truncated bacteriophytochrome from *R. palustris* (RpBphP1) resulted in a non-fluorescent holoprotein, removal of the PAS domain together with the knot lasso yielded fluorescent holoprotein [47].

Previous studies identified a short highly conserved sequence in the PAS domain designated here as ‘IQ’ motif based on 100 % conservation of the glutamine residue. Residues in the ‘IQ’ motif were mapped onto a mini β-strand (HLI: 24-26) and a conserved loop (QPH: 27-29) at the knot in the crystal structure of Cph1 [25]. The mini β-strand (GYI, 112-114) and the adjacent loop, s-loop1 (QPH: 115-117) were also conserved in At.phyB structure. A turn in Cph1 and the cap-helix in

At.phyB were found N-terminal to the conserved mini β-strand [68]. The cap helix (104-110) in At.phyB is particularly interesting as it was identified as the molecular interface for direct interaction with At.PIF3 [115]. Saturation mutagenesis of 'IQ' motif and the cap helix should prove useful in understanding how the biochemical role of these regions relates to the basic structure and function of plant phytochromes. The 'IQ' motif, present only in knotted phytochromes, might mediate the cross-talk between the PAS domain and the GAF domain during folding, chromophore binding and photoconversion.

4.2.2.4 The Y291H mutant of Sb.phyB

There is currently no structural information on the active signaling state of plant phytochrome. The constitutively active YHB mutant might provide a novel structural insight into this signaling state. To investigate the structural differences between the Pr state and the presumably constitutively active mutant, three different mutants (Y291HNPGP, Y291HPGP & Y291HPG) were produced and characterized. As expected all Y291H mutants exhibited impaired photochemistry (Figs. 3.73 & 3.75). Y291H mutants after FR-irradiation faithfully imitated the spectroscopic and circular dichroic properties of WT proteins after R-irradiation (Fig. 3.78 & 3.85). Unpredictably, the Y291HPGP mutant uniquely showed a small absorbance peak near 720 nm after saturating FR-light irradiation (Fig. 3.78). Absorption and excitation spectra were indistinguishable for all YH mutant proteins (Figs. 3.80 - 3.82 B). On the other hand, the excitation spectra were clearly broader than the absorption spectra for all wild-type constructs resolving chromophore conformational heterogeneity, otherwise hidden in the absorption spectra (Figs. 3.80 - 3.82 A). Y291H mutant derived from PG appeared to have lower fluorescent than other Y291H mutants generated (Figure 3.84). This was consistent with the result reported for similar construct of Cph1 albeit with PCB [61].

Preparative SEC in this study showed the peak for Y291HNPGP was broader than that of NPGP (Figure 3.88) suggesting protein conformational heterogeneity. High proteolytic susceptibility of Y291HNPGP mutant (Figure 3.89) implied extensive structural rearrangements in the dark adapted state. Temperature induced unfolding transition of Y291HNPGP showed very gradual changes between 25 ^{o}C and 50 ^{o}C (Figure 3.86) in CD spectra in contrast to closely spaced spectra observed for the wild-type (Figure 3.36 A). This could be due to structural rearrangement before thermal unfolding in Y291HNPGP. Taken together, SEC, limited proteolysis and temperature-

induced unfolding indicated phytochromes carrying the Y291H mutation exhibited conformational polymorphism. Conformational polymorphism is a prevalent phenomenon in many gain-of-function mutations [177] and apo-proteins such as apo-hemoglobin [178].

Fischer et al. (2005) performed saturation mutagenesis of the Y176 residue in Cph1. These mutations can be grouped into four classes. The first group is the group of Y176H and Y176Q mutations with R/B ratios of 2.5 compared to 3.3 for the wild-type [60, 61]. Su and Lagarias (2007) demonstrated the equivalent mutations in At.phyB (Y276H and Y276Q) exhibited COP phenotype [62]. Y291H mutant of Sb.NPGP showed R/B ratio of 2.5 consistent with the reports by Fischer et al. (2005). The second group is that of Y176S, Y176N and Y176T with $2 \geq R/B \geq 0.9$ [61]. The third group contains mutations like Y176I with R/B ratio of 0.5. These mutations have no COP phenotype but can complement phyB null mutant plants [61, 62]. Fischer et al. (2005) also showed group four mutations. The Y176R mutants of Cph1 and Y276R mutants of At.phyB bind endogenous porphyrins. phyB Y276R mutant *Arabidopsis* showed neither COP phenotype nor the ability to complement the phyB null mutant under red-light [61, 62].

Fischer et al. (2005) proposed the presence of heterogeneous chromophore conformations during substitutions of Y176 with amino acids which have both hydrogen bond donor and acceptor atoms in their side chains [61]. However, there was no evidence to support chromophore conformational heterogeneity. Conformational protein polymorphism inferred here might be due to different interaction of the protein with presumably identical chromophore conformation. Y291H mutants might assume multiple nearly isoenergetic conformational sub-states interacting differently with the same chromophore conformation. Volkman et al. (2001) proposed a population shift model to explain the functional selection of a particular conformer by protein-protein interaction in conformationally heterogeneous proteins like YHB [179].

4.2.3 Crystallization of Sb.phyB constructs

Based on At.PGP crystal structure solved by Burgie et al. (2014) and HDX map reported by von Horsten et al. (2016), the most variable regions in the PSM of plant phytochromes include the NTE (1-102), f-loop1 (142-155), f-loop2 (377-394) and the tongue of the PHY domain (555-600) [68, 172]. The difficulty of phytochrome crystallization seems to increase with loop length and heterogeneity in the Pr state. In this study, NPGP, PGP and PG holoproteins from Sb.phyB were successfully produced (Figure 3.15) and purified (Figure 3.52). PGP purified by the original

purification protocol showed heterogeneity (Fig. 3.22). Aggregates do occasionally rearrange to form crystals, but usually the aggregate is the final state of the system. This might explain the cause of crystallization problem from the original purification protocol (Fig. 3.41).

NPGP, PGP and PG purified by the improved purification protocol have been successfully crystallized. Crystals from each construct often diffract X-rays (Figs. 3.44 & 3.59). The diffraction qualities of both NPGP and PGP crystals exhibited variability from crystal to crystal using the same cryoprotecting solution. Working with freshly-grown crystals from freshly-prepared protein samples was found to increase both the probability and quality of diffraction. Diffraction experiments indicated that the NTE did not abolish crystallization but affected the diffraction quality of the resulting crystals. NPGP crystals were deeply rooted in the skin of denatured phytochromes exacerbating problems during harvesting (Figure 3.42). Surface Entropy Reduction (Section 3.4), chromophore replacement (Figure 3.47), lysine methylation, crystals from the additive screening (Figure 3.45) and additional constructs such as HPGP (Figure 3.5) which include conserved aromatic residues in the NTE need to be tested to improve the diffraction quality of Sb.phyB constructs. C-terminal truncation of the PHY domain produced needle-like PG crystals (Figure 3.58) that diffracted X-rays to atomic resolution (Figure 3.59). Apart from this, Y291HNPGP, Y291HPGP and Y291HPG were all proven recalcitrant to immense crystallization screening efforts (Section 3.5.4).

Diffractions from PGP crystals were strongly anisotropic (Figure 3.44). A common source of anisotropy that is presumably relevant to phytochrome crystals is the overall movement of the whole molecules or domains of the molecules as rigid bodies within the crystal lattice [181]. Different sources of anisotropy will require careful refinement strategies. Essen et al. (2008) successfully rescaled and solved anisotropic diffraction dataset from Cph1Δ2 crystals [25]. On the other hand, the reflection intensities obtained for PG crystals were very weak and the spots were highly irregular (Figure 3.59). This irregularity might be caused by diffraction from more than one crystal, physical or radiation induced damage to crystals. PG crystals grew as stack of needles hampering single crystal harvesting without physical damage to the crystals.

X-rays from synchrotron radiation could also induce substantial damage to cryocooled protein crystals. The most radiation-sensitive residues within a protein are presumably active-site residues [182]. PG construct does not have the NTE and the PHY domain (Figure 3.48) to seal the

chromophore binding pocket completely. In the absence of NTE and PHY domain, direct interaction between the ionizing radiation and electrons in chromophore might cause radiation damage. Using X-ray crystallography and single crystal spectroscopic techniques, Li et al. (2015) suggested protonation of BV chromophore due to radiation damage in the PAS-GAF construct of Dr.BphP [183]. Advanced data collection strategies, radiation damage scavengers, optimized dose rate and optimized photon energies should be tested to mitigate the effect of X-rays induced radiation damage to PG crystals. Additive screening and seeding strategies should be tested to improve the morphology and quality of PG crystals.

4.3 PIFs in *Sorghum*

4.3.1 Production and purification of putative Sb.PIF1

A mechanistic model for how light induced conformational changes in the photosensory module of phyB propagate through the activated phytochrome B binding site (APB) into the DNA-binding (DBD) domain of PIF3, a key step in light regulated gene expression, remains to be established. Cloning and expression of CDS fragments of PIF1 from *Sorghum bicolor* are described here: they provide typical examples of the problems associated with cloning and over-expressing eukaryotic transcription factors in *E. coli.* The full-length Sb.PIF1 accumulated as insoluble inclusion bodies and aggregated strongly under all conditions investigated (Figs. 3.97 & 3.103). The *Sb.APB* construct was not produced despite several attempts under many different conditions (Figure 3.98).

The soluble fraction of Sb.PIF1 showed very little affinity to Ni-NTA column perhaps due to soluble aggregation/misfolding. Purification of Sb.PIF1 was possible via Ni-NTA-affinity chromatography under native and denaturing conditions. The purified soluble fraction was eluted in the void volume during SEC, implying extensive aggregation (Figure 3.103). Attempt to co-purify Y291HNPGP and PIF1 complex by SEC was unsuccessful (Figure 3.107). The most likely reason for the absence of detectable interaction is misfolding and intramolecular aggregation of *E. coli* produced Sb.PIF1. A conclusive statement should be made only after Sb.PIF1 is produced as major non-void peak. The difficulties with Sb.PIF1 experienced in this project provide examples of the technical problems thwarting studies of PIF structure and function.

4.3.2 Characterization of Sb.PIF1

PIFs are known to be directly involved in retrograde signaling by negatively regulating chlorophyll biosynthesis in the model plant *Arabidopsis* [73, 184]. Different tetrapyrroles such as heme [185], Mg-protoporphyrinIX [75, 186] and bilins [187] were each considered to be involved in plastid-to-nucleus retrograde signaling. When the full-length Sb.PIF1 codon-optimized for production and purification in *Escherichia coli* was purified, the protein solution was amber colored (Figure 3.102) showing a weak zinc-induced fluorescence after SDS-PAGE (Figure 3.101). The absorbance spectrum of the purified protein is typical of heme proteins (Figure 3.104). The binding of a likely porphyrin to recombinant Sb.PIF1 in *E. coli* revealed in the present project might thus be physiologically relevant in light of retrograde signaling. It is not unlikely that currently unknown ligands are important in PIF function. On the other hand, the nature of the putative porphyrin and its mode of binding were not clearly defined in the present project, nor was it established that a similar ligand is associated with PIF proteins in plants. Hence the nature of the correct tetrapyrrole and its role in the activity of Sb.PIF1 remains to be established.

The rewiring of transcriptional networks is an important source of evolutionary novelty [188]. The genome of *Chlamydomonas,* a classical model system for unicellular algae, has eight bHLH genes and no phytochrome gene [133, 189]. In contrast, there are ten bHLH genes and two phytochrome genes in the draft genome sequence of *Klebsormidium flaccidum*; a filamentous charophyte lacking differentiation of specialized cells [129]. The presence of a putative PIF gene in *K. flaccidum* genome suggested that phytochrome-mediated transcriptional regulation might have evolved prior to the divergence of *Klebsormidium* and the lineage leading to land plants [131, 132]. The phy-PIF system might have evolved in terrestrial multicellular photosynthetic organisms to coordinate the assembly of chlorophyll biosynthetic pathway with the ambient conditions of light, oxygen and temperature. PIFs act as a set of signaling hubs that play central roles in plant growth and development [190].

5 Summary

The main aim of this study was to investigate the structural differences between the photosensory modules of plant phytochromes and cyanobacterial phytochromes. To this end, three constructs of Sb.phyB, namely NPGP, PGP and PG, were successfully cloned, produced, purified and crystallized. Unlike prokaryotic phytochromes such as Cph1, obtaining diffraction-quality crystals for plant phytochromes is extremely challenging, indeed, no structure for a complete phyB photosensory module has been published to date. In this study, however, it proved possible to generate both NPGP and PGP crystals diffracting to 6 Å. PG crystals diffracting to atomic resolution were also generated and a complete dataset was collected, albeit with some anomalous pattern. SER mutation and additive screening efforts were also completed to mitigate difficult crystal morphology and improve the diffraction quality of NPGP and PG crystals. Additional work is still required to solve the structure of these fragments.

The NTE of Sb.phyB, which is rich in charged and polar amino acids, seems to function as a molecular shield to cover the chromophore. Like the extensively studied PHY domain, the NTE was shown here to be important for the spectral integrity of Pfr. It is thus not unlikely for NTE to exhibit rudimentary PHY like structure and function. Conversely, the NTE of plant phytochromes could be remnant of an N-terminal ancestral fusion protein. Complete removal of the PAS domain in Sb.GP affected proper folding and chromophore ligation. Limited proteolysis, SEC and temperature-induced unfolding suggest conformational heterogeneity in the previously identified gain-of-function mutation in the GAF domain (Y291H). No crystallization condition was identified for Y291HNPGP, Y291HPGP and Y291HPG constructs from Sb.phyB. Like the NTE, the use of PΦB instead of PCB might increase conformational plasticity in plant phytochromes. The half-life of Pfr form was found to be remarkably longer for NPGP assembled with PCB than that with PΦB.

Phytochromes regulate numerous light responses via different signaling partners, such as phytochrome interacting factors (PIFs). PIFs contain a region of low sequence similarity to PAS domains, many of which are known to bind cofactors such as heme. In this study, recombinant putative Sb.PIF1 expressed in *E. coli* was unexpectedly found to bind a likely porphyrin cofactor. It would be worthwhile to characterize the putative PAS domain in Sb.PIF1 with its likely ligand and its interaction with light-activated phytochrome at the structural level.

6 References

1. Rockwell, N.C. and J.C. Lagarias, *A brief history of phytochromes.* ChemPhysChem, 2010. **11**(6): p. 1172-1180.
2. Borthwick, H.A., et al., *A reversible photoreaction controlling seed germination.* Proceedings of the National Academy of Sciences, 1952. **38**(8): p. 662-666.
3. Sharrock, R.A. and P.H. Quail, *Novel phytochrome sequences in Arabidopsis thaliana: structure, evolution, and differential expression of a plant regulatory photoreceptor family.* Genes & Development, 1989. **3**(11): p. 1745-1757.
4. Hughes, J., et al., *A prokaryotic phytochrome.* Nature, 1997. **386**(6626): p. 663-663.
5. Giraud, E., et al., *Bacteriophytochrome controls photosystem synthesis in anoxygenic bacteria.* Nature, 2002. **417**(6885): p. 202-205.
6. Rottwinkel, G., I. Oberpichler, and T. Lamparter, *Bathy phytochromes in rhizobial soil bacteria.* Journal of bacteriology, 2010. **192**(19): p. 5124-5133.
7. Yeh, K.-C., et al., *A cyanobacterial phytochrome two-component light sensory system.* Science, 1997. **277**(5331): p. 1505-1508.
8. Blumenstein, A., et al., *The Aspergillus nidulans phytochrome FphA represses sexual development in red light.* Current Biology, 2005. **15**(20): p. 1833-1838.
9. Yu, Z., O. Armant, and R. Fischer, *Fungi use the SakA (HogA) pathway for phytochrome-dependent light signalling.* Nature Microbiology, 2016. **1**: p. 16019.
10. Fry, K.T. and F.E. Mumford, *Isolation and partial characterization of a chromophore-peptide fragment from pepsin digests of phytochrome.* Biochemical and Biophysical research communications, 1971. **45**(6): p. 1466-1473.
11. Lagarias, J.C. and D.M. Lagarias, *Self-assembly of synthetic phytochrome holoprotein in vitro.* Proceedings of the National Academy of Sciences, 1989. **86**(15): p. 5778-5780.
12. Lagarias, J.C. and H. Rapoport, *Chromopeptides from phytochrome. The structure and linkage of the Pr form of the phytochrome chromophore.* Journal of the American Chemical Society, 1980. **102**(14): p. 4821-4828.
13. Davis, S.J., J. Kurepa, and R.D. Vierstra, *The Arabidopsis thaliana HY1 locus, required for phytochrome-chromophore biosynthesis, encodes a protein related to heme oxygenases.* Proceedings of the National Academy of Sciences, 1999. **96**(11): p. 6541-6546.
14. Kohchi, T., et al., *The Arabidopsis HY2 gene encodes phytochromobilin synthase, a ferredoxin-dependent biliverdin reductase.* The Plant Cell, 2001. **13**(2): p. 425-436.
15. Wu, S.-H. and J.C. Lagarias, *Defining the bilin lyase domain: lessons from the extended phytochrome superfamily.* Biochemistry, 2000. **39**(44): p. 13487-13495.
16. Butler, W.L., et al., *Detection, assay, and preliminary purification of the pigment controlling photoresponsive development of plants.* Proceedings of the National Academy of Sciences, 1959. **45**(12): p. 1703-1708.
17. Montgomery, B.L. and J.C. Lagarias, *Phytochrome ancestry: sensors of bilins and light.* Trends in plant science, 2002. **7**(8): p. 357-366.
18. Zhao, K.-H., et al., *Photochromic biliproteins from the cyanobacterium Anabaena sp. PCC 7120: lyase activities, chromophore exchange, and photochromism in phytochrome AphA.* Biochemistry, 2004. **43**(36): p. 11576-11588.
19. Rüdiger, W., et al., *Isolation of the phytochrome chromophore. The cleavage reaction with hydrogen bromide.* Zeitschrift für Naturforschung C, 1980. **35**(9-10): p. 763-769.

20. Thümmler, F., T. Brandlmeier, and W. Rüdiger, *Preparation and properties of chromopeptides from the Pfr form of phytochrome.* Zeitschrift für Naturforschung C, 1981. **36**(5-6): p. 440-449.
21. Thümmler, F., et al., *Chromopeptides from Phytochrome and Phycocyanin. NM R Studies of the Pfr and Pr Chromophore of Phytochrome and E, Z Isomeric Chromophores of Phycocyanin.* Zeitschrift für Naturforschung C, 1983. **38**(5-6): p. 359-368.
22. Ulijasz, A.T., et al., *Structural basis for the photoconversion of a phytochrome to the activated Pfr form.* Nature, 2010. **463**(7278): p. 250-254.
23. Song, C., et al., *The D-ring, not the A-ring, rotates in Synechococcus OS-B' phytochrome.* Journal of Biological Chemistry, 2014. **289**(5): p. 2552-2562.
24. Hahn, J., et al., *Probing protein–chromophore interactions in Cph1 phytochrome by mutagenesis.* Febs Journal, 2006. **273**(7): p. 1415-1429.
25. Essen, L.-O., J. Mailliet, and J. Hughes, *The structure of a complete phytochrome sensory module in the Pr ground state.* Proceedings of the National Academy of Sciences, 2008. **105**(38): p. 14709-14714.
26. Yang, X., J. Kuk, and K. Moffat, *Crystal structure of Pseudomonas aeruginosa bacteriophytochrome: photoconversion and signal transduction.* Proceedings of the National Academy of Sciences, 2008. **105**(38): p. 14715-14720.
27. Lamparter, T., B. Esteban, and J. Hughes, *Phytochrome Cph1 from the cyanobacterium Synechocystis PCC6803.* European Journal of Biochemistry, 2001. **268**(17): p. 4720-4730.
28. Quail, P.H., et al., *Phytochromes: photosensory perception and signal transduction.* Science, 1995. **268**(5211): p. 675.
29. Schneider-Poetsch, H.A., et al., *Phytochromes and bacterial sensor proteins are related by structural and functional homologies Hypothesis on phytochrome-mediated signal-transduction.* FEBS letters, 1991. **281**(1-2): p. 245-249.
30. Ni, M., J.M. Tepperman, and P.H. Quail, *PIF3, a phytochrome-interacting factor necessary for normal photoinduced signal transduction, is a novel basic helix-loop-helix protein.* Cell, 1998. **95**(5): p. 657-667.
31. Ni, M., J.M. Tepperman, and P.H. Quail, *Binding of phytochrome B to its nuclear signalling partner PIF3 is reversibly induced by light.* Nature, 1999. **400**(6746): p. 781-784.
32. Rösler, J., I. Klein, and M. Zeidler, *Arabidopsisfhl/fhy1 double mutant reveals a distinct cytoplasmic action of phytochrome A.* Proceedings of the National Academy of Sciences, 2007. **104**(25): p. 10737-10742.
33. Jaedicke, K., et al., *A phytochrome–phototropin light signaling complex at the plasma membrane.* Proceedings of the National Academy of Sciences, 2012. **109**(30): p. 12231-12236.
34. Hughes, J., *Phytochrome cytoplasmic signaling.* Annual review of plant biology, 2013. **64**: p. 377-402.
35. Hughes, J., *Phytochrome three-dimensional structures and functions.* Biochemical Society Transactions, 2010. **38**(2): p. 710-716.
36. Nagano, S., *From photon to signal in phytochromes: similarities and differences between prokaryotic and plant phytochromes.* Journal of plant research, 2016. **129**(2): p. 123-135.
37. Lagarias, D.M., S.-H. Wu, and J.C. Lagarias, *Atypical phytochrome gene structure in the green alga Mesotaenium caldariorum.* Plant molecular biology, 1995. **29**(6): p. 1127-1142.
38. Taylor, B.L. and I.B. Zhulin, *PAS domains: internal sensors of oxygen, redox potential, and light.* Microbiology and Molecular Biology Reviews, 1999. **63**(2): p. 479-506.
39. Möglich, A., R.A. Ayers, and K. Moffat, *Structure and signaling mechanism of Per-ARNT-Sim domains.* Structure, 2009. **17**(10): p. 1282-1294.

40. Gilles-Gonzalez, M.-A. and G. Gonzalez, *Signal transduction by heme-containing PAS-domain proteins.* Journal of Applied Physiology, 2004. **96**(2): p. 774-783.
41. Cusanovich, M.A. and T.E. Meyer, *Photoactive yellow protein: a prototypic PAS domain sensory protein and development of a common signaling mechanism.* Biochemistry, 2003. **42**(17): p. 4759-4770.
42. Ho, Y.S.J., L.M. Burden, and J.H. Hurley, *Structure of the GAF domain, a ubiquitous signaling motif and a new class of cyclic GMP receptor.* The EMBO Journal, 2000. **19**(20): p. 5288-5299.
43. Lamparter, T., et al., *Phytochrome from Agrobacterium tumefaciens has unusual spectral properties and reveals an N-terminal chromophore attachment site.* Proceedings of the National Academy of Sciences, 2002. **99**(18): p. 11628-11633.
44. Wagner, J.R., et al., *A light-sensing knot revealed by the structure of the chromophore-binding domain of phytochrome.* Nature, 2005. **438**(7066): p. 325-331.
45. Anders, K., et al., *Structure of the cyanobacterial phytochrome 2 photosensor implies a tryptophan switch for phytochrome signaling.* Journal of Biological Chemistry, 2013. **288**(50): p. 35714-35725.
46. Takala, H., et al., *Signal amplification and transduction in phytochrome photosensors.* Nature, 2014. **509**(7499): p. 245.
47. Rumyantsev, K.A., et al., *Minimal domain of bacterial phytochrome required for chromophore binding and fluorescence.* Scientific reports, 2015. **5**.
48. Narikawa, R., et al., *A novel photoactive GAF domain of cyanobacteriochrome AnPixJ that shows reversible green/red photoconversion.* Journal of molecular biology, 2008. **380**(5): p. 844-855.
49. Rockwell, N.C., et al., *Diverse two-cysteine photocycles in phytochromes and cyanobacteriochromes.* Proceedings of the National Academy of Sciences, 2011. **108**(29): p. 11854-11859.
50. Wagner, J.R., et al., *High resolution structure of Deinococcus bacteriophytochrome yields new insights into phytochrome architecture and evolution.* Journal of Biological Chemistry, 2007. **282**(16): p. 12298-12309.
51. Cornilescu, G., et al., *Solution structure of a cyanobacterial phytochrome GAF domain in the red-light-absorbing ground state.* Journal of molecular biology, 2008. **383**(2): p. 403-413.
52. Cornilescu, C.C., et al., *Dynamic structural changes underpin photoconversion of a blue/green cyanobacteriochrome between its dark and photoactivated states.* Journal of Biological Chemistry, 2014. **289**(5): p. 3055-3065.
53. Li, H., et al., *Quaternary organization of a phytochrome dimer as revealed by cryoelectron microscopy.* Proceedings of the National Academy of Sciences, 2010. **107**(24): p. 10872-10877.
54. Burgie, E.S., et al., *Crystallographic and electron microscopic analyses of a bacterial phytochrome reveal local and global rearrangements during photoconversion.* Journal of Biological Chemistry, 2014. **289**(35): p. 24573-24587.
55. Song, C., et al., *Two ground state isoforms and a chromophore D-ring photoflip triggering extensive intramolecular changes in a canonical phytochrome.* Proceedings of the National Academy of Sciences, 2011. **108**(10): p. 3842-3847.
56. Song, C., et al., *A red/green cyanobacteriochrome sustains its color despite a change in the bilin chromophore's protonation state.* Biochemistry, 2015. **54**(38): p. 5839-5848.
57. Björling, A., et al., *Ubiquitous Structural Signaling in Bacterial Phytochromes.* The journal of physical chemistry letters, 2015. **6**(17): p. 3379-3383.

58. Bhoo, S.H., et al., *Phytochrome photochromism probed by site-directed mutations and chromophore esterification.* Journal of the American Chemical Society, 1997. **119**(48): p. 11717-11718.
59. Mailliet, J., et al., *Spectroscopy and a high-resolution crystal structure of Tyr263 mutants of cyanobacterial phytochrome Cph1.* Journal of molecular biology, 2011. **413**(1): p. 115-127.
60. Fischer, A.J. and J.C. Lagarias, *Harnessing phytochrome's glowing potential.* Proceedings of the National Academy of Sciences of the United States of America, 2004. **101**(50): p. 17334-17339.
61. Fischer, A.J., et al., *Multiple roles of a conserved GAF domain tyrosine residue in cyanobacterial and plant phytochromes.* Biochemistry, 2005. **44**(46): p. 15203-15215.
62. Su, Y.-s. and J.C. Lagarias, *Light-independent phytochrome signaling mediated by dominant GAF domain tyrosine mutants of Arabidopsis phytochromes in transgenic plants.* The Plant Cell, 2007. **19**(7): p. 2124-2139.
63. Buckley, C.E., et al., *Reversible Optogenetic Control of Subcellular Protein Localization in a Live Vertebrate Embryo.* Developmental cell, 2016. **36**(1): p. 117-126.
64. Jeong, A.-R., et al., *New Constitutively Active Phytochromes Exhibit Light-Independent Signaling Activity.* Plant Physiology, 2016: p. pp. 00342.2016.
65. Scheer, H., *Biliproteins.* Angewandte Chemie International Edition in English, 1981. **20**(3): p. 241-261.
66. Yang, X., J. Kuk, and K. Moffat, *Conformational differences between the Pfr and Pr states in Pseudomonas aeruginosa bacteriophytochrome.* Proceedings of the National Academy of Sciences, 2009. **106**(37): p. 15639-15644.
67. Burgie, E.S., J. Zhang, and R.D. Vierstra, *Crystal Structure of Deinococcus Phytochrome in the Photoactivated State Reveals a Cascade of Structural Rearrangements during Photoconversion.* Structure, 2016. **24**(3): p. 448-457.
68. Burgie, E.S., et al., *Crystal structure of the photosensing module from a red/far-red light-absorbing plant phytochrome.* Proceedings of the National Academy of Sciences, 2014. **111**(28): p. 10179-10184.
69. Chamovitz, D.A., et al., *The COP9 complex, a novel multisubunit nuclear regulator involved in light control of a plant developmental switch.* Cell, 1996. **86**(1): p. 115-121.
70. Chory, J., et al., *Arabidopsis thaliana mutant that develops as a light-grown plant in the absence of light.* Cell, 1989. **58**(5): p. 991-999.
71. Endo, M., et al., *PHYTOCHROME-DEPENDENT LATE-FLOWERING accelerates flowering through physical interactions with phytochrome B and CONSTANS.* Proceedings of the National Academy of Sciences, 2013. **110**(44): p. 18017-18022.
72. Chory, J., *Light signal transduction: an infinite spectrum of possibilities.* The Plant Journal, 2010. **61**(6): p. 982-991.
73. Huq, E., et al., *Phytochrome-interacting factor 1 is a critical bHLH regulator of chlorophyll biosynthesis.* Science, 2004. **305**(5692): p. 1937-1941.
74. Kim, J., et al., *Functional characterization of phytochrome interacting factor 3 in phytochrome-mediated light signal transduction.* The Plant Cell, 2003. **15**(10): p. 2399-2407.
75. Strand, Å., et al., *Chloroplast to nucleus communication triggered by accumulation of Mg-protoporphyrinIX.* Nature, 2003. **421**(6918): p. 79-83.
76. Deng, X.-W., T. Caspar, and P.H. Quail, *cop1: a regulatory locus involved in light-controlled development and gene expression in Arabidopsis.* Genes & development, 1991. **5**(7): p. 1172-1182.

77. Hershey, H.P., et al., *Molecular cloning of cDNA for Avena phytochrome.* Proceedings of the National Academy of Sciences, 1984. **81**(8): p. 2332-2336.
78. Hunt, R.E. and L.H. Pratt, *Phytochrome immunoaffinity purification.* Plant physiology, 1979. **64**(2): p. 332-336.
79. Vierstra, R.D. and P.H. Quail, *Native phytochrome: Inhibition of proteolysis yields a homogeneous monomer of 124 kilodaltons from Avena.* Proceedings of the National Academy of Sciences, 1982. **79**(17): p. 5272-5276.
80. Hunt, R.E. and L.H. Pratt, *Physicochemical differences between the red-and the far-red-absorbing forms of phytochrome.* Biochemistry, 1981. **20**(4): p. 941-945.
81. Litts, J., J.M. Kelly, and J.C. Lagarias, *Structure-function studies on phytochrome. Preliminary characterization of highly purified phytochrome from Avena sativa enriched in the 124-kilodalton species.* Journal of Biological Chemistry, 1983. **258**(18): p. 11025-11031.
82. Koornneef, M., E. Rolff, and C. Spruit, *Genetic control of light-inhibited hypocotyl elongation in Arabidopsis thaliana (L.) Heynh.* Zeitschrift für Pflanzenphysiologie, 1980. **100**(2): p. 147-160.
83. Osterlund, M.T., et al., *Targeted destabilization of HY5 during light-regulated development of Arabidopsis.* Nature, 2000. **405**(6785): p. 462-466.
84. Paterson, A.H., et al., *The Sorghum bicolor genome and the diversification of grasses.* Nature, 2009. **457**(7229): p. 551-556.
85. Reed, J.W., et al., *Phytochrome A and phytochrome B have overlapping but distinct functions in Arabidopsis development.* Plant physiology, 1994. **104**(4): p. 1139-1149.
86. Furuya, M., *Molecular properties and biogenesis of phytochrome I and II.* Advances in biophysics, 1989. **25**: p. 133-167.
87. Parks, B.M. and P.H. Quail, *hy8, a new class of arabidopsis long hypocotyl mutants deficient in functional phytochrome A.* The Plant Cell, 1993. **5**(1): p. 39-48.
88. Dehesh, K., et al., *Arabidopsis HY8 locus encodes phytochrome A.* The Plant Cell, 1993. **5**(9): p. 1081-1088.
89. Nagatani, A., J. Chory, and M. Furuya, *Phytochrome B is not detectable in the hy3 mutant of Arabidopsis, which is deficient in responding to end-of-day far-red light treatments.* Plant and Cell Physiology, 1991. **32**(7): p. 1119-1122.
90. Shinomura, T., et al., *Action spectra for phytochrome A-and B-specific photoinduction of seed germination in Arabidopsis thaliana.* Proceedings of the National Academy of Sciences, 1996. **93**(15): p. 8129-8133.
91. Nagatani, A., J.W. Reed, and J. Chory, *Isolation and initial characterization of Arabidopsis mutants that are deficient in phytochrome A.* Plant Physiology, 1993. **102**(1): p. 269-277.
92. Shinomura, T., K. Uchida, and M. Furuya, *Elementary processes of photoperception by phytochrome A for high-irradiance response of hypocotyl elongation in Arabidopsis.* Plant Physiology, 2000. **122**(1): p. 147-156.
93. Fankhauser, C., et al., *PKS1, a substrate phosphorylated by phytochrome that modulates light signaling in Arabidopsis.* Science, 1999. **284**(5419): p. 1539-1541.
94. Choi, G., et al., *Phytochrome signalling is mediated through nucleoside diphosphate kinase 2.* Nature, 1999. **401**(6753): p. 610-613.
95. Chen, D., et al., *Antagonistic basic helix-loop-helix/bZIP transcription factors form transcriptional modules that integrate light and reactive oxygen species signaling in Arabidopsis.* The Plant Cell, 2013. **25**(5): p. 1657-1673.
96. Bae, G. and G. Choi, *Decoding of light signals by plant phytochromes and their interacting proteins.* Annu. Rev. Plant Biol., 2008. **59**: p. 281-311.

97. Deng, X.-W., et al., *COP1, an Arabidopsis regulatory gene, encodes a protein with both a zinc-binding motif and a G β homologous domain.* Cell, 1992. **71**(5): p. 791-801.
98. Sakamoto, K. and A. Nagatani, *Nuclear localization activity of phytochrome B.* The Plant Journal, 1996. **10**(5): p. 859-868.
99. Nagatani, A., *Light-regulated nuclear localization of phytochromes.* Current opinion in plant biology, 2004. **7**(6): p. 708-711.
100. Martínez-García, J.F., E. Huq, and P.H. Quail, *Direct targeting of light signals to a promoter element-bound transcription factor.* Science, 2000. **288**(5467): p. 859-863.
101. Ni, W., et al., *Multisite light-induced phosphorylation of the transcription factor PIF3 is necessary for both its rapid degradation and concomitant negative feedback modulation of photoreceptor phyB levels in Arabidopsis.* The Plant Cell, 2013. **25**(7): p. 2679-2698.
102. Christians, M.J., et al., *The light-response BTB1 and BTB2 proteins assemble nuclear ubiquitin ligases that modify phytochrome B and D signaling in Arabidopsis.* Plant physiology, 2012. **160**(1): p. 118-134.
103. Ni, W., et al., *A mutually assured destruction mechanism attenuates light signaling in Arabidopsis.* Science, 2014. **344**(6188): p. 1160-1164.
104. Wagner, D., et al., *Chromophore-bearing NH2-terminal domains of phytochromes A and B determine their photosensory specificity and differential light lability.* Proceedings of the National Academy of Sciences, 1996. **93**(9): p. 4011-4015.
105. Park, C.-M., S.-H. Bhoo, and P.-S. Song. *Inter-domain crosstalk in the phytochrome molecules*. in *Seminars in cell & developmental biology*. 2000. Elsevier.
106. Chen, M., et al., *Regulation of phytochrome B nuclear localization through light-dependent unmasking of nuclear-localization signals.* Current Biology, 2005. **15**(7): p. 637-642.
107. Vierstra, R.D. and P.H. Quail, *Proteolysis alters the spectral properties of 124 kdalton phytochrome from Avena.* Planta, 1982. **156**(2): p. 158-165.
108. Baron, O. and B.L. Epel, *Differences in the physical properties of native and partially degraded phytochrome as probed by their differential sensitivity to permanganate oxidation.* Plant physiology, 1983. **73**(2): p. 471-474.
109. Stockhaus, J., et al., *Serine-to-alanine substitutions at the amino-terminal region of phytochrome A result in an increase in biological activity.* Genes & development, 1992. **6**(12a): p. 2364-2372.
110. Wagner, D., M. Koloszvari, and P.H. Quail, *Two small spatially distinct regions of phytochrome B are required for efficient signaling rates.* The Plant Cell, 1996. **8**(5): p. 859-871.
111. Sweere, U., et al., *Interaction of the response regulator ARR4 with phytochrome B in modulating red light signaling.* Science, 2001. **294**(5544): p. 1108-1111.
112. Yeh, K.-C. and J.C. Lagarias, *Eukaryotic phytochromes: light-regulated serine/threonine protein kinases with histidine kinase ancestry.* Proceedings of the National Academy of Sciences, 1998. **95**(23): p. 13976-13981.
113. Matsushita, T., N. Mochizuki, and A. Nagatani, *Dimers of the N-terminal domain of phytochrome B are functional in the nucleus.* Nature, 2003. **424**(6948): p. 571-574.
114. Oka, Y., et al., *Functional analysis of a 450–amino acid N-terminal fragment of phytochrome B in Arabidopsis.* The Plant Cell, 2004. **16**(8): p. 2104-2116.
115. Kikis, E.A., et al., *Residues clustered in the light-sensing knot of phytochrome B are necessary for conformer-specific binding to signaling partner PIF3.* PLoS Genet, 2009. **5**(1): p. e1000352.

116. Oka, Y., et al., *Mutant screen distinguishes between residues necessary for light-signal perception and signal transfer by phytochrome B.* PLoS Genet, 2008. **4**(8): p. e1000158.
117. Pao, C.-I. and P.W. Morgan, *Genetic Regulation of Development in Sorghum bicolar I. Role of the Maturity Genes.* Plant physiology, 1986. **82**(2): p. 575-580.
118. Yamaguchi, M., et al., *Sorghum Dw1, an agronomically important gene for lodging resistance, encodes a novel protein involved in cell proliferation.* Scientific Reports, 2016. **6**.
119. Pratt, L.H., et al., *Sorghum expressed sequence tags identify signature genes for drought, pathogenesis, and skotomorphogenesis from a milestone set of 16,801 unique transcripts.* Plant physiology, 2005. **139**(2): p. 869-884.
120. Morgan, P.W., L.W. Guy, and C.-I. Pao, *Genetic regulation of development in Sorghum bicolor III. Asynchrony of thermoperiods with photoperiods promotes floral initiation.* Plant Physiology, 1987. **83**(2): p. 448-450.
121. Childs, K.L., L.H. Pratt, and P.W. Morgan, *Genetic Regulation of Development in Sorghum bicolor VI. The ma3R Allele Results in Abnormal Phytochrome Physiology.* Plant physiology, 1991. **97**(2): p. 714-719.
122. Childs, K.L., et al., *Genetic Regulation of Development in Sorghum bicolor VII. ma3r Flowering Mutant Lacks a Phytochrome that Predominates in Green Tissue.* Plant physiology, 1992. **99**(2): p. 765-770.
123. Childs, K.L., et al., *The sorghum photoperiod sensitivity gene, Ma3, encodes a phytochrome B.* Plant Physiology, 1997. **113**(2): p. 611-619.
124. Halliday, K.J., et al., *poc1: An Arabidopsis mutant perturbed in phytochrome signaling because of a T DNA insertion in the promoter of PIF3, a gene encoding a phytochrome-interacting bHLH protein.* Proceedings of the National Academy of Sciences, 1999. **96**(10): p. 5832-5837.
125. Xu, X., et al., *Illuminating progress in phytochrome-mediated light signaling pathways.* Trends in plant science, 2015. **20**(10): p. 641-650.
126. Tepperman, J.M., et al., *Multiple transcription-factor genes are early targets of phytochrome A signaling.* Proceedings of the National Academy of Sciences, 2001. **98**(16): p. 9437-9442.
127. Shin, J., et al., *Phytochromes promote seedling light responses by inhibiting four negatively-acting phytochrome-interacting factors.* Proceedings of the National Academy of Sciences, 2009. **106**(18): p. 7660-7665.
128. Khanna, R., et al., *A novel molecular recognition motif necessary for targeting photoactivated phytochrome signaling to specific basic helix-loop-helix transcription factors.* The Plant Cell, 2004. **16**(11): p. 3033-3044.
129. Hori, K., et al., *Klebsormidium flaccidum genome reveals primary factors for plant terrestrial adaptation.* Nature communications, 2014. **5**.
130. Li, F.-W., et al., *Phytochrome diversity in green plants and the origin of canonical plant phytochromes.* Nature communications, 2015. **6**.
131. Catarino, B., et al., *The stepwise increase in the number of transcription factor families in the Precambrian predated the diversification of plants on land.* Molecular Biology and Evolution, 2016: p. msw155.
132. Inoue, K., et al., *Phytochrome signaling is mediated by PHYTOCHROME INTERACTING FACTOR in the liverwort Marchantia polymorpha.* The Plant Cell, 2016: p. tpc. 01063.2015.
133. Feller, A., et al., *Evolutionary and comparative analysis of MYB and bHLH plant transcription factors.* The Plant Journal, 2011. **66**(1): p. 94-116.

134. Zhu, L., et al., *A Negative Feedback Loop between PHYTOCHROME INTERACTING FACTORs and HECATE Proteins Fine-Tunes Photomorphogenesis in Arabidopsis.* The Plant Cell, 2016. **28**(4): p. 855-874.
135. Pires, N. and L. Dolan, *Origin and diversification of basic-helix-loop-helix proteins in plants.* Molecular biology and evolution, 2010. **27**(4): p. 862-874.
136. Leung, D.W., et al., *Genetically encoded photoswitching of actin assembly through the Cdc42-WASP-Arp2/3 complex pathway.* Proceedings of the National Academy of Sciences, 2008. **105**(35): p. 12797-12802.
137. Box, G.E. and D.W. Behnken, *Some new three level designs for the study of quantitative variables.* Technometrics, 1960. **2**(4): p. 455-475.
138. Laemmli, U.K., *Cleavage of structural proteins during the assembly of the head of bacteriophage T4.* nature, 1970. **227**: p. 680-685.
139. Berkelman, T.R. and J.C. Lagarias, *Visualization of bilin-linked peptides and proteins in polyacrylamide gels.* Analytical biochemistry, 1986. **156**(1): p. 194-201.
140. Bradford, M.M., *A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding.* Analytical biochemistry, 1976. **72**(1-2): p. 248-254.
141. Ellman, G.L., *Tissue sulfhydryl groups.* Archives of biochemistry and biophysics, 1959. **82**(1): p. 70-77.
142. Garcia-Ruiz, J.M., et al., *Granada Crystallisation Box: a new device for protein crystallisation by counter-diffusion techniques.* Acta Crystallographica Section D: Biological Crystallography, 2002. **58**(10): p. 1638-1642.
143. Heras, B. and J.L. Martin, *Post-crystallization treatments for improving diffraction quality of protein crystals.* Acta Crystallographica Section D: Biological Crystallography, 2005. **61**(9): p. 1173-1180.
144. Newman, J., *A review of techniques for maximizing diffraction from a protein crystal in stilla.* Acta Crystallographica Section D: Biological Crystallography, 2006. **62**(1): p. 27-31.
145. Bowler, M.W., et al., *Automation and experience of controlled crystal dehydration: Results from the european synchrotron hc1 collaboration.* Crystal Growth & Design, 2015. **15**(3): p. 1043-1054.
146. Derewenda, Z.S. and P.G. Vekilov, *Entropy and surface engineering in protein crystallization.* Acta Crystallographica Section D: Biological Crystallography, 2006. **62**(1): p. 116-124.
147. Zhang, Y., *I-TASSER server for protein 3D structure prediction.* BMC bioinformatics, 2008. **9**(1): p. 40.
148. Dale, G.E., C. Oefner, and A. D'Arcy, *The protein as a variable in protein crystallization.* Journal of structural biology, 2003. **142**(1): p. 88-97.
149. Slabinski, L., et al., *The challenge of protein structure determination—lessons from structural genomics.* Protein Science, 2007. **16**(11): p. 2472-2482.
150. Kelley, L.A. and M.J. Sternberg, *Protein structure prediction on the Web: a case study using the Phyre server.* Nature protocols, 2009. **4**(3): p. 363-371.
151. Linding, R., et al., *Protein disorder prediction: implications for structural proteomics.* Structure, 2003. **11**(11): p. 1453-1459.
152. Dosztányi, Z., et al., *IUPred: web server for the prediction of intrinsically unstructured regions of proteins based on estimated energy content.* Bioinformatics, 2005. **21**(16): p. 3433-3434.
153. Villalobos, A., et al., *Gene Designer: a synthetic biology tool for constructing artificial DNA segments.* BMC bioinformatics, 2006. **7**(1): p. 1.

154. Gustafsson, C., S. Govindarajan, and J. Minshull, *Codon bias and heterologous protein expression.* Trends in biotechnology, 2004. **22**(7): p. 346-353.
155. Boël, G., et al., *Codon influence on protein expression in E. coli correlates with mRNA levels.* Nature, 2016. **529**(7586): p. 358-363.
156. Deforce, L., et al., *In vitro assembly of apophytochrome and apophytochrome deletion mutants expressed in yeast with phycocyanobilin.* Proceedings of the National Academy of Sciences, 1991. **88**(23): p. 10392-10396.
157. Wahleithner, J.A., L. Li, and J.C. Lagarias, *Expression and assembly of spectrally active recombinant holophytochrome.* Proceedings of the National Academy of Sciences, 1991. **88**(23): p. 10387-10391.
158. Rao, K.J., C.-H. Kim, and S.-K. Rhee, *Statistical optimization of medium for the production of recombinant hirudin from Saccharomyces cerevisiae using response surface methodology.* Process Biochemistry, 2000. **35**(7): p. 639-647.
159. Mailliet, J., et al., *Dwelling in the dark: procedures for the crystallography of phytochromes and other photochromic proteins.* Acta Crystallographica Section D: Biological Crystallography, 2009. **65**(11): p. 1232-1235.
160. Sparta, K.M., et al., *XDSAPP2. 0.* Journal of Applied Crystallography, 2016. **49**(3).
161. Strong, M., et al., *Toward the structural genomics of complexes: crystal structure of a PE/PPE protein complex from Mycobacterium tuberculosis.* Proceedings of the National Academy of Sciences, 2006. **103**(21): p. 8060-8065.
162. Mueller, U., et al., *Facilities for macromolecular crystallography at the Helmholtz-Zentrum Berlin.* Journal of synchrotron radiation, 2012. **19**(3): p. 442-449.
163. Gräslund, S., et al., *The use of systematic N-and C-terminal deletions to promote production and structural studies of recombinant proteins.* Protein expression and purification, 2008. **58**(2): p. 210-221.
164. Goldschmidt, L., et al., *Toward rational protein crystallization: A Web server for the design of crystallizable protein variants.* Protein Science, 2007. **16**(8): p. 1569-1576.
165. Shen, H., et al., *Light-induced phosphorylation and degradation of the negative regulator PHYTOCHROME-INTERACTING FACTOR1 from Arabidopsis depend upon its direct physical interactions with photoactivated phytochromes.* The Plant Cell, 2008. **20**(6): p. 1586-1602.
166. Kumar, I., et al., *Evolutionary divergence of phytochrome protein function in Zea mays PIF3 signaling.* Journal of experimental botany, 2016: p. erw217.
167. Al-Sady, B., et al., *Photoactivated phytochrome induces rapid PIF3 phosphorylation prior to proteasome-mediated degradation.* Molecular cell, 2006. **23**(3): p. 439-446.
168. Perez-Iratxeta, C. and M.A. Andrade-Navarro, *K2D2: estimation of protein secondary structure from circular dichroism spectra.* BMC structural biology, 2008. **8**(1): p. 1.
169. Greenfield, N.J., *Using circular dichroism spectra to estimate protein secondary structure.* Nature protocols, 2006. **1**(6): p. 2876-2890.
170. Jiao, Y., et al., *Conservation and divergence of light-regulated genome expression patterns during seedling development in rice and Arabidopsis.* The Plant Cell, 2005. **17**(12): p. 3239-3256.
171. Hershey, H.P., et al., *Nucleotide sequence and characterization of a gene encoding the phytochrome polypeptide from Avena.* Gene, 1987. **61**(3): p. 339-348.
172. von Horsten, S., et al., *Mapping light-driven conformational changes within the photosensory module of plant phytochrome B.* Scientific Reports, 2016. **6**.
173. Kneip, C., et al., *Effect of chromophore exchange on the resonance Raman spectra of recombinant phytochromes.* FEBS letters, 1997. **414**(1): p. 23-26.

174. Schmidt, P., et al., *Chromophore-protein interaction controls the complexity of the phytochrome photocycle.* Journal of Photochemistry and Photobiology B: Biology, 1996. **34**(1): p. 73-77.
175. Shin, A.-Y., et al., *Evidence that phytochrome functions as a protein kinase in plant light signalling.* Nature communications, 2016. **7**.
176. Neduva, V. and R.B. Russell, *Linear motifs: evolutionary interaction switches.* FEBS letters, 2005. **579**(15): p. 3342-3345.
177. Badger, J., et al., *The c-Abl Tyrosine Kinase Adopts Multiple Active Conformational States in Solution.* Biochemistry, 2016.
178. Wang, F. and X.-j. Tang, *Conformational heterogeneity and stability of apomyoglobin studied by hydrogen/deuterium exchange and electrospray ionization mass spectrometry.* Biochemistry, 1996. **35**(13): p. 4069-4078.
179. Volkman, B.F., et al., *Two-state allosteric behavior in a single-domain signaling protein.* Science, 2001. **291**(5512): p. 2429-2433.
180. Rockwell, N.C. and J.C. Lagarias, *The structure of phytochrome: a picture is worth a thousand spectra.* The Plant Cell, 2006. **18**(1): p. 4-14.
181. Sawaya, M.R., *Methods to refine macromolecular structures in cases of severe diffraction anisotropy.* Structural Genomics: General Applications, 2014: p. 205-214.
182. Weik, M., et al., *Specific chemical and structural damage to proteins produced by synchrotron radiation.* Proceedings of the National Academy of Sciences, 2000. **97**(2): p. 623-628.
183. Scheer, H., U. LINSENMEIER, and C. Krauss, *Studies on Plant Bile Pigments, II. Chemical and Photochemical Oxygenation of a Phytochrome P r Chromophore Model Pigment to Purpurins.* Hoppe-Seyler´ s Zeitschrift für physiologische Chemie, 1977. **358**(1): p. 185-196.
184. Moon, J., et al., *PIF1 directly and indirectly regulates chlorophyll biosynthesis to optimize the greening process in Arabidopsis.* Proceedings of the National Academy of Sciences, 2008. **105**(27): p. 9433-9438.
185. Woodson, J.D., J.M. Perez-Ruiz, and J. Chory, *Heme synthesis by plastid ferrochelatase I regulates nuclear gene expression in plants.* Current Biology, 2011. **21**(10): p. 897-903.
186. Woodson, J.D., et al., *Ubiquitin facilitates a quality-control pathway that removes damaged chloroplasts.* Science, 2015. **350**(6259): p. 450-454.
187. Duanmu, D., et al., *Retrograde bilin signaling enables Chlamydomonas greening and phototrophic survival.* Proceedings of the National Academy of Sciences, 2013. **110**(9): p. 3621-3626.
188. Sayou, C., et al., *A promiscuous intermediate underlies the evolution of LEAFY DNA binding specificity.* Science, 2014. **343**(6171): p. 645-648.
189. Merchant, S.S., et al., *The Chlamydomonas genome reveals the evolution of key animal and plant functions.* Science, 2007. **318**(5848): p. 245-250.
190. Chaiwanon, J., et al., *Information Integration and Communication in Plant Growth Regulation.* Cell, 2016. **164**(6): p. 1257-1268.

7 Table of Figures

8 List of Tables

9 Abbreviation used in the text

2D	2 dimensional
3D	3 dimensional
Å	Angstrom (= 0.1 nm)
A	anti conformation of a single bond in a bilin
Aa	Amino acids
Abs.	Absorbance
AC	Adenylyl cyclase
AnPixJ	CBCR from Anabaena (Nostoc) sp.
Amp	Ampicillin
AmS	Ammonium sulfate
APS	Ammonium persulfate
Bathy-BphP	Bathy-bacteriophytochrome
BLAST	Basic Local Alignment Tool
Bp	Base pairs
BSA	Bovine serum albumin
BV	Biliverdin
ca	Circa
CBCR	Cyanobacteriochrome
CCD	Charge-coupled device -detector
CD	Circular dichroism spectroscopy
C-di-GMP	Cyclic diguanylate
CDS	Coding sequence
Chla	Chlorophyll a
COP1	Constitutive photomorphogenic 1
cR	continuous red light
CV	Column volume
DHp	Dimerisation and histidine phosphorylation domain
DNA	Deoxyribonucleic acid
DTT	Dithiothreitol.
E	*trans* configuration of a double bond in a bilin.

EDTA	Ethylenediaminetetraacetic acid
e.g.	*exempli gratia*
EM	electron microscopy
ESRF	European synchrotron radiation facility (Grenoble, France)
etc.	*et cetera*
et al.	*et alii*
Fig.	Figure
FL	Full length
FR	Far-red light (λmax = 720 nm± 30 nm)
FTIR	Fourier-transform-Infra-red-spectroscopy
f/w	forward (primer)
FWHM	Full width at half maximum
GAF	cGMP phosphodiesterase/adenylyl cyclase/FhlA domain
G	Gram
h	Hours
H-bond	Hydrogen bond
HIR	High irradiance response
HKRD	Histidine kinase related domain
HO	Heme oxygenase
HOS	2-helix-output sensor
HY2	biliverdin reductase
HY5	Long hypocotyls transcription factor
i.e.	*id est*
IEX	Ion-exchange chromatography.
IMAC	Immobilized-metal-affinity-chromatography
IPTG	isopropyl 1-thio-□-D-galactopyranoside
Kan	Kanamycin
kDa	Kilo Dalton
Λ	Wavelength
λ_{max}	Wavelength of maximal absorbance
LB	Lysogeny broth
LED	Light-emitting diode

LFR	Low fluence response
LOV	Light oxygen voltage domain
MAS	Magic-angle spinning
MES	morphlino-4-ethanesulfonic
Milli-Q	Deionized and purified water (18.2 MΩ.cm)
Mg	Microgram
Mg	Milligram
Min	Minute
NDPK2	Nucleoside phosphate dikinase 2
Ni-NTA	nickel nitrilotriacetic acid
NLS	Nuclear import signal
Nm	Nanometer
NMR	Nuclear magnetic resonance
NPGP	NTE-PAS-GAF-PHY
NTE	N-terminal extension
OD	Optical density
OPM	output module
ORF	Open Reading Frame
ΦF	Fluorescence quantum yield
PAS	Period/Arnt/Singleminded domain
PCB	Phycocyanobilin
PcyA	Ferredoxin oxidoreductase
PCR	Polymerase chain reaction
PDB	Protein data bank
PEB	Phycoerythrobilin.
PEG	Polyethylene glycol
Pfr	Phytochrome in its far-red light absorbing state
PHY	Phytochrome-specific domain
phyA	holophytochrome A
phyB	holophytochrome B
pI	Isoelectric point.
PIFs	Phytochrome interacting factors

PIL	PIF3-like protein
PKS1	phytochrome kinase substrate 1
PMSF	phenylmethylsulfonyl fluoride
PΦB	Phytochromobilin
PPIXa	protoporphrin IXa
Pr	Phytochrome in its red light absorbing state
PRD	PAS related domain
PSM	photosensing module
PYP	Photoactive yellow protein
QM/MM	Quantum mechanics (QM)/molecular mechanics (MM)
R	Red light (λ_{max} = 660 ± 20 nm)
RE	resitriction endonuclease
Rev	reverse (primer)
RT	Room temperature (21°C)
RR	Resonance Raman spectroscopy
S	*syn* conformation of a single bond in a bilin
ssNMR	Solid-state-NMR
SAR	Specific absorbance ratio
SAXS	Small-angle X-ray scattering
SDS	Sodium dodecyl sulphate
SEC	Size-exclusion chromatography
SER	Surface entropy reduction
TB	Terrific broth medium
TBE	Tris Borate EDTA buffer
T-DNA	transfer DNA
TE	Tris EDTA buffer
TEMED	Tetramethylethylenediamine
TES-β	Tris-EDTA-Salt-β-mecaptoethanol
TLM	Transmitter-like module
Tm	Melting temperature of a primer (°C).
TRIS	Tri(hydroxymethyl)-aminomethane
TSI-I10	Tris-Salt-IDA-Imidazole

T(K)	Temperature in Kelvin
UV-VIS	Ultraviolet-visible spectroscopy
VLFR	Very low fluence response
Vol	Volume
v/v	volume per volume
WT	Wild-type
w/o	Without
Z	*cis* configuration of a double bond in a bilin.

10 Appendix

10.1 Synthetic DNAs

10.1.1 SbNPGP_ synthetic

ggtacc*gaattc*attaaagaggagaaattaa**ccATGG**CTTCTGGTTCTCGTGCTACCCCGACCCGTTCTCCGTCTTCTGCTCGTCCGGAAGCTCC
GCGTCACGCTCACCATCACCATCACCATCACTCTCAGTCTTCTGGTGGTTCTACCTCTCGTGCTGGTGGTGGCGGTGGTGGCGGTGG
TGGCGGTGGTGGCACCGCTGCTACCGCTACCGCTACCGCTACCGCTACCGAATCTGTTTCTAAAGCTGTTGCTCAGTACACCCTGGACGCTCG
TCTGCACGCTGTTTTCGAACAGTCTGGTGCTTCTGGTCGTTCTTTCGACTACTCTCAGTCTCTGCGTGCTCCGCCGACCCCGTCTTCTG
AACAGCAGATCGCTGCTTACCTGTCTCGTATCCAGCGTGGTGGTCACATCCAGCCGTTCGGTTGCACCCTGGCTGTTGCTGACGACT
CTTCTTTCCGTCTGCTGGCTTTCTCTGAAAACGCTGCTGACCTGCTGGACCTGTCTCCGCACCACTCTGTTCCGTCTCTGGACTCTGCT
GCTCCGCCGCCGGTAAGCCTGGGTGCTGACGCTCGTCTGCTGTTCTCTCCGTCTTCTGCTGTTCTGCTGGAACGTGCGTTCGCTGCTC
GTGAAATCTCTCTGCTGAACCCGCTGTGGATCCACTCTCGTGTTTCTTCTAAACCGTTCTACGCTATCCTGCACCGTATCGACGTTGGT
GTTGTTATCGACCTGGAACCGGCT**CGTACCGAAGACCC**GGCTCTGTCTATCGCTGGTGCTGTTCAGTCTCAGAAACTGGCTGTTCGT
GCTATCTCTCGTCTGCAGGCTCTGCCGGGTGGTGACATCAAACTGCTGTGCGACACCGTTGTTGAACACGTTCGTGAACTGACCGGT
TACGACCGTGTTATGGTTTACCGCTTCCACGAAGACGAACACGGTGAAGTTGTTGCTGAATCTCGTCGTGACAACCTGGAACCGTAC
CTGGGTCTGCACTACCCGGCTACCGACATCCCGCAGGCTTCTCGTTTCCTGTTCCGTCAGAACCGTGTTCGTATGATCGCTGACTGCC
ACGCTACCCCGGTTCGTGTTATCCAGGACCCGGGTATGTCTCAGCCGCTGTGCCTGGTTGGTTCTACCCTGCGTGCTCCGCACGGTT
GCCACGCTCAGTACATGGCTAACATGGGTTCTATCGCTTCTCTGGTTATGGCTGTTATCATCTCTTCTGGTGGTGACGACGAACAGAC
CGGTCGTGGTGGTATCTCTTCTGCTATGAAACTGTGGGGTCTGGTTGTTTGCCACCACACCTCTCCGCGTTGCATCCCGTTCCCGCTG
CGTTACGCTTGCGAATTTCTGATGCAGGCGTTCGGTCTGCAGCTGAACATGGAACTGCAGCTGGCTCACCAGCTGTCTGAAAAACAC
ATCCTGCGTACCCAGACCCTGCTGTGCGACATGCTGCTGCGTGACTCTCCGACCGGTATCGTTACCCAGTCTCCGTCTATCATGGACC
TGGTTAAATGCGACGGTGCTGCTCTGTACTACCACGGTAAATACTACCCGCTGGGTGTTACCCCGACCGAATCTCAGATCAAAGACA
TCATCGAATGGCTGACCGTTTGCCACGGTGACTCTACCGGTCTGTCTACCGACTCTCTGGCTGACGCTGGTTACCTGGGTGCTGCTG
CTCTGGGTGACGCTGTTTGCGGTATGGCTGTTGCTTACATCACCCCGTCTGACTACCTGTTCTGGTTCCGTTCTCACACCGCTAAAGA
AATCAAATGGGGTGGTGCTAAACACCACCCGGAAGACAAAGACGACGGTCAGCGTATGCACCCGCGTTCTTCTTTCAAAGCGTTCCT
GGAAGTTGTTAAATCTCGTTCTCTGCCGTGGGAAAACGCTGAAATGGACGCTATCCACTCTCTGCAGCTGATCCTGCGTGACTCTTTC
CGTGACGCTGCTGAAGGTACCTCTAACTCTAAAGCTATCGTTAACGGTCAGGTTCAGCTGGGT**CATCATCATCATCATCACTAAAAG**
CTTCCCTAC

10.1.2 Sb.PIF1_synthetic

GAATTCA**TTAAAGAGGAGAAA**AC**CATATG**CATCATCATCATCACCACAACCAGTTTGTGCCGGATTGGAGCAACATGGGCGATACA
AGTCGCCCTCTGGGCGAAGATGACGACCTGATCGAACTGCTGTGGTGCAATGGCCATGTGGTTATGCAGAGCCAGACCCATCGTAA
AGTTCCGCCGCGTCCGGAGAAACAGGCCGCAGTTGTTGCACCTGCCCCTCCGGCACCTGCAAGTGTTCCGCAGGAAGATGAGGGCG
GCCTGTGGTTTCCGTTCGCACTGGCCGACAGCCTGGACAAGGACATTTTTAGCGAATTTTTCTATGAAGCACCGGCCCCTGCAGCCG
TTGCAGCAGCAGCAGCACCGGTGGCACCGGTTGCAAGTGGTGGCGTGGGTACCGAAACCGACGCAGGTAAGAGCTGTCGTGATGA
CGTTCCGGCAGGTGCAGATCGCCGTGGCGCCTGTCCTGTTGTGAGCGAAGACCCGTGTGATCTGATGCCTCCGCCGAAAAGTACCC
CTGCAAGCTGTAGCCGCCAGCAGACCATGAGCCTGCCGAATGGTGGTGGTGATAACGCCGGTGATCTGAGCGATCTGGTTGTGCGT
GCAGGCAGTGCCGGCAAAAGTGCAGCCGCAGCCGCAGCAGTTGCAGAAGCAGGCGCCAGTAGTATGCTGAGCGCAATCGGTAGC
AGCATTTGCGGCAGCAATCAGGTGCTGGTTCAGCGTGCAGTTGGTGCACCGGGCCGTGCCAGTGGTAGTGGCAGTGGTACCGCCA
ACGCCAATGCTATGGGTGGTGGTCGCGGCAACGAAGCAAGCAGCAGTGGTCGTAGTACCTACTGTTTCGGTACCGCAACAACCACA
ACCACCACCACAACCACCACAGAACCTACCAGTACCAGCAACCGCAGCAGTAAGCGCAAACGCCTGGACACCGAGGATAGCGAAA
GTCCGAGTGAAGACGCCGAGAGCGAAAGCGCAGCCATGCTGGCACGTAAGCCGCCGCAGAAGATGACCACCGCACGTCGCAGCC
GCGCAGCAGAAGTTCATAATCTGAGCGAACGCCGTCGCCGCGATCGCATTAACGAAAAAATGCGCGCCCTGCAGGAACTGATCCCG
CACTGCAACAAAACCGACAAGGCCAGCATGCTGGATGAAGCCATTGAATATCTGAAAAGTCTGCAACTGCAGGTGCAAATGATGTG
GATGGGCAGCGGCATTGCAGCCCCTCCGGCAGTTATGTTTCCGGGCGTGCACCAGTATCTGAGCCGCATGGGCGTGGGTATGGGTC
CTGCAGCCGCCATGCCGAGTATGCCGCGTCTGCCGTTTATGGCAGCCCCTCAACCGGTGGTTCCGCCGAATGCACAGGTGAATCCG
GTGCCGGGTTATCGCGGCCATCATCACATGCCGGCCGCCGTTGGTATGGCAGAGCCGTATGGCCACTACCTGGGCGTGAATCACTT
ACAGCCGCCGCCGAGTCAGCATTATGCCCAGGGCGTTGGCTATTATCCGCCGCCTCTGGGCGCCAAAGCAGTTCAGCAGCAGGCCC
CGGAACTGCATCATGTTCCGGGTCCTGGCGCAAGCATGCCTGCAGGTGCAGGTGCAGCAGCCCCTGGCGTTCTGCTGCCGGAAAGT
GCACCTAGCCGTGGTCCGGGTTAAAA**CTCGAG**GATCC

10.2 Proteins

10.2.1 NTE of different phytochromes

>PHYB_ARATH
MVSGVGGSGGGRGGGRGGEEEPSSSHTPNNRRGGEQAQSSGTKSLRPRSNTESMSKAIQQYTVDARLHAVFEQSGESGKSFDYSQSLK
TTTYGSSVPEQQIT
>D7L2U2_ARALL
MVSGVGGSGRGGEAEPSSSHTTNNRRGGEQAQSSGTKSLRPRESMSKAIQQYTVDARLHAVFEQSGESGKSFDYSQSLKTTTYGSSVPE
QQIT
>Q5G889_ARALY
GVGGSGRGGEAEPSSSHTTNNRRGGEQAQSSGTKSLRPRESMSKAIQQYTVDARLHAVFEQSGESGKSFDYSQSLKTTTYGSSVPEQQIT
>M4E095_BRARP
MVSGVGGGGGSGGSGRGGGRGGEESSSTHRREQAQSSGTKSLRPQSQPQTESISKAIQQYTVDARLHAVFEQSGESGRSFDYSQSLKTT
TYGSSVPEQQIT
>C4TGE0_9BRAS
MVSGGGGGSGSARGGDGGEEASSSHRVDNILHTTHPRREQAQSSGTKSLRPQQNQPQSHTVSMSKAIQQYTVDARLHAVFEQSGGSG
KSFDYSQSLKTTTYGSSVPEQQIT
>C4TGD9_CARNO
MVSGGGGGSGSARGGGGGEEASSSHRVDNILHTTHPRREQAQSSGTKSLRPQQNQPQSHTVSMSKAIQQYTVDARLHAVFEQSGGSG
KSFDYSQSLKTTTYGSSVPEQQIT
>B9U4G7_VITRI
MSSGNRGTQSHHQAQSSGTSNLRVYHTDSMSKAIAQYTMDARLHAVYEQSGESGKSFDYSQSVRTTTQSVPEQQIT
>F6H723_VITVI
MSSGNRGTQSHHQAQSSGTSNLRVYHTDSMSKAIAQYTMDARLHAVYEQSGESGKSFDYSQSVRTTTQSVPEQQIT
>B9RZR1_RICCO
MASGGRNSQQQQQQQQQQRYVHQPTTAQSSGTSNLRAHNTESMSKAIAQYTVDAQLHAVFEQSGGSGKSFDYSQSVRTTNQSIAEQ
QIT
>PHYB_SOLTU
MASGSRTKHSHHSSSQAQSSGTSNVNYKDSISKAIAQYTADARLHAVFEQSGESGKFFDYSQSVKTTTQSVPERQIT
>PHYB_TOBAC
MASGSRTKHSHQSGQGQVQAQSSGTSNVNYKDSISKAIAQYTADARLHAVFEQSGESGKSFDYSQSIKTTTQSVVPEQQIT
>O24117_NICPL
MASGSRTKHSHQSGQGQGQVQAQSSGTSNVNYKDSISKAIAQYTADARLHAVFEQSGESGKSFDYSQSVKTTTQSVVPEQQIT
>B9HIN5_POPTR
MASQSQRQSNQRQHQNQAAQSSGTSNMRQHHHATESVSKAIAQYTVDAQLHAVFEQSGGSGKSFDYSQSVRTTSQSVPEEQIT
>J7MDP2_FRAAN
MASASGKAAAAQSSSGTSNFRSAGQQQNRNSTAAAESVSKAVAQYTVDARLHAVFEQSESGKSFDYSQSMRSTKDSVPEKQIT
>PHYB1_SOLLC
MASGSRTKHSYHNSSQGQAQSSGTSNMNYKDSISKAIAQYTADARLHAVFEQSGESGKSFDYSQSVKTTTQSVPERQIT
>G7IFW3_MEDTR
MASTSRGEKTEPNTSSSSSQRRHQLQITTTKEPEQQNNNNNKSMKKAIAQYTEDARLHAVFEQSGDSFDYSQSIRLTTAAHSQSVPEQQI
T
>G1FMD6_POPTN
MASQSQRQSNQRQHQNQAAQSSGTSNMRQHHHATESVSKAIAQYTVDAQLHAVFEQSGGSGKSFDYSQSVRTTSQSVPEEQIT
>Q9FPQ2_POPTR
MASQSQRQSNQPVHNQAQSSGTSNMRQHHHATESVSKAIAQYTVDAQLHAVFEQSGGTGRSFDYSKSVRTTNQSVPEQQIT
>Q334B4_POPTN
MASQSQRQSNQPAHNQAQSSGTSNMRQHHHPTESVSKAIAQYTVDAQLHAVFEQSGGTGKSFDYSKSVRTPNQSVPEEQIT
>Q334C1_POPTN
MASQSQRQSNQPAHNQAQSSGTSNMRQHHHPTESVSKAIAQYTVDAQLHAVFEQSGGTGKSFDYSKSVRTPNQS
>Q9SWS6_SOLLC
MASGSGSRGKHDRNHQPKNQSQFSGTSNTNALSKAVAQYTTDARLHAAFEQSGESGKNFDYSQSVRNSTESVTEHQIT
>Q334D4_POPTN
MASQSQRQSNQPAHNQAQSSGTSNMRQHHHPTESVSKAIAQYTVDAQLHAVFEQSGGTGKSFDYSKSVRTPNQSVPEEQIT
>Q334B1_POPTN
MASQSQRQSNQPAHNQAQSSGTSNMRQHHHPTESVSKAIAQYTVDAQLHAVFEQSGGTGKSFDYSKSVRTPNQSVPEEQIT
>Q334A2_POPTN

MASQSQRQSNQPAHNQAQSSGTSNMRQHHHPTESVSKAIAQYTVDAQLHAVFEQSGGTGKSFDYSKSVRTPNQS
>Q333Z4_POPTN
MASQSQRQSNQPAHNQAQSSGTSNMRQHHHPTESVSKAIAQYTVDAQLHAVFEQSGGTGKSFDYSKSVRTPNQSVPEEQIT
>R4J043_9ROSI1
MASQSQRQSNQPVHNQAQSSGTSNMRQHHNATESVSKAIAQYTVDAQLHAVFEQSGGTGRSFDYSKSVRTTNQSVPEQQIT
>R4IZB2_9ROSI
MASQSQRQNNQPVHNQAQSSGTSNMRQHHHATESVSKAIAQYTVDAQLHAVFEQSGGTGRSFDYSKSVRTTNQSVPEQQIT
>M1CLI4_SOLTU
MASGSGSRGKYGRNHQPKNQSQFSVTSNNALSKAVAQYTTDARLHAAFEQSGESFDYSQSVRNTTESVTEHQITSFDYSKSVRTTNQSV
PEQQIT
>R4IZI9_9ROSI
MASQSQRQSNQPVHNQAQSSGTSNMRQHHHATESVSKAIAQYTVDAQLHAVFEQSGGTGRSFDYSKSVRTTNQSVPEQQIT
>Q334C3_POPTN
MASQSQRQSNQPAHNQVQSSGTSNMRQHHHPTESVSKAIAQYTVDAQLHAVFEQSGGTGKSFDYSKSVRTPNQSVPEEQIT
>R4IZN9_9ROSI
MASQSQRQSNQPVHNQAQSSGTSNMRQHHHATESVSKAIAQYTVDAQLHAVFEQSGGTGRSFDYSKSVRTTNQSVPEQQIT
>R4IZA2_9ROSI
MASQSQRQSNQPVHNQAQSSGTSNMRQHHHATESVSKAIAQYTVDAQLHAVFEQSGGTGRSFDYSKSVRTTN
>Q333Z2_POPTN
MASQSQRQSNQPAHNQAQSSGTSNMRQHHHPTESVSKAIAQYTVDAQLHAVFEQSGGTGKSFDYSKSVRTPNQSVPEEQIT
>Q334C7_POPTN
MASQSQRQSNQPAHNQAQSSGTSNMRQHHHPTESVSKAIAQYTVDAQLHAVFEQSGGTGKSFDYSKSVRTPNQS
>PHYB_SOYBN
MLQQAERRIPPFRRRKSTPHEQRLSHHSSNNNNNIDSMSKAIAQYTEDGVHAVFEQSGESGRSFNYSESIRIASESVPEQQIT
>N0DKN4_LOTJA
MASASRRGGGGGGSGSGSVVERGKEEEKLKRGGGGGGGESVMMRKAIAQYTEDARLHAVYEQSGESFDYSHSLRVTVESVPEQQIT
>B9S180_RICCO
MGFQNARERGTTTLSSSAASSMKPFATNSENTATIAQYNADAGLLAEFEQSGVSGKSFNYSRSVLSAPHNVPEEQIT
>PHYB_ORYSI
MASGSRATPTRSPSSARPAAPRHQHHHSQSSGGSTSRAGGGGGGGGGGGGGGAAAAESVSKAVAQYTLDARLHAVFEQSGASGRSFDY
TQSLRASPTPSSEQQIA
>I3RUI1_ORYRU
MASGSRATPTRSPSSARPAAPRHQHHHSQSSGGSTSRAGGGGGGGGGGGGGGAAAAESVSKAVAQYTLDARLHAVFEQSDASGRSFDY
TQSLRASPTPSSEQQIA
>I3RUE1_ORYSJ
MGSGSRATPTRSPSSARPAAPRHQHHHSQSSGGSTSRAGGGGGGGGGGGGGGAAAAESVSKAVAQYTLDARLHAVFEQSGASGRSFDY
TQSLRASPTPSSEQQIA
>I3RUC7_ORYSJ
MASGSRATPTRSPSSARPAAPRHQHHHSQSSGGSTSRAGGGGGGGGGGGGGGAAAAESVSKAVAQYTLDARLHAVFEQSGASGRSFDY
TQSLRASPTPSSEQQIA
>I3RUE5_ORYSJ
MGSGSRATPTRSPSSARPAAPRHQHHHSQSSGGSTSRAGGGGGGGGGGGGGGAAAAESVSKAVAQYTLDARLHAVFEQSGASGRSFDY
TQSLRASPTPSSEQQIA
>I3RUG1_9ORYZ
MASGSRATPTRSPSSARPAAPRHQHHHSQSSGGSTSRAGGGGGGGGGGGGGGAAAAESVSKAVAQYTLDARLHAVFEQSGASGRSFDY
TQSLRASPTPSSEQQIT
>I3RUC3_ORYSJ
MASGSRATPTRSPSSARPAAPRHQHHHSQSSGGSTSRAGGGGGGGGGGGGGGAAAAESVSKAVAQYTLDARLHAVFGQSGASGRSFD
YTQSLRASPTPSSEQQIA
>I1PAN5_ORYGL
MASGSRATPTRSPSSARPAAPRHQHHHSQSSGGSTSRAGGGGGGGGGGGGAAAAESVSKAVAQYTLDARLHAVFEQSGASGRSFDYTQ
SLRASPTPSSEQQIA
>I3RUD3_ORYSJ
MGSGSRATPTRSPSSARPAAPRHQHHHSQSSGGSTSRAGGGGGGGGGGGGGGAAAAESVSKAVAQYTLDARLHAVFEQSGASGRSFYY
TQSLRASPTPSSEQQIA
>Q6S527_SORBI
MASGSRATPTRSPSSARPEAPRHAHHHHHHHSQSSGGSTSRAGGGGGGGGGGGGTAATATATATESVSKAVAQYTLDARLHAVFEQS
GASGRSFDYSQSLRAPPTPSSEQQIA

>Q6S514_SORBI
MASGSRATPTRSPSSARPEAPRHAHHHHHNHSQSSGGSTSRAGGGGGGGGGGGGTAATATATATESVSKAVAQYTLDARLHAVFEQS
GASGRSFDYSQSLRAPPTPSSEQQIA
>Q6S513_SORBI
MASGSRATPTRSPSSARPEAPRHAHHHHHHSQSSGGSTSRAGGGGGGGGGGGGTAATATATATESVSKAVAQYTLDARLHAVFEQSG
ASGRSFDYSQSLRAPPTPSSEQQIA
>Q6XFQ3_MAIZE
MASGSRATPTRSPSSARPEAPRHAHHHHHSQSSGGSTSRAGGGAAATESVSKAVAQYTLDARLHAVFEQSGASGRSFDYSQSLRAPPTP
SSEQQIA
>F2DN15_HORVD
MASGSRATPTPTRSPSSARPSAPHQHHTQSSGGSTSRAGGGGGGAAGSAAATESVSKAVAQYTLDAGLHAVFEQSGASGRSFDYSQSLL
APPTPSSEQQIA
>A9JR06_WHEAT
MASGSRATPTRSPSSARPAAPHQNHTQSSGGSTSRAGGGGGGAAGSAAATESVSKAVAQYTLDAGLHAVFEQSGASGRSFDYSQSLLA
PPTPSSEQQIA
>Q6XFQ2_MAIZE
MASDSRPPKRSPSARRVAPRHAHHHHSQSSGGSTSRAGAGGGGGGAAATESVSKAVAQYNLDARLHAVFEQSGASGRSFDYSQSLRAP
PTPSSEQQIA
>K4A524_SETIT
MASGSRATPTRSPSSARPAAPRHAHHHHHSQSSGGSTTSQTGGGGGGGGGGGGAATAATESVSKAVAQYNLDARLHAVFEQSGASG
RSFDYSQSLRAPPTPSSEQQIA
>M0ZFB2_HORVD
MASGSRATPTPTRSPSSARPSAPHQHHTQSSGGSTSRAGGGGGGAAGSAAATESVSKAVAQYTLDAGLHAVFEQSGASGRSFDYSQSLL
APPTPSSEQQIA

10.2.2 Sb.phyB derived constructs

>Sb.phyB_NPGP
MASGSRATPTRSPSSARPEAPRHAHHHHHHHSQSSGGSTSRAGGGGGGGGGGGGTAATATATATESVSKAVAQYTLDARLHAVFEQS
GASGRSFDYSQSLRAPPTPSSEQQIA**AYLSRIQRGGHIQPFGCTLAVADDSSFRLLAFSENAADLLDLSPHHSVPSLDSAAPPPVSLGADA
RLLFSPSSAVLLERAFAAREISLLNPLWIHSRVSSKPFYAILHRIDVGVVIDLEPA**RTEDPALSIAGAVQSQKLAVRAISRLQALPGG**DIKLLC
DTVVEHVRELTGYDRVMVYRFHEDEHGEVVAESRRDNLEPYLGLHYPATDIPQASRFLFRQNRVRMIADCHATPVRVIQDPGMSQPL
CLVGSTLRAPHGCHAQYMANMGSIASLVMAVIISSGGDDEQTGRGGISSAMKLW*GLVVCHH*TSPRCIPFPLRYACEFLMQAFGLQLN
MELQLAHQLSEKH**ILRTQTLLCDMLLRDSPTGIVTQSPSIMDLVKCDGAALYYHGKYYPLGVTPTESQIKDIIEW*LTVCHGD*STGLSTDS
LADAGYLGAAALG*DAVCGMA*VAYITPSDYLFWFRSHTAKEIKWGGAKHHPEDKDDGQRMHPRSSFKAFLEVVKSRSLPWENAEMD
AIHSLQLILRDSFRDAAEGT*SNSKAIVNGQVQLGHHHHHH***
>Sb.phyB_HPGP
MTESVSKAVAQYTLDARLHAVFEQSGASGRSFDYSQSLRAPPTPSSEQQIA**AYLSRIQRGGHIQPFGCTLAVADDSSFRLLAFSENAADLL
DLSPHHSVPSLDSAAPPPVSLGADARLLFSPSSAVLLERAFAAREISLLNPLWIHSRVSSKPFYAILHRIDVGVVIDLEPA**RTEDPALSIAGA
VQSQKLAVRAISRLQALPGG**DIKLLCDTVVEHVRELTGYDRVMVYRFHEDEHGEVVAESRRDNLEPYLGLHYPATDIPQASRFLFRQNR
VRMIADCHATPVRVIQDPGMSQPLCLVGSTLRAPHGCHAQYMANMGSIASLVMAVIISSGGDDEQTGRGGISSAMKLWGLVVCHH
TSPRCIPFPLRYACEFLMQAFGLQLNMELQLAHQLSEKH**ILRTQTLLCDMLLRDSPTGIVTQSPSIMDLVKCDGAALYYHGKYYPLGVTP
TESQIKDIIEWLTVCHGDSTGLSTDSLADAGYLGAAALGDAVCGMAVAYITPSDYLFWFRSHTAKEIKWGGAKHHPEDKDDGQRMH
PRSSFKAFLEVVKSRSLPWENAEMDAIHSLQLILRDSFRDAAEGT*SNSKAIVNGQVQLGHHHHHH***
>Sb.phyB_PGP
**MAAYLSRIQRGGHIQPFGCTLAVADDSSFRLLAFSENAADLLDLSPHHSVPSLDSAAPPPVSLGADARLLFSPSSAVLLERAFAAREISLL
NPLWIHSRVSSKPFYAILHRIDVGVVIDLEPA**RTEDPALSIAGAVQSQKLAVRAISRLQALPGG**DIKLLCDTVVEHVRELTGYDRVMVYRF
HEDEHGEVVAESRRDNLEPYLGLHYPATDIPQASRFLFRQNRVRMIADCHATPVRVIQDPGMSQPLCLVGSTLRAPHGCHAQYMAN
MGSIASLVMAVIISSGGDDEQTGRGGISSAMKLWGLVVCHHTSPRCIPFPLRYACEFLMQAFGLQLNMELQLAHQLSEKH**ILRTQTLLC
DMLLRDSPTGIVTQSPSIMDLVKCDGAALYYHGKYYPLGVTPTESQIKDIIEWLTVCHGDSTGLSTDSLADAGYLGAAALGDAVCGMA
VAYITPSDYLFWFRSHTAKEIKWGGAKHHPEDKDDGQRMHPRSSFKAFLEVVKSRSLPWENAEMDAIHSLQLILRDSFRDAAEGT*SN
SKAIVNGQVQLGHHHHHH***
>Sb.phyB_PG
**MAAYLSRIQRGGHIQPFGCTLAVADDSSFRLLAFSENAADLLDLSPHHSVPSLDSAAPPPVSLGADARLLFSPSSAVLLERAFAAREISLL
NPLWIHSRVSSKPFYAILHRIDVGVVIDLEPA**RTEDPALSIAGAVQSQKLAVRAISRLQALPGG**DIKLLCDTVVEHVRELTGYDRVMVYRF
HEDEHGEVVAESRRDNLEPYLGLHYPATDIPQASRFLFRQNRVRMIADCHATPVRVIQDPGMSQPLCLVGSTLRAPHGCHAQYMAN
MGSIASLVMAVIISSGGDDEQTGRGGISSAMKLWGLVVCHHTSPRCIPFPLRYACEFLMQAFGLQLNMELQLAHQLSEKHHHHHHH**

10.2.3 bHLHs from representative PIFs

>ZmPIF3
SERRRRDRINEKMRALQELIPNCNKIDKASMLDEAIEYLKTLQLQVQ
>ZmPIF1
SERRRRDRINEKMRALQELIPHCNKTDKASMLDEAIEYLKSLQLQVQ
>ZmPIF6
SERKRRDRINQKMKTLQKLVPNSNKTDKASMLDEVIDYLKQLQAQVQ
>ZmPIF8
SERKRRDRINQKMKTLQKLVPNSNKTDKASMLDEVIDYLKQLQAQVQ
>SlPIF1
SERKRRDRINEKMKALQELIPCCNKSDKASMLDEAIEYLKSLQLQVQ
>SlPIF4
SERRRRDRINEKMKALQELLPHSTKTDKASMLDEAIEYLKSLQMQLQ
>SlPIF7
SERRRRDRINQRMKALQRLVPNASKTDKASMLDEVIDYLKQLQAQVQ
>SlPIF8
SERKRRDKINQRLKTLQKLVPTSSKTDTASMLDEVIEYLKQLRAQVK
>AtPIF1
SERKRRDRINERMKALQELIPRCNKSDKASMLDEAIEYMKSLQLQIQ
>AtPIF3
SERRRRDRINEKMRALQELIPNCNKVDKASMLDEAIEYLKSLQLQVQ
>AtPIF4
SERRRRDRINERMKALQELIPHCSKTDKASILDEAIDYLKSLQLQLQ

>AtPIF5
SERRRRDRINERMKALQELIPHCSRTDKASILDEAIDYLKSLQMQLQ
>AtPIF6
PERNQRNDINKKMRTLQNLLPNSHKDDNESMLDEAINYMTNLQLQVQ
>AtPIF7
SERRRRDRINQRMRTLQKLLPTASKADKVSILDDVIEHLKQLQAQVQ
>SbPIF3
SERRRRDRINEKMRALQELIPNCNKIDKASMLEEAIEYLKTLQLQVQ
>SbPIF1
SERRRRDRINEKMRALQELIPHCNKTDKASMLDEAIEYLKSLQLQVQ
>SbPIF3i2
SERRRRDRINEKMRALQELIPNCNKIDKASMLDEAIEYLKTLQLQVQ
>SbPIF4
SERRRRDRINEKMRALQELIPHCNKTDKASILDETIEYLKSLQMQVQ
>AtPIF8
SERKRRDKINQRMKTLQKLVPNSSKTDKASMLDEVIEYLKQLQAQVS
>SbPIF5
SERRRRDRINEKMRALQELVPHCNKTDKASILDEAIEYLKSLQMQVQ
>KfPIF1
SERRRRDRINDRMRALQELIPNSNKTDKASVLDEAIEYLKNLQTQLQ
>MpPIF1
SERRRRDRINEKMKALQELIPNSNKTDKASMLDEAIEYLKMLQLQLQ

10.2.4 PIFs in *Sorghum*

>Sb.PIF1
MNQFVPDWSNMGDTSRPLGEDDDLIELLWCNGHVVMQSQTHRKVPPRPEKQAAVVAPAPPAPASVPQEDEGGLWFPFALADSLDK
DIFSEFFYEAPAPAAVAAAAAPVAPVASGGVGTETDAGKSCRDDVPAGADRRGACPVVSEDPCDLMPPPKSTPASCSRQQTMSLPNGG
GDNAGDLSDLVVRAGSAGKSAAAAAAVAEAGASSMLSAIGSSICGSNQVLVQRAVGAPGRASGSGSGTANANAMGGGRGNEASSSGR
STYCFGTATTTTTTTTTTEPTSTSNRSSKRKRLDTEDSESPSEDAESESAAMLARKPPQKMTTARRSRAAEVHNLSERRRRDRINEKMRALQ
ELIPHCNKTDKASMLDEAIEYLKSLQLQVQMMWMGSGIAAPPAVMFPGVHQYLSRMGVGMGPAAAMPSMPRLPFMAAPQPVVPP

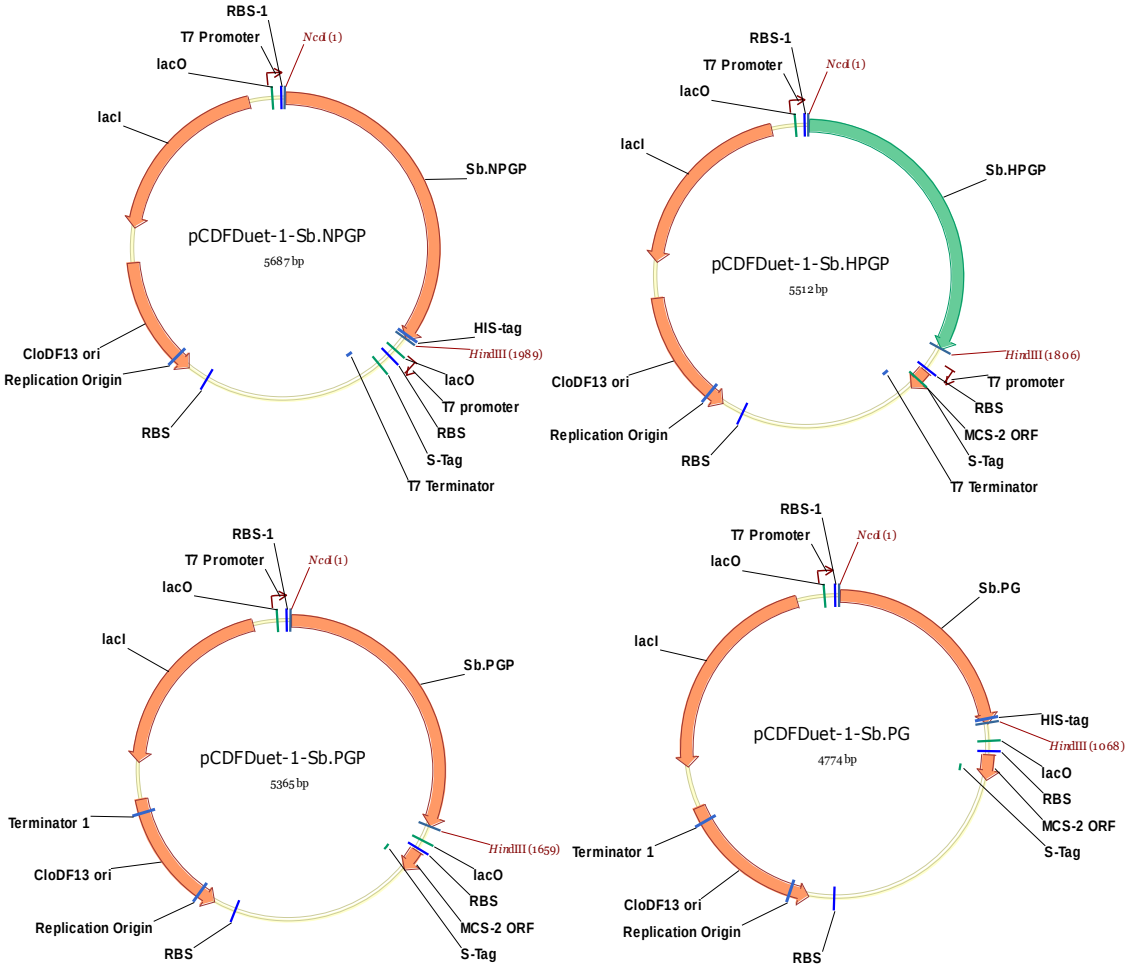
NAQVNPVPGYRGHHHMPAAVGMAEPYGHYLGVNHLQPPPSQHYAQGVGYYPPPLGAKAVQQQAPELHHVPGPGASMPAGAGAAA
PGVLLPESAPSRGPG
>Sb.PIF3
MGVRYLGASGSRKWLLAMDGGERPATTSHKAPHPFRRPDGELVELLWQDGAVVAHSQAQAHHHHHHHYRRPLVQVAAGNTGASGV
TGEQAPALPWLPCSGGAMGGDVYSQLWQSIAQADGRVVGADAVAGAPRPPARSGNSGVGSSRTAGQEVGSSFSGSNLVAAAMHLD
DDIDDVGVAAALPVPLDDPATATGAGAGASTSSGWNSNAPLHKRSRDEFDEDADLDTVDETPPSSRDRRPASNKRRTRAAEVHNMSER
RRRDRINEKMRALQELVPHCNKTDKASILDEAIEYLKSLQMQVQIMWMSTGMAPMMIPGAHQLMPPMTMGLNSARMPPPAVQFLS
QMQRVPPHFMNNPLLNQMPQMLPPPTNAPNVTDQAQSNRMALPRNPFLHPNDSDALTTPHQVPSLFGYGPQMVQVNEIQELLTST
AAPALGTDLPSSSDGTGA
>Sb.PIF3i2
MSDSNEFAELLWENGQAVVHGRRKQTQAAFPPFTCGAASSSRAQEKQPGTDPVALLKTGLFGDFSSGLDATRDNVDLDDTVPWINYPI
VEEDTAAPALAESYSPDFFSELHAAAVAAAATNLSALPPPVQHTGNNNNRSTPVATTSRGPEPAKETHRIPIAGPATRPEPQAEFAATRKP
RPEGGGEGLMNFSLFSRPAAMARASLQRPPQTGTDKASNVTTSTRVESTVLQSASGQRTAPVFADQRTPWSQSKEVRFSCAAAPTTAT
AGNLQQEMPRDMTPQKKVETRKASEVAVATSSVCSGNGAGIGNDESWRQQKRKSQAECSASQDDDLDDDSGGMRRSGSRGTKRSRT
AEVHNLSERRRRDRINEKMRALQELIPNCNKIDKASMLDEAIEYLKTLQLQVQMMSMGSGLCIPPMLLPPTMQHLQIPPIAHFPHLGMG
LGYGMGVLDMNSTAAVPFPPMPGAHFPCSMIPGTPPQGLGMPGRNTMPMFGVPGQAIHPSASSIQPFPSLAGLPVRPNLAPQASAV
MANMVQEQQQGATQQQQSLNNEAQQGANTGDPQLQTTVQAENQHFSVPYSAQTESNQFLDGGGNRTDTTGRNEAET
>Sb.PIF4
MDGNARSSAANQKKPIVADDDLIELLWHNGSVVAQPQAHHRPAPAPAPPPSDRDRPAGASGLTGEETAAWFPDTLDDALEKDLYTHL
WYSTIADAAPQHEGTLLATPQPSPPPPLPVGSSGVESSWAGDICSTFCGSNQVPRTPAGTSRKDAALQSAGTSGAGGHDGTSSSGGSGS
NYGGSGLPSDSVHVHKRKGRCRDDSDSPSEDAECEEATEETKPSRRHGPKRRTRAAEVHNLSERRRRDRINEKMRALQELIPHCNKTDKA
SILDETIEYLKSLQMQVQIMWMTSGMAPMMFPGAHQFMPPMALGMNSGCIPAAQGLSQMPRLPYMNHPLPNHIPLNSSPAMNPLN
AANQMMQNGHLREASNHFLHLDGRTAVVPQVPGPHVYGHQIAQAEEPNSILEVAASTVVPTSRAGQPPSFDGV
>Sb.PIF5
MPVSISKICRTGPGDELAELLWDHGPALRRPPPPFQPFTCSAAGSSRSHELKRHPVPVPVPLGMHTHGAGLGLGGLPVHDDEDDAVPWL
HCPVVVDDGDADTAPLPPEYCASLLSEYSEAAAPAATASHAAAVVPAAPPPPEAAVANKLAPPSATGSGVGEGVLNFTFFSRPLQRPQAA
SAAASHPVESTVVQAATNRLRSTPLFSEQRMAWLQPPKAPRATAAAAATPPPPPPPPAPPQLAPMLPDTRHGETVAQPRSQPDARPP
DAAAVTASSVCSGNGDRSQLKRSCHLAADCSVSPDEDMDDEPGATRRSAARSAKRSRTAEVHNMSERRRRDRINEKMRALQELIPNCN
KIDKASMLEEAIEYLKTLQLQVQMMSMGTGLCMPPAAMLLPAMHQHMQALHHHHHHPMAHFPHLGMGLGYGMGMGMGATAAA
GFDMLPRVAAGAHFPCPPMAMPPPPGAAMFGVPGTGMATLPAAFAHMAGAAPAGQMEAGAGDAAPAAAAAAAAAAPARPEEAE
HPPVRGVAQGDQEAQQQQHP
>Sb.PIF6
MNQCVPSWDLDDPALVVAAGGGSGGGLNHHQVSTTIAGGAPHRGLQVSAGGGAFAPVVLPMSDQYYEVAELTWEKGNISSHGLGLL
NRPVPVQQHKYAPPPSSSQLQAIGGGGGGGSSGDRETLEAVVGEAATRSSHFLSQRHPVPWVGVGGAAVDAVARSAADALVPCAAAR
VHDEAAAAEGGDAGAGPGSRRKRARVVGEDGGLVCASQGSTAAAPGPRRRGESALLTLDYACCGTGGADDVCGFTTTTTNNSTSLERD
DKGSPDTENTSIGGGASDSRCFSRRSQASYL

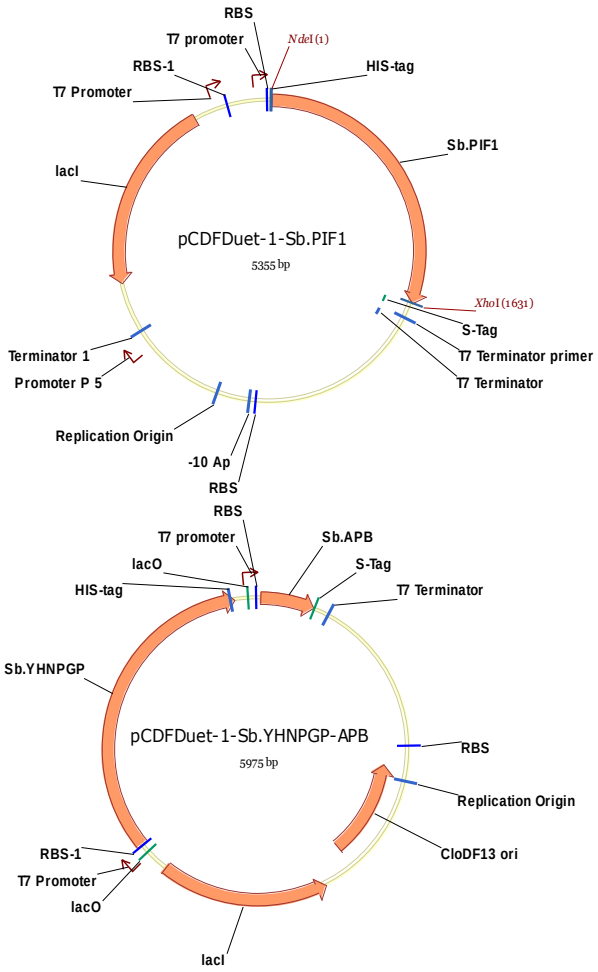
10.2.5 Sb.PIF1 derived constructs

>Sb.PIF1
MNQFVPDWSNMGDTSRPLGEDDDLIELLWCNGHVVMQSQTHRKVPPRPEKQAAVVAPAPPAPASVPQEDEGGLWFPFALADSLDK
DIFSEFFYEAPAPAAVAAAAAPVAPVASGGVGTETDAGKSCRDDVPAGADRRGACPVVSEDPCDLMPPPKSTPASCSRQQTMSLPNGG
GDNAGDLSDLVVRAGSAGKSAAAAAAVAEAGASSMLSAIGSSICGSNQVLVQRAVGAPGRASGSGSGTANANAMGGGRGNEASSSGR
STYCFGTATTTTTTTTTTEPTSTSNRSSKRKRLDTEDSESPSEDAESESAAMLARKPPQKMTTARRSRAAEVHNLSERRRRDRINEKMRALQ
ELIPHCNKTDKASMLDEAIEYLKSLQLQVQMMWMGSGIAAPPAVMFPGVHQYLSRMGVGMGPAAAMPSMPRLPFMAAPQPVVPP
NAQVNPVPGYRGHHHMPAAVGMAEPYGHYLGVNHLQPPPSQHYAQGVGYYPPPLGAKAVQQQAPELHHVPGPGASMPAGAGAAA
PGVLLPESAPSRGPG
>APB
MNQFVPDWSNMGDTSRPLGEDDDLIELLWCNGHVVMQSQTHRKVPPRPEKQAAVVAPAPPAPASVPQEDEGGLWFPFALADSLDK
DIFSEFFYEAPAPAAVAAAAAP
>LEA
MNQFVPDWSNMGDTSRPLGEDDDLIELLWCNGHVVMQSQTHRKVPPRPEKQAAVVAPAPPAPASVPQEDEGGLWFPFALADSLDK
DIFSEFFYEAPAPAAVAAAAAPVAPVASGGVGTETDAGKSCRDDVPAGADRRGACPVVSEDPCDLMPPPKSTPASCSRQQTMSLPNGG
GDNAGDLSDLVVRAGSAGKSAAAAAAVAEAGASSMLSAIGSSICGSNQVLVQRAVGAPGRASGSGSGTANANAMGGGRGNEASSSGR
STYCFGTATTTTTTTTTTEPTSTSNRSSKRKRLDTEDSESPSE
>bHLH

MNQFVPDWSNMGDTSRPLGEDDDLIELLWCNGHVVMQSQTHRKVPPRPEKQAAVVAPAPPAPASVPQEDEGGLWFPFALADSLDK
DIFSEFFYEAPAPAAVAAAAAPVAPVASGGVGTETDAGKSCRDDVPAGADRRGACPVVSEDPCDLMPPPKSTPASCSRQQTMSLPNGG
GDNAGDLSDLVVRAGSAGKSAAAAAAVAEAGASSMLSAIGSSICGSNQVLVQRAVGAPGRASGSGSGTANANAMGGGRGNEASSSGR
STYCFGTATTTTTTTTTTEPTSTSNRSSKRKRLDTEDSESPSEDAESESAAMLARKPPQKMTTARRSRAAEVHNLSERRRRDRINEKMRALQ
ELIPHCNKTDKASMLDEAIEYLKSLQLQVQMMWMG

10.2.6 Plasmid maps





10.3 Box-Behnken Design

	Induction OD	IPTG	Temperature	OD_{600} at harvest	Lysis buffer (ml)
No	X1	X2	X3		
1	0.5	0.5	28°C	2.56	6.4
2	1	0.5	28°C	1.96	4.9
3	0.5	2	28°C	2.56	6.4
4	1	2	28°C	1.48	3.7
5	0.5	1	18°C	1.84	4.6
6	1	1	18°C	1.8	4.5
7	0.5	1	37°C	3.6	9.0
8	1	1	37°C	3.04	7.6
9	0.75	0.5	18°C	1.76	4.4
10	0.75	2	18°C	1.76	4.4
11	0.75	0.5	37°C	3.2	8.0
12	0.75	2	37°C	3.6	9.0

13	0.75	1	28°C	3.2	8.0
14	0.75	1	28°C	2.72	6.8
15	0.75	1	28°C	2.64	6.8

10.4 Crystallization strategy

Screening

- 8 plates were screened using 15 mg/ml sbPGP protein in 5 mM Tris-HCl pH 7.8, 15 mM NaCl and 0.25 mM TCEP
- The screened plates were all from JCSG in sitting drop vapor diffusion format. Check the file Marxtal_SorBi_PhyB_PGP2

Results from Crystallization screening

Hit	Buffer	Major Precipitant	Salts	Others
1	0.1 M CHES pH 9.5	10% PEG 8000	0.2 M NaCl	
2	0.1 M CAPSO pH 9.5	30% PEG 400	0.1 M NaCl	0.1 M MgCl2
3	0.1 M Tris-HCl pH 8.5	30% PEG 4000	0.2 M MgCl2	
4	0.1 M HEPES pH 7.5	30% PEG 400	0.2 M MgCl2	
5	0.1 M HEPES pH 7.5	1.26 M Ammonium sulfate	0.2 M MgCl2	

1. Classification (CL plates)

- Preliminary studies to define the concentration of major precipitants for hanging-drop
- Grid screen 4 precipitants and 6 levels $4x6 = 24$ experiments

Results from preliminary classification

Buffer	Major Precipitant	Salts
0.1 M CHES pH 9.5	8% PEG 8000	0.2 M NaCl
0.1 M CAPSO pH 9.5	23% PEG 400	0.2 M MgCl2
0.1 M Tris-HCl pH 8.5	23% PEG 4000	
0.1 M HEPES pH 7.5	1.00 M Ammonium sulfate	

2. Dimension reduction (DR plates)

- Translation of sitting drop vapor diffusion into hanging-drop vapor diffusion by using complete factorial statistical design
- Factor screening at identified precipitant concentrations
- To this end 4 buffers, 4 major precipitants and 2 additives will be screened
- Total 32 experiments, each experiment in triplicate per cover slide
- $4x4x2 = 32$

Factor screening result: Identification of critical factors for translation

Experiment	Buffer	Major Precipitant	Salts
8	0.1 M CAPSO pH 9.5	8% PEG 8000	0.2 M NaCl

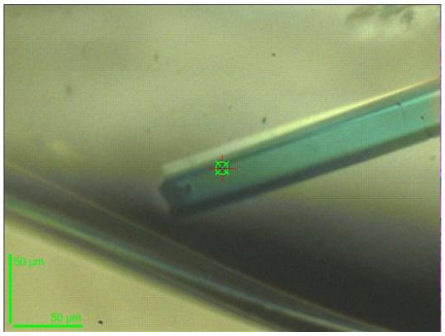
- Other results can be iteratively optimized here after

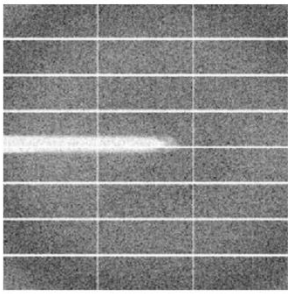
3. Optimization (OP plates)

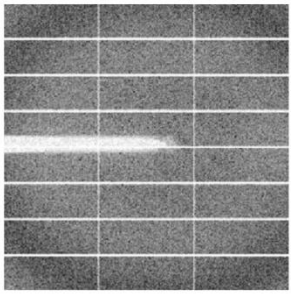
- Application of Box-Behnken statistical design (4 factors and 3 levels) to fine tune major factor levels
- To this end 3 blocks and 8 experiments per block are required
- $3x8 = 24$
- The center point is replicated separately: 27

Factor	-1	0	1
0.1 M CAPSO pH	9.00	9.25	9.5
PEG 8000	6%	7%	8%
NaCl	1 M	1.5 M	2 M
Protein	20 mg/ml	25 mg/ml	30 mg/ml

4. Data collection for PGP (Courtesy of Sophie Franz)







10.5 Accepted manuscript

Structural communication between the chromophore binding pocket and the N-terminal extension in plant phytochrome phyB

Francisco Velázquez Escobar[1]*, David Buhrke[1], Maria Fernandez Lopez[1], Sintayehu Manaye Shenkutie[2], Silke von Horsten[3], Lars-Oliver Essen[3,4], Jon Hughes[2], and Peter Hildebrandt[1]

1 Technische Universität Berlin, Institut für Chemie, Sekr. PC14, Straße des 17. Juni 135, D- 10623 Berlin, Germany

2 Plant Physiology, Justus-Liebig University Gießen, Senckenbergstr.3, D-35390 Giessen, Germany

3. Philipps-Universität Marburg, Fachbereich Chemie, Strukturbiochemie, Hans-Meerwein-Strasse 4, D-35032 Marburg, Germany. LOEWE Center for Synthetic Microbiology, Philipps-Universität, D 35032 Marburg, Germany

Correspondence

P. Hildebrandt, Technische Universität Berlin, Institut für Chemie, Sekr. PC14, Straße des 17.

Juni 135, D-10623 Berlin, Germany Fax: +49-30-31421122 Tel: +49-30-31421419

E-mail: hildebrandt@chem.tu-berlin.de

Abstract

The N-terminal extension (NTE) of plant phytochromes has been suggested to play a functional role in signaling photoinduced structural changes. Here, we use resonance Raman spectroscopy to study the effect of the NTE on the chromophore structure of B-type phytochromes from two evolutionarily distant plants. NTE deletion seems to have no effect on the chromophore in the inactive Pr state, but alters the torsion of the C-D ring methine bridge and the surrounding hydrogen bonding network in the physiologically active Pfr state. These changes are accompanied by a shift of the conformational equilibrium between two Pfr sub-states, which might affect the thermal isomerization rate of the C-D double bond and, thus, account for the effect of the NTE on the dark reversion kinetics.

Keywords

Plant phytochrome, chromophore structure, N-terminal extension, resonance Raman spectroscopy

Abbreviations

A-B/C-D, C=C stretching of the AB/CD methine bridge; ARR, Arabidopsis response regulator; GAF, cGMP phosphodiesterase/adenylyl cyclase/FhlA; HDX, H/D exchange with mass spectrometry; HOOP, hydrogen-out-of-plane; N-H ip, N-H in-plane bending; NTE, N-terminal extension; PAS, Per/Arnt/Sim; PCB, phycocyanobilin; PHY, phytochrome-specific; phyA/B, phytochrome A/B; PIF, phytochrome interaction factor; Pr/Pfr, red/far-red absorbing form of phytochrome; PSM, photosensory module; PΦB, phytochromobilin; RR, resonance Raman.

Introduction

Light-dependent development in plants is regulated principally by phytochromes, a family of photoreceptors harboring the tetrapyrrole phytochromobilin (PϕB) as the light-absorbing chromophore [1-3] (Fig. 1). A photoinduced reaction cascade, starting with a double bond isomerization of the cofactor, leads to the conversion of the red-adsorbing Pr state to the far-red absorbing Pfr state which is the physiologically active form. Important insight into the molecular processes of the photosensor module (PSM) has been obtained from bacterial and cyanobacterial representatives of this photoreceptor family [1-3]. However, despite similar reaction steps and structural changes in the chromophore binding pocket, the signaling

process by which the PSM transduces its structural changes to the cell is quite different in prokaryotic and plant phytochromes [4-6]. In particular, whereas most prokaryotic phytochromes are light-repressed histidine protein kinases signaling from the C-terminal transmitter module, plant phytochromes are active in the Pfr state, signaling from the PSM directly. The domain structure of plant phytochromes is also more complex, with an additional PAS-PAS module separating the PSM and the transmitter-like module. Moreover, whereas prokaryotic phytochromes carry an N-terminal extension (NTE) of only about 20 residues, in plant phytochromes this is much larger, comprising ca. 110 and 70-75 residues in B- and other types, respectively (Fig. 1). In phyB the NTE is likely to interact with the PSM since it affects the dark reversion from the Pfr to the Pr state [7]. Furthermore, recent H/D exchange probed by mass spectrometry (HDX) provided evidence for Pr/Pfr-dependent differences [8]. In Arabidopsis the NTE is bound by the response regulator-like protein ARR4 which also thereby regulates the rate of reversion [9,10]. Furthermore, it contains phosphorylation sites that affect the spectral characteristics as well as the binding of the phytochrome interaction factor 3 (PIF3) [11,12]. 3D structural information about the NTE and its interaction with the PSM are not yet available. Despite its limited resolution, the recently-determined PSM crystal structure for *Arabidopsis* phyB in the Pr state [7] reveals far-reaching similarities in particular to Cph1 from the cyanobacterium *Synechocystis 6803* [13] on account of their similar protein-chromophore linkage. However, crystallized variants of phyB and Cph1 are devoid of any structural information for NTE.

The present work focuses on the role of the NTE in phyB using resonance Raman (RR) spectroscopy, a method that selectively probes the structure of the chromophore and its interactions with the molecular environment [14-16]. By comparing phyB from evolutionarily distant plants, the dicotyledon *Arabidopsis thaliana* (AtphyB) and the monocotyledon *Sorghum bicolor* (SbphyB) and their deletion variants lacking the NTE (*At*NTE, *Sb*NTE) with phytochrome A3 from *Avena sativa* (*As*phyA), we provide insight into structural details of the chromophore binding pocket of plant phytochromes and the impact of the NTE on the photoinduced formation of the functional Pfr state.

Materials and Methods

Protein expression and purification. *At*phyB with and without the NTE was expressed, purified, and assembled with PCB as described previously [8]. The photosensory module *Sb*phyB with or without the NTE was produced from pCDFDuet-based constructs together with PφB or PCB in *E. coli* to form the holoprotein *in vivo* following procedures similar to those established for prokaryotic phytochromes [17, 18]. Both recombinant phytochromes carried a C-terminal His6-tag for Ni-affinity purification, followed by size-exclusion chromatography to give a Qy/protein absorbance ratio of ca. 1.4 and 1.15 for the variants with and without the NTE, respectively (Fig. S1). In the case of the PCB adducts, close similarities are apparent between the equivalent PSM constructs from *Sorghum* and *Arabidopsis* with and without the NTE. In the case of *Sb*phyB the native PφB adduct showed the expected bathochromic shifts in the red and far-red peaks of Pr and Pfr concomitant with the π electron system extending into the D-ring side chain.

Resonance Raman spectroscopy. Samples were concentrated in aqueous TS50 buffer (50 mM Tris-Cl, 300 mM NaCl, 5 mM EDTA, pH = 7.8). Pr-to-Pfr (Pfr-to-Pr) conversion was achieved by continuous (2 min, LED) light exposure at 670 nm (Pr) or 750 nm (Pfr) at 20 °C. RR measurements of phyB phytochrome were performed with an RFS-100 FT-Raman spectrometer (Bruker Optics), equipped with an 1064 nm Nd-YAG laser (670 mW) and a Linkam (Resultec) cryostat. Measurements were performed at −140 °C. Each spectrum is based on 1000 - 2000 single interferometer scans, corresponding to accumulation times between 30 - 60 minutes per spectrum. All spectra processing was performed using the OPUS suite. Further experimental details are given elsewhere [14-16].

Results

The RR spectra of the Pr and Pfr parent states of PφB and PCB adducts of *Avena* (oat) phyA3 (*As*phyA) have been analysed in detail previously [15,16] and now serve as a reference for the comparison with the corresponding data for phyB with and without the larger NTE. Here we focus on three spectral regions characteristic of the specific structural parameters of the chromophore and its interactions with the immediate molecular environment. In the region between 1520 and 1680 cm^{-1}, the spectra of the Pr state are dominated by three peaks originating from the N-H in-plane bending of rings B and C (N-H ip), and the C=C stretching modes of the CD and AB methine bridges (Fig. 2A; see Fig. 1C for ring notation). In A*s*-phyA, these modes are found at 1575 (N-H ip), 1626 (C-D), and 1646 cm^{-1} (A-B), reflecting the ZZZssa

geometry of the fully protonated PϕB chromophore (Fig. 1) [15]. Between 750 and 850, the spectra display less intense bands including the C-H out-of-plane deformation coordinates (HOOP). In *As*phyA, the most pronounced band at 802 cm^{-1} is a HOOP mode of the CD methine bridge and is accompanied by a shoulder at 795 cm^{-1}. Finally, the carbonyl stretching modes of the terminal pyrrole rings give rise to weak RR bands at 1712 and 1729 cm^{-1}. On the basis of previous IR studies [19], these bands are readily assigned to the C=O stretching of ring D and A, respectively.

The Pr state of phyB. The spectra of the PϕB adducts of SbphyB and *As*phyA (Fig. 2B) are nearly identical, indicating that even structural details of the chromophore and its interactions with the protein environment are closely similar. For *As*phyA, structural models revealed very similar chromophore geometries for PϕB and PCB, implying that the different substitution at ring D (vinyl vs. ethyl) does not perturb the interactions with the protein [15, 16,20]. Thus, the differences in the RR spectra between the *As*phyA adducts can mainly be attributed to the intrinsic effect of the side chain difference on the cofactor modes, including the frequency upshift of the C-D (1626 to 1637 cm^{-1}) and A-B (1646 to ca. 1650 $^{cm-1}$) stretching modes in the PCB adduct due to the reduced length of the delocalized π-electron system that is also reflected by the UV-vis absorption spectra (Fig. S1). The notable differences between the PCB adducts of *Sb*phyB and *As*phyA (Fig. 2C, E) are, therefore, quite surprising. In *Sb*phyB, a distinctly larger upshift of the C-D stretching to 1643 cm^{-1} is observed, such that the putatively invariant A-B stretching mode (ca. 1650 cm^{-1}) is largely obscured. This finding points to a structural change at the CD methine bridge leading to a decrease of the N(C)-C-C-C(D) dihedral angle, according to previous theoretical analyses [14,15]. Most likely, this distortion originates from the specific structure of the chromophore binding pocket in *Sb*phyB in the vicinity of ring D, which stabilizes a similar PϕB structure to that in *As*phyA but, unlike in the latter, small geometrical adjustments are necessary to accommodate PCB. Subtle structural alterations compared to *As*phyA seem to occur also in *At*phyB (Fig. 2D). Here, however, the C-D stretching remains unchanged whereas the 6-cm^{-1} upshift of the CD HOOP mode compared to *As*phyA may be taken as a small decrease of the C(C)-C-C-N(D) dihedral angle. Thus, we conclude that the chromophore structure differs only in small details between oat phyA and the two phyB orthologs. These small differences also include the hydrogen bonding network as reflected by minor differences between the C=O stretching frequencies of the rings A and D. In the Pr state,

the RR spectra show no significant differences between the *Sb*phyB and AtPhyB variants in regard to the presence of an NTE.

The Pfr state of phyB. Both the Pr and the Pfr state of *As*phyA display a structural heterogeneity in the chromophore binding site [21, 16], but RR spectroscopy only distinguishes a temperature-dependent equilibrium between sub-states in the case of Pfr (Pfr-I, Pfr-II) [16]. Each of these Pfr substates adopts a protonated ZZEssa chromophore structure. However, the torsion between the rings C and D probably differs by ca. 10^{o} as implied by two bands originating from the HOOP mode, separated by 16 and 8 cm^{-1} in the PϕB and PCB adducts, respectively (Figs. 4A, 3A). The intensity ratio of these bands can be taken as a measure for the relative populations of the sub-states. This conformational heterogeneity is also reflected in the region between 1600 and 1620 cm^{-1}, albeit less clearly visible due to the superposition of four components, i.e., for each conformer a strongly RR-active C=C stretching mode of the CD methine bridge at lower frequency and a C=C stretching mode of ring D a higher frequency, which exhibits weaker RR intensity [16]. Based on isotopic labelling and temperature-dependent measurements, it could be shown that Pfr-I is associated with the low frequency HOOP mode and the high frequency C-D stretching component, whereas the high-frequency HOOP and low-frequency C-D stretching originate from Pfr-II. This interpretation may also be extended to the *Sb*phyB and *At*phyB variants including the NTE, which show a similar HOOP mode ratio and only small adduct-related frequency differences (Figs. 3, 4). Also the other marker bands between 1500 and 1800 cm^{-1} are similar within the experimental accuracy, and the most pronounced frequency shifts refers to the ring D C=O stretching, which is lowered by 7 cm^{-1} in the PCB adducts of the phyB PSM.

The effect of the NTE on the chromophore binding pocket. Unlike Pr, the deletion of the NTE segment has a remarkable effect on the RR spectra of Pfr. For the PCB adducts, we note a downshift of the high-frequency HOOP component from 816 cm^{-1} in *Sb*phyB (*At*phyB: 814 cm^{-1} in) to 811 cm^{-1}, accompanied by a reversal of the intensity ratio between the low and high-frequency components, which reflects an increase of the Pfr-I/Pfr-II population ratio (Fig. 3B – E). Accordingly, this shift of the conformational equilibrium is accompanied by a growth of the high-frequency C-D stretching component (vide supra), which can in fact account for the intensity increase of the 1618-cm^{-1} shoulder. Concomitant to the redistribution among the

sub-state population, we note a small downshift of the N-H ip mode from 1556 to 1554 cm^{-1}, indicating an alteration of the hydrogen bond interactions in the chromophore pocket [22]. The same interpretation may hold for the 5-cm^{-1} upshift of the C=O stretching of ring D from 1705 to 1710 cm^{-1}. Interestingly, upon deletion of the NTE, the RR spectra become very similar to that of the PSM of cyanobacterial phytochrome Cph1 2, including the distribution among the conformational sub-states (Fig. 3F, Fig. 4D) [16, 23]. In contrast, such far-reaching similarities with Cph1 2 do not exist for the RR spectra of the Pr state, neither for phyA nor for phyB [15, 22, 23]. For the PΦB adducts the effect of NTE deletion is smaller than for the PCB adducts, but with respect to the changes of the sub-state populations the same tendency can be derived from intensity re-distribution of the composite peak of the two HOOP modes (Fig. 4B, C).

Discussion

The present results demonstrate state-dependent interaction between the chromophore and the NTE in phyB. This interaction is similar for phyB from *A. thaliana* and *S. bicolor*, suggesting that it represents a general characteristic of phyB. The impact of these interactions on the chromophore binding pocket is small and only observed in the Pfr state. It comprises a perturbation of the hydrogen bonding network and the CD portion of the chromophore, reflected by a shift of the conformational Pfr-I/Pfr-II equilibrium. Both effects, which most likely constitute the molecular basis for the hypsochromic shift of the electronic absorption in NTE-deletion variants (Fig. S1) [7, 8], are presumably linked although no conclusions can be drawn about possible cause-effect relationships between them. The NTE undergoes structural changes upon Pr → Pfr photoconversion as demonstrated by recent HDX experiments [8]. These studies identified several regions within the NTE that display a slower exchange in Pfr (compared to Pr), indicating a restructuring of the NTE.

One region close to the AB rings of the chromophore that displays changes in the HDX rates links the C-terminus of the NTE with the start of the PAS domain (Y104-I108). However, the altered dynamics of this region does not affect the geometry at the AB rings of the chromophore upon Pr → Pfr transition, consistent with the RR data. All other regions in the NTE with HDX rates altered by photoconversion probably lack direct contact with the chromophore itself [8]. This implies that only indirect interactions are able to explain the observed effect of the NTE on the structural changes of the bilin chromophore. Our packing

model for the NTE/PSM interactions suggested that in the Pfr state a stretch with comparably slow intrinsic HDX rates (P28-Q36) may interact with the β1 strand of the GAF domain and its adjacent loop [8]. A prominent part of this strand is Y276, which forms crucial interactions with the chromophore and which upon mutation to histidine forms a strongly-fluorescent, Pfr-mimicking signaling state in darkness [24]. Interactions between the β1GAF strand and the NTE could hence affect the CD ring surroundings and account for the RR spectroscopic differences between the wildtype and NTE deletion variants. A second candidate for an indirect interaction site for the NTE is the tongue region of the PHY domain, which is thought to undergo major structural changes next to the chromophore upon photoconversion. Overall, it appears that indirect interactions of the CD ring surroundings with the NTE suffice to affect the chromophore geometry and thereby its photochemistry. The conformational coupling of the NTE segment and the chromophore binding pocket are probably of functional importance. It has been shown previously that dark reversion kinetics of phyB is substantially accelerated upon removal of the NTE segment [7]. This may in fact be related to the structural differences compared to the complete PSM, reflected by the shift of the conformational equilibrium between the sub-states Pfr-I/Pfr-II which mainly differ by CD methine bridge torsion and the hydrogen bond network around the chromophore. This conformational heterogeneity has only been observed for cyanobacterial and prototypical bacteriophytochromes that, like plant phytochromes, thermally revert to the Pr state. In contrast, bathy bacteriophytochromes exhibit a highly homogeneous chromophore structure in the Pfr state that does not decay to the Pr state [14]. It was therefore concluded that thermal CD methine bridge isomerization, which is presumably the rate-limiting step for dark reversion, runs via one of the Pfr sub-states. In phyB, NTE-deletion increases the population of the sub-state associated with a larger C(14)-C(15)-C(16)-N(D) dihedral angle [14], which might lower the energy barrier for the rotation around the C(15)-C(16) bond as one important parameter promoting E → Z isomerization. Also the changes of the hydrogen bond interactions induced by NTE-deletion might contribute to the acceleration of the dark reversion via favoring the transient formation of the enol form of the chromophore as the essential mechanistic step of thermal double bond isomerization [25].

Acknowledgements

The authors acknowledge financial support by the Deutsche Forschungsgemeinschaft (SFB1078 B6 to PH, HU702/8 and /9 to JH and ES152/10 to LOE).

Author Contributions statement:

PH, LOE, and JH conceived and supervised the study and wrote the manuscript; FVE, DB, and MFL performed the spectroscopic experiments; SMS and SvH prepared the samples.

References

1 Briggs WR, Spudich JL (2005). *Handbook of Photosensory Receptors.* Wiley Verlag: Weinheim, Germany.

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

3 Auldridge ME, Forest KT (2011). Bacterial phytochromes: more than meets the light. *Crit. Rev. Biochem. Mol. Biol.* **46**, 67–88.

4 Zhang J, Stankey RJ, Vierstra RD (2013). Structure-Guided Engineering of Plant Phytochrome B with Altered Photochemistry and Light Signaling. *Plant Physiol.* **161**, 1445–1457.

5 Burgie ES, Vierstra RD (2014). Phytochromes: an atomic perspective on photoactivation and signaling. *Plant Cell* **26**, 4568–4583.

6 Hughes J. (2013). Phytochrome Cytoplasmic Signaling. *Ann. Rev. Plant Biol.* **64**, 377– 402.

7 Burgie ES, Bussell AN, Walker JM, Dubiel K, Vierstra RD (2014). Crystal structure of the photosensing module from a red/far-red light-absorbing plant phytochrome. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 10179–10184.

8 von Horsten S, Straß S, Hellwig N, Gruth V, Klasen R, Mielcarek A, Linne U, Morgner N, Essen LO (2016). Mapping light-driven conformational changes within the photosensory module of plant phytochrome B. *Sci. Rep.* **6**, 34366.

9 Sweere U, Eichenberg K, Lohrmann J, Mira-Rodado V, Bäurle I, Kudla J, Nagy F, Schäfer E, Harter K (2001). Interaction of the response regulator ARR4 with phytochrome B in modulation red light signaling. *Science* **294**, 1108–1111.

10 Fankhauser C (2002). Light perception in plants: cytokinins and red light join forces to keep phytochrome B active. *Trends Plant Sci.* **7**, 143–145.

11 Medzihradszky M, Bindics J, Ádám E, Viczián A, Klement E, Lorrain S, Gyula P, Mérai Z, Fankhauser C, Medzihradszky KF, Kunkel T, Schäfer E, Nagy F (2013). Phosphorylation of Phytochrome B Inhibits Light-Induced Signaling via Accelerated Dark Reversion in Arabidopsis. *Plant Cell* **25**, 535–544

12 Kazumasa N, Wong CCL, Yates III JR, Chory J (2013). Tyrosine Phosphorylation Regulates the Activity of Phytochrome Photoreceptors. *Cell Rep.* **3**, 1970–1979.

13 Essen LO, Maillet J, Hughes J (2008). The structure of a complete phytochrome sensory module in the Pr ground state. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 14709–14714.

14 Salewski J, Velazquez F, Kaminski S, von Stetten D, Keidel A, Rippers Y, Michael N, Scheerer P, Piwowarski P, Bartl F, Frankenberg-Dinkel N, Ringsdorf S, Gärtner W, Lamparter T, Mroginski MA, Hildebrandt P (2013). The structure of the biliverdin cofactor in the Pfr state of bathy and prototypical phytochromes. *J. Biol. Chem.* **288**, 16800–16814.

15 Mroginski M.A, Kaminski S, von Stetten D, Ringsdorf S, Gärtner W, Essen LO, Hildebrandt P (2011). The structure of the chromophore binding pocket in the Pr state of plant phytochrome phyA. *J. Phys. Chem. B* **115**, 1220–1231.

16 Velazquez Escobar F, von Stetten D, Günther M, Keidel A, Michael N, Yang Y, Heyne K, Lamparter T, Essen LO, Hughes J, Gärtner W, Mroginski MA, Hildebrandt P (2015). Conformational heterogeneity of the Pfr chromophore in plant and cyanobacterial phytochromes. *Front. Mol. Biosci.* 2, 37.

17 Landgraf FT, Forreiter C, Hurtado Picó A, Lamparter T, Hughes J (2001). Recombinant holophytochrome in *Escherichia coli*. *FEBS Lett.* **508**, 459-462

18 Song C, Psakis G, Kopycki J, Lang C, Matysik J, Hughes J (2013) The D-ring, not the A-ring, rotates in Synechococcus OS-B' phytochrome. *J. Biol. Chem.* **289**, 2552-2562.

19 Schwinté P, Foerstendorf H, Gärtner W, Mroginski MA, Hildebrandt P, Siebert F (2008). Fourier transform infrared studies of the photoinduced processes of phytochrome phyA using isotopically labelled chromophores and density functional theory calculations. *Biophys. J.* **95**, 1256–1267.

20 Kneip C, Mozley D, Hildebrandt P, Gärtner W, Braslavsky SE, Schaffner K (1997). Effect of Chromophore Exchange on the Resonance Raman Spectra of Recombinant Phytochromes. *FEBS Lett.* **414**, 23–26.

21 Rohmer, T, Lang, C, Hughes J, Essen LO, Gärtner, W, Matysik, J (2008) Light-induced chromophore activity and signal transduction in phytochromes observed by ^{13}C and ^{15}N magic-angle spinning NMR. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 15229–15234.

22 Mroginski MA, von Stetten D, Velazquez Escobar F, Strauss HM, Kaminski S, Scheerer P, Günther M, Murgida DH, Schmieder P, Bongards C, Gärtner W, Maillet J, Hughes J, Essen LO, Hildebrandt P (2009). Chromophore structure of cyanobacterial phytochrome Cph1 in the Pr state: reconciling structural and spectroscopic data by QM/MM calculations. *Biophys. J.* **96**, 4153–4163.

23 Velazquez Escobar F, Lang C, Takiden A, Schneider C, Balke J, Hughes J, Alexiev U, Hildebrandt P, Mroginski MA (2017). Protonation-dependent structural heterogeneity in the chromophore binding site of cyanobacterial phytochrome Cph1 2. *J. Phys. Chem. B* 121, 47–57.

24 Su Y, Lagarias JC (2007). Light-Independent Phytochrome Signaling Mediated by Dominant GAF Domain Tyrosine Mutants of Arabidopsis Phytochromes in Transgenic Plants. *Plant Cell* 19, 2124–2139.

25 Velazquez Escobar F, Piwowarski P, Salewski J, Michael N, Fernandez Lopez M, Rupp A, Qureshi BM, Scheerer P, Bartl F, Frankenberg-Dinkel N, Siebert F, Mroginski MA, Hildebrandt P (2015). A protonation-coupled feedback mechanism controls the signaling process in bathy phytochromes. *Nat. Chem.* 7, 423–430.

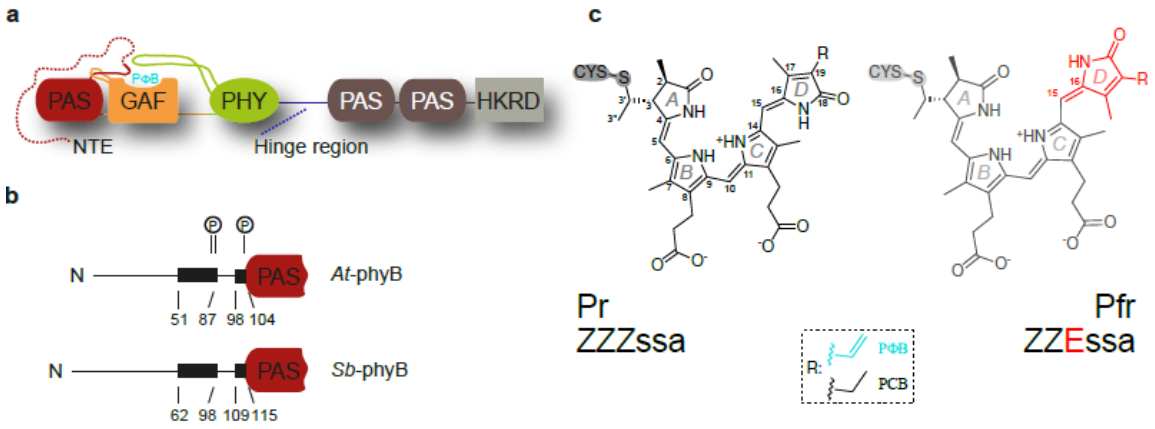


Fig. 1. (a) Domain organisation of phyB; PΦB denotes the chromophore binding site (adapted from ref. 8), (**b**) NTEs in *Sb*-phyB and *At*-phyB ("P" refers to putative phosphorylation sites and the black boxes mark conserved regions), **(c)** structure of the bilin chromophores PΦB (R = vinyl) and PCB (R = ethyl) in the *ZZEssa* configuration of the Pfr state.

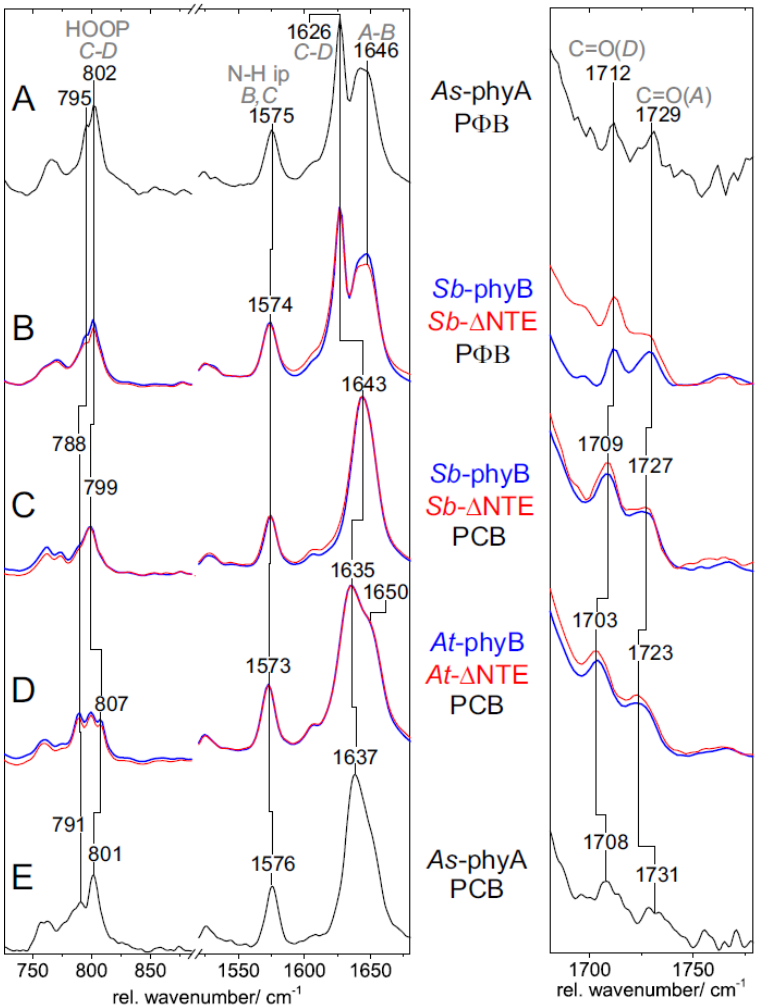


Fig. 2. RR spectra of the Pr state of **(A)** As-phyA and **(B)** *Sb*-phyB including the PΦB chromophore, and **(C)** *Sb*-phyB, **(D)** *At*-phyB, and **(E)** *As*-phyA including the PCB chromophore. The red traces in **(B)**, **(C)**, and **(D)** refer to the variants lacking the NTE segments (NTE). The spectra in **(A)** and **(E)** have been published earlier [15].

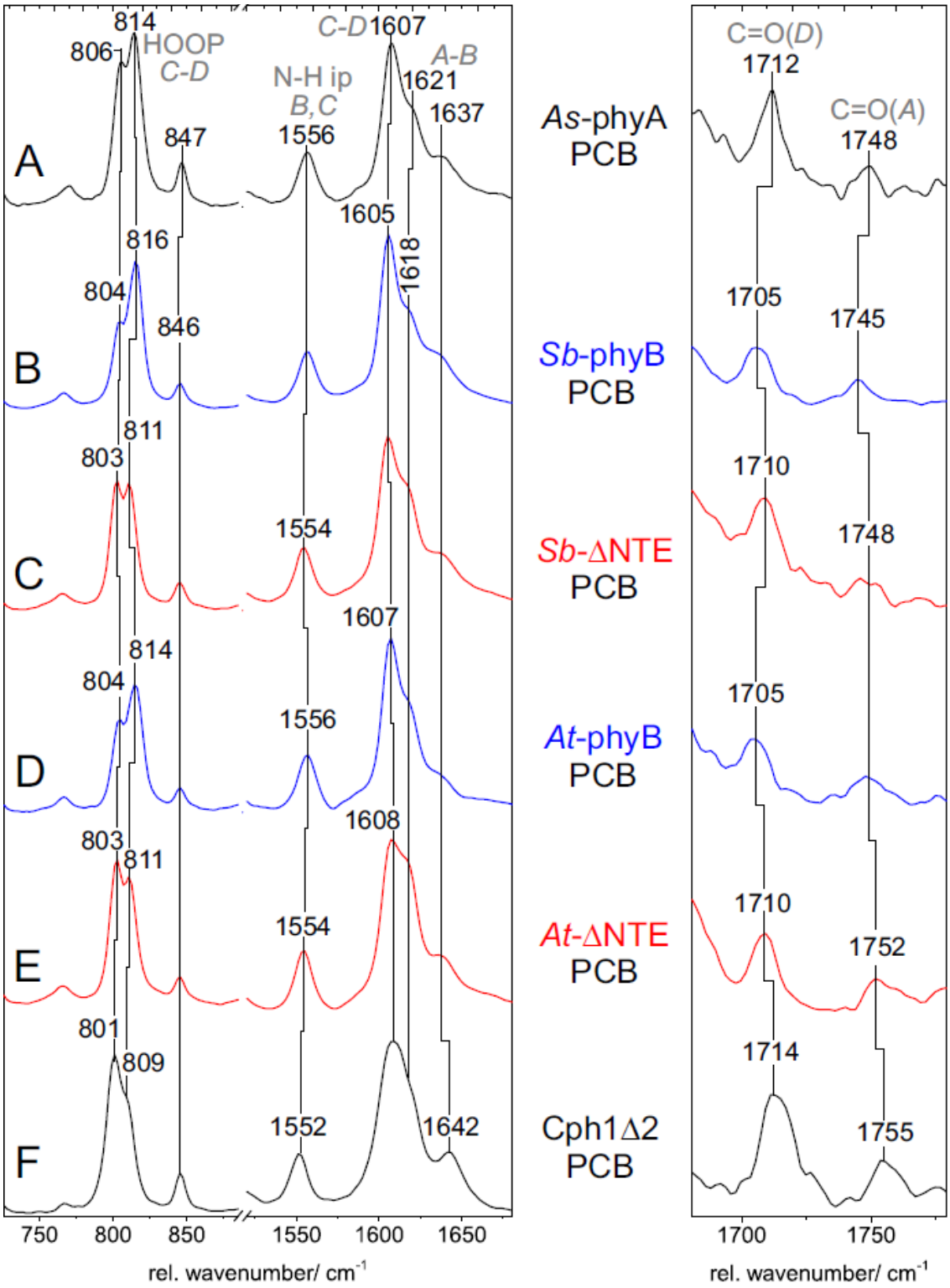


Fig. 3. RR spectra of the Pfr state of phytochrome adducts including the PCB chromophore. (A) As-phyA, (B) Sb-phyB, (C) Sb- NTE, (D) At-phyB, (E) At- NTE, and (F) Cph1 2. The spectra in (A) and (F) have been published earlier [16].

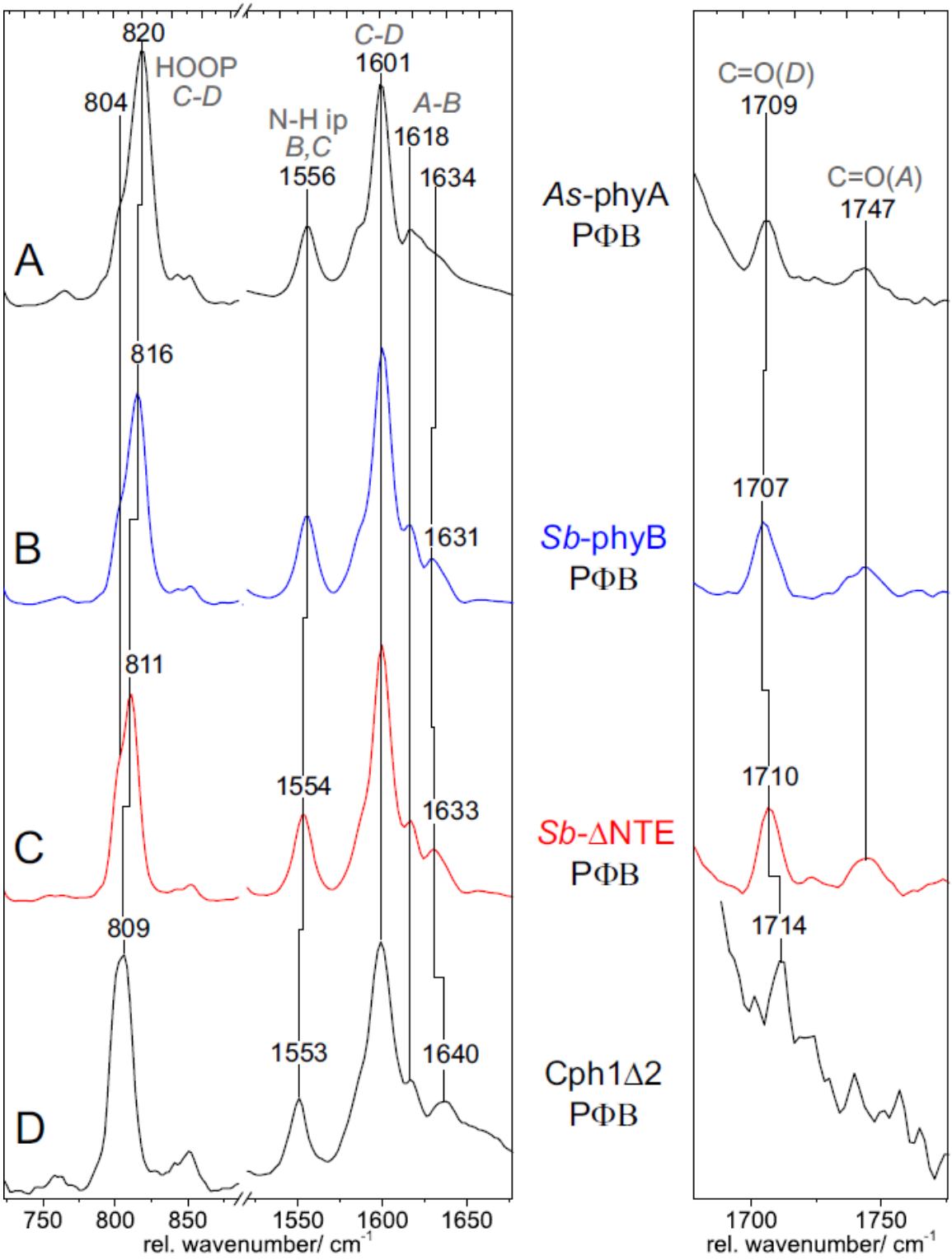


Fig. 4. RR spectra of the Pfr state of phytochrome adducts including the PΦB chromo phore. (A) As-phyA, (B) Sb-phyB, (C) Sb- NTE, and (D) Cph1 2. The spectra in (A) and (D) have been published earlier [16].

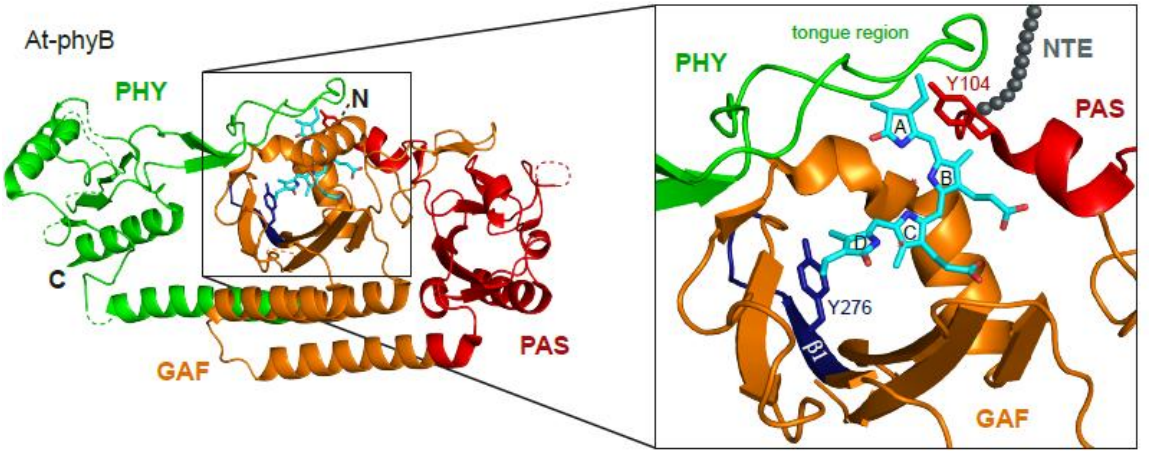


Fig. 5. Structure of At-phyB (PDB code: 4OUR); the PAS, GAF and PHY domains are depicted in red, orange and green, respectively. The chromophore, phytochromobilin (PΦB), is colored in cyan. The inlet enlarges the chromophore binding site. The β1 strand of the GAF domain is highlighted in blue and the N-terminal extension (NTE) in grey spheres.

11 Acknowledgements

I would like to thank Professor Jon Hughes for trusting my abilities and giving me the opportunity to work on this high-risk high-reward research project. I would try to consolidate his constant demand for 'controls', attention to details, critical thinking and 'wording' into all my future experiments. I am very grateful for Anna Lena Ermert, Christina Lang and Melanie Bingel for supporting me both technically and emotionally. I appreciate the cooperation of Professor Lars-Oliver Essen and all members of his group. I am indebted to Ralf Poeschke for conducting all crystallization screening in the dark. Kuba, Klaus Ibrahim, Silke and Soshi have been very nice colleagues and great friends. It is with the highest gratitude that I express my appreciation to several people that have helped me to live and work in Germany. I am aware of how difficult I can sometimes be. So I am very grateful for all members of AG Hughes and AG Zeilder for putting up with me over the years.

Selbständigkeitsversicherung

Ich erkläre: Ich habe die vorgelegte Dissertation selbständig und ohne unerlaubte fremde Hilfe und nur mit den Hilfen angefertigt, die ich in der Dissertation angegeben habe. Alle Textstellen, die wörtlich oder sinngemäß aus veröffentlichten Schriften entnommen sind, und alle Angaben, die auf mündlichen Auskünften beruhen, sind als solche kenntlich gemacht. Bei den von mir durchgeführten und in der Dissertation erwähnten Untersuchungen habe ich die Grundsätze guter wissenschaftlicher Praxis, wie sie in der „Satzung der Justus-Liebig-Universität Gießen zur Sicherung guter wissenschaftlicher Praxis“ niedergelegt sind, eingehalten.“

Ort und Datum Unterschrift