Understanding plant phytochrome 3D structure and signalling mechanisms

Dissertation

zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

am Fachbereich Biologie und Chemie der

Justus-Liebig-Universität Gießen

vorgelegt von

Kaoling Guan

aus

Shandong, China

Gießen 2022

Contents

Abstract

The main aim of this study was to expand our understanding of plant phytochrome action by illuminating the mysteries behind Pr-Pfr photoconversion, Pfr nuclear translocation, and phytochrome-PIF interaction via combinations of structural, biochemical, biophysical and *in vivo* cell biological techniques.

To this end, various constructs of *Glycine max* (soybean) phyA and phyB were produced, purified, characterised and subjected to crystallisation trials. After great effort in the crystallisation screening, condition optimisation, X-ray diffraction trials, data processing, model building and refinement, two structures of plant phytochrome in the Pr state were determined at high resolution, namely for a PGP (nPAS-GAF-PHY) construct of phyB and a PG (nPAS-GAF) construct of phyA from *Glycine max*. The latter represents the first crystal structure of A-type plant phytochrome, and surprisingly showed R/FR photochromicity in the absence of the PHY domain. Together with extensive spectroscopic data, this work provides new insight into plant phytochrome structure and function.

To determine plant phytochrome Pfr structures, various Pfr-stabilised and the Pfr-mimic mutants were studied similarly. Two of them, R549A and Y242H mutants of *Gm*phyA(PGP) crystallised successfully but have yet to yield useful diffraction data. Our studies provide valuable hints for further conditions optimisations of these and novel Pfr crystallisation trials of other constructs. In addition, *Gm*phyA(PG) crystallised under continuous or pulsed orange light with crystal packing pattern and conformation of the N-terminus of the nPAS domain quite distinct from those of the Pr crystals. The photoconverted *Gm*phyA(PG) crystals diffract to 2 Å, providing suitable materials for further X-ray free electron laser (XFEL) studies.

Nuclear translocation of phyA depends on the function of FHY1, a carrier protein bearing both nuclear localisation and nuclear export signals. Remarkably however, a KRKR motif representing a putative Class I NLS, was identified in the 380s loop of the phyA subfamily. We showed that KRKR promotes nuclear accumulation of YFP in onion epidermal cells, while a single mutation of the KRKR motif abolished this activity. Earlier work showed phyA:GFP as cytoplasmic foci in mutants lacking FHY1 function; here we found that

some onion epidermal cells transfected with *PHYA:GFP* bearing the KRKR→AAAA mutation showed similar foci.

We are eager to understand how phytochromes interact with and thereby regulate the activity of the PIF family of transcription factors at the structural level. In preparation, the interaction between *Gm*phyB and *Arabidopsis* PIF6 was investigated using size exclusion chromatography. We detected Pfr-dependent interaction between the NPGP (NTE-nPAS-GAF-PHY) construct of *Gm*phyB and PIF6 as well as light-independent interaction in the case of the Y272H phytochrome mutant, supporting the hypothesis that the Y272H mutant mimics the Pfr signalling conformation constitutively. Moreover, we found the N-terminal extension is particularly important for this interaction as no interaction was detected between the PGP construct and PIF6. This conclusion is consistent with experiments which showed that PIF6 suppresses dark reversion of NPGP but not PGP. Also, the PHY domain seems to be dispensable for phytochrome interaction with PIFs, since the Y272H mutant of the NPG construct still bound PIF6. This is consistent with earlier *in planta* studies which had shown that an NPG construct could trigger normal photomorphogenesis when it is dimerised and localised in the nucleus. Our studies emphasis the indispensable role of NTE in phyB(NPGP)-PIF interaction which should be taken into consideration in crystallisation trials of the complexes.

Zusammenfassung

Hauptziel dieser Studie war es, unser Verständnis der Phytochrom-Wirkung von Pflanzen zu erweitern, indem wir die Geheimnisse hinter der Pr-Pfr-Photokonversion, der Pfr-Kerntranslokation und der Phytochrom-PIF-Interaktion durch eine Kombination struktureller, biochemischer, biophysikalischer und *in vivo* zellbiologischer Techniken entschlüsseln.

Zu diesem Zweck wurden verschiedene Konstrukte von *Glycine max* (Soyabohne) phyA und phyB hergestellt, gereinigt, charakterisiert und Kristallisationsversuchen unterzogen. Nach großen Anstrengungen beim Kristallisationsscreening, der Optimierung der Bedingungen, den Röntgenbeugungsversuchen, der Datenverarbeitung, der Modellbildung und -verfeinerung wurden zwei Strukturen des pflanzlichen Phytochroms im Pr Zustand mit hoher Auflösung bestimmt, nämlich für ein PGP (nPAS-GAF-PHY) phyB-Konstrukt sowie ein PG (nPAS-GAF) phyA-Konstrukt von *Glycine max*. Letztere stellt die erste Kristallstruktur des pflanzlichen Phytochroms vom A-Typ dar und zeigte überraschenderweise R/FR-Photochromie in Abwesenheit der PHY-Domäne. Zusammen mit umfangreichen spektroskopischen Daten bietet diese Arbeit neuen Einblick in die Struktur und Funktion des pflanzlichen Phytochroms.

Zur Bestimmung der Strukturen des pflanzlichen Phytochroms Pfr wurden verschiedene Pfr-stabilisierte und Pfr-mimische Mutanten ähnlich untersucht. Zwei von ihnen, die R549A- und Y242H-Mutanten von *Gm*phyA(PGP), kristallisierten erfolgreich, haben aber noch keine brauchbaren Beugungsdaten geliefert. Unsere Studien liefern wertvolle Hinweise für weitere Optimierungen der Bedingungen für diese sowie weitere Pfr-Kristallisationsversuche anderer Konstrukte. Darüber hinaus kristallisierte *Gm*phyA(PG) unter kontinuierlichem oder gepulstem orangefarbenem Licht mit einem Kristallpackungsmuster und einer Konformation des N-Terminus der nPAS-Domäne, die sich deutlich von denen der Pr-Kristalle unterscheiden. Die photokonvertierten *Gm*phyA(PG)-Kristalle beugen sich immer noch bei 2 Å und bieten somit geeignete Materialien für weitere XFEL-Studien.

Die Kerntranslokation von phyA hängt von der Funktion von FHY1 ab, ein Shuttle-Protein das sowohl Kern-Import als auch -Export Signale trägt. In der 380s-Schleife der

phyA-Subfamilie wurde jedoch ein KRKR-Motiv identifiziert, das ein mutmaßliches Klasse-I-NLS darstellt. Wir konnten zeigen, dass KRKR die nukleäre Akkumulation von YFP in epidermalen Zwiebelzellen fördert, während eine einzige Mutation des KRKR-Motivs diese Aktivität aufhob. In früheren Studien wurden phyA:GFP zytoplasmatische Foci bei einer FHY1-defizienzten Mutante gezeigt; hier fanden wir, dass in einigen Zellen, die phyA:GFP mit der KRKR→AAAA-Mutation exprimierten, ähnliche Foci zeigten.

Wir sind sehr daran interessiert zu verstehen, wie Phytochrome mit der PIF-Familie von Transkriptionsfaktoren auf struktureller Ebene interagieren und damit ihre Aktivität regulieren. In Vorbereitung wurde die Interaktion zwischen *Gm*phyB und *Arabidopsis* PIF6 mittels SEC untersucht. Wir konnten eine Pfr-abhängige Interaktion zwischen *Gm*phyB(NPGP) und PIF6 sowie eine lichtunabhängige Interaktion bei der Y272H-Mutante nachweisen: dies unterstützt die Hypothese, dass die Y272H-Mutante die Pfr-Signalkonformation konstitutiv nachahmt. Darüber hinaus haben wir festgestellt, dass die N-terminale Verlängerung für diese Interaktion besonders wichtig ist, da zwischen dem PGP-Konstrukt und PIF6 keine Interaktion festgestellt wurde. Diese Schlussfolgerung steht im Einklang mit Experimenten, die zeigten, dass PIF6 die dunkle Reversion von NPGP, aber nicht von PGP unterdrückt. Außerdem scheint die PHY-Domäne für die Interaktion von Phytochrom mit PIFs entbehrlich zu sein, da die Y272H-Mutante des NPG-Konstrukts immer noch PIF6 bindet. Dies steht im Einklang mit früheren *in planta*-Studien, die gezeigt hatten, dass ein NPG-Konstrukt eine normale Photomorphogenese auslösen kann, wenn es dimerisiert und im Zellkern lokalisiert ist. Unsere Studien unterstreichen die unverzichtbare Rolle von NTE in der phyB(NPGP)-PIF-Interaktion, die bei Kristallisationsversuchen der Komplexe berücksichtigt werden sollte.

1. Introduction

1.1 Plant phytochromes

As well as being an energy source, light is also a signal that controls many aspects of plant growth and development. To monitor their ever-changing light environment, plants evolved a group of appropriate photoreceptors - including phytochromes that sense red and far-red light.

Dark-grown seedlings are etiolated: dicots show long hypocotyls, closed cotyledons, apical hooks (Fig 1.1, adapted from Sullivan *et al*. [1]); analogous growth and development are seen in monocots. This developmental strategy is called skotomorphogenesis. Light-grown seedlings, on the other hand, display photomorphogenesis: short hypocotyls, expanded cotyledons, open apical hooks, *etc*. Photomorphogenesis is one of the major responses regulated by phytochromes, others include germination, shade avoidance and flowering. These enormous physiological effects are brought about through changes in gene expression encompassing nearly a quarter of the genome [2].

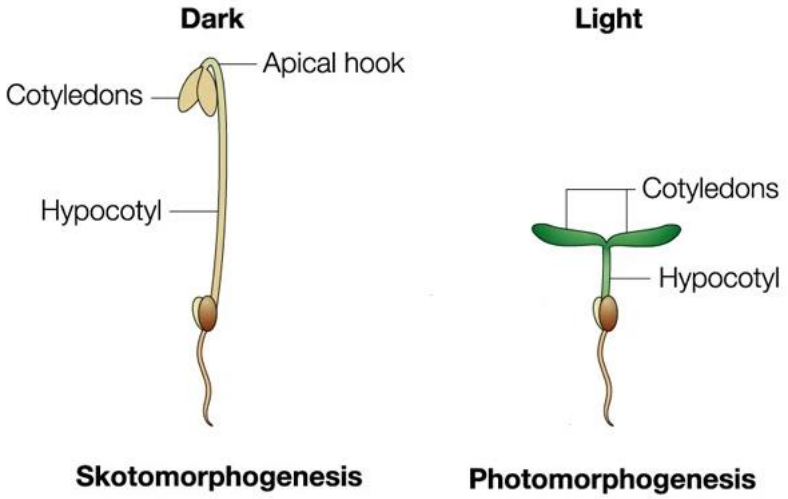


Figure 1.1 Skotomorphogenesis and photomorphogenesis. (Adapted from Sullivan *et al.* [1])

Phytochromes are red/far-red photochromic photoreceptors first discovered in plants [3], and also have been found in various microorganisms [4-7]. Phytochromes photoconvert reversibly between red-light absorbing (Pr) and far-red-light absorbing (Pfr) states after irradiation with red or far-red light. In most cases, Pfr reverts to the lowest energy Pr state in a slow thermal process known as dark reversion.

Phytochromes use bilin (linear tetrapyrrole) cofactors (chromophores) to absorb light. Pr→Pfr photoconversion involves a double bond isomerization and rotation of the chromophore *D* ring that is associated with conformational change of the protein. In plants, Pfr is the physiologically active signalling state. Pr is restricted to the cytoplasm, but upon photoconversion, Pfr translocates into the nucleus where it controls gene expression through light-dependent interaction with various signalling partners. Plant phytochromes have both similar and specific features regarding their domain structure, chromophore type, photoconversion mechanism, signal transduction and three-dimensional structure relative to their prokaryotic homologs. These features are discussed below.

1.1.1 Phytochrome gene family and domain structure

Higher plants possess several phytochromes, for example, the *Arabidopsis thaliana* genome encodes five phytochromes (phyA-phyE) classified into two groups [8]. phyA is a type I phytochrome, labile in light, while phyB-phyE are light stable type II phytochromes[9]. phyA accumulates to relatively high levels in dark-grown seedlings but is rapidly degraded in light, while phyB-phyE amounts are very low and little affected by light. phyB is the most abundant phytochrome in light, whereas in darkness phyA predominates [10,11].

Plant phytochromes exist as dimers with each monomer of about 120 kDa consisting of an N-terminal photosensory module (PSM) and a C-terminal histidine kinase-like module (HKM) separated by a PAS (Period/ARNT/Singleminded) domain repeat. The PSM comprises four domains: the NTE (N-terminal extension), nPAS (N-terminal PAS), GAF (cGMP phosphodiesterase / adenylate cyclase / FhlA) and PHY (phytochrome-specific) domains. The HKM comprises DHP (dimerization and histidine phosphoreceptor-like) and CAT (catalytic ATPase-like) domains. Typical domain structures of plant phytochromes are shown in Fig 1.2 (phyA and phyB from *Glycine max* as examples). For plant phytochromes, the conserved cysteine to which the chromophore is covalently attached locates in the GAF domain. The NTE and PAS-repeat are peculiar to plant phytochromes, whereas the nPAS, GAF, PHY, DHP, and CAT domains are also represented in many prokaryotic phytochromes [12]. The serine/threonine-rich NTE is

critical to Pfr stability and seems to regulate plant phytochrome activity through phosphorylation [13,14].

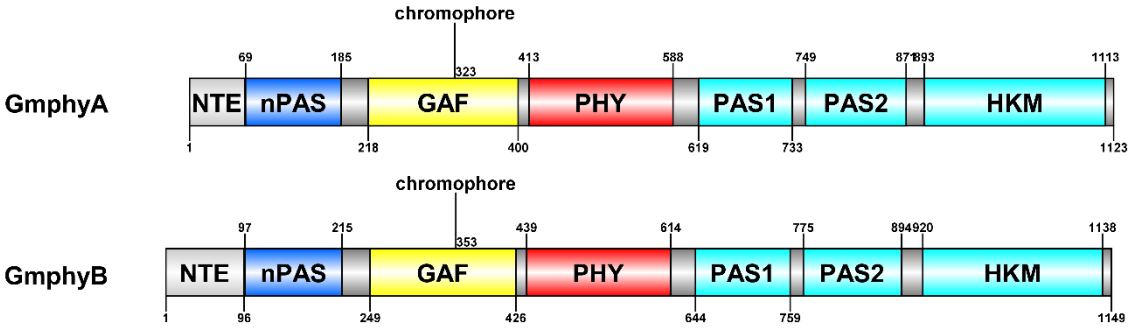


Figure 1.2 The domain structure of *Glycine max* phyA and phyB. NTE: the N-terminal extension. PAS: Period/ARNT/Singleminded, GAF: cGMP phosphodiesterase/adenylate cyclase/FhlA, PHY: phytochrome-specific domain, HKM: histidine kinase-like module. The numbers indicate the positions of the residues.

1.1.2 Chromophore and photoconversion

The bilin chromophore is characterized by an open chain of four pyrrole rings (Fig 1.3). In plant phytochromes and cyanobacterial phytochrome Cph1, the bilin is covalently attached to a conserved cysteine in the GAF domain via a thioether linkage (Fig 1.2). Plant phytochromes use phytochromobilin (PΦB), while Cph1 and bacteriophytochromes use phycocyanobilin (PCB) and biliverdin IXα (BV) as chromophores, respectively. For PCB and PΦB, covalent linkage forms between a cysteine and the $C3^1$ of ring *A*, while it is the $C3^2$ of ring *A* in the case of BV (Fig 1.3). Circular dichroism (CD) spectra of Cph1,

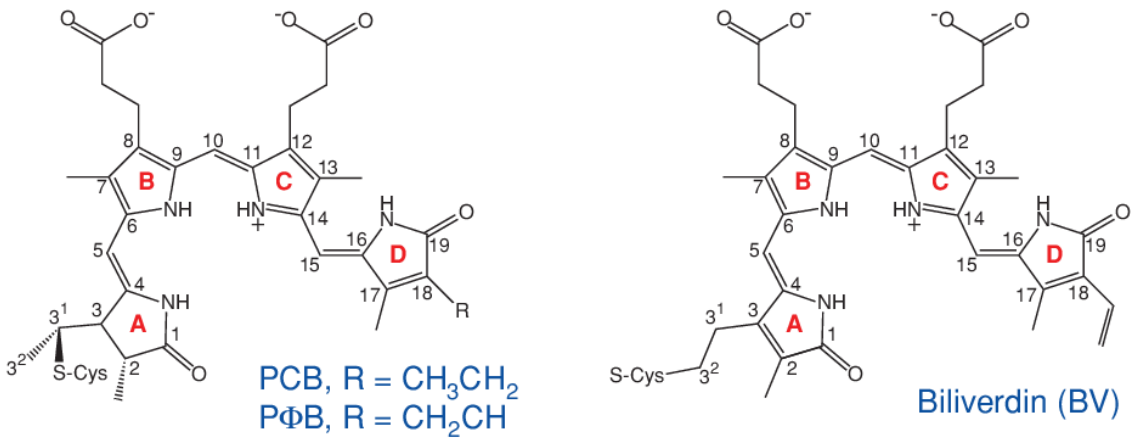


Figure 1.3 Chemical structures of PCB, PΦB and biliverdin shown in the *ZZZssa* configuration (C5–Z, syn C10–Z, syn C15–Z, anti) Pr state. The numbers indicate the carbon atoms. (Adapted from Rockwell *et al*. [20]).

plant phytochromes and *Deinococcus radiodurans* bacteriophytochrome (*Dr*BphP) exhibit negative bands in the red region in Pr, which become weakly negative in Pfr for *Dr*BphP, but switch to positive for Cph1 and plant phytochrome Pfr. The negative and positive red absorbance bands imply α- and β-facial disposition, respectively, of ring *D*[15], whereby ring *D* is above the *B* and *C* plane in the α- but below it in the β-facial disposition.

PΦB is synthesized in the chloroplast through a series of enzymatic reactions. Heme oxygenase catalyzes the oxidation of heme to produce BV, which is then reduced to PΦB by phytochromobilin synthase [16]. In *Arabidopsis* these enzymes are encoded by *HY1* and *HY2,* respectively [17,18]. Although PΦB is the native chromophore of plant phytochromes, they can also attach PCB, resulting in a blue-shifted absorbance spectrum compared with the PΦB adducts as a result of the slightly shorter conjugated π-orbital system [19].

The chromophore adopts a twisted *ZZZssa* configuration (C5–Z, syn; C10–Z, syn; C15–Z, anti) in Pr state (Fig 1.3) [20]. Photoconversion from Pr to Pfr involves *Z-E* (*cis-trans*) isomerization of the C15=C16 double bond and the rotation of ring *D*, followed by side chain movements and protein conformational changes. UV-Vis spectroscopic studies reveal several intermediate states between Pr and Pfr (Fig 1.4) [21,22]. In detail, photon absorption initiates the *Z-E* isomerization to generate the lumi-R intermediate within picoseconds. Lumi-R relaxes to form the meta-Ra and meta-Rc intermediates successively in the micro- to millisecond timescale. The chromophore is temporarily deprotonated in meta-Rc in company with the absorbance bleaching but re-protonated to

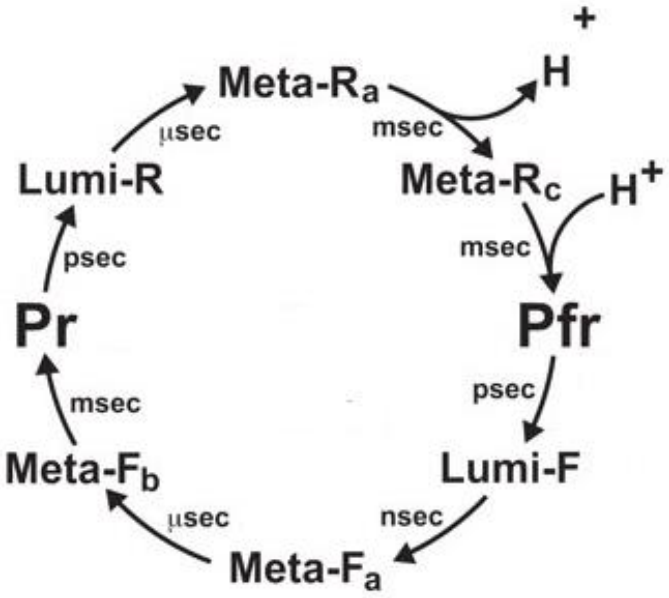


Figure 1.4 Photocycle of phytochrome (Adapted from Wagner *et al.* [21])

form Pfr [23]. Pfr can be converted back to Pr by FR irradiation via lumi-F, meta-Fa, and meta-Fb [24], or by slow thermal processes.

1.1.3 Signalling mechanisms

A domain-swapping analysis showed that the N-terminal photosensory module of *Arabidopsis* phyA and phyB determines the respective physiological specificity and light lability [25]. Indeed, the isolated PSM of *Arabidopsis* phyB was able to trigger light-dependent hypocotyl growth inhibition when fused to a dimerization domain and a nuclear localisation signal [26]. Further study suggested that the PHY domain too was unnecessary for this photoresponse [27]. The PSM of oat phyA3 also showed partial physiological activity when dimerised and localised in the nucleus [28]. Remarkably, the first 406 amino acids (NTE, nPAS, and GAF domains) of *Arabidopsis* phyA as dimer targeted to the nucleus triggered a constitutive photomorphogenic-like response [29]. A recent study showed that the N-terminal photosensory domain of phyB binds to PIF3 and inhibits its transactivation activity [30]. On the other hand, the C-terminal of plant phytochrome is biologically active too. It is required for phytochrome dimerization, and PIF3 degradation [31]. Intriguingly, missense mutations in both phyA and phyB displaying long hypocotyl under light are concentrated in a 160-residue segment between the PAS1 and PAS2 domains (Quail box) [32]. How these residues function is not clear.

Plant phytochromes reside in the cytoplasm as Pr in darkness. A key event initializing phytochrome-mediated light signalling is the translocation of Pfr to the nucleus, a process that is far from being fully understood for any of the plant phytochrome types. In the case of phyB, it was suggested that a nuclear localization signal (NLS) located in the PAS-repeat region is unmasked in Pfr [33]. No such NLS is apparent, however. More recently, it was suggested that phytochrome interacting factors (PIFs) which carry NLS motifs and interact with phyB Pfr might act as carriers [34]. On the other hand, although phyA interacts with PIF3 as Pfr, its nuclear import requires FHY1 function [35-38]. FHY1 carries both an NLS and a nuclear export signal (NES). The NLS is recognised by importin α (IMPα), mediating the translocation of the complex to the nucleus [39]. FHY1 releases phyA in the nucleus, returning to the cytoplasm with the help of its NES, allowing it to shuttle between the cytoplasm and nucleus [39,40].

In the nucleus, plant phytochromes involve light-dependent interaction with various signalling partners, for example, the PIFs (phytochrome interacting factors). PIFs belong to the basic helix-loop-helix (bHLH) transcription factor family, binding to a conserved DNA motif CACGTG known as the G-box, thereby modulating the expression of the target genes [41-43]. *Arabidopsis* genome encodes eight PIFs (PIF1-8), the first of which, PIF3, was identified as a phyB interactor in a yeast two-hybrid screen [44]. Indeed, its interaction with phyB is reversibly induced by light [45]. Most of the PIFs possess a so-called active phyB-binding domain (APB) near their N-terminus, while PIF1 and PIF3 possess an additional active phyA-binding domain (APA) downstream of the APB (Fig 1.5) [46]. Mutagenic studies identified mutations at four positions in the PSM of

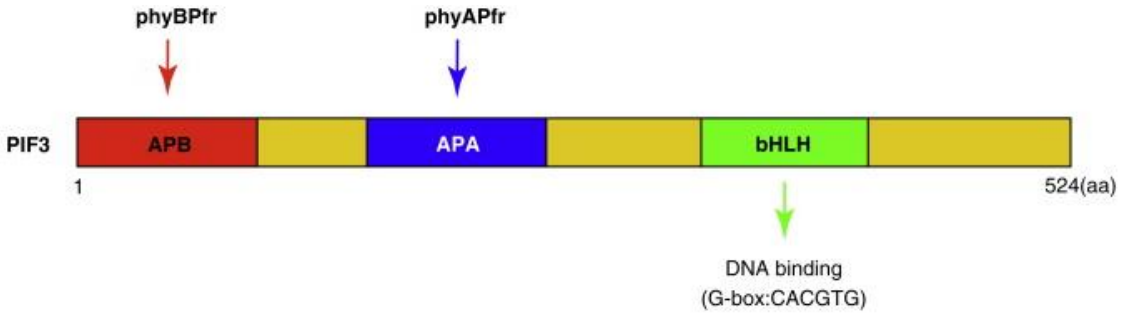


Figure 1.5 PIF3 domain structure. APA/APB: binding sites for photoactivated phyA (APA) and phyB (APB). bHLH: DNA binding basic helix-loop-helix domain. (From Leivar. & Quail [46]).

Arabidopsis phyB (R110, G111, G348, R352) that disrupt PIF3 binding [47,48]. PIFs promote skotomorphogenesis by activating etiolation gene expression in darkness. Upon light exposure, active phytochromes (Pfr) in the nucleus interact with PIFs, leading to the rapid phosphorylation, ubiquitination, and degradation of PIFs. Although they include a histidine kinase-like module, plant phytochromes are probably not histidine kinases as the crucial histidine residue is not conserved. Nevertheless, there is evidence that phytochrome phosphorylates PIFs directly through its own supposed serine/threonine kinase activity [49]. However, subsequent work showed that PIF3 phosphorylation is more likely caused by MUT9-like Ser/Thr kinases (MLKs) - also called photoregulatory protein kinases (PPKs) - that interact with PIF3 and phyB Pfr [50]. Ubiquitination involves LRBs (Light-Response Bric-a-Brack/Tramtrack/Broad E3 ubiquitin ligases) or CRLs (Cullin ring ubiquitin ligases) [51]. For example, LRBs mediate ubiquitination of both phyB and PIF3 [52], while $CUL^{COP1/SPA}$ is responsible for the light-induced ubiquitination of PIF1[53]. In summary, phytochrome-PIF interaction results in the phosphorylation,

ubiquitination, and degradation of PIFs and phytochromes, therefore inhibiting etiolation gene expression and initiating photomorphogenesis.

In addition to PIFs, phytochromes also interact with the ubiquitin E3 ligase CONSTITUTIVELY PHOTOMORPHOGENIC 1 (COP1) which acts as a repressor of photomorphogenesis [54]. The ubiquitin ligase activity of COP1 *in vivo* depends on the SUPPRESSOR OF PHYA (SPA) protein. In darkness, COP1/SPA mediate the ubiquitination of ELONGATED HYPOCOTYL 5 (HY5), a photomorphogenesis-promoting transcription factor, inducing its degradation. In light, phytochromes interact with and inactivate COP1/SPA, releasing their suppression of HY5, whereby HY5 activates gene expression to promote de-etiolation [55].

Besides photomorphogenesis, phytochrome regulates numerous aspects of plant growth and development, such as germination, shade avoidance, flowering, *etc* [54]. In detail, regulation of gene expression by PIF1 leads to low GA/ABA (gibberellic acid/abscisic acid) ratios which strongly inhibits germination, whereas Pfr inhibits PIF1 activity in light to promote germination. When plants grow in shade, the reduced R/FR ratio reduces the proportion of phyB Pfr, whereby repression of PIFs is reduced. PIFs, especially PIF7, enhance the shade-avoidance response by promoting auxin biosynthesis. Additionally, phytochromes are involved in photoperiodic flowering by regulating the abundance of FLOWERING LOCUS T (FT), which directly induces flowering [54,55].

1.2 Prokaryotic phytochromes

Phytochromes were thought to be restricted to plants, however, discoveries of Cph1 in cyanobacteria and other phytochromes in non-photosynthetic bacteria and fungi clearly showed a wider distribution in microorganisms [4-7,56].

PCB and biliverdin IXα (BV) are the natural chromophores of Cph1 and bacteriophytochromes, respectively. Cyanobacteria use *pcyA* (ferredoxin-biliverdin oxidoreductase) to catalyze the reduction of BV to phycocyanobilin (PCB) [57]. Unlike plant phytochromes and Cph1, the chromophore attachment site in bacteriophytochromes is a cysteine residue at the N-terminus of the nPAS domain.

Cph1 is closely similar to plant phytochromes in terms of the domain organization in the PSM (nPAS, GAF and PHY domains) and type and attachment position of the

chromophore. However, Cph1 has a much shorter NTE (only about 20 residues). Also, Cph1, like many bacteriophytochromes, is a light-regulated histidine kinase that transduces the light signal via its C-terminal histidine kinase domain to a response regulator [12,15,16]. Strangely, it signals as Pr [58], rather than as Pfr like plant phytochromes.

In addition to prototypical phytochromes, a group of bacteriophytochromes generates Pfr as the dark-adapted state, undergoing dark reversion from Pr to Pfr [59]. They were termed bathy-type because the absorbance maximum of their resting state is bathochromically (red) shifted.

The discovery of prokaryotic phytochromes greatly facilitated biophysical work as they could be produced and purified much easier than plant homologs. Indeed, they opened to door to 3D structural studies, as had been expected.

1.3 Structural studies of phytochromes

Structural biology strives to understand the function of biological macromolecules at the (near) atomic level by determining their 3D structures. A century's research has developed three main methods to this end, namely nuclear magnetic resonance (NMR), X-ray crystallography and, most recently, cryogenic electron microscopy (cryo-EM). X-ray crystallography was used extensively in the present study. The specific procedures include sample production, purification, crystallisation, condition optimisation, X-ray diffraction measurements at synchrotron facilities, data processing, and model building / refinement. The details are described in the Methods section.

3D structures not only unlock the secrets of conformation, function and interaction of macromolecules but also offer novel approaches in protein engineering. Structural biology has already revealed numerous mysteries of phytochrome, such as the chromophore conformation and how it interacts with its binding pocket, the remarkable knot around the nPAS domain, and what conformational changes happen during photoconversion. These achievements are discussed below.

1.3.1 Prokaryotic phytochrome structures

The first phytochrome crystal structure was of the chromophore binding domain (nPAS and GAF bidomain) of *Dr*BphP assembled with biliverdin (PDB: 1ZTU) [60]. This pioneering work revealed the overall conformation of the nPAS-GAF region as well as

protein-chromophore interactions in the chromophore binding pocket. It showed a little-expected *ZZZssa* configuration of the chromophore and the thioether linkage between chromophore and cysteine 259. Moreover, it revealed that the N-terminal part of the nPAS domain passes through a loop region of the GAF domain to form a figure-of-eight knot. The chromophore is, however, exposed to the solvent in this structure and is locked in the Pr state. The reason why this construct is unable to form Pfr is still unclear, but is certainly associated with the absence of the PHY domain.

Subsequently, two structures of the complete PSM (nPAS, GAF, and PHY domains) were determined, namely Cph1 from *Synechocystis* 6803 assembled with phycocyanobilin (PCB) in the dark-adapted Pr state (PDB: 2VEA) [61] and a bathy-type bacteriophytochrome from *Pseudomonas aeruginosa* (*Pa*BphP) assembled with biliverdin in the dark-adapted Pfr state (PDB: 3C2W) [62]. Cph1 crystallised as an

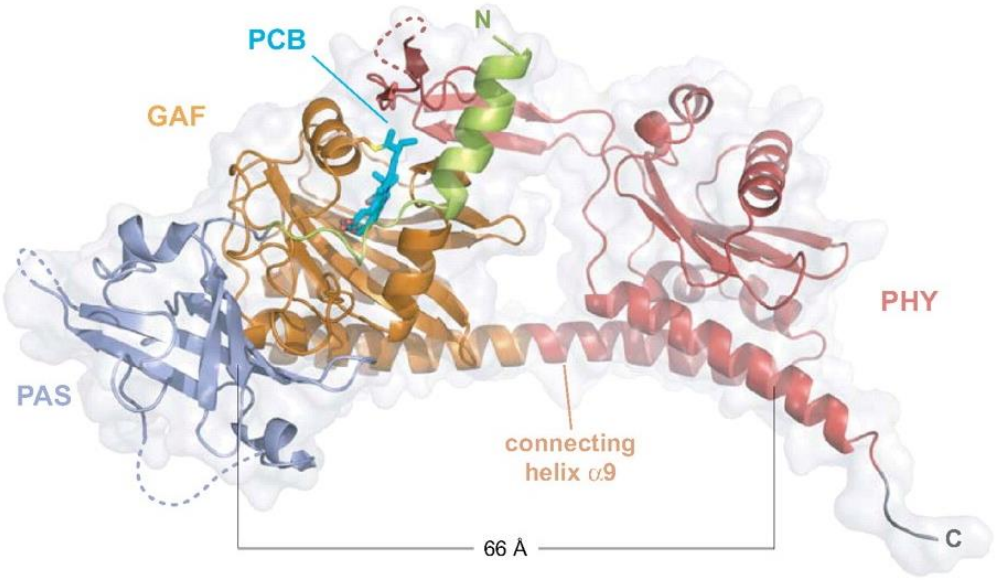


Figure 1.6 The overall structure of the photosensory domain of Cph1. nPAS (blue), GAF (orange), PHY (red), PCB chromophore (cyan), and molecular surface (grey). (From Essen *et al*. [61]).

antiparallel dimer: the overall architecture of the protomer is shown in Fig 1.6. The nPAS-GAF lobe is similar to that of *Dr*BphP in 1ZTU. The GAF and PHY domains are connected via a long α-helix but also through the tongue, a hairpin-like loop extending from the PHY domain to contact the GAF domain, thereby sealing the chromophore binding pocket. The PHY domain showed structural similarity to GAF domains but with the addition of the tongue. Intriguingly, the tongue shows rather different structures in the

two nPAS-GAF-PHY structures, more exact details of which in *Pa*BphP were revealed in a later paper [63].

Structural details of the chromophore binding pocket of Cph1 are shown in Fig 1.7. The cofactor adapts a *ZZZssa* configuration similar to that in 1ZTU with ring *A* attached to Cys259 via a single-carbon thioether linkage. Two conserved argin5ines are close to the propionate side chain of ring *B*. H260, the *A*-*B*-*C* ring nitrogen triad, seems to play a key role in chromophore protonation. It forms hydrogen bonding network with propionate chain of ring *C*, pyrrole water and nitrogen atoms of ring *A*, *B* and *C*. H290 forms a hydrogen bond with the carbonyl group of ring *D*. This hydrogen bond is supposed to be broken due to rotation of the ring *D* during Pr→Pfr photoconversion. The conserved residues Y176, Y203 and Y263 form a hydrophobic cavity around the *D*-ring. Interestingly, the Y176H mutation of Cph1 is unable to form Pfr absorbance, instead showing strong fluorescence [64]. Remarkably, the homologous mutation in *Arabidopsis* phyB (Y276H) showed a constitutively photomorphogenic (*cop*) phenotype [65], suggesting

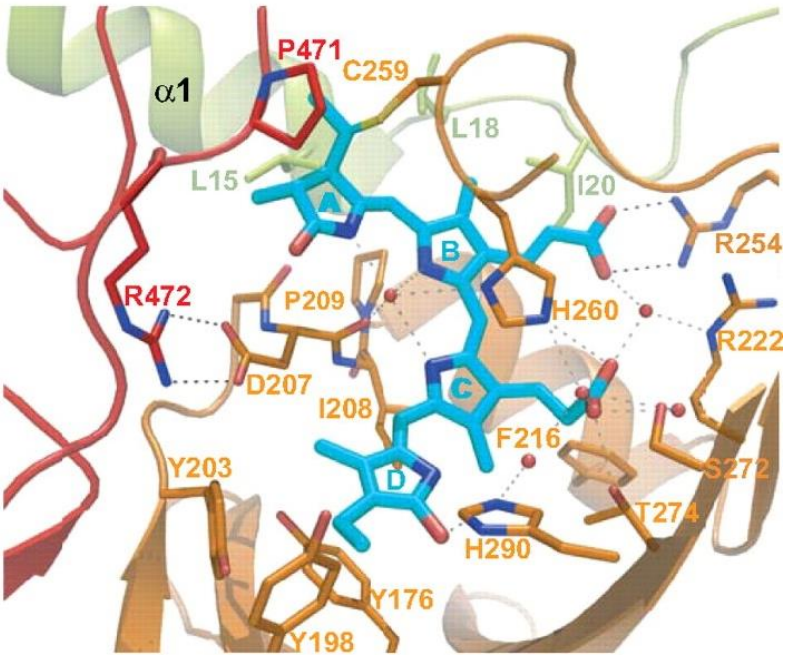


Figure 1.7 The chromophore binding pocket of Cph1 from the 2vea crystal structure. (From Essen *et al*. [61]).

that it might mimic the Pfr signalling conformation. The tongue region contains several conserved motifs, such as PRxSF and WGG. R472 from the PRxSF motif forms a salt

bridge with D207, whose main chain oxygen forms hydrogen bonds with the *A*-*B*-*C* ring nitrogens and the pyrrole water.

A comparison of 2VEA and 3C2W revealed differences in the hydrogen bonding network of the chromophore binding pocket, the positions of the conserved tyrosines around the *D*-ring, and different conformation of the tongue. However, it remained problematical to make conclusions about Pr→Pfr photoconversion by comparing these two structures derived from two very different phytochromes [12].

Magic-angle spinning nuclear magnetic resonance spectroscopy (MAS-NMR) data of Cph1 identified two isoforms in the Pr state (Pr-I and Pr-II), characterised by different charge distributions and hydrogen bonding networks [66]. The Pr-II isoform closely matched the 2VEA crystal structure. Moreover, this study showed *ZZEssa* configuration of chromophore and β-facial disposition of ring *D* in Pfr of Cph1, contrasting with the α-facial disposition seen in *Pa*BphP Pfr and correlating with CD studies [15].

Subsequently, improved structures of *Pa*BphP were published, including freeze-trapped intermediate states during Pfr-Pr photoconversion [63]. The new Pfr structure (PDB: 3NHQ) showed that the tongue forms a helix in contrast to the anti-parallel sheets seen in 2VEA. The intermediate structures captured at different temperatures suggested that light induces structural changes beginning with the *E*-*Z* isomerization (photoflip) of chromophore ring *D* and extending further to ring *B* and *C*. Essen suggested that the tongue refolds, and two tryptophans (from the WGG and WxE motifs) swap at the tongue-GAF interface [67]. Meanwhile, the Ser replaces the Arg from the PRxSF motif to form a salt bridge with the Asp from the DIP motif of the GAF domain.

The problem remained, however, that two different phytochromes were being compared, not just Pr and Pfr. The dilemma was finally resolved when Takala *et al.* reported the crystal structures of the PSM of *Dr*BphP in both dark-adapted (Pr: 4O0P) and illuminated (mixed Pr/Pfr: 4O01) states [68]. Crucially, the 4O01 structure showed the tongue to be partially helical as in the *Pa*BphP Pfr structure, whereas the Pr structure, like that of Cph1, showed anti-parallel sheets. As predicted [67], the tongue is shortened in the illuminated state structure because of the sheet-to-helix refolding, pulling the PHY and GAF domains together. In the dimer, this results in the paired PHY domains being pulled apart. However, it is under debate if the open and close form derived from crystal structure represents the

change in solution because the small-angle X-ray scattering (SAXS) data does not fit with the crystal structures. Additionally, the distances between different spin label positions remain unchanged during photoconversion in full-length *Agrobacterium* phytochrome 1 (Agp1) [69].

Subsequently, the F469W mutant of *Dr*BphP in which Pfr→Pr thermal reversion is much reduced was crystallised as a homogeneous Pfr/Pfr homodimer (PDB: 5C5K) [70]. Consistent with Takala *et al.*, this structure showed a helical conformation of the tongue. 5C5K also revealed more accurate differences between the chromophore and its pocket relative to the wide-type Pr structure (PDB: 4Q0J) [71], including the partner swap between ring *B* propionate and R254/R222, the position change of Y176 and F203 (Y in Cph1 and plant phytochromes) side chain, and the replacement of the R466 - D207 salt bridge by a hydrogen bond from S468 of the tongue to D207 (Fig 1.8).

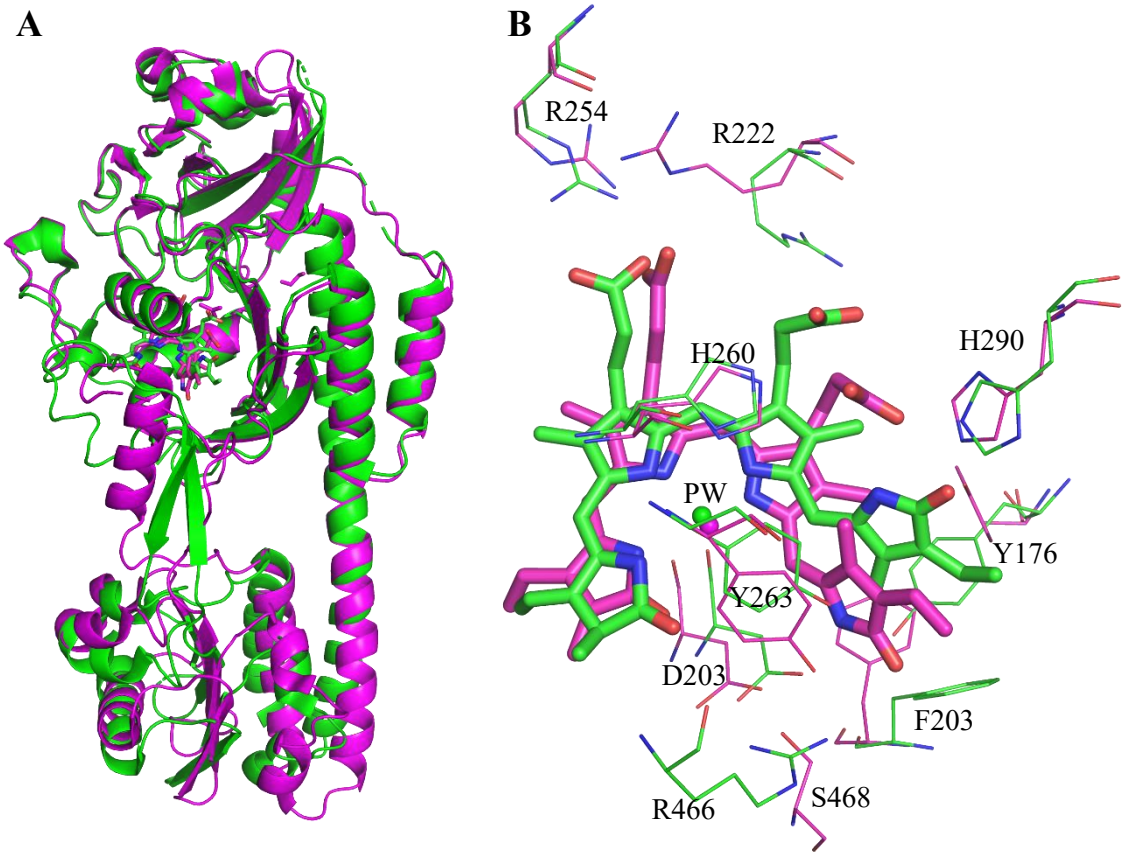


Figure 1.8 Superimposition of *Dr*BphP photosensory module wide-type structure in Pr (green, PDB: 4Q0J) and F469W mutant structure in Pfr (magenta, PDB: 5C5K). A: overall structure. B: relative positions of biliverdin and selected residues. (PW: pyrrole water)

1.3.2 Plant phytochrome structures

The first published crystal structure of a plant phytochrome was that of the PSM of *Arabidopsis thaliana* phyB (PDB: 4OUR) assembled with phytochromobilin (PΦB) in the Pr state (Fig 1.9) [72].

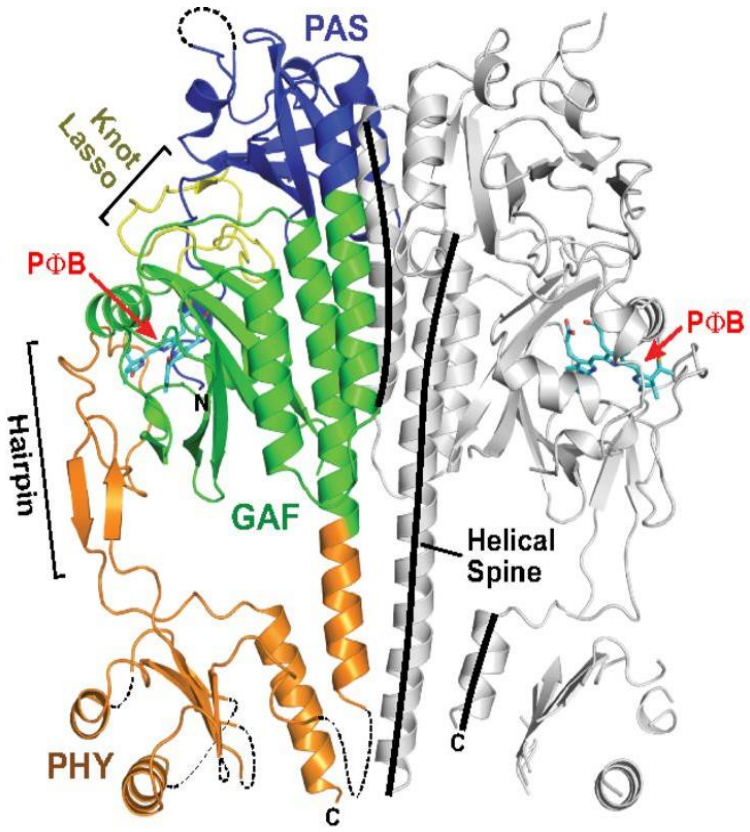


Figure 1.9 Crystal structure of the photosensory module from *Arabidopsis* phyB. nPAS (blue), GAF (green), PHY (orange) in subunit A. PΦB (cyan). Unresolved loops (dashed lines). (From Burgie *et al*. [72]).

Although the 3.4 Å 4OUR structure shows strong homology to its prokaryotic homologs, several differences were apparent. For example, two loops comprising residues 145-155 (the "150s loop") and 379-393 (the "380s loop") are much longer in plant phytochrome sequences than in prokaryotes and were thus poorly resolved probably due to their high mobility. Also, the propionate side chain of ring *C* is parallel to that of the ring *B*, and contacts with R322, whereas it points away from the side chain of ring *B* to connect to a histidine in prokaryotic phytochromes [61]. Substitution of several tongue residues affected the rate of thermal reversion. Specifically, R582A slowed reversion, while S584A and S584E dramatically accelerated it. Substitution of the first glycine of the WGG motif,

G564E, also profoundly slowed thermal reversion. Based on the data from prokaryotic phytochromes (see above), it was suggested that plant phytochrome photoconversion mechanism is consistent with the ideas derived from prokaryotes [72].

As in the case of Cph1, limited structural information for plant phytochromes as Pr and Pfr was provided by MAS-NMR [73]. The data showed for the NPGP construct of oat phyA3 a *ZZEssa* chromophore configuration as expected, but with ring *D* below (β-facial) the A-B-C plane, similar to the situation in Cph1 but in contrast to bacteriophytochromes in Pfr. Moreover, the MAS-NMR data indicated a helical conformation of the tongue. Interestingly, the data implied that the NTE is less mobile in Pfr than Pr, interactions with the *A*-ring being apparent in Pfr but not Pr.

Previous studies of phytochrome 3D structure focused on the ~70 kDa N-terminal photosensory module or even smaller fragments, while the ~120 kDa full-length plant phytochrome protomer forming a stable dimer in solution, is large enough be a suitable candidate for cryo-EM. This technique is revolutionary, relying upon images of individual molecules to determine 3D structures without the need of crystallisation. Recent advances in both hardware and software in cryo-EM have allowed structural determination of proteins at near-atomic resolution [74]. A preliminary 17 Å resolution cryo-EM structure of full-length *Arabidopsis* phyA was published in 2021, and the structure was interpreted as a head-to-head dimer [75], rather as expected. Very recently, however, a quite different cryo-EM based 3.3 Å structure of full-length *Arabidopsis* phyB (PDB: 7RZW) was published (Fig 1.10) [76]. There the PSMs are arranged head-to-tail together with the PAS2 domains located between them and together with the HKM as a head-to-head dimer forming a

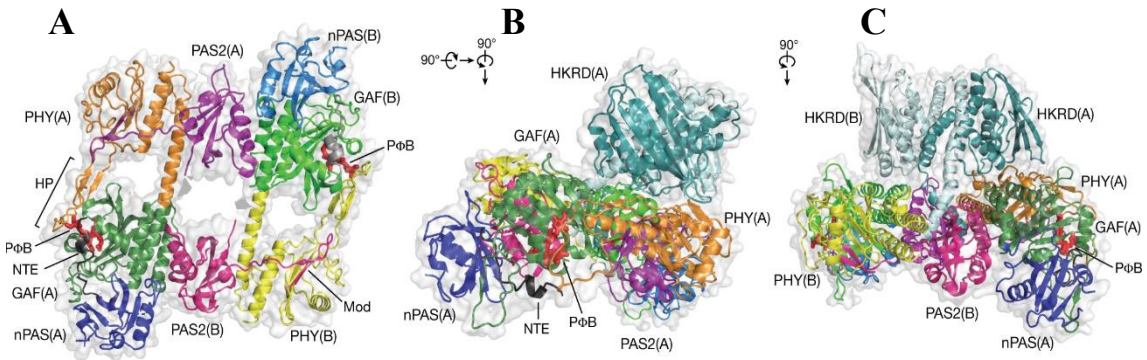


Figure 1.10 3.3 Å Cryo-EM model of dimeric phyB in basal (A), front (B) and side (C) views. Domains in protomers A and B are coloured individually. PΦB is shown as red sticks. (From Li *et al*. [76]).

parallelogram-shaped platform. The latter would seem appropriate to allow interactions with various signalling partners. The structure is not complete, however: the first 101 residues are missing as is the PAS1 domain [76].

1.4 Aim of this study

Despite decades of research, plant phytochrome signalling mechanisms remain only partially understood. The connection between photon absorption and plant developmental regulation requires in-depth research to acquire a comprehensive understanding of the signalling mechanisms involved. First of all, high-resolution structures of plant phytochromes both as Pr and Pfr are the starting points for understanding photoconversion. Secondly, as a key event in the early stage of phytochrome signalling, how Pfr nuclear translocation is achieved is an interesting research topic. Studies have shown that phyA relies on FHY1 for this, but how phyB accesses the nucleus remains unclear. Thirdly, phytochromes transduce light signals via protein-protein interactions in the nucleus, but how exactly phytochromes interact with PIFs is unknown. Given that plant phytochromes are distinct from their prokaryotic homologs, and that phyA and phyB represent two types of plant phytochromes that differ from each other in various aspects, this project aimed to determine high-resolution crystal structures of plant phytochrome A and B as Pr, and more importantly, as Pfr, the physiologically active state. In addition, we investigated how plant phytochromes translocate into the nucleus and subsequently interact with PIFs using biochemical, structural and *in vivo* cell biological techniques. We found that both phyA and phyB from *Glycine max* could be produced as holoproteins in *E.coli* with high yield and good stability, thereafter allowing effective purification, characterisation, crystallisation trials and structural determination of different constructs. Crystallisation trials of the Pfr states were also performed by generating Pfr-stabilised and Pfr-mimic mutants. Phytochrome-PIF interactions were explored *in vitro* by size-exclusion chromatography using phyB from *Glycine max* and PIF6 from *Arabidopsis*. High-resolution crystal structures of both phyA and phyB as Pr, together with spectral and biochemical studies, provided novel insights into the structure/functional relationships of plant phytochromes. Crystallisation and preliminary X-ray diffraction data for the Pfr-stabilised and Pfr-mimic mutants of *Gm*phyA provided valuable hints for further investigations. *Gm*phyA(PG) illuminated structure suggested potential dynamics

of the NTE during photoconversion. Photoconversion of *Gm*phyA(PG) *in crystallo* also provided appropriate materials for X-ray free-electron laser (XFEL) studies kinetically. In addition, the function of a putative nuclear localisation signal in phyA was investigated.

2. Materials and Methods

2.1 Materials

2.1.1 Instruments and consumables

Instruments

Instruments used in this study are described in the Methods section.

Consumables

Crystallisation materials:	MRC 2 well crystallisation plate (Jena Bioscience)
	24 wells VDX plate with sealant (Hampton Research)
	Crystal Direct Plate (Jena Bioscience)
	22 mm siliconized cover slides (Hampton Research)
	Crystal clear sealing tape (Hampton Research)
Crystal handling materials:	Dual Thickness Microloops (MiTeGen)
	CryoLoops (Hampton Research)
	Goniometer Base B5 with CryoVial (Jena Bioscience)
	Crystal Crusher (Hampton Research)
	Seed Bead (Hampton Research)
DNA purification kits:	QIAquick PCR purification kit (Qiagen)
	QIAprep Spin Miniprep kit (Qiagen)
Enzymes:	Homemade Taq DNA polymerase
	Phusion high-fidelity DNA polymerase (NEB)
	T4 polynucleotide kinase (NEB)
	T4 DNA ligase (NEB)
	DpnI (NEB)
Protein purification:	Ni-NTA Superflow (Qiagen)

	Superdex 200 16/60 prep grade (GE Healthcare)

2.1.2 Chemicals

Chemicals were purchased at analytical quality from AppliChem, Merck and Sigma-Aldrich unless otherwise stated.

2.1.3 Buffers and solutions

For DNA

5x TBE buffer:	500 mM Tris pH 8.0, 425 mM Boric acid, 50 mM EDTA.
6x DNA 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
Agarose gel:	1 % (w/v) agarose gel prepared in TBE buffer.
DNA ladder:	1kb DNA ladder (NEB)
EtBr staining solution:	1 : 20000 EtBr in 0.5 % TBE buffer

For cell culture

Antibiotic stock (1000X):	100 mg/ml Ampicillin, 50 mg/ml Spectinomycin. Sterile filtered.
IPTG stock solution:	1 M IPTG dissolved in Milli-Q, Sterile filtered.
LB medium:	1 % (w/v) tryptone, 0.5 % (w/v) yeast extract and 1 % (w/v) NaCl. Autoclaved.
LB agar plates:	1 % (w/v) tryptone, 0.5 % (w/v) yeast extract, 1 % (w/v) NaCl 1.5 % agar (w/v) prepare in Milli-Q water. Autoclaved and cooled before adding antibiotics and pour into petri dishes.
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.

Terrific broth (TB) medium:	12 g tryptone, 24 g yeast extract, and 400 µl glycerol. Prepared in 900 ml Milli-Q water, autoclaved. Mixed with 100ml TB buffer to make 1 L TB medium.
TB medium buffer:	170 mM KH_2PO_4, 720 mM K_2HPO_4. Autoclaved.

For protein

Lysis buffer:	50 mM HEPES pH 7.8, 200 mM NaCl, 10 mM imidazole, 5% glycerol, 1 mM β- mercaptoethanol.
Elution buffer:	50 mM HEPES pH 7.8, 200 mM NaCl, 250 mM imidazole, 5% glycerol, 1 mM β-mercaptoethanol.
SEC buffer:	20 mM HEPES pH 7.8, 1 mM EDTA, 200 mM NaCl, 5% glycerol, 5 mM dithiothreitol. Sterile filtered and degassed.
Ammonium sulfate buffer:	50 mM HEPES pH 7.8, 3.3 M $(NH_4)_2SO_4$.
Acrylamide solution (30 %):	30 % (w/v) of 37.5 : 1 (w/w) (Acrylamide : Bisacrylamide).
10 % APS:	10 % (w/v) Ammonium persulphate (APS)
4X stacking gel buffer:	0.5 M Tris pH 6.8, 0.6 % (w/v) SDS.
4X separating gel buffer:	1.5 M Tris pH 8.8, 0.6 % (w/v) SDS.
Stacking gel:	260 µl Acrylamide solution (30 %), 500 µl 4X stacking gel buffer, 1.2 ml Milli-Q, 15 µl 10% APS, 5 µl TEMED.
12 % Separating gel:	2.4 ml Acrylamide solution (30 %), 1.5 ml 4X separating gel buffer, 2.1 ml Milli-Q, 50 µl 10% APS, 10µl TEMED.
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.
SDS-PAGE running buffer:	25 mM Tris, 192 mM glycine, 0.1 % (w/v) SDS.
Zinc staining buffer:	1 mM Zinc acetate
Coomassie staining:	0.5 % (w/v) Coomassie Brilliant Blue R-250, 25 %,

	(v/v) Isopropanol and 10 % (v/v) Acetic acid.
De-staining solution:	10 % (v/v) Acetic acid
Crystallisation buffer:	Crystal Screen, PEGRx, SaltRx, PEG/Ion (Hampton Research) Morpheus, ProPlex, PGA Eco, MIDAS (Molecular Dimensions) JBScreen JCSG (Jena Bioscience) Cryo Kit, Low Ionic Strength Kit (Sigma Aldrich)
Additives:	Additive screen kit (Hampton Research)
Cryoprotectant kit:	CryoProtX MD1-61 (Molecular Dimensions)

2.1.4 Primers, plasmids and *E.coli* strains

Primers and plasmids used in this study are shown in Table 2.1 and Table 2.2, respectively. *E.coli* strain XL1-Blue or DH5α was used for cloning and plasmids amplification, while BL21(DE3) was used for protein production.

Table 2.1: Primers used in this study.

Name	Sequence(5'-3')	Description
AK7	CATGGTATATCTCCTTATTAAAGTTAAACAAATTATTTC	pCDFDuet-1-reverse
AK8	ATCAGTGGCACCGCAGATG	*Gm*phyA NTE deletion
AK11	GTGACCAGCGAAAGCGTG	*Gm*phyB NTE deletion
AK38	CAGTTCAATCTCCTTGTTAACGTG	*Gm*phyA PHY deletion-reverse
AK39	GCCAGCAGCTTCAAGGCCTTTCTG	*Gm*phyB R575A-forward
AK40	CGGATGCATGCGCTGACCGTC	*Gm*phyB R575A-reverse
AK41	AAGGTGCCAAACATCATCCGGAGGA	*Gm*phyB G557E-forward
AK42	CCCATTTAATCTCCTTTGCGGTGTGACT	*Gm*phyB G557E-reverse
AK43	GGAAGGCCTTTCTGGAGGTTGTTAAAAG	*Gm*phyB F578W-forward
AK44	AGCTGCTGCGCGGATGCATG	*Gm*phyB F578W-reverse
AK45	GCCAGTAGCTTCAAAGTGTTTTT	*Gm*phyA R549A-forward
AK46	CGGGTGCATGCGACGGCC	*Gm*phyA R549A-reverse

AK57	CACAAATTTCATGAAGACGATCATGGTG	*Gm*phyA Y242H-forward
AK58	GGCCATAACGCGATCATAGCCGG	*Gm*phyA Y242H-reverse
AK61	CATCGCTTTCACGAAGACGAGCACG	*Gm*phyB Y272H-forward
AK62	CACCATCACGCGGTCATAGCCGGT	*Gm*phyB Y272H-reverse
TR05	AAGAGATAAAGCGGCCGCGACTCTAGAATTCCA	Insert KRKR to YFP construct
TR06	TCTCTTCTTGTACAGCTCGTCCATGCCGAGAG	Insert KRKR to YFP construct
TR07	AAGAGAAAGGTCTAAAGCGGCCGCGACTCTAGAATTCCA	Insert PKKKRKV to YFP construct
TR08	CTTCTTGGGCTTGTACAGCTCGTCCATGCCGAGAG	Insert PKKKRKV to YFP construct
TR11	AAGAGATAAAGCGGCCGCGACTCTAGAATTCCA	Insert KAKR to YFP construct
TR12	GGCCTTCTTGTACAGCTCGTCCATGCCGAGAG	Insert KAKR to YFP construct
AK102	GCAAAGAGACTATGGGGTTTAGTGGTTTGT	KRKR to AAAA in *At*phyA-forward
AK103	CTTTTGAGGCTGTGTAGTAGCATCAGGAGC	KRKR to AAAA in *At*phyA-reverse

Table 2.2: Plasmids used in this study.

Name	Backbone	Antibiotic	Promoter	Tag	Description
p171		Ampicillin	T5	No	PCB production for *in vivo* assembly
pKG001	pCDFDuet-1	Spectinomycin	T7	C-ter His tag	*Gm*phyA(NPGP) production
pKG002	pCDFDuet-1	Spectinomycin	T7	C-ter His tag	*Gm*phyB(NPGP) production
pKG005	pCDFDuet-1	Spectinomycin	T7	C-ter His tag	*Gm*phyA(PGP) production
pKG011	pCDFDuet-1	Spectinomycin	T7	C-ter His tag	*Gm*phyB(PGP) production
pKG013	pCDFDuet-1	Spectinomycin	T7	C-ter His tag	*Gm*phyA(PG) production
pKG015	pCDFDuet-1	Spectinomycin	T7	C-ter His tag	*Gm*phyB(NPGP)-R575A production
pKG016	pCDFDuet-1	Spectinomycin	T7	C-ter His tag	*Gm*phyB(NPGP)-G557Eproduction
pKG017	pCDFDuet-1	Spectinomycin	T7	C-ter His tag	*Gm*phyB(NPGP)-F578W production
pKG019	pCDFDuet-1	Spectinomycin	T7	C-ter His tag	*Gm*phyA(PGP)-R549A production
pKG018	pCDFDuet-1	Spectinomycin	T7	C-ter His tag	*Gm*phyB(NPGP)-Y272H production
pKG021	pCDFDuet-1	Spectinomycin	T7	C-ter His tag	*Gm*phyA(PGP)-Y242H production
pKG023	pCDFDuet-1	Spectinomycin	T7	C-ter His tag	*Gm*phyB(NPG) production
pKG024	pCDFDuet-1	Spectinomycin	T7	C-ter His tag	*Gm*phyB(NPG)-Y272H production

pKG022	pCDFDuet-1	Spectinomycin	T7	C-ter His tag	*Gm*phyA(PG)-Y242H production
pHB111		Ampicillin	T7	C-ter His tag	GFP-*At*PIF6(1-100aa) production

2.2 Methods

2.2.1 Plasmid construction

The codon-optimised coding sequences of the NPGP (NTE, nPAS, GAF and PHY domains) of phyA and phyB of *Glycine max* bearing a C-terminal His6 tag were synthesised and cloned into the pCDFDuet vector by GENEWIZ (South Plainfield, New Jersey). Further truncations and site-directed mutations were generated via back-to-back PCR:

1. A linear plasmid was amplified using Phusion high-fidelity DNA polymerase (NEB) according to the manufacturer's instructions. The product size was checked by agarose gel electrophoresis.
2. The PCR product was purified using QIAquick PCR purification kit (Qiagen) and treated with T4 polynucleotide kinase (NEB) subsequently before the T4 DNA ligase treatment (NEB), which cyclises the linear plasmid.
3. The product was heat-inactivated at 65 ℃ for 10 min before being treated with DpnI (NEB) for the removal of methylated parental plasmid DNA.
4. The product was transformed into *E.coli* XL1-Blue or DH5α competent cells for amplification. All plasmids were verified by sequencing (Eurofins Genomics).

For transient protein expression in onion epidermal cells, a YFP expression cassette using the 35S promoter from cauliflower mosaic virus (CaMV) and the nopaline synthase (NOS) terminator from *Agrobacterium tumefaciens* was employed. Coding sequences of various short peptides: PKKKRKV (the nuclear localization signal from the Simian virus 40 large T antigen), KRKR (from wild-type phyA) and KAKR (mutated from phyA) were inserted at the C-terminus of YFP by back-to-back PCR as described above. mCherry fused to the nuclear localisation signal (NLS) of the VirD2 protein of *Agrobacterium* was used as a nuclear marker. The phyA:GFP construct was a gift from Dr. Mathias Zeilder (JLU Giessen), further mutations were generated via back-to-back PCR as described above. All novel constructs were verified by sequencing.

2.2.2 Transformation of *E.coli*

Plasmids were cloned by transformation into *E.coli* XL1-Blue or DH5α competent cells, while *E.coli* BL21 (DE3) was used for protein production. Frozen aliquots of competent cells were thawed on ice before being mixed with plasmid DNA and incubated on ice for 30 min. The cells were then heat-shocked at 42 ℃ for 45 s and then cold-shocked on ice for 1 min. Finally, the cells were shaken at 120 rpm, 37 ℃ for 40 min in 300 µl of SOC medium before pour-plating onto LB agar medium containing 100 µg/mL ampicillin or 50 µg/mL spectinomycin for selection and incubated overnight at 37 ℃.

2.2.3 Protein production

Plasmids for apophytochrome production were co-transformed into *E.coli* BL21 (DE3) competent cells together with either p171 or p183 for the biosynthesis of phycocyanobilin (PCB) or phytochromobilin (PΦB), respectively [77]. p171 carries two genes from *Synechocystis*, namely heme oxygenase (*HO*) and ferredoxin oxidoreductase (*pcyA*) for the synthesis of PCB from heme. p183 carries the heme oxygenase (*HO*) from *Synechocystis* and phytochromobilin synthase (*HY2*) from *Arabidopsis thaliana* for PΦB synthesis. After transformation, the competent cells were plated with 100 µg/mL ampicillin and 50 µg/mL spectinomycin and a single colony grown in 100 ml TB medium with antibiotics overnight at 37 °C. 7 ml overnight culture was transferred to a 2 L ribbed flask containing 700 ml TB medium with proper antibiotics, shaken at 120 rpm and 37 °C (Innova 42/42R, Eppendorf) until the OD_{600} reached 0.4. The culture was then swiftly cooled on wet ice and isopropyl β-D-1-thiogalacto-pyranoside (IPTG) was added to 200 µM for holoprotein production overnight at 16 °C with vigorous shaking. Production and purification procedures for GFP-*At*PIF6 (1-100aa) were as described [78].

2.2.4 Protein purification

2.2.4.1 Cell harvest and lysis

After 16-18 h of incubation at 16 °C, the cells were harvested by centrifugation at 6000 rpm for 10 min at 4 °C. The pellet was washed (resuspended in Lysis buffer (50 mM HEPES pH 7.8, 200 mM NaCl, 10 mM imidazole, 5% glycerol, 1 mM β-

mercaptoethanol), then pelleted again at 3900 rpm for 25 minutes at 4 °C) and resuspended in Lysis buffer.

The cells were cracked by 2-3 passages through a pre-cooled 1cm piston diameter French pressure cell with a force of about 21000 N (PP60KN, Watz Hydraulik). The lysate was then clarified by centrifuging at 15000 rpm for 40 min at 4 °C. The supernatant was removed carefully and mixed with ammonium sulfate (AmS) to a final concentration 60% (v/v) to salt-out the protein. The pellet was spun down at 12000 rpm for 10 minutes at 4 °C, then resuspended in Lysis buffer and the sample clarified at 12000 rpm for 10 minutes at 4 °C. The final supernatant was subjected to affinity chromatography.

2.2.4.2 Ni-NTA affinity chromatography

The C-terminal His6 tag allows Ni^{2+}-affinity purification. An appropriate volume of Ni-NTA beads was poured as a small gravity column and equilibrated with Lysis buffer before the protein supernatant was loaded. The column was then washed with Lysis buffer to remove non-specifically bound impurities before the column-bound protein was eluted with the Elution buffer (50 mM HEPES pH 7.8, 200 mM NaCl, 250 mM imidazole, 5% glycerol, 1 mM β-mercaptoethanol). The column was washed with Milli-Q water and stored in 20 % (v/v) ethanol for re-use.

2.2.4.3 Size exclusion chromatography (SEC)

The eluate from affinity chromatography purification was purified and analysed by size-exclusion chromatography (ÄKTA Purifier, GE Healthcare). The Superdex 200 16/60 prep grade column was pre-cooled to 4 °C with a water jacket and equilibrated with SEC buffer (20 mM HEPES pH 7.8, 1 mM EDTA, 200 mM NaCl, 5% glycerol, 5 mM dithiothreitol). Phytochrome sample in a syringe was irradiated with a far-red LED (730 nm) before injection into the Superdex column in darkness. Phytochrome sample was eluted with SEC buffer with absorbance monitoring at 280, 645, and 700 nm, while GFP-PIF6 protein was monitored at 280 and 488 nm. The SEC column was calibrated using marker proteins with specific molecular weight (by Christina Lang). Appropriate fractions corresponding to the absorption peak were collected and analysed by UV-Vis absorbance spectroscopy and SDS-PAGE.

2.2.5 UV-Vis spectroscopy

Protein samples were centrifuged at 50000 g, 4 ℃ for 10 min (Biofuge stratos, Heraeus) before absorbance measurement at room temperature using quartz fluorescence microcuvettes (600 µl, Hellma) in an 8453 diode array detector spectrophotometer (Agilent) equipped with a cuvette holder providing R (638 nm) and FR (745 nm) actinic light from each side.

2.2.6 SDS-PAGE

Protein samples were mixed with the proper volume of SDS loading dye and incubated at 95 °C for 5 min. After centrifugation at 14000 rpm for 5 min, the supernatant sample was loaded into a lane of a 12 % SDS gel. The gel was run at 220V constant voltage until the dye front reached the end of the gel. The gel was incubated in 1 mM zinc acetate for 10 minutes and photographed under UV light to show phytochrome holoprotein bands and then stained with Coomassie brilliant blue to reveal all protein bands.

2.2.7 Dark reversion measurement

SEC-purified phytochrome samples were irradiated with saturating 590 nm orange light before being distributed into a 96-well "black" plate (Brand GmbH) to avoid artefacts from scattered light. The absorbance at 720 and 800 nm was recorded at appropriate intervals at room temperature using a Synergy 2 plate reader (BioTek). Absorbance at 800 nm was treated as control, and dark reversion characterised by measuring the decrease of absorbance at 720 nm. The measurement was repeated three times independently for the same sample.

2.2.8 Crystallisation screening and conditions optimisation

The SEC-purified holoprotein samples were concentrated, centrifuged at 50000 g at 4 ℃ and irradiated with FR light (730 nm). Initial crystallisation screening was performed using the sitting-drop vapor-diffusion method. Plates were set up under dim green safelight using a pipetting robot (Honeybee 963 robotic system (Genomic Solutions) or Oryx 4 (Douglas Instruments)), and 96-well MRC plates (Jena Bioscience). 60 µl buffer was added as the reservoir. Droplets comprised 0.2 µl buffer from reservoir and an equal volume of the SEC-purified protein sample at a concentration of 10-15 mg/ml. The

crystallisation plates were sealed with Crystal Clear tape (Hampton Research), wrapped in aluminium foil to exclude light, and incubated at 10 °C in darkness.

A lead condition found to support crystal growth for a particular construct need not represent the optimum. Conditions were therefore optimised to obtain crystals showing the best diffraction in tests at the BESSY II synchrotron in Berlin. Generally, buffers with different pH, different types and concentrations of the salts and precipitants were investigated. Optimisations were performed using either the sitting-drop or hanging-drop method. The 24-well VDX plates with silicone grease (Hampton research) were used for the hanging-droop method. Each hanging drop comprised 1-2 µl protein and an equal volume of crystallisation buffer. Another 500 µl crystallisation buffer was added to the reservoir. All the experiments were carried out under dim green safelight.

Besides optimisation, various post-crystallisation treatments can also help improve the diffraction quality of crystals. These include cryo-annealing, additive screening, seeding, and dehydration.

Additive screening is a method to evaluate the ability of various additional compounds to improve crystallisation. The Additive Screen Kit (Hampton Research) containing 96 unique reagents was used in screens with crystallisation conditions that had been found to support the crystal growth. The procedure was the same as that used in the sitting-drop method, except for the addition of the additive to the protein-buffer droplet.

During the crystallisation process, the metastable zone represents conditions under which spontaneous homogeneous nucleation is rare but crystal growth is favourable. Under such conditions, crystals can often be grown easily when protein seeds are added. Crystal seeds were prepared according to the instructions provided by Hampton Research. Briefly, a seed stock solution was made by crushing the crystals in their mother liquor. Serial dilutions of the stock were made by dilution at different ratios. For crystallisation trials, the seed solution was mixed with the protein sample and buffer to set up sitting-drop or hanging-drop plates.

Protein crystals typically contain a high percentage of water. It is well established that reduction of solvent content may produce better ordered and packed crystals, thereafter,

improving the resolution of X-ray diffraction. Crystal dehydration was performed according to Heras *et al* [79].

2.2.9 Crystal harvest and freezing

Flash-freezing of crystals often leads to dramatic loss of diffraction due to ice crystal formation. Cryoprotectants are compounds that help to prevent the formation of ice during the flash-cooling of the crystals. Typically cryoprotectants include glycerol, ethylene glycol, 1,2-propanediol, diethylene glycol, 2-methyl-2,4-pentanediol (MPD), polyethylene glycol (PEG), paraffin oil, mineral oil and others. Commercial pre-mixed cryoprotectants kits are also available. To find out which cryoprotectant works best for specific crystals, one needs to test different cryoprotectants in the initial experiment. The ideal concentration of cryoprotectant was decided by flash-freezing 3-5 µl of the solutions within 10 µl tips in liquid nitrogen. The cryoprotectant should remain clear and not form ice in the tips when frozen in liquid nitrogen. Thus, crystals mounted in loops or mesh pins were transferred to crystallisation solution to which appropriate cryoprotectants had been added, soaked for 2 to 4 seconds, and then flash-frozen in liquid nitrogen. Normally, crystals were firstly transferred into crystallisation solution containing a lower concentration (half of the ideal concentration) of cryoprotectants to help crystals adapt to the environmental changes.

2.2.10 X-ray diffraction, datasets collection and processing

X-ray diffraction was performed at the Berlin synchrotron (BESSY II) with the help of beamline scientists. Crystals mounted in loops or mesh pins are centered in the X-ray beam and rotated via a goniometer. Crystal quality is determined initially by collecting two diffraction images, the resolution being estimated from the maximal diffraction angle. Diffraction datasets are then collected using an appropriate strategy suggested by iMosflm. Cryo-annealing, which involves warming crystals to room temperature followed by a renewed flash-cooling sometimes reduces cooling-induced disorder, thereby increasing the diffraction quality of protein crystals [79]. Cryo-annealing was performed using the cryo-annealing shutter of the beamline to block the cryostream for a short time. This method was essential for the structural determination of Cph1 [61]. The diffraction data are then indexed, integrated and scaled using XDS and AIMLESS from CCP4 package [80-82].

The phase problem was solved using molecular replacement using a homologous model and PHASER [83]. The structure was refined by COOT [84] and REFMAC5 [85]. Molecular images were created using the PyMOL molecular graphics system (Schrödinger, LLC).

2.2.11 SEC assay for protein interaction analysis

Size-exclusion chromatography is a valuable method for detecting protein-protein interactions. It separates proteins by their difference in size, therefore, a protein complex resulting from protein-protein interaction can be distinguished from its components based on a shifted elution peak. For phytochrome-PIF6 interaction studies, three samples (the SEC-purified phytochrome sample, GFP-*At*PIF6(1-100aa), and a mixture of the first two) with the same volume were irradiated with R or FR and injected into the Superdex 200 10/30 column in line at a flow rate of 0.5 ml/min. Proteins were monitored at 488, 645 and 700 nm and absorbance peaks were compared between these three samples.

2.2.12 Transient expression of YFP-fusion constructs in onion epidermal cell

Biolistic transfection of onion epidermis cells was achieved using a helium-driven biolistic device also as described [86]. Briefly, 2 µg each of the phyA:GFP (or YFP fusion) plasmid and the mCherry-NLS plasmid were coated onto gold particles in the presence of $CaCl_2$ and spermidine. The transfected material was kept in darkness for 2 days before observation and documentation with a DM6000b fluorescence microscope equipped with a DFC 7000 T camera (Leica) using the manufacturer's software (Leica Application Suite X). Standard bandpass filters for yellow and red fluorescent protein were used.

2.2.13 Fluorescence measurement

Phytochrome fluorescence spectra were measured using a Fluoromax4 spectrofluorometer (Horiba Jobin Yvon) as described [87,88]. Holophytochrome samples at approximately 0.1 mg/ml were used. Emission spectra were obtained by exciting the sample at 610 nm with 0.3 nm slit bandwidth and a 610 nm interference filter to minimize photoconversion. Emission spectra 630–750 nm were measured with a 10 nm slit bandwidth and a 610 nm cut-off filter. Excitation spectra were obtained by measuring fluorescence at 720 nm with 10 nm slit bandwidth as a function of excitation 580-700 nm

(1 nm slit bandwidth). A grey filter and a 690 nm cut-off filter were used to minimize photoconversion and avoid interference from scattered light, respectively.
To determine the fluorescence quantum yield (Φ_F), pheophytin A with a defined $\Phi_F = 0.2$ was used as a reference, together with the formula below, whereby F is the fluorescence and T is the transmittance of the sample:

$$\Phi F = \frac{F}{F^{pheo}} \times \frac{(1-T^{pheo}) \times \Phi F^{pheo}}{(1-T)}$$

2.2.14 *In silico* analysis

The annealing temperature for Phusion PCR was decided by Tm calculator (NEB). The protein sequences were acquired from UniProt under the specific accession number (*Gm*phyA: B4YB07; *Gm*phyB: I1MGE5). Molecular weight and other physical and chemical parameters of proteins were calculated by ProtParam. Protein sequence alignments were performed using CLC sequence viewer (QIAGEN) or Clustal Omega. Published structural models were taken from the Protein Data Bank (PDB).

3. Results

3.1. Crystal structures of plant phytochrome as Pr

3.1.1 Production and characterisation of various constructs of *Glycine max* phyA and phyB

The N-terminal extension (NTE) might be an obstacle to crystallisation due to its high mobility (experience acquired from Dr. Soshichiro Nagano and Dr. Sintayehu Manaye Shenkutie), therefore, PGP constructs of phyA and phyB were generated and investigated in the preliminary crystallisation trials. Holophytochrome was produced in *E.coli* followed by lysis, clarification, Ni-NTA affinity and size-exclusion chromatography purification (see 2.2.3 and 2.2.4). The SEC purification results and SDS-PAGE analysis of the peak fractions are shown in Fig 3.1. Sample absorbances were monitored at 280, 647 and 700 nm. The absorption peaks of *Gm*phyA(PGP) and *Gm*phyB(PGP) are located at 79 and 82 ml (Fig 3.1), respectively. A comparison between the SEC column calibration data and the calculated molecular weight from the protein sequence suggested both samples probably exist as monomers. SDS-PAGE results show specific bands which fluoresced in UV/B-Zn^{2+} and after Coomassie staining locating between the 45 kDa and 66 kDa markers (Fig 3.1), as expected from the molecular weight of *Gm*phyA(PGP) of 60 kDa and *Gm*phyB(PGP) of 58 kDa. These results show successful production and purification of these PGP constructs. UV-Vis absorbance spectra of the PGP constructs show the expected R/FR photochromicity, the absorbance peaks locating at about 645 nm and 705 nm for Pr and Pfr, respectively (Fig 3.2).

In addition, a shorter phyA construct without the PHY domain, *Gm*phyA(PG), was produced and characterised since no *Gm*phyA(PGP) crystals were obtained despite various crystallisation trials (see section 3.1.2). The SEC purification result for the smaller construct is shown in Fig 3.1. Here too the SEC profile and SDS-PAGE bands indicate successful production and purification of holophytochrome. Normally, phytochromes do not form *bona fide* Pfr when the PHY domain is deleted, displaying little Pfr absorbance after R illumination. This is reflected in the *Sorghum bicolor* phyB(PG) construct (Figure 3.3) and the *Arabidopsis* phyB equivalent [72]. Surprisingly, however, *Gm*phyA(PG) still

showed a strong absorbance in the far-red region after R light irradiation (Figure 3.3) which is characteristic of Pfr. This absorbance peak locates at 690 nm, about 15 nm blue-shifted compared to the Pfr absorbance of *Gm*phyA(PGP).

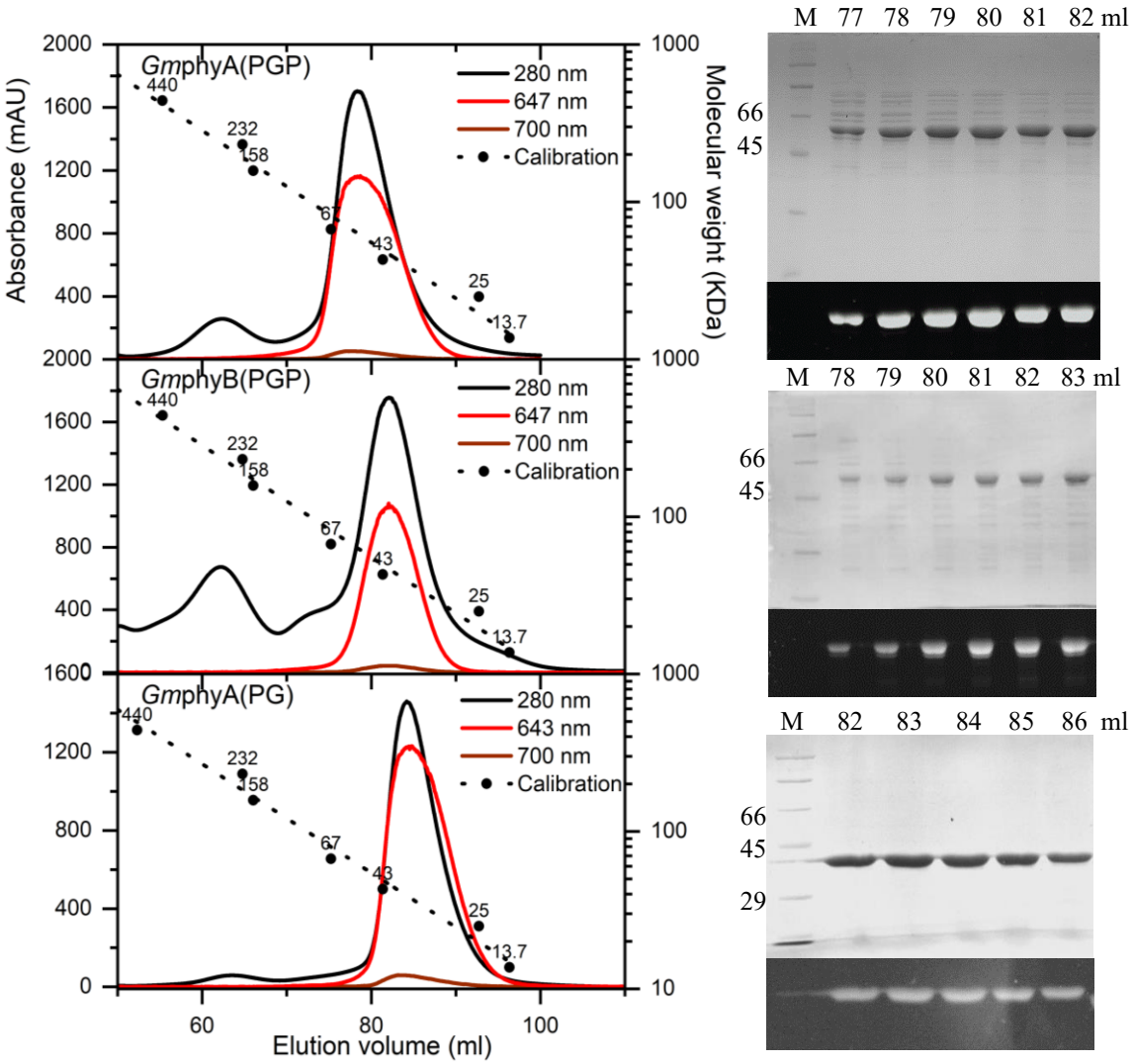


Figure 3.1 Left: SEC profiles of *Gm*phyA(PGP), *Gm*phyB(PGP) and *Gm*phyA(PG). Samples were monitored at 280 nm (black), 643 nm (or 647 nm, red) and 700 nm (dark red). The dotted line shows the SEC column calibration from the seven marker proteins (sizes labelled in kDa). Right: Coomassie and Zn^{2+} visualization of SDS-PAGE. 66, 45 and/or 29 kDa marker bands and elution volumes of fractions are shown.

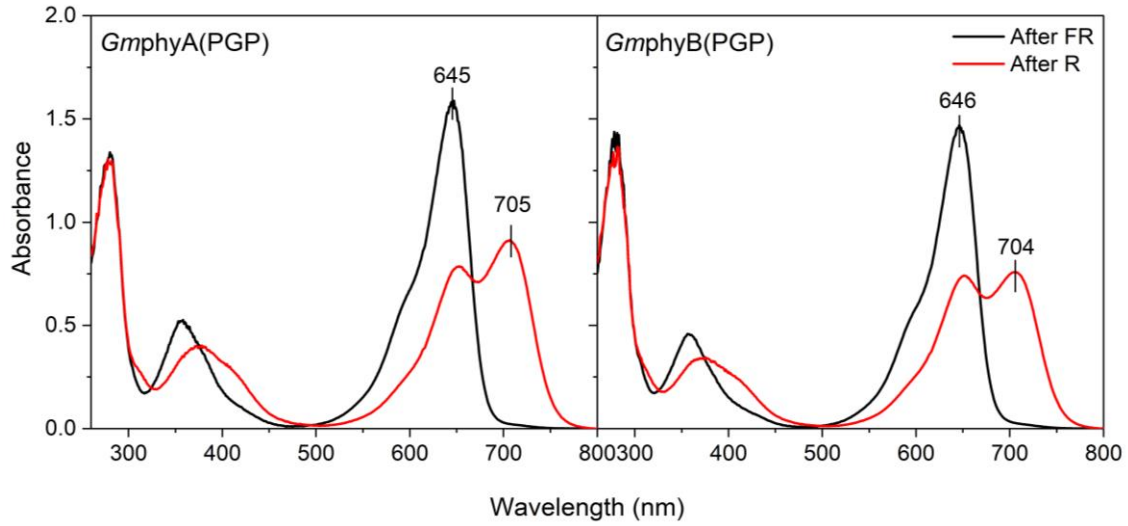


Figure 3.2 UV-Vis absorbance spectra of *Gm*phyA(PGP) and *Gm*phyB(PGP). Absorbance after far-red and red light irradiation are shown in black and red, respectively.

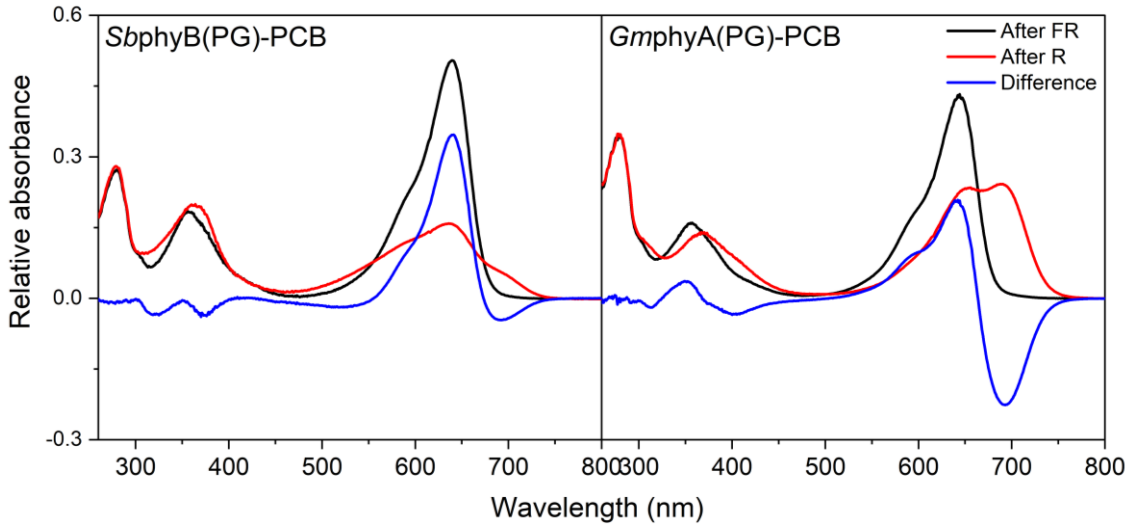


Figure 3.3 UV-Vis absorbance spectra of *Sb*phyB(PG) and *Gm*phyA(PG). Absorbance after far-red and red light irradiation are shown in black and red, respectively. Calculated difference spectra (Pr-Pfr) are shown in blue. The phyB construct was created by Dr. Soshichiro Nagano.

Most of the constructs could be produced in *E.coli* with high yield and good stability except for *Gm*phyA(NPGP) which was produced insolubly. However, the first 19 residues of the NTE in *Gm*phyA are predicted to be disordered and unstructured, while the 19-38 aa part is predicted to be helical (PredictProtein server: https://predictprotein.org/). Deletion of the first 19 residues from *Gm*phyA(NPGP) successfully produced soluble *Gm*phyA(HPGP) in *E.coli*, in harmony with the secondary structure prediction. SDS-PAGE results show *Gm*phyA(NPGP) remains in the pellet fraction, while HPGP could be

extracted as a soluble sample and purified normally. The UV-Vis absorbance spectrum showed the expected R/FR photochromicity. (Fig 3.4).

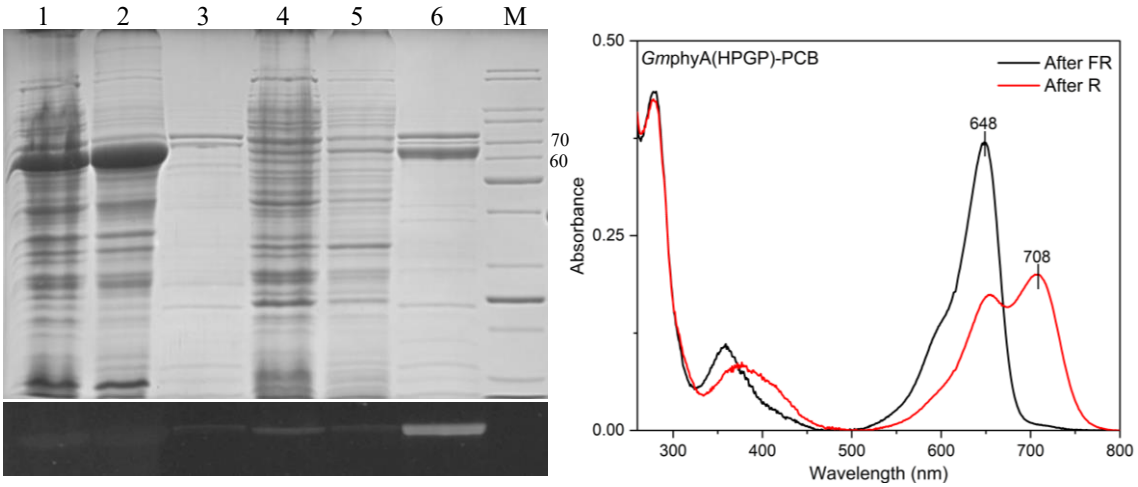


Figure 3.4 Purification and characterization of *Gm*phyA(HPGP). Left: SDS-PAGE of *Gm*phyA(NPGP) and *Gm*phyA(HPGP) Upper: Coomassie staining. Lower: Zn^{2+} staining. 1,2,3: lysate, pellet and elution samples of affinity purification of *Gm*phyA(NPGP). 4,5,6: lysate, pellet and elution samples of affinity purification of *Gm*phyA(HPGP). M: marker bands with 70 and 60 kDa labelled. Right: UV-Vis absorbance spectrum of *Gm*phyA(HPGP)-PCB.

3.1.2 Crystallisation and structural characterisation of *Gm*phyB(PGP) and *Gm*phyA(PG)

The SEC-purified *Gm*phyA(PGP) and *Gm*phyB(PGP) holoprotein samples were concentrated to 10-15 mg/ml, then centrifuged at 50000 g at 4 °C and irradiated with FR (see 2.2.8). Crystallisation plates were set up in darkness and occasional dim green light (495 nm) using 96-well sitting-drop plates and a pipetting robot (see 2.2.8). Plates were incubated at 10 °C in darkness and were regularly inspected under dim green safelight.

Crystals of *Gm*phyB(PGP) (Fig 3.5) appeared within one week in three lead conditions from the PGA Eco Screen kit (Molecular Dimensions), namely:

E5: 0.1 M MES pH 6.5, 0.2 M KBr, 0.2 M K thiocyanate, 3% w/v γ-PGA, 5% w/v PEG 3350.

E6: 0.1M MES pH 6.5, 0.2 M KBr, 0.2 M K thiocyanate, 3% w/v γ-PGA, 5% w/v PEG 4000.

E7: 0.1 M MES pH 6.5, 0.2 M KBr, 0.2 M K thiocyanate, 3% w/v γ-PGA, 5% w/v PEG 8000.

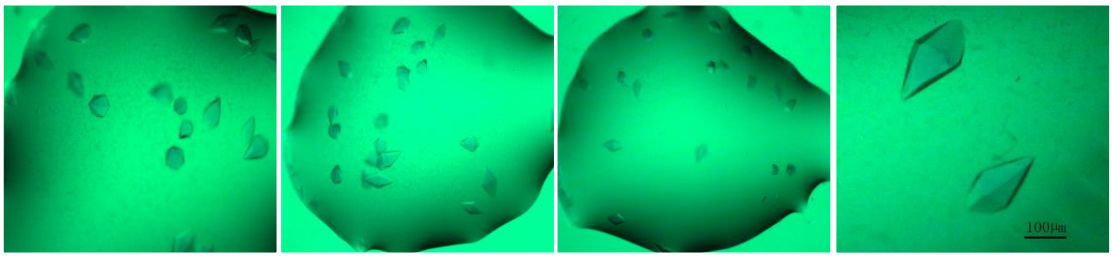


Figure 3.5 Crystals of *Gm*phyB(PGP) in Pr growing in lead conditions. (From left to right: E5, E6, E7 in sitting drop plate and E5 in a hanging drop plate). Pictures were taken under blue-green safe light to minimize photoconversion.

These crystals could be reproduced by the hanging-drop method (see 2.2.8). *Gm*phyB(PGP) crystals were mounted in loops and frozen in liquid nitrogen before the diffraction experiments. Various cryoprotectants were tried, such as ethylene glycol, glycerol, PEGs, paraffin or mineral oil as well as several cryoprotectant mixes from the CryoProtX kit (Molecular Dimensions). Blue crystals grown in hanging drops and mounted in a loop are shown in Fig 3.6. *In crystallo* optical spectroscopy (icOS) study of

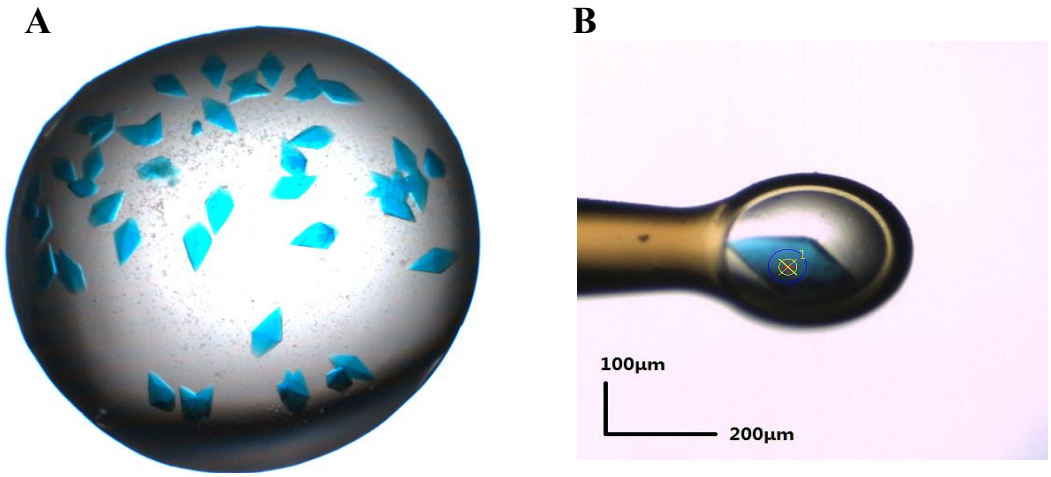


Figure 3.6 *Gm*phyB(PGP) crystals grown in a hanging drop (A) and mounted in a loop (B).

the crystals using the 100 K cryobench at BESSY II showed typical Pr absorbance characteristics, but with small differences typical of cryogenic spectra, such as a sharper red peak and more obvious shoulder at around 600 nm (Fig 3.7, measured by Prof. Jon

Hughes). Fortunately, the first diffraction trial at the BESSY II synchrotron yielded three datasets with estimated resolution of 3.5-3.7 Å. The best diffracting crystal was from the E5 condition with Cryo Mix 5 as cryoprotectant.

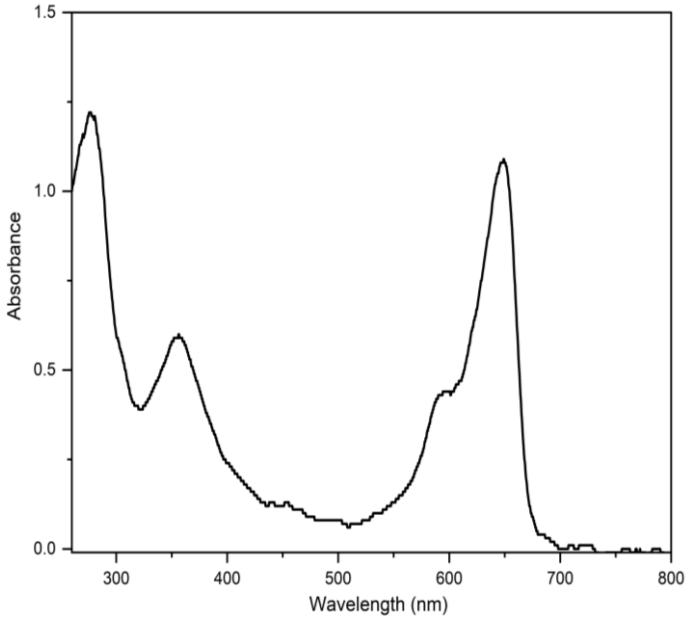


Figure 3.7 icOS absorbance spectrum of a *Gm*phyB(PGP) crystal (measured by Prof. Jon Hughes).

The lead conditions were optimised thoroughly to improve resolution. Specifically, buffer pH was varied from 6.5 to 5.5 to 9.5, KBr was replaced by NaBr, KCl, LiBr or LiCl, and additives from the Additive screening kit (Hampton Research) were tried. Further diffraction tests showed that the resolution had improved to *ca*. 3 Å when the salt was changed to LiCl or NaBr. Similar resolution improvement was noticed when NaBr was used as an additive, underlining the significance of LiCl or NaBr in *Gm*phyB(PGP) crystallisation.

The dataset collected from the best diffracting crystal was indexed, integrated and scaled as described (see 2.2.10). The phase problem was solved by molecular replacement in PHASER using the *Sorghum bicolor* phyB(PG) structure (6TBY) [89] and Cph1 (2VEA) [61] as search models. The model was built and refined manually in COOT and automatically by Refmac and Phenix.refine. These steps were carried out by Dr. Soshichiro Nagano.

The final model was determined at 2.9 Å with the R_{work}/R_{free} of 25.7% and 28.2%, respectively. The model was deposited as PDB 6TL4. The crystallography statistics are summarised in Table 7.1.

In contrast to the parallel (head-to-head) dimeric structure of *Arabidopsis* phyB (4OUR) and the antiparallel dimer of Cph1 (2VEA), *Gm*phyB(PGP) crystallised as a staggered complex (Fig 3.8A). The overall structure of *Gm*phyB(PGP) is shown in Fig 3.8B. Although the *Gm*phyB(PGP) construct comprises residues 85-615, the start at the N-terminus (85-94 aa), the 150s loop (138-152 aa) and the 380s loop (376-385 aa) are not resolved in the structure probably because of their high mobility.

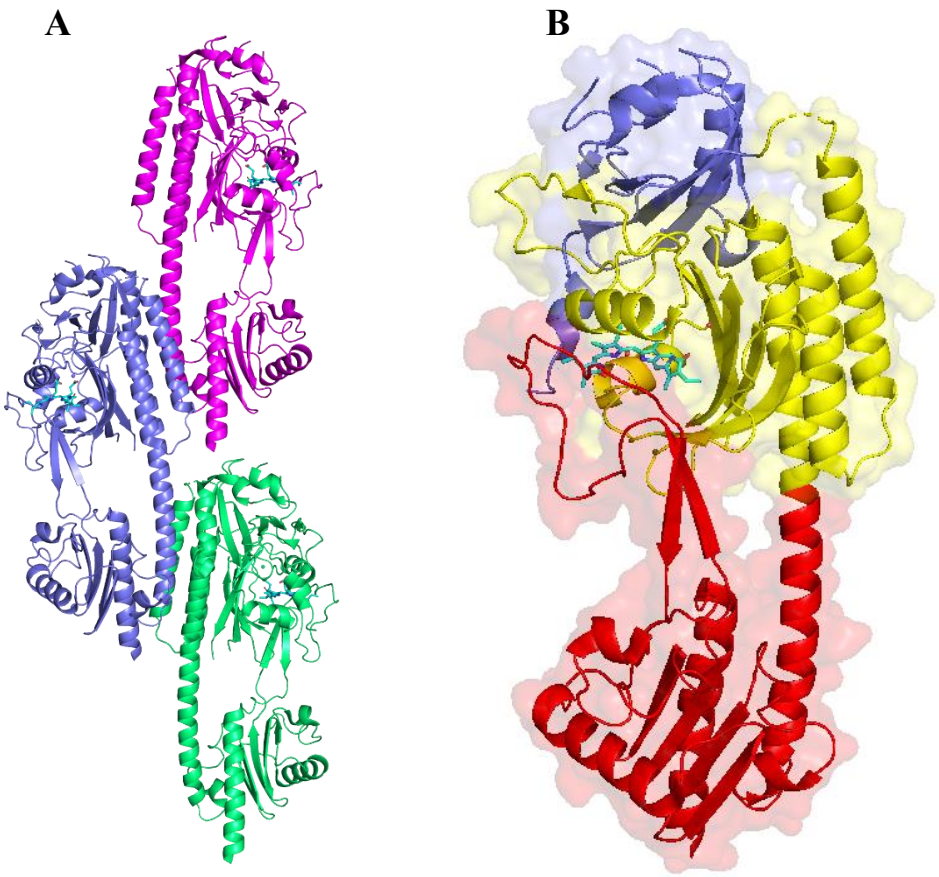


Figure 3.8 2.9 Å crystal structure of *Gm*phyB(PGP) (6TL4). A: Crystal packing as a staggered complex (each protomer shown in a different color). B: Overall structure of *Gm*phyB(PGP). nPAS, slate; GAF, yellow; PHY, red. Chromophores are shown as cyan sticks.

The chromophore and its interactions with neighbouring residues and water molecules is shown in Fig 3.9. The cofactor adapts a *ZZZssa* geometry with ring *A* attached to C353. The *A–B*, *B–C* and *C–D* ring tilts are 11°, 3° and 51° (measured by Dr. Soshichiro Nagano), respectively. As expected, the chromophore interacts extensively with nearby residues. Rings *A*, *B* and *C* form hydrogen bonds with both the D303 main chain oxygen and the pyrrole water. R348 forms a salt bridge to the propionic side chain of ring *B*, while

R318 is connected to both propionates. The histidine above the chromophore, H354, forms hydrogen bonds with the propionic side chain of ring *C* and the pyrrole water. The side chains of the three tyrosines around ring *D* (Y272, Y299 and Y357) show similar orientations to those in prokaryotic phytochromes. The hydrogen bond between the carbonyl group of ring *D* and H396 stabilises the Pr configuration of the chromophore. R575 from the PRxSF motif of the tongue forms a salt bridge to D303.

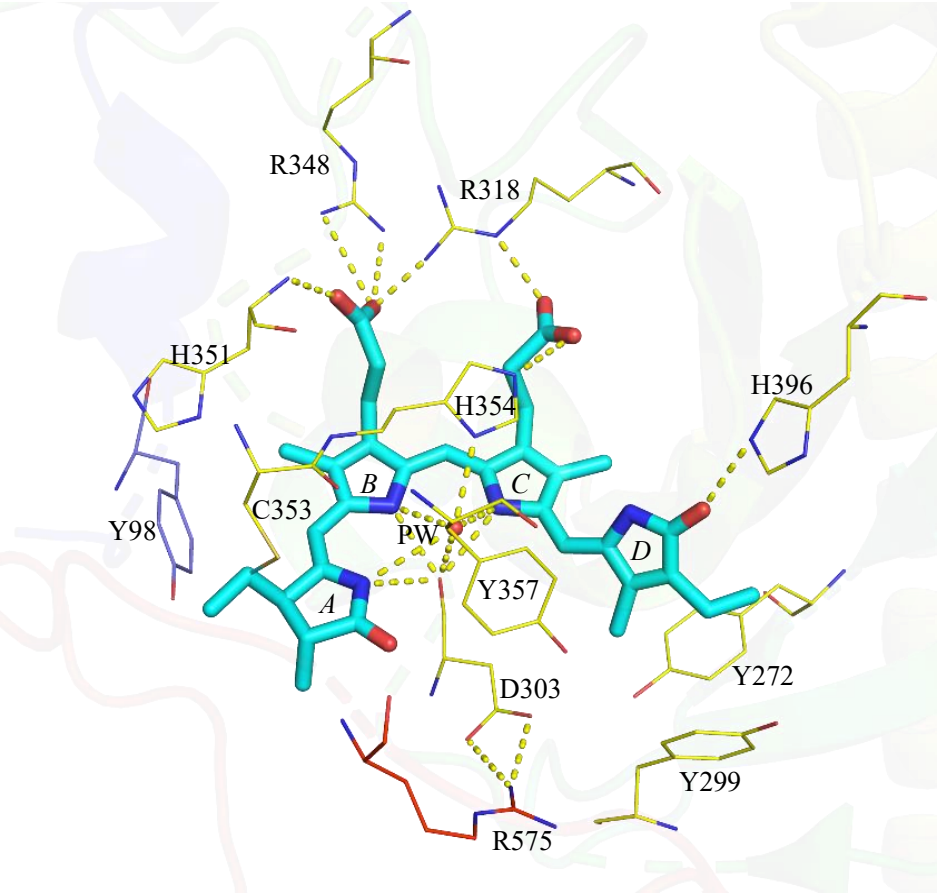


Figure 3.9 Detail of the chromophore pocket in *Gm*phyB(PGP) . The chromophore is shown as sticks with cyan carbons. Otherwise, carbons, nitrogens and oxygens are shown in yellow, blue and red. Pyrrole rings, residues and pyrrole water are labelled. Hydrogen bonds are shown as yellow dashed lines.

We were eager to determine the structure of *Gm*phyA(PGP) for comparison but, unfortunately, no crystals were obtained despite much effort (for example, various incubation temperatures (4 °C, 10 °C, 18 °C), seeding methods and even constructs different affinity tag positions (N- instead of C-terminal His tag). However, a shorter construct in which the PHY domain was deleted, *Gm*phyA(PG), crystallised successfully. Four buffer kits were used in the crystallisation screening of *Gm*phyA(PG), namely

Crystal screen (Hampton Research), JBScreen JCSG (Jena Bioscience), Morpheus and PGA Eco Screen (Molecular Dimensions). Crystals formed under many of the Morpheus kit conditions within one week (see Table 3.1). Crystals are shown in Fig 3.10.

Table 3.1 Lead conditions of *Gm*phyA(PG).

Kit	Well	Condition
Morpheus	A10	0.06 M Divalents, 0.1 M Buffer System 1 pH 6.5, 30% v/v Precipitant Mix 3
Morpheus	B3	0.09 M Halogens, 0.1 M Buffer System 1 pH 6.5, 30% v/v Precipitant Mix 3
Morpheus	B7	0.09 M Halogens, 0.1M Buffer System 2 pH 7.5, 30% v/v Precipitant Mix 3
Morpheus	C3	0.09 M NPS, 0.1 M Buffer System 1 pH 6.5, 30% v/v Precipitant Mix 3
Morpheus	C6	0.09 M NPS, 0.1 M Buffer System 2 pH 7.5, 30% v/v Precipitant Mix 2
Morpheus	C7	0.09 M NPS, 0.1 M Buffer System 2 pH 7.5, 30% v/v Precipitant Mix 3
Morpheus	C10	0.09 M NPS, 0.1 M Buffer System 3 pH 8.5, 30% v/v Precipitant Mix 2
Morpheus	F3	0.12 M Monosaccharides, 0.1 M Buffer System 1, pH 6.5, 30% v/v Precipitant Mix 3
Morpheus	F7	0.12 M Monosaccharides, 0.1 M Buffer System 2 pH 7.5, 30% v/v Precipitant Mix 3
Morpheus	F10	0.12 M Monosaccharides, 0.1M Buffer System 3, pH 8.5, 30% v/v Precipitant Mix 2
Morpheus	G7	0.1 M Carboxylic acids, 0.1 M Buffer System 2 pH 7.5, 30% v/v Precipitant Mix 3
Morpheus	G10	0.1 M Carboxylic acids, 0.1 M Buffer System 3 pH 8.5, 30% v/v Precipitant Mix 2
Morpheus	H3	0.1 M Amino acids, 0.1 M Buffer System 1 pH 6.5, 30% v/v Precipitant Mix 3
Morpheus	H11	0.1 M Amino acids, 0.1 M Buffer System 3 pH 8.5, 30% v/v Precipitant Mix 3

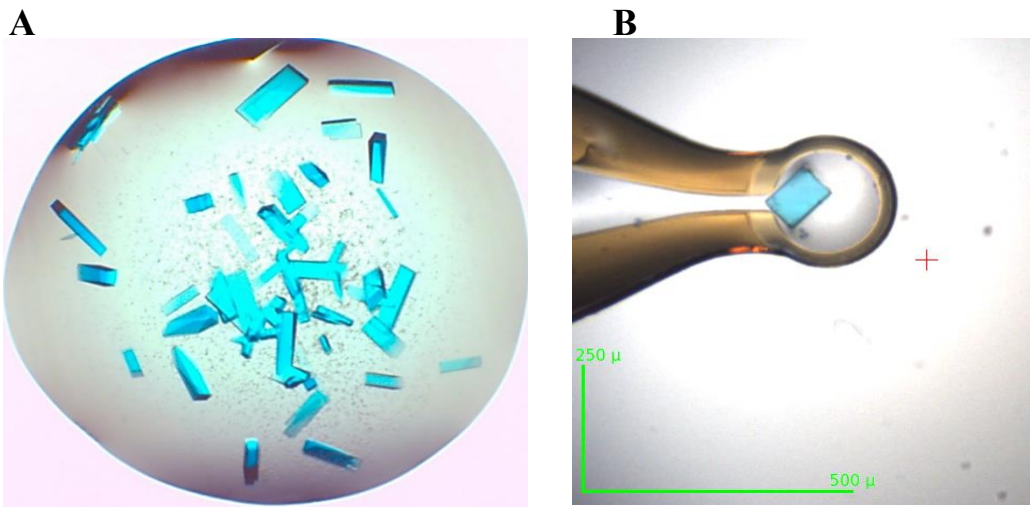


Figure 3.10 *Gm*phyA(PG) crystals grown in sitting-drop (A) and mounted in loop (B).

*Gm*phyA(PG) crystals were fished from the original sitting drops and frozen in liquid nitrogen directly since the lead conditions already provided cryoprotection. In addition, paraffin oil or mineral oil was used as an extra cryoprotectant for some crystals. Initial diffraction was examined at beamline 14.1 at BESSY II. Three datasets were collected, the best (from H3 with paraffin oil as cryoprotectant) showing diffraction to *ca*. 3 Å resolution. To further improve resolution, crystallisation conditions were optimized by reshuffling these six additives (Divalents, Halogens, NPS, Monosaccharides, Carboxylic, Amino acids), three buffers with different pH (pH 6.5, 7.5, 8.5) and two precipitants (Precipitant Mix 2 and 3). In addition, additive screening and seeding methods were tried. Crystals were fished and examined at BESSY II. Fortunately, the resolution reached 2.1 Å in the second trial. The crystal was from H11 as a sitting drop with the mother liquor as cryoprotectant.

The dataset collected was processed as described in 2.2.10. The final model was determined at 2.1 Å with the R_{work}/R_{free} of 18.9% and 24.2%, respectively. The crystallographic statistics are summarised in Table 7.2. Data processing and structure determination were done by Dr. Soshichiro Nagano.

The 2.1 Å *Gm*phyA(PG) structure (PDB 6TC7) represents the first crystal structure of an A-type plant phytochrome. Two almost identical (rmsd of 0.5 Å for Cα atoms) protomers form a head-to-head dimer. Similarly to the *Gm*phyB(PGP) structure (6TL4) described above, several residues at the N-terminus (51-66 aa), the 150s loop (111-118 aa), and the 380s loop (345-359 aa) were not resolved on account of their high mobility. Although the side chains of the other residues could be modelled nicely with clear electron density, the chromophore position and geometry remained uncertain because of the poor electron density in that region (Fig 3.12). To improve upon this, crystallisation conditions of *Gm*phyA(PG) were furtherly optimised, several new conditions being identified via cross-seeding. Luckily, crystals grown under these new conditions diffracted to a higher resolution. For example, a crystal obtained from the E9 condition of the Morpheus kit (0.12 M Ethylene glycols, 0.1 M Buffer System 3 pH 8.5, 30% v/v Precipitant Mix 1) diffracted to 1.8 Å. However, even at this resolution, no improvement was seen for the chromophore.

The weak density of the chromophore might be caused by its intrinsic dynamic features,

and the chromophores might even remain mobile in crystals even in the 100 K stream. To examine if a lower temperature could stabilise the chromophore, the BESSY MX team kindly provided a helium cryostream to cool the crystals to *ca*. 30 K during dataset collection. Two datasets were collected from the same crystal at 30 and 100 K, respectively. 1.58 Å models were deprived from both, and, interestingly, both showed clear electron density for the chromophore, enabling highly confident modelling of that region.

In general, the 1.58 Å structure resembles 6TC7 (Fig 3.11; rmsd of 0.3 Å for Cα atoms). However, the clear electron density of the cofactor (see $2F_O$-F_C map in Fig 3.12) allowed much more accurate modelling. In particular, the ring *D* ethyl group modelled towards the β face in 6TC7, is rotated about 120° to point to the α surface in the new model. During structure refinement with the new data, the F_O-F_C map using the previous model showed a positive density region above the ethyl group that was eliminated in the new model (Fig 3.13).

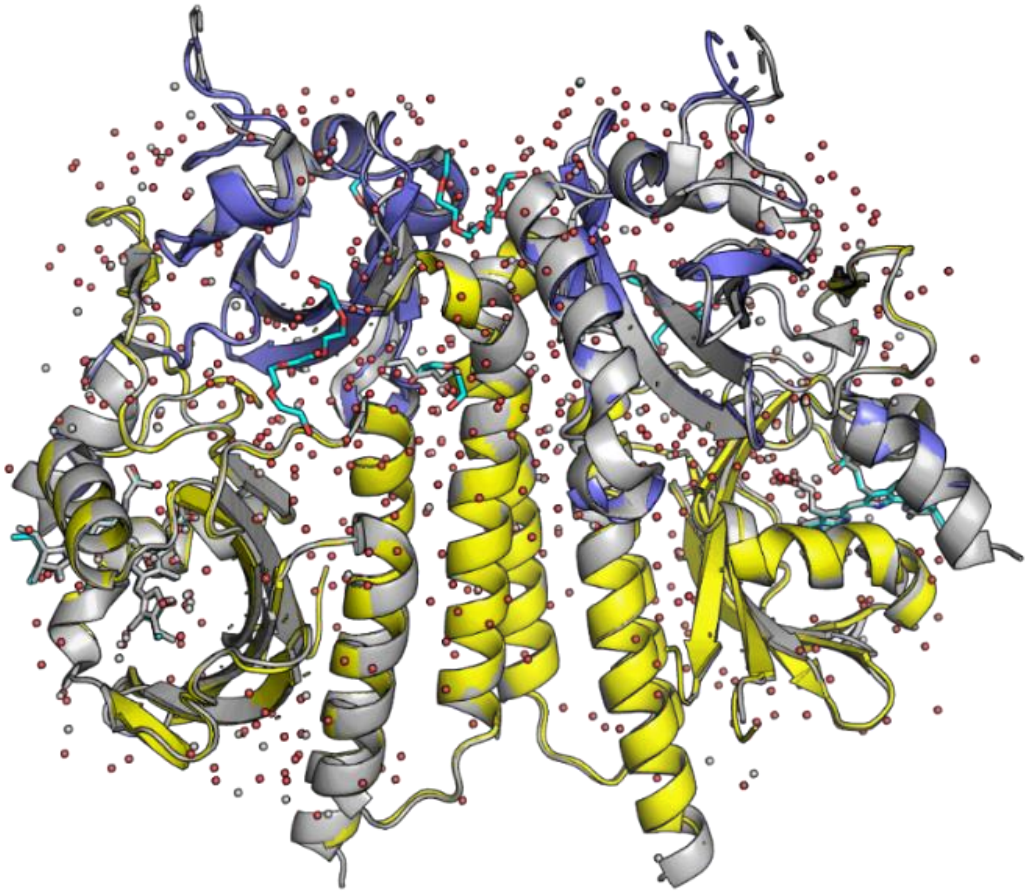


Figure 3.11 The 1.58 Å structure of *Gm*phyA(PG) (nPAS: slate, GAF: yellow, Water: red dots) superposed with 2.1 Å 6TC7 (grey).

The high-resolution structure reveals details of the protein-chromophore interactions in the chromophore binding pocket (Fig 3.14), most of which are homologous to those seen in other phytochromes. PCB is attached to C323 via a thioether bond. A group of hydrogen bonds, salt bridges and hydrophobic interactions involve many conserved amino acids and water molecules (for example, R318, R288, H324, H370, D273 and the pyrrole water (PW)). The main chain oxygen of D273, PW and the nitrogen atoms of PCB rings *A*, *B* and *C* have extensive contact with each other, while the propionate chains of rings *B* and *C* hydrogen bond with R318, R288 and several conserved water molecules. Moreover, a hydrogen bond between H370 and the carbonyl group of ring *D* apparently stabilises the *Za* configuration in Pr. H324 acts as a bridge between PW and the propionate chain of ring *C*. The three tyrosine residues around ring *D* (Y242, Y269 and Y327) provide space for ring *D* rotation upon photoisomerization. In addition, S278, S333 and S338 contact indirectly with PCB through water molecules.

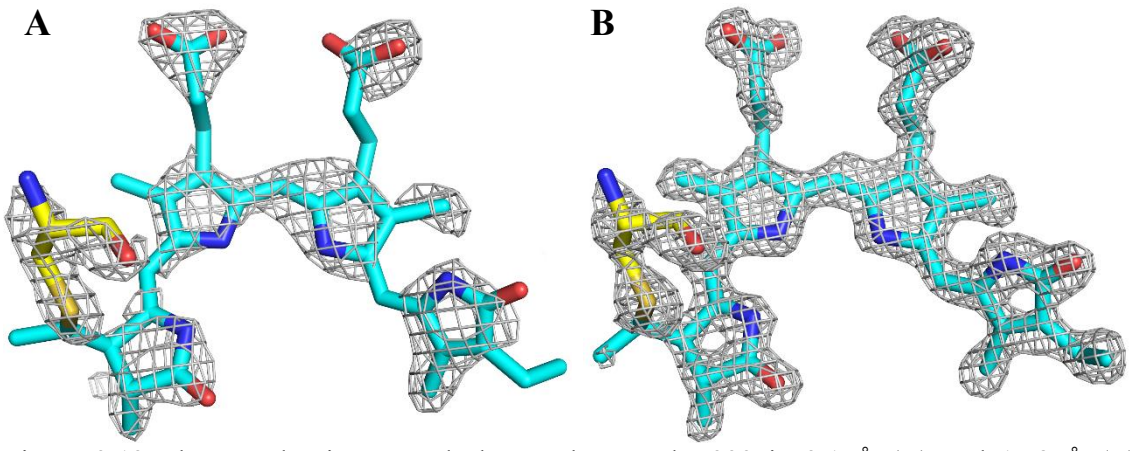


Figure 3.12 Electron density around chromophore and C323 in 2.1 Å (A) and 1.58 Å (B) *Gm*phyA(PG) structure. 2Fo-Fc map contour at 1.5 σ (grey mesh). PCB: cyan, C323: yellow.

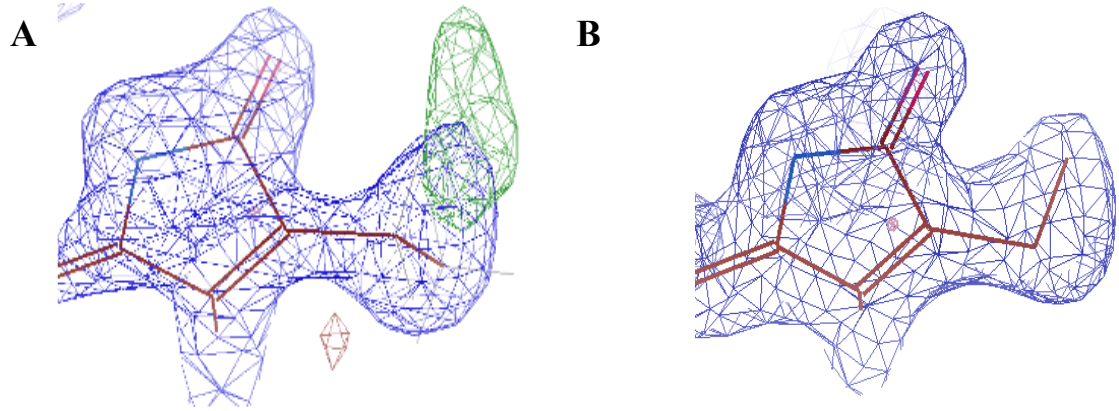


Figure 3.13 The ring *D* ethyl group position changes from β-facial (A) to α-facial (B) indicated by the Fo-Fc map contour at 3 σ (green mesh), and 2Fo-Fc map contour at 1 σ (blue mesh).

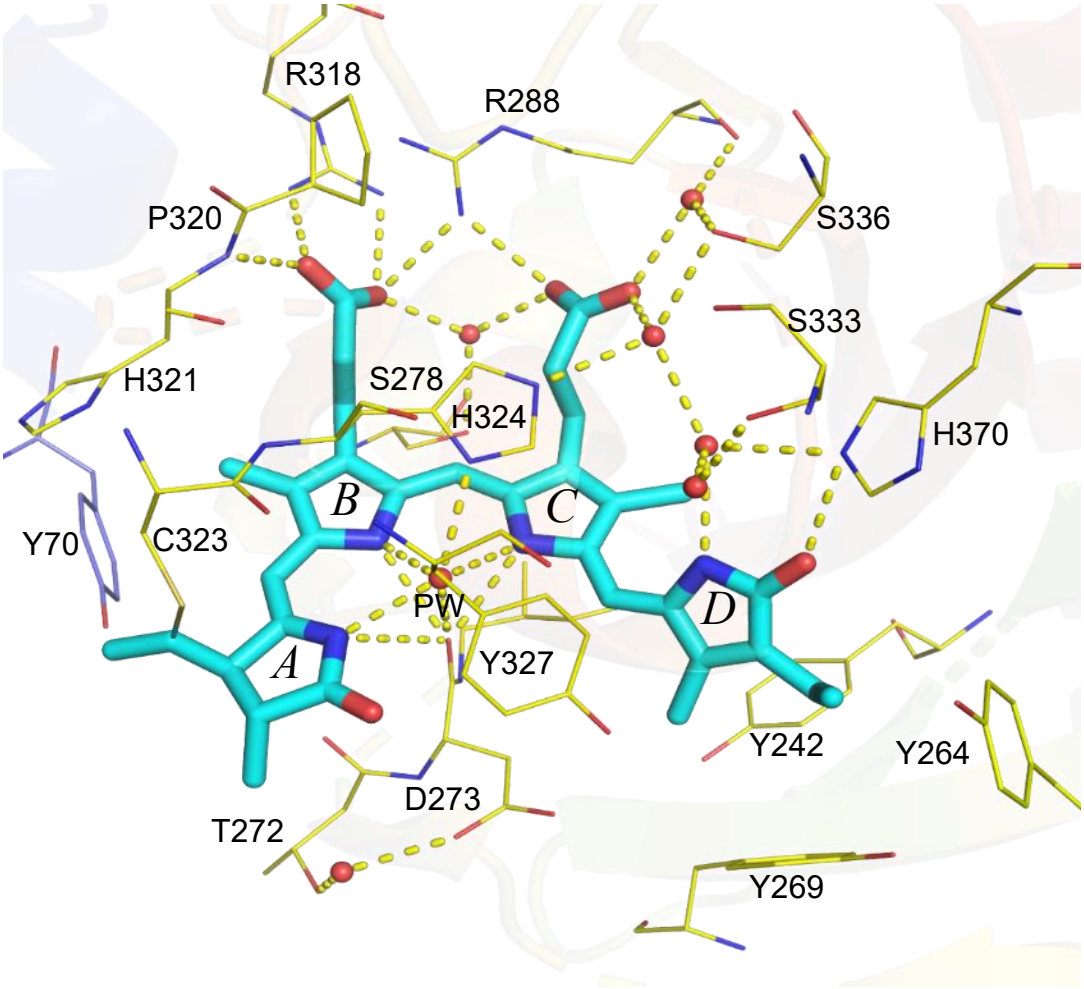


Figure 3.14 Protein-chromophore interactions in the chromophore binding pocket of *Gm*phyA(PG). Residues of nPAS and GAF domain (slate and yellow sticks, respectively), water molecules (red dots), and hydrogen bonds (yellow dashed lines) are shown.

3.2. Attempts to determine plant phytochrome Pfr structure

3.2.1 Production and characterisation of Pfr-stabilised and -mimic mutants

Structures of plant phytochrome in the Pfr state are particularly interesting as they would not only help us to understand the photoactivation mechanism but also the translocation and cellular signalling mechanisms. However, crystallisation of plant phytochrome as Pfr is a tough project. First of all, Pfr reverts to Pr via a thermal mechanism that is likely to be much faster than any crystallisation. Secondly, phytochromes exist as Pr-Pfr mixtures after R light irradiation on account of their overlapping spectral absorbance. Such sample heterogeneity hinders crystallisation. On the other hand, previous research that described

the *Dr*BphP structure in the illuminated form (4O01) suggested that periodic R light pulses might allow crystallisation of phytochrome as Pfr by maintaining a relatively high Pfr proportion despite dark reversion [68]. However, R light leads to cycling between Pr and Pfr which might prevent the formation of Pfr crystals entirely. Interestingly, however, several tongue mutations that stabilise Pfr by attenuating dark reversion have been reported, for example, PRxSF→PAxSF or PRxSW and WGG→WEG. Indeed, the crystal structure of *Dr*BphP as Pfr (5C5K) represents the PRxSF→PRxSW mutant rather than the wild-type [70]. A further and rather intriguing route to the signalling structure is presented by the Y276H mutant of *Arabidopsis* phyB. This variant triggers photomorphogenesis in darkness in transgenic plants [65], suggesting that it mimics the Pfr conformation although its absorbance remains typical of Pr. Several routes to the Pfr structure were therefore investigated.

To gain insight into plant phytochrome Pfr structures, the three Pfr-stabilised mutants and the Pfr-mimic mutant of various *Gm*phyA and *Gm*phyB constructs were studied. It should be noted that the NTE stabilises Pfr state of plant phytochromes, its deletion greatly accelerating dark reversion [72]. Consequently, the NPGP construct would be a more appropriate candidate than the PGP construct in terms of crystallisation trials of Pfr.

The Pfr-stabilised mutants (R575A, G557E, F578W) and Pfr-mimic mutant (Y272H) of *Gm*phyB(NPGP) construct were generated, produced and purified following the protocol described in the Methods section 2.2.3 and 2.2.4. *Gm*phyB(NPGP) WT and mutants were obtained with high yield and purity as shown by SEC profiles and SDS-PAGE of the peak fractions (Fig 3.15). Typical R/FR photochromicity with almost identical absorption maxima to those of the wild type was observed for the Pfr-stabilised mutants, whereas, the Pfr-mimic mutant did not show Pfr-typical absorbance following R light irradiation (Fig 3.16), similar to its homologs in Cph1 and *Arabidopsis* phyB [65,90].

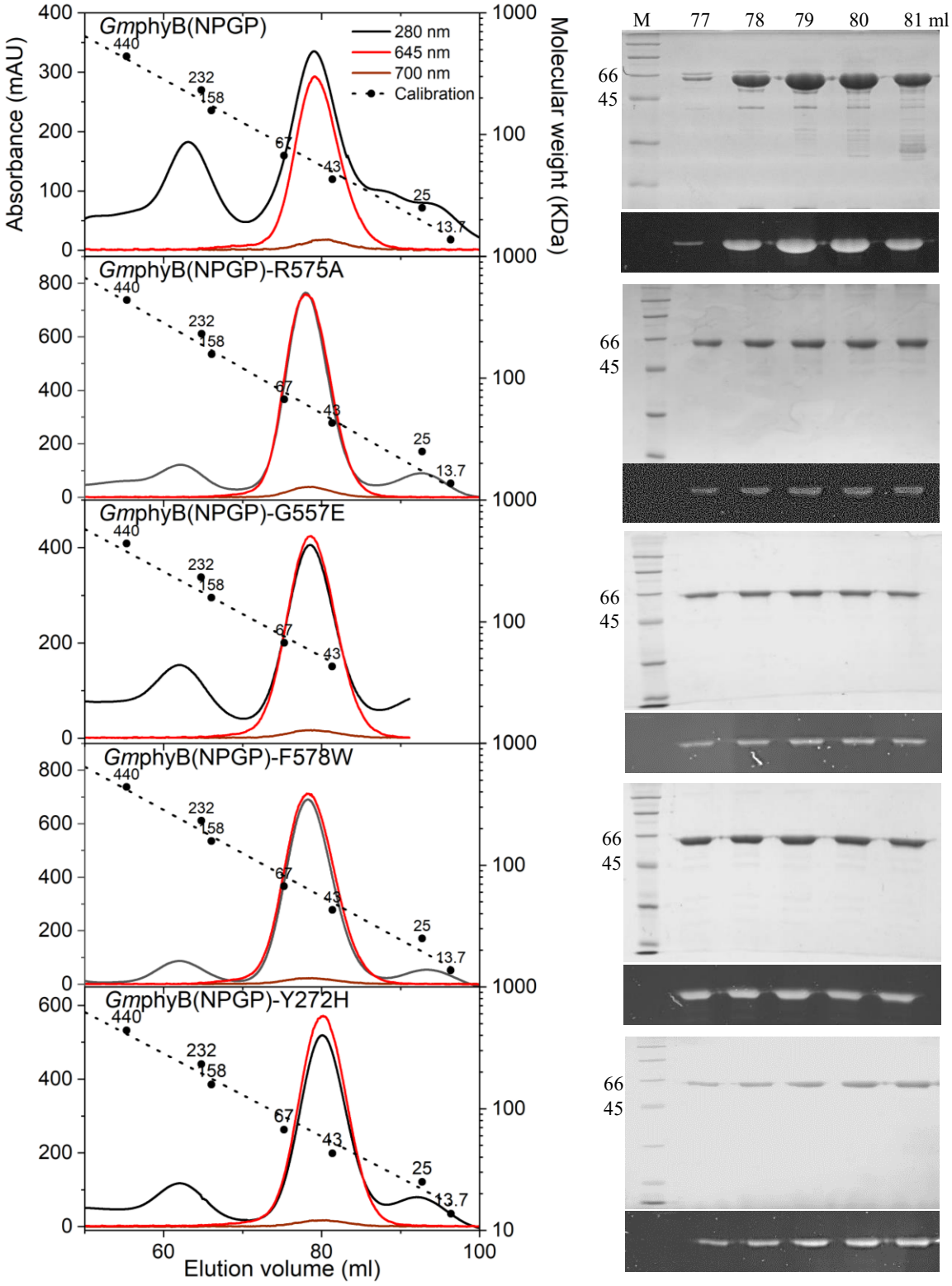


Figure 3.15 Left: SEC profiles of *Gm*phyB(NPGP) WT and mutants. Protein monitored at 280, 645 and 700 nm, dotted line show the SEC column calibration. Right: Coomassie and Zn^{2+} visualization of SDS-PAGE. 66 and 45 kDa marker band and elution volume of fractions are labelled.

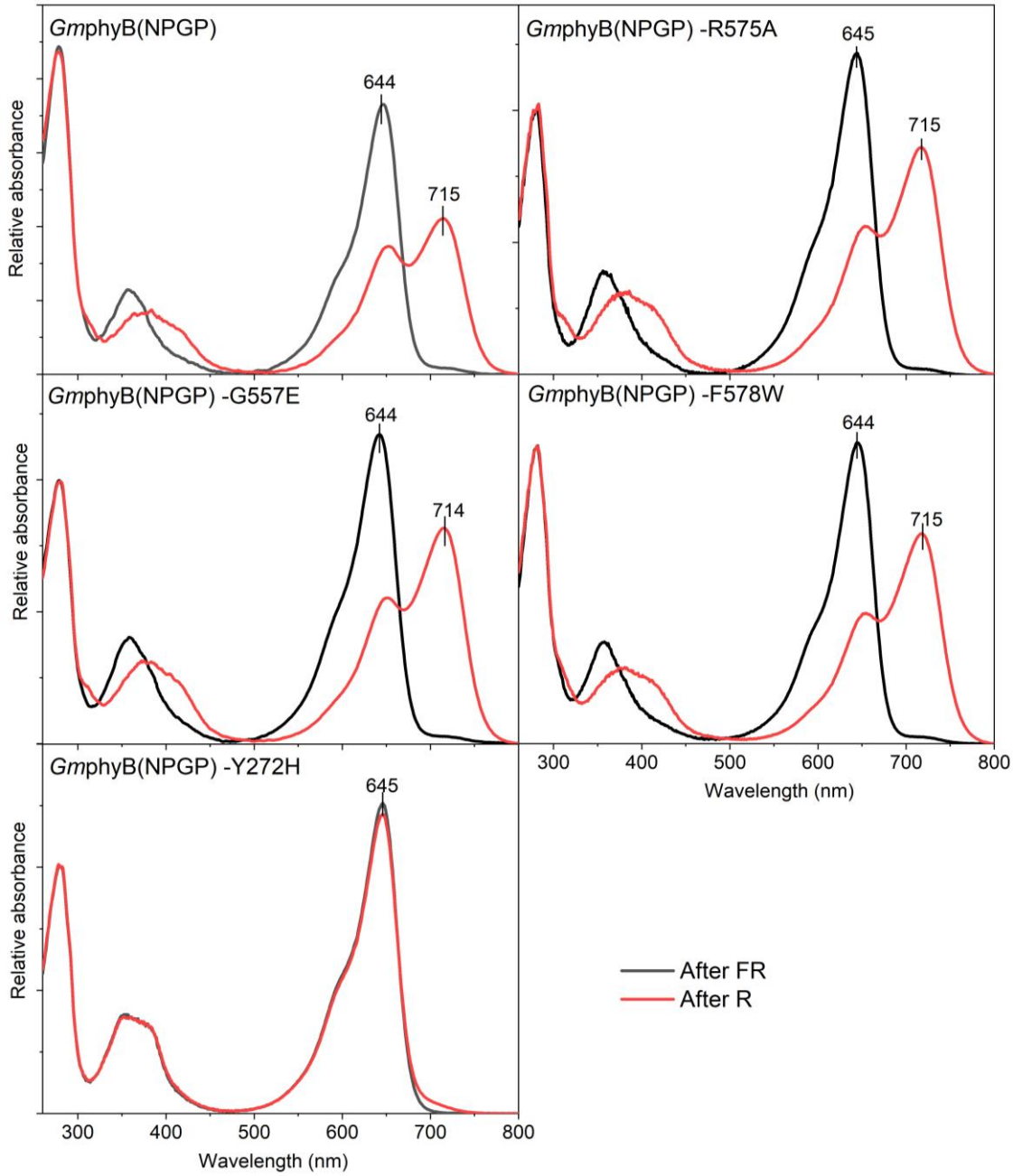


Figure 3.16 Absorbance spectra of *Gm*phyB(NPGP) WT and mutants. Absorbance after far-red and red light irradiation are shown in black and red, respectively.

To examine if these mutations indeed stabilised Pfr, the dark reversion rate was measured with the help of a plate reader and compared to WT (see 2.2.7). As expected, all three mutants showed substantially slower dark reversion than the *Gm*phyB(NPGP) wild-type parent (Fig 3.17). Interestingly, whereas the R575A and F578W mutant data could be fitted by a single exponential decay model, with a half-life of about 1290 min, data for

the wild type and the G557E mutant required a biexponential, with a half-life of 20 min and 348 min for WT, 28 min and 745 min for G557E.

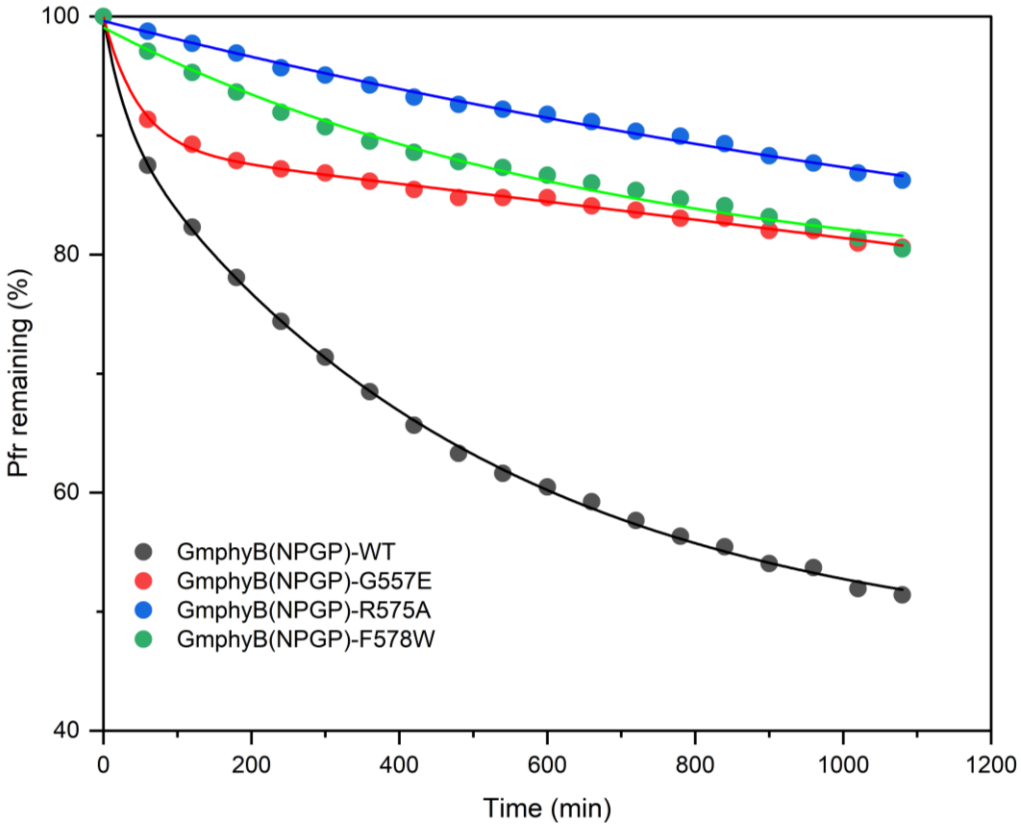


Figure 3.17 Dark reversion measurement of WT and three Pfr-stabilising mutants of *Gm*phyB(NPGP). The solid lines show the exponential decay fits.

Since we were for some unknown reason not able to produce *Gm*phyA(NPGP) as soluble protein from *E. coli*, the Pfr-stabilised mutant (R549A) and Pfr-mimic mutant (Y242H) were generated based on *Gm*phyA(PGP) construct. The elution peak of both mutants on SEC column was at about 78 ml (Fig 3.18), suggesting a hydrodynamic radius of 64 kDa protein according to the calibration, close to the calculated molecular weight of the monomer (60 kDa). The peak fractions collected from the SEC were analyzed by SDS-PAGE (Fig 3.18), the data indicating that the holophytochromes had been expressed and purified successfully.

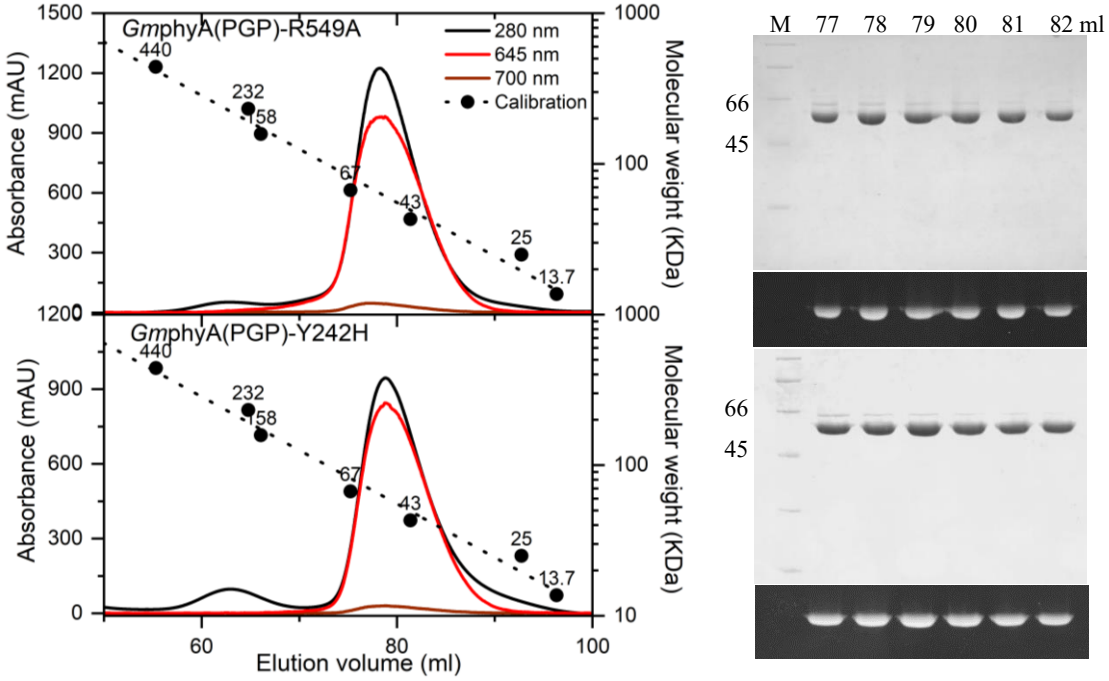


Figure 3.18 SEC profiles and SDS-PAGE of the peak fractions of *Gm*phyA(PGP)-R549A and -Y242H mutants. SEC of the protein sample monitored at 280, 645, 700 nm. The dotted line shows the SEC column calibration with marker proteins (kDa). Coomassie and Zn^{2+} visualization of SDS-PAGE. 66 and 45 kDa marker band and elution volume of fractions are labelled.

The absorbance spectra of the SEC-purified proteins are shown in Fig 3.19. After FR irradiation *Gm*phyA(PGP)-R549A and -Y242H mutants showed typical absorption peaks at *ca.* 645 nm. After R irradiation, *Gm*phyA(PGP)-R549A showed an additional absorption peak at about 705 nm, similar to that of the *Gm*phyA(PGP) wild type (Fig.

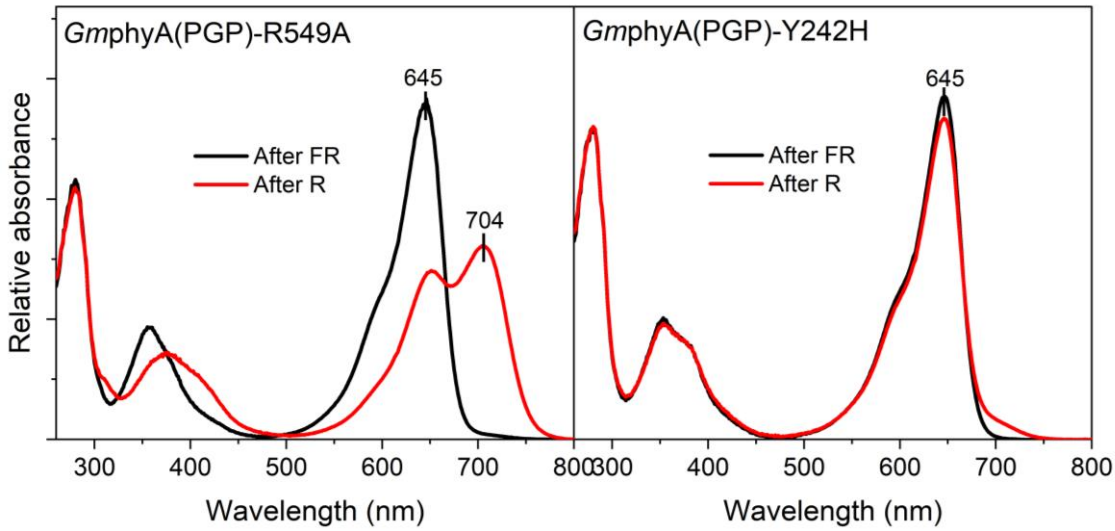


Figure 3.19 UV-Vis absorbance of *Gm*phyA(PGP)-R549A and -Y242H mutants. Absorbance after FR and R light irradiation shown in black and red, respectively.

3.2), while the *Gm*phyA(PGP)-Y242H mutant showed only a weak absorbance shoulder in the FR, as expected.

Similar to the Pfr-stabilised mutants of *Gm*phyB, the R549A mutant of *Gm*phyA(PGP) showed a much slower dark reversion rate than the WT. The dark reversion data of both could be fitted by a two-phase exponent (Fig 3.20), with half-life of 60 min and 600 min for WT, 60 min and 1846 min for R549A.

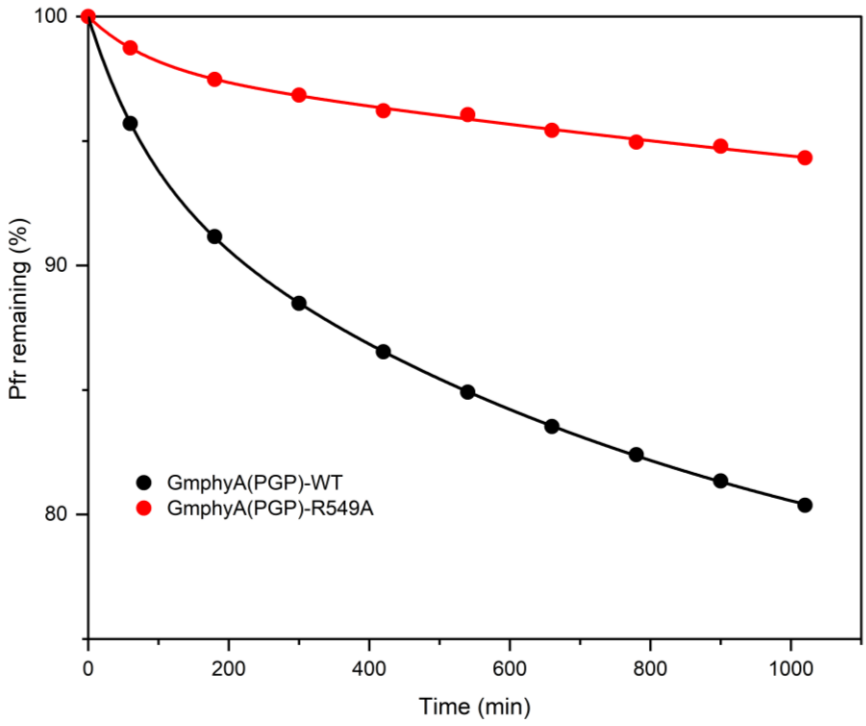


Figure 3.20 Dark reversion of *Gm*phyA(PGP) WT and R549A mutant. The black and red lines show the biphasic exponential decay fit.

The corresponding substitution of *Gm*phyA-Y242H in Cph1 (Y176H) is strongly fluorescent, thus fluorescence in phyA mutants was studied spectroscopically. For *Gm*phyA(PGP)-Y242H, the excitation and emission spectra showed maxima at 647 and 669 nm, respectively, while 649 and 672 nm were recorded for *Gm*phyA(PG)-Y242H (Fig 3.21). The excitation spectra corresponded well with the absorbance spectra. The measured fluorescence quantum yield of *Gm*phyA(PGP)-Y242H was 13.2 %, slightly higher than that of *Gm*phyA(PG)-Y242H (10.7 %). (Fig 3.21. These analyses were carried out together with Dr. Soshichiro Nagano)

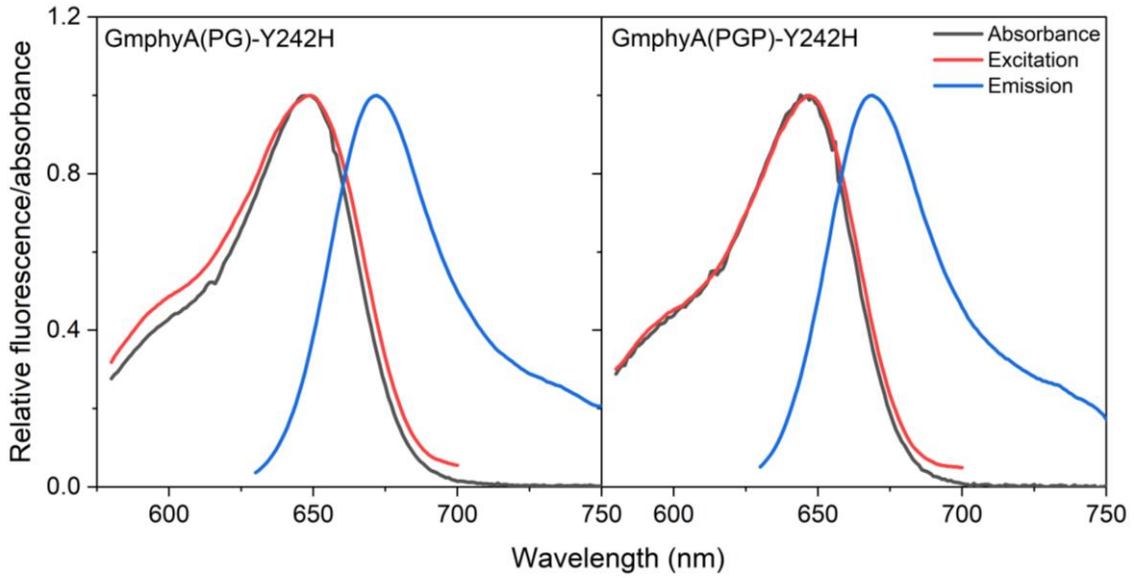


Figure 3.21 Absorbance (black line), excitation (red line) and emission (blue line) spectra of *Gm*phyA(PG)-Y242H and *Gm*phyA(PGP)-Y242H.

3.2.2 Crystallisation and X-ray diffraction of *Gm*phyA(PGP)-R549A and Y242H mutants

The SEC-purified Pfr-stabilised mutants of *Gm*phyB(NPGP) and *Gm*phyA(PGP) were concentrated to about 12 mg/ml and irradiated with 590 nm light from an LED before crystallisation screening experiments. The crystallisation plates were incubated at 10 °C in a light box providing periodic 590 nm light pulses (three seconds every four hours) to maintain a high Pfr proportion, while plates of the Pfr-mimic mutants were incubated at 10 °C in darkness. Although mutants of *Gm*phyB(NPGP) yielded no crystals to date, both *Gm*phyA(PGP)-R549A and Y242H mutants crystallised within two weeks. *Gm*phyA(PGP)-R549A crystals formed in a lead condition containing 0.1 M Bicine pH 9.0, 10 % (v/v) 2-Methyl-2,4-pentanediol. *Gm*phyA(PGP)-Y242H crystallised under two conditions, namely in 0.1 M sodium acetate pH 4.5, 0.2 M zinc acetate, 10 % (w/v) polyethylene glycol 3000 and in 0.1 M magnesium formate, 50% (v/v) glycerol (see Fig 3.22). It should be noted that both *Gm*phyA(PGP)-R549A and *Gm*phyA(PGP)-Y242H crystals have poor reproducibility while seeding (see 2.2.8) improved this situation.

Diffraction of *Gm*phyA(PGP)-Y242H and R549A crystals was checked at BESSY II. Unfortunately however, none of the crystals has shown useful diffraction to date despite many efforts in optimising crystallisation conditions and post-crystallisation treatments

such as cryo-annealing and dehydration. Moreover, the Y242H crystals failed to diffract even *in situ* at room temperature in our recent preliminary measurements (performed by Prof. Jon Hughes and Dr. Soshichiro Nagano with the help of beamline scientists), implying that the cryo-protection/freezing procedure is not the problem.

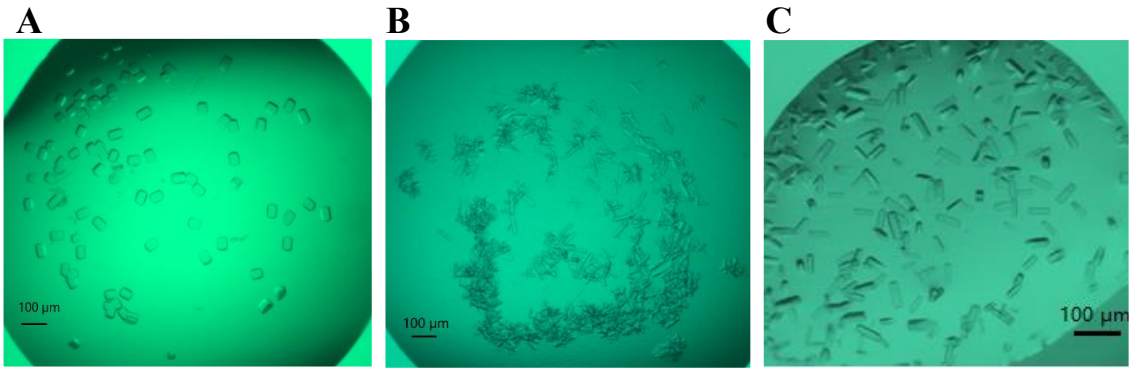


Figure 3.22 *Gm*phyA(PGP)-R549A crystals (A) and *Gm*phyA(PGP)-Y242H crystals grown in two different conditions (B and C). Scale bar: 100 µm.

3.2.3 Photoconversion of Pr crystals and crystallisation of *Gm*phyA(PG) illuminated state

In addition to crystallisation trials of Pfr-stabilised and Pfr-mimic mutants, direct photoconversion of Pr crystals might be another way to determine plant phytochrome Pfr structure through subtraction of the Pr state information from a Pr/Pfr mixture.

The absorbance spectrum of a *Gm*phyB(PGP) crystal measured at the BESSY II cryobench showed a Pfr absorbance peak after white light irradiation at ambient temperature, suggesting that crystals could be photoconverted (Fig 3.23, measured by Prof. Jon Hughes). However, the crystals had completely lost diffraction after photoconversion. On the other hand, the *Gm*phyA(PG) crystals not only showed extensive photochromicity, but diffraction was also retained. To maximise the Pfr proportion in these crystals, orange (590 nm) light which could penetrate the crystals deeply was used to irradiate the crystals before freezing in liquid nitrogen.

Further experiments at BESSY II confirmed that orange-light-irradiated *Gm*phyA(PG) crystals diffracted to a similar resolution to that seen prior to irradiation. However, the structures derived from datasets from photoconverted crystals closely resembled the Pr

structure derived from datasets from photoconverted crystals closely resembled the Pr structure except that electron density corresponding to the chromophore was completely missing.

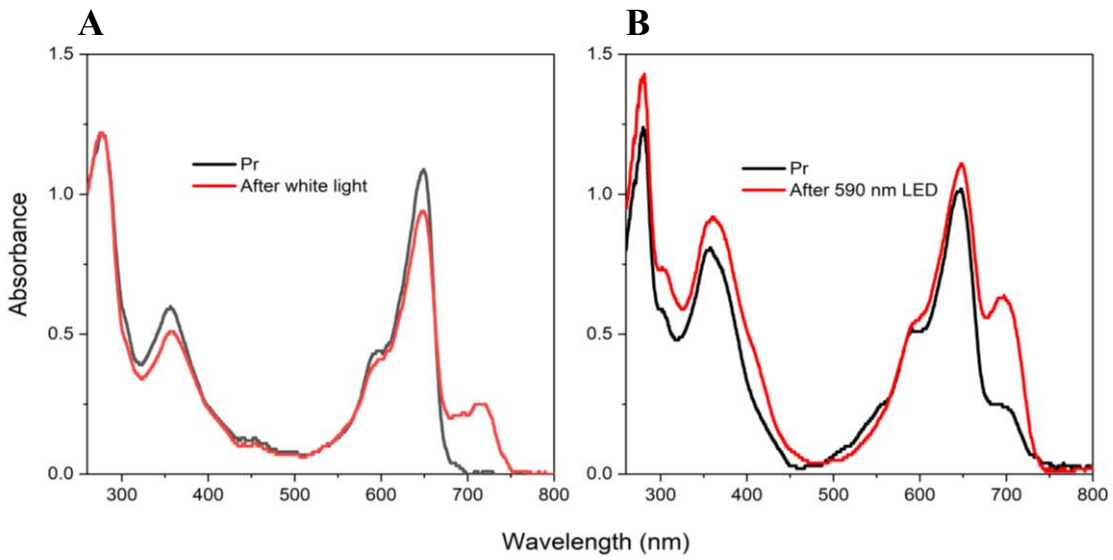


Figure 3.23 icOS absorbance spectra of *Gm*phyB(PGP) (A) and *Gm*phyA(PG) (B) crystal as Pr and after light irradiation (measured by Prof. Jon Hughes).

Photoconversion of the Pr crystals therefore does not seem to be a straightforward way to determine the Pfr structure. To crystallise Pfr or a Pr/Pfr mixture of *Gm*phyA(PG), the holoprotein was irradiated with 590 nm orange light before crystallisation plates were set up, and subsequently with pulses (three seconds every four hours). Crystals appeared in the majority of the conditions of the Morpheus kit within one week. Spectrum of the crystal clearly shows a Pfr-like absorbance (Fig 3.24B, measured by Prof. Jon Hughes) Interestingly, although some crystallisation conditions overlapped with those that generated Pr crystals, others were novel. Datasets were collected, followed by data processing and model building. The packing pattern is distinct from the Pr crystals (see Table 7.3 and Table 7.4). Firstly, the space group was found to be P 4_3 2 2, instead of P 2_1 for Pr crystals. Secondly, the unit cell dimensions were 55, 55, 266Å/90, 90, 90°, different from the 55, 111, 68 Å/90, 92.3, 90° of Pr crystals. Thirdly, each asymmetry unit contained one protein molecule, rather than the head-to-head dimer of the Pr structure.

Although the illuminated structure is similar to that of 1.58 Å Pr structure, the conformation of the N-terminus (residues 68-78) is different. In detail, it was modelled

as an unstructured loop rather than a short helix in Pr structure (Fig 3.24A). It would be interesting to observe possible position changes of the chromophore and key residues around it, but, rather surprisingly, only minimal effects were apparent. In particular, the chromophore is not resolved due to the poor electron density, a similar problem to was observed in the 2.1 Å Pr structure 6TC7. The latter problem was solved by minimising the green safelight dose (Fig 3.12). Since the illuminated crystals were frozen under orange light, chromophore cycling is likely to have caused the missing electron density. To further examine if darkness could stabilise chromophore, new batches of crystals were fished under orange light but frozen after a few seconds of darkness. Indeed, this resulted in a better density quality around chromophore (Fig 3.24C).

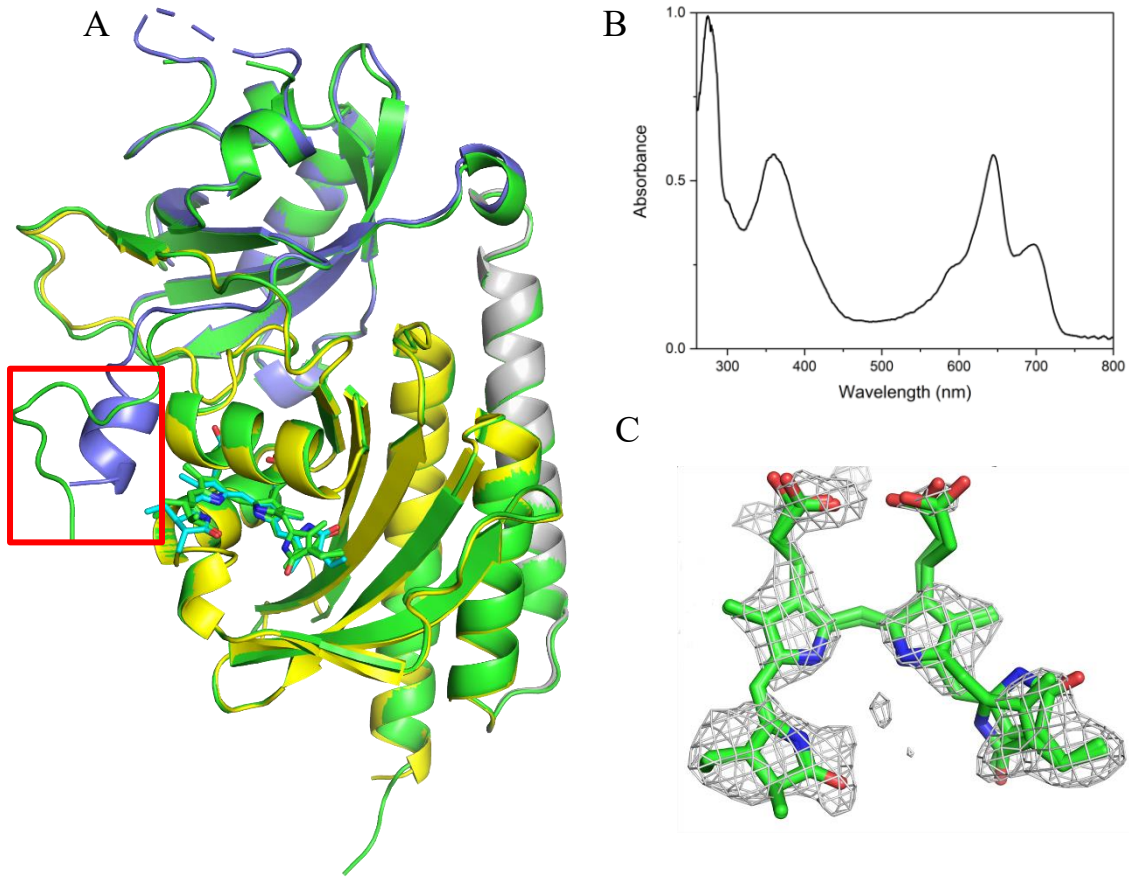


Figure 3.24 A: superimposition of illuminated structure (green) and Pr structure (nPAS: slate, GAF: yellow) of *Gm*phyA(PG). conformational changes indicated by red square. B: absorbance spectrum of illuminated crystal (by Prof. Jon Hughes). C: chromophore shown as green sticks and 2Fo-Fc map contour at 1 σ (grey mesh).

3.3 Plant phytochrome signal transduction

3.3.1 A putative nuclear localisation signal in phyA: KRKR motif

Plant phytochromes upon photoconversion to Pfr translocate into the nucleus to transduce light signals. The nuclear translocation of phyA is strictly FHY1/FHL-dependent, whereas phyB translocation does not involve FHY1.

Interestingly, we discovered four exposed basic residues (KRKR) by chance in the 380s loop region of *Gm*phyA, constituting a typical nuclear localisation signal that had gone unnoticed since Hershey's cloning of the phyA gene in 1984 [91]. Sequence alignment showed that the consensus K(R/K)K(R/K) is conserved in and exclusive to the phyA lineage (Fig 3.25). KRKR represents a Class I NLS [92] and is also flagged in the phyA sequences at high probability by various NLS prediction algorithms. Moreover, several publications already showed that KRKR conferred nuclear translocation of proteins, whereas mutation of the sequence abolished the effect [93,94]. The actual function of the

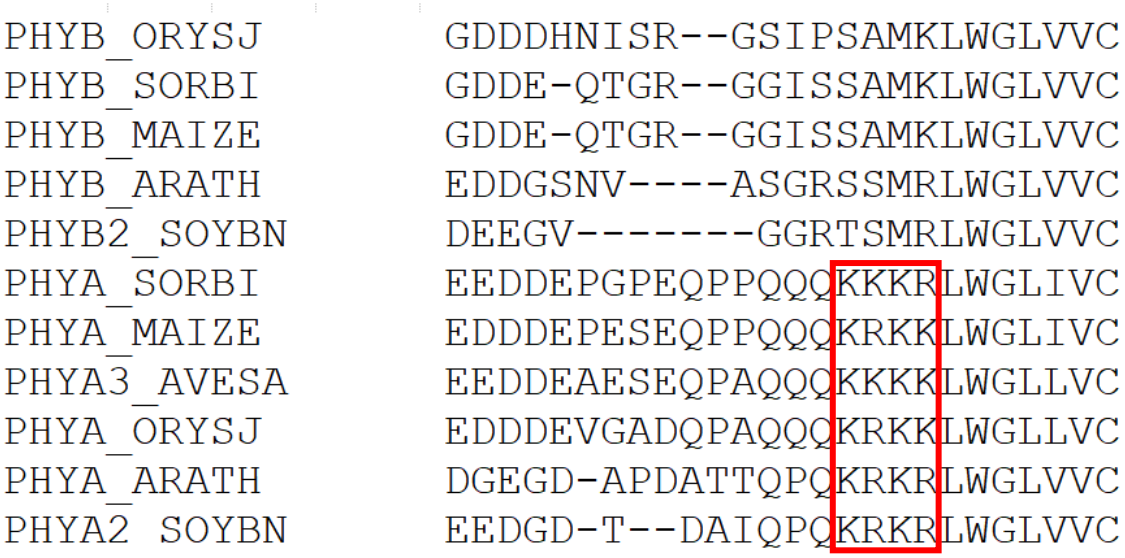


Figure 3.25 Sequence alignment of various plant phytochromes showed conserved K(R/K)K(R/K) in phyA. ORYSJ: *Oryza sativa*, SORBI: *Sorghum bicolor*, MAIZE: *Zea mays*, ARATH: *Arabidopsis thaliana*, AVESA: *Avena sativa*, SOYBN: *Glycine max*.

K(R/K)K(R/K) motif in type-A phytochromes is of considerable interest, given that their nuclear translocation absolutely requires the Pfr state and FHY1 function. To examine whether KRKR functions as an NLS in plants, particle bombardment was performed to transfect plasmids encoding YFP either alone, with the KRKR motif or with the motif

mutated to KAKR into onion epidermal cells. The coding sequence of the classical SV 40 large T antigen NLS (PKKKRKV) attached to YFP as a positive control alongside the mCherry-NLS plasmid which we routinely include as a nuclear marker (here the NLS is from the VirD2 protein of *Agrobacterium*, see 2.2.1). Part of this work was done by undergraduate Tim Ruder under my supervision.

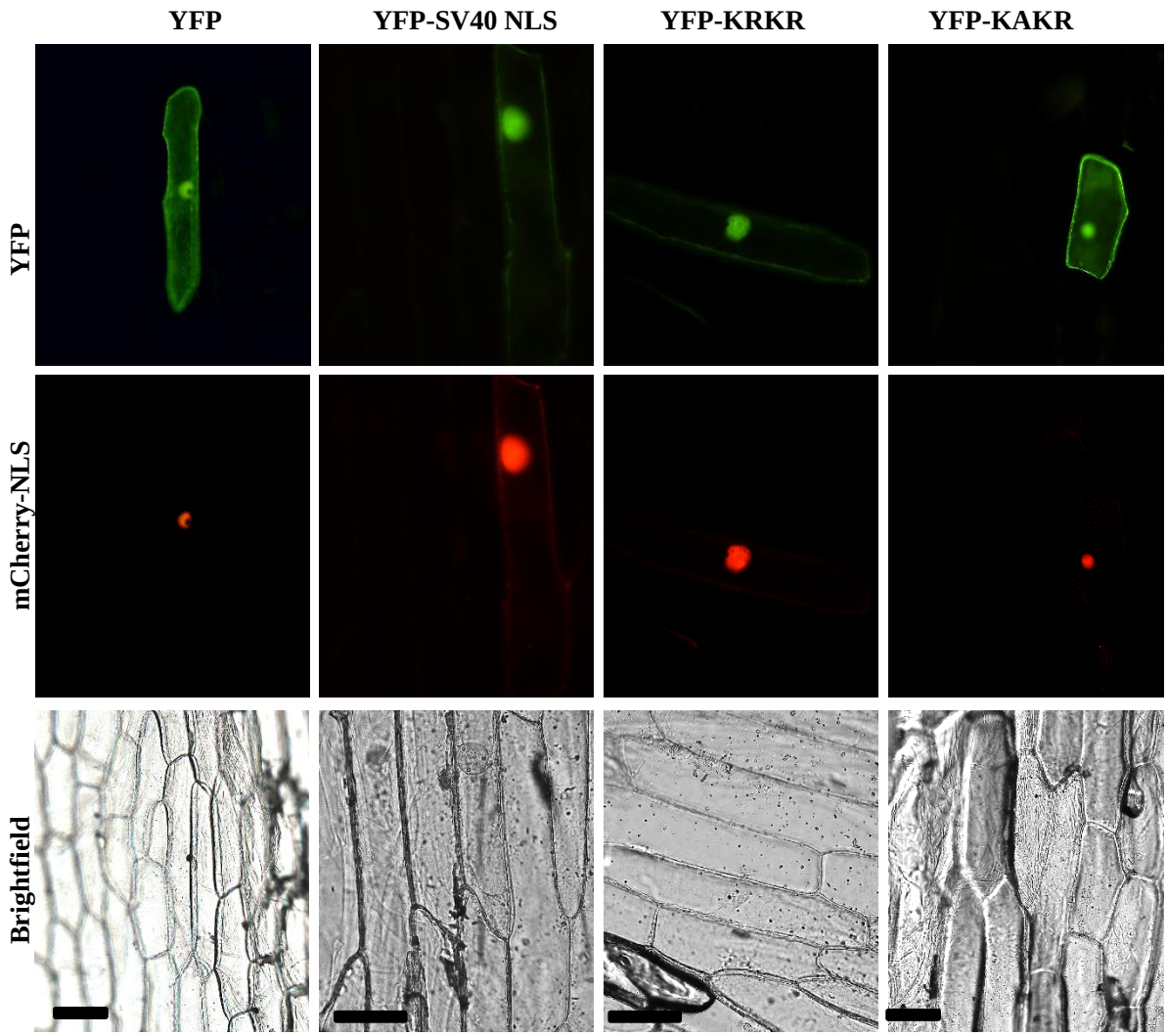


Figure 3.26 Subcellular localisation of various YFP fusion constructs in onion epidermal cells. Scale bar: 100 µm

The YFP wild type showed both cytoplasmic and nuclear localisation as expected, probably due to the passive diffusion through the nuclear pore complex. The presence of KRKR at the C-terminus markedly increased YFP nuclear accumulation similar to the effects of the SV40 NLS. Moreover, a single mutation KRKR→KAKR completely

abolished nuclear translocation (Fig 3.26). These results show that KRKR can act as an NLS, as expected.

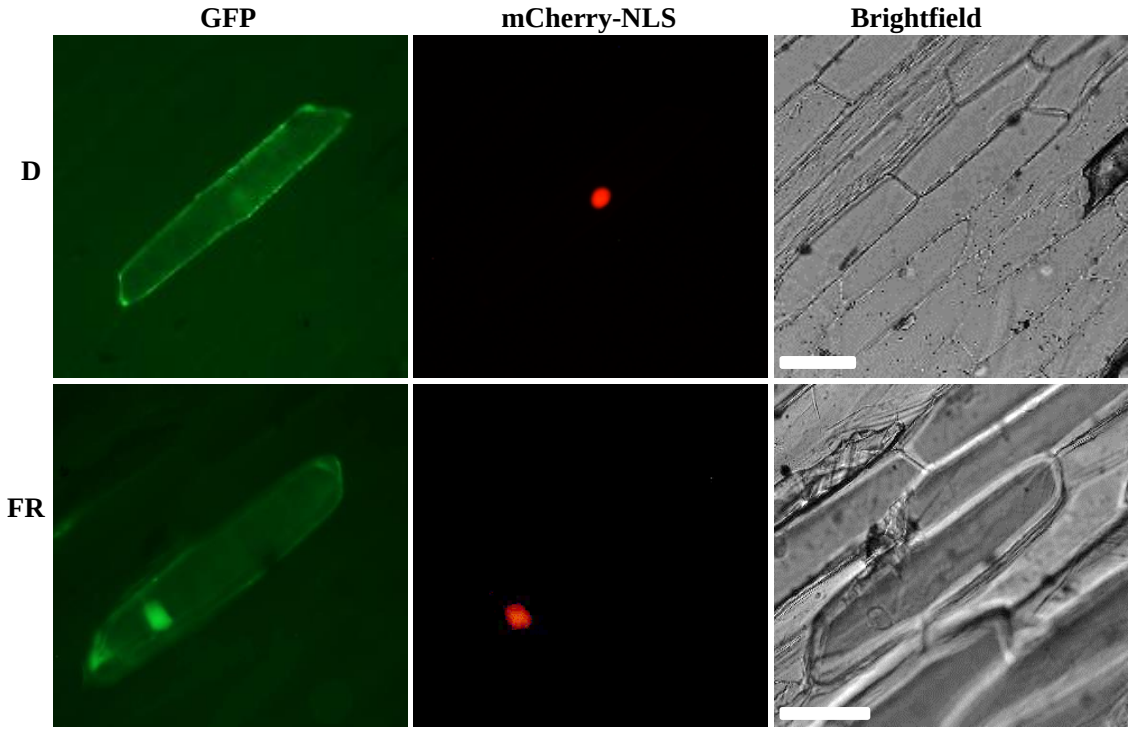


Figure 3.27 FR-dependent nuclear localisation of phyA:GFP in onion epidermal cells. Scale bar: 100 μm.

This finding, on the other hand, contradicts the established notion that phyA nuclear translocation obligately requires FHY function. To study KRKR involvement in nuclear transport of phyA, a wild-type phyA:GFP construct and an equivalent mutant bearing a KRKR→AAAA mutation were transformed into onion epidermal cells. As seen in previous studies [95], phyA:GFP showed clear nuclear fluorescence only following FR irradiation, corresponding to Pfr production under high-irradiance responses (HIR) conditions (Fig 3.27). Interestingly, although the phyA:GFP construct bearing the KRKR →AAAA mutation still showed light-dependent nuclear localisation, some cells showed cytoplasmic foci (Fig 3.28) resembling those reported in the *fhy1/fhl* mutant [96]. However, this was not observed when the experiment was repeated.

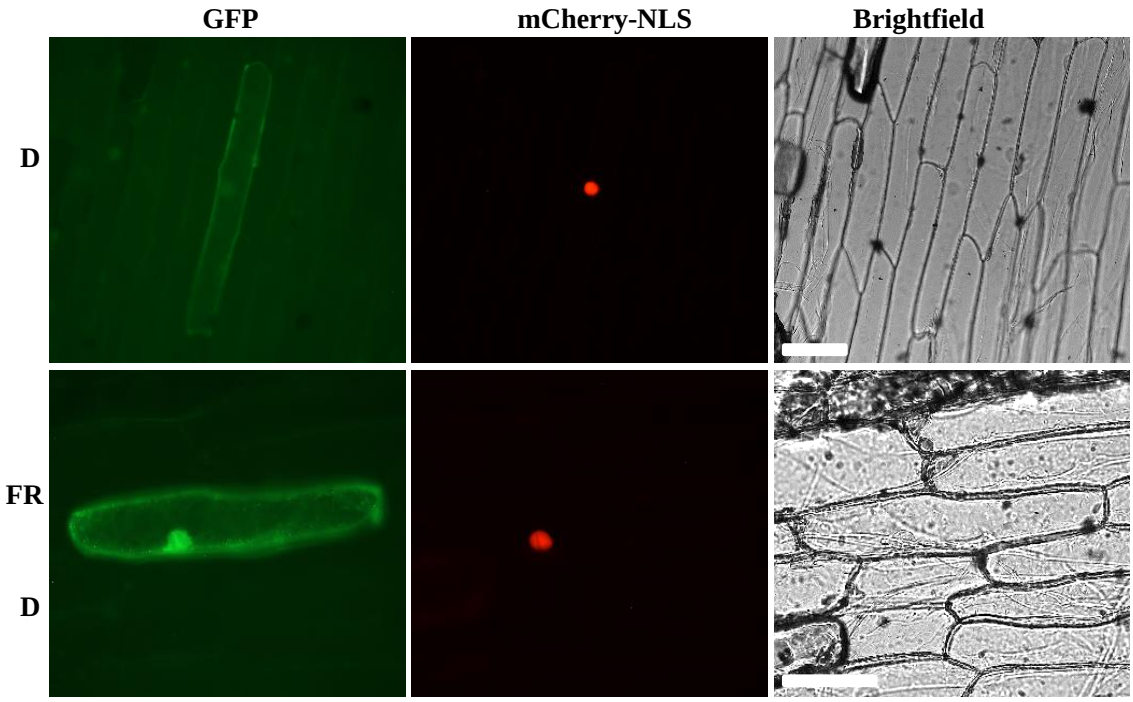


Figure 3.28 FR-dependent nuclear localisation of phyA (KRKR → AAAA):GFP in onion epidermal cells. Scale bar: 100 μm.

3.3. 2 Characterisation of phyB-PIF6 interaction

In the nucleus, plant phytochromes interact with various signalling partners, in particular the PIF transcription factors, to transduce light signals. How this interaction happens remains unclear although various biochemical studies identified residues both in phytochromes and PIFs that are important for their interaction. 3D structural information about the complex would help enormously to understand the molecular mechanism of phytochrome-PIF interaction. Previous work based on SEC mobility showed that *Arabidopsis* phyB NPGP interacts with a GFP-fused *Arabidopsis* PIF6 fragment (1-100 aa) [97,98]. Thus, the first 100 residues of *At*PIF6 contain at least part of the binding sequence, namely the APB motif that is required for PIF interaction with phyB [99]. To crystallise the phytochrome-PIF complex, we first investigated GFP-PIF6 (1-100 aa) interaction with *Gm*phyB(NPGP). This work was done together with Dr. Soshichiro Nagano.

The interaction between GFP-PIF6 and *Gm*phyB(NPGP) was examined by SEC as described in 2.2.11. Firstly, after FR irradiation *Gm*phyB(NPGP) eluted at 13.7 ml, while the elution peak was slightly shifted to a later position (14.0 ml) after R irradiation (Fig 3.29), probably due to its conformational change. Secondly, GFP-PIF6 was injected into

the SEC column eluted at 14.8 ml. Thirdly, a mixture of *Gm*phyB(NPGP) and GFP-PIF6 was injected after either FR or R irradiation. Whereas after FR treatment, both remained in their previous elution positions, they comigrated at 13.2 ml after R light irradiation (Fig 3.29). These results suggest that *Gm*phyB(NPGP) as Pfr but not Pr forms a large complex with GFP-PIF6 via direct interaction.

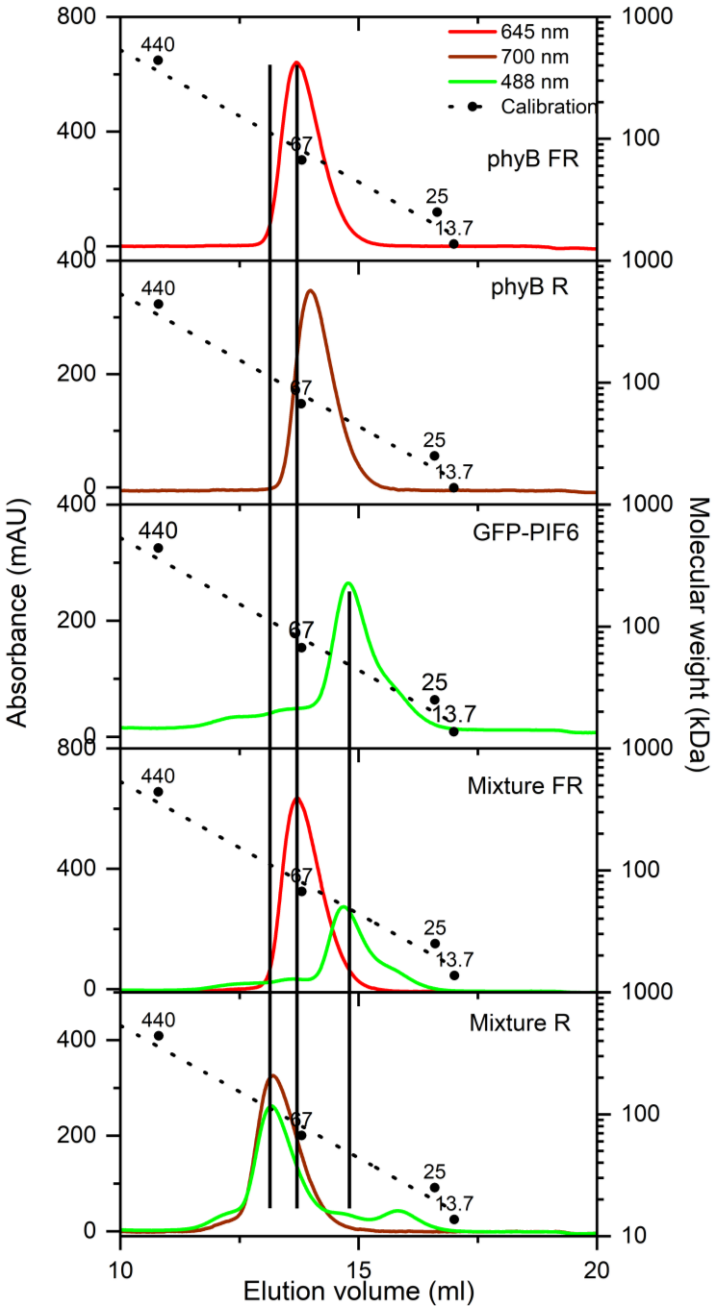


Figure 3.29 Analysis of interaction between *Gm*phyB(NPGP) and GFP-AtPIF6(1-100 aa) on SEC. Samples were monitored at 645 nm (red lines, for phytochrome in Pr), 700 nm (dark red lines, for phytochrome in Pfr) and 488 nm (green lines, for GFP-AtPIF6). Dotted line shows the SEC column calibration with marker proteins (kDa)

The YH Pfr-mimic mutant of plant phytochrome probably simulates the Pfr signalling conformation constitutively. Thus its interaction with GFP-PIF6 was also examined by SEC. Here again, *Gm*phyB(NPGP)-Y272H comigrated with GFP-PIF6 to an earlier position on SEC column, but even in darkness, suggesting it indeed interacts with GFP-PIF6 in a light-independent manner (Fig 3.30A). Interestingly, deletion of the NTE abolished the interaction, since no peak shift was observed as both *Gm*phyB(PGP) and GFP-PIF6 retain previous elution positions on SEC (Fig 3.30B). This finding is significant, as NTE removal accelerates Pfr dark reversion to Pr in the wild type and might therefore interfere with the assay, whereas this is unlikely in the case of the Y272H Pfr mimic mutant which acts independently of light and therefore cannot dark revert. Moreover, further investigation showed that the PHY domain is not necessary for the interaction, since *Gm*phyB(NPG)-Y272H also interacts with GFP-PIF6 (Fig 3.30C).

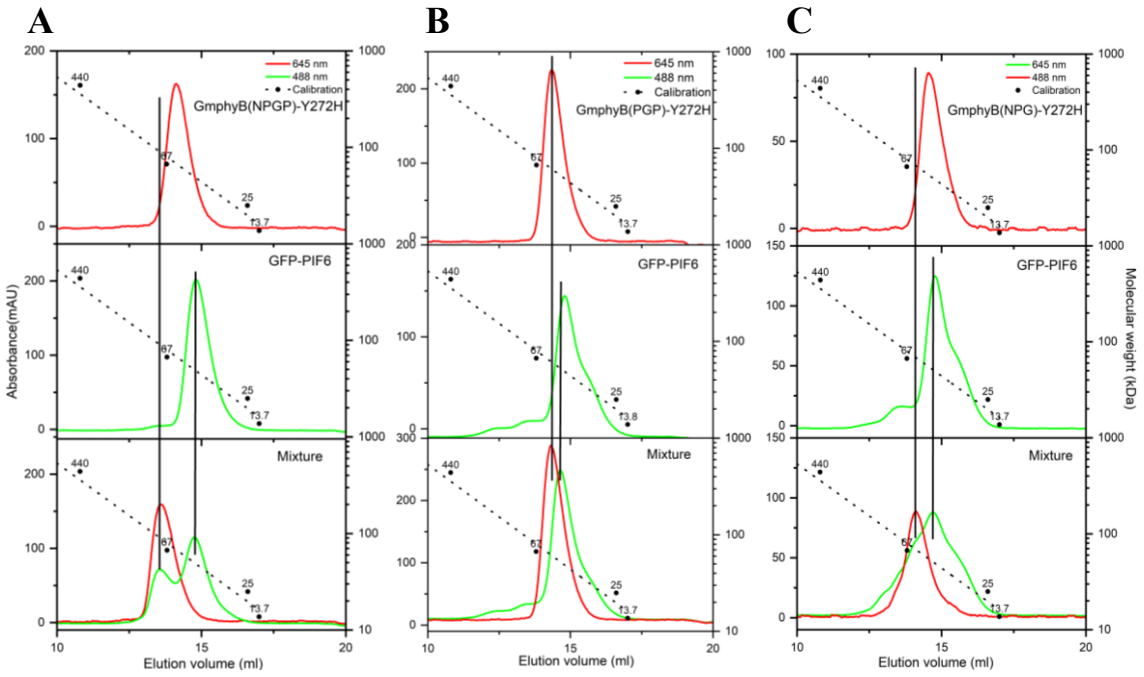


Figure 3.30 Analysis of interaction between GFP-AtPIF6(1-100 aa) and Y→H mutant of different constructs of *Gm*phyB on SEC (A: *Gm*phyB(NPGP)-Y272H, B: *Gm*phyB(PGP)-Y272H, C: *Gm*phyB(NPG)-Y272H). Samples were monitored at 645 nm (red lines, for phytochrome) and 488 nm (green lines, for GFP-AtPIF6). Dotted line showed the SEC column calibration with marker proteins (kDa).

Although the possibility that GFP might interact with phyB nonspecifically is not ruled out here, a publication already indicated that fluorescence tag does not interact with

Arabidopsis phyB NPGP [98]. Taken together, these results indicate that *Gm*phyB (NPGP) interacts with GFP-PIF6 in a light-dependent manner, while the Y272H mutant binds GFP-PIF6 constitutively. Moreover, the NTE is crucial while the PHY domain is dispensable for phyB(NPGP)-PIF6 interaction.

To further confirm the interaction results, we explored GFP-PIF6 interaction on dark reversion of *Gm*phyB as a previous study indicated that PIF6 binding to *Arabidopsis* phyB retarded dark reversion [100]. Consistently with those findings, PIF6 slowed the dark reversion of *Gm*phyB(NPGP) but did not affect the dark reversion of the PGP construct (Fig 3.31), suggesting GFP-PIF6 interacts with the NTE of *Gm*phyB.

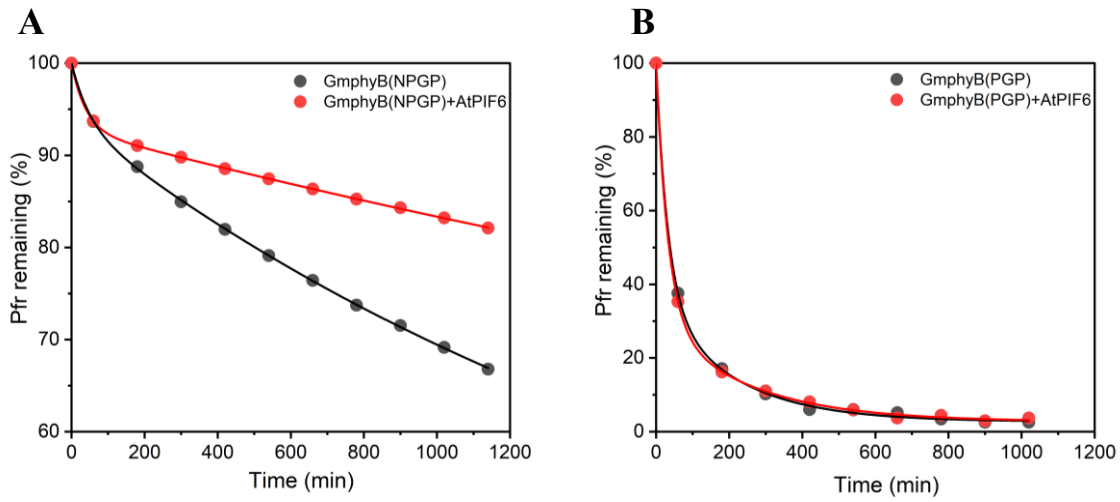


Figure 3.31 Effects of GFP-PIF6 on the dark reversion of *Gm*phyB(NPGP) (A) and *Gm*phyB(PGP) (B). The percentage of Pfr is represented by black and red dots. The black and red lines show the biphasic exponential decay fitting (half-lives: 33 and 1391 min for *Gm*phyB(NPGP), 27 and 3476 min for *Gm*phyB(NPGP) in presence of GFP-PIF6, 22 and 171 min for *Gm*phyB(PGP), 22 and 141 min for *Gm*phyB(PGP) in presence of GFP-PIF6).

4. Discussion

4.1 Characterisation of various *Glycine max* phyA/B constructs

The discovery of phytochromes in prokaryotic organisms provided access to structural studies of phytochrome since they could be produced in *E. coli* and purified much more easily than their eukaryotic homologs. Indeed, investigation of the 3D structure of phytochromes still focuses on prokaryotic phytochromes, such studies of plant phytochrome structures lagging behind. Nevertheless, developments in molecular biology techniques and advances in the phytochrome research field itself now make it possible to produce high-quality plant phytochrome samples for structural characterisation. Indeed, several 3D structures of plant phytochromes in the Pr state have been reported in the past decade, in particular from our laboratory. The all-important Pfr signalling state remains poorly understood at the structural level, however.

Building on the work of Dr. Soshichiro Nagano and Dr. Sintayehu Manaye Shenkutie, different constructs of phyA and phyB from *Glycine max* were produced, characterised and subjected to crystallisation trials in this study (see 3.1 and 3.2). Although most of the constructs could be produced with high yield and good stability, the *Gm*phyA(NPGP) construct remained insoluble in *E.coli* (Fig 3.4). Neither adding sorbitol to the growth medium nor using chaperone-containing competent cells improved its solubility. However, deletion of the first 19 residues helped: *Gm*phyA(HPGP) proved to be soluble (Fig 3.4). Furthermore, whereas the *Gm*phyA(NPG) construct was produced solubly, it was proteolytically degraded. Since the His tag was C-terminal and Ni^{2+} purification was possible, the degradation probably occurs in the N-terminal region. This result correlates with early studies of phytochrome (phyA) extracted directly from etiolated oat seedlings: the N-terminus was proteolytically degraded to 118 or 114 kDa (so-called "large phytochrome"): only in 1983 was the native 124 kDa polypeptide isolated [101]. Further analysis suggested that the 6 or 8 kDa degradation happened at N-terminus [102].

Most of the *Gm*phyA and *Gm*phyB constructs described here exhibited the R/FR photochromicity typical of phytochromes. Absorbance peaks of Pr and Pfr of *Gm*phyB(NPGP) were at about 644 nm and 715 nm, respectively, deletion of the NTE (in of *Gm*phyB(NPGP)) resulting in about 10 nm blue-shift of the Pfr absorption peak (from

715 to 705 nm) with minimal effect on Pr (Fig 3.2 and Fig 3.16). The PHY domain was believed to be essential for Pfr formation, and indeed the *Sb*phyB(PG) construct showed almost no photochromicity (Fig 3.3). Remarkably, however, the *Gm*phyA(PG) construct shows strong Pfr-like absorbance after R light treatment (Fig 3.3). Compared to *Gm*phyA(PGP), the Pfr-like peak of the PG construct was 15 nm blue-shifted (from 705 to 690 nm). To our knowledge, there are only two other phytochromes that still show R/FR photochromicity without the PHY domain, namely Cph2 from *Synechococcus* OS-B [103,104] and a triple mutant (P145S, L311E, and L314E) of *Dr*BphP [105]. The reason why the *Gm*phyA(PG) photoproduct still shows Pfr-like absorbance despite the absence of the PHY domain remains elusive. In the phytochrome photocycle, the weakly-absorbing MetaRc intermediate represents a temporally deprotonated state that is then reprotonated to generate Pfr. Raman spectroscopy data from our collaborators showed the chromophore is protonated in the *Gm*phyA(PG) photoproduct [89]. Therefore, the photoproduct of *Gm*phyA(PG) might represent an intermediate between MetaRc and Pfr, which we suggested be termed MetaRd. A more detailed discussion is given in Nagano *et al*. [89].

4.2 Plant phytochrome Pr structures

Successful protein production and purification paved the way for crystallisation and structural determination. Whereas *Gm*phyA(PGP) has still failed to crystallise in numerous trials (at least 20 96-well plates, see 3.1.2), both *Gm*phyB(PGP) and *Gm*phyA(PG) crystallise within a week.

Dark-blue crystals of *Gm*phyB(PGP) look similar to those of Cph1. The initial resolution of 3.5 Å could be improved to 2.9 Å after thorough optimisation, in particular following replacement of potassium bromide with lithium chloride or sodium bromide (see 3.1.2).

Despite the evolutionary divergence, our *Gm*phyB(PGP) structure generally resembles that of *Arabidopsis* phyB (4OUR), especially for the nPAS, GAF domains and the chromophore binding pocket (Fig 4.1A). It should be noted that the PΦB model was wrong in 4OUR, this was corrected in Nagano *et al* [89]. nPAS, GAF and PHY domains are almost linearly arranged, the chromophore sitting in the binding pocket in GAF domain. The N-terminal part of nPAS domain passes through a loop of GAF domain to form a figure-8 knot (Fig 4.1B), as in prokaryotic phytochromes [60]. How the knot forms during

protein folding is unclear. One study identified mutations locating in the knot might disrupt binding with PIF3 [48].

GAF and PHY domain are connected by a 68 Å long helical spine as well as by a tongue-like hairpin loop extending from the PHY domain to the chromophore binding pocket, as in prokaryotic phytochromes – where the tongue is known to undergo sheet-to-helix refolding upon Pr→Pfr photoconversion.

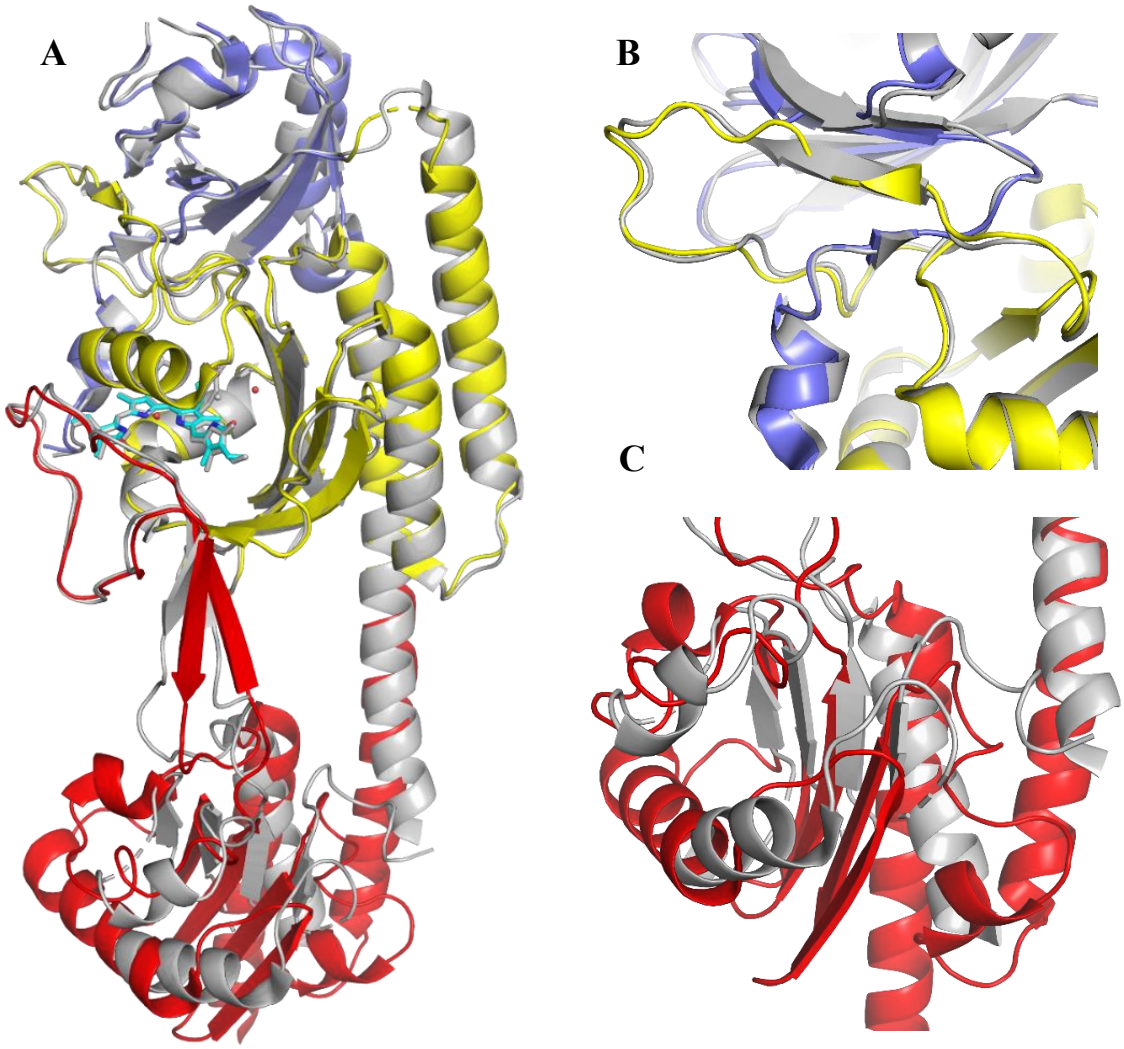


Figure 4.1 Superimposition of *Gm*phyB(PGP) (nPAS: slate, GAF: yellow, PHY: red) and AtphyB(PGP) (4OUR, grey). A: Overall structure. B: the knot. C: PHY domain.

Whereas the PHY domain in 4OUR was poorly resolved with several loops missing entirely, our 6TL4 structure was much more satisfactory in this region, the loops in particular being modelled successfully (Fig 4.1C). In addition, the PHY domain shows different orientations in *Gm*phyB(PGP), 4OUR and Cph1 2VEA because the helical spine

bends at different angles. Since *Gm*phyB(PGP), *At*phyB(PGP) and Cph1 crystallise as a staggered complex, a parallel dimer and an antiparallel dimer, respectively, the different orientations are likely to result from crystal packing artefacts deriving from the inherent flexibility of the connecting helix.

The structure of the tongue is quite different in plant and prokaryotic phytochromes. In Cph1, the tongue forms an additional anti-parallel sheet at the tip to seal the chromophore cavity (Fig 4.2C), while it forms a lasso leaving the chromophore exposed to the solvent via a tunnel in *Gm*phyB(PGP) (Fig 4.2A). This is apparent in 4OUR too (Fig 4.2B). The chromophore in Pfr is oxidized 10-fold faster than that in Pr according to permanganate oxidation experiments, suggesting that the chromophore is more exposed in Pfr than that in Pr [106,107]. It was suggested that the tongue opens and closes solvent access to the chromophore in Pfr and Pr, respectively.

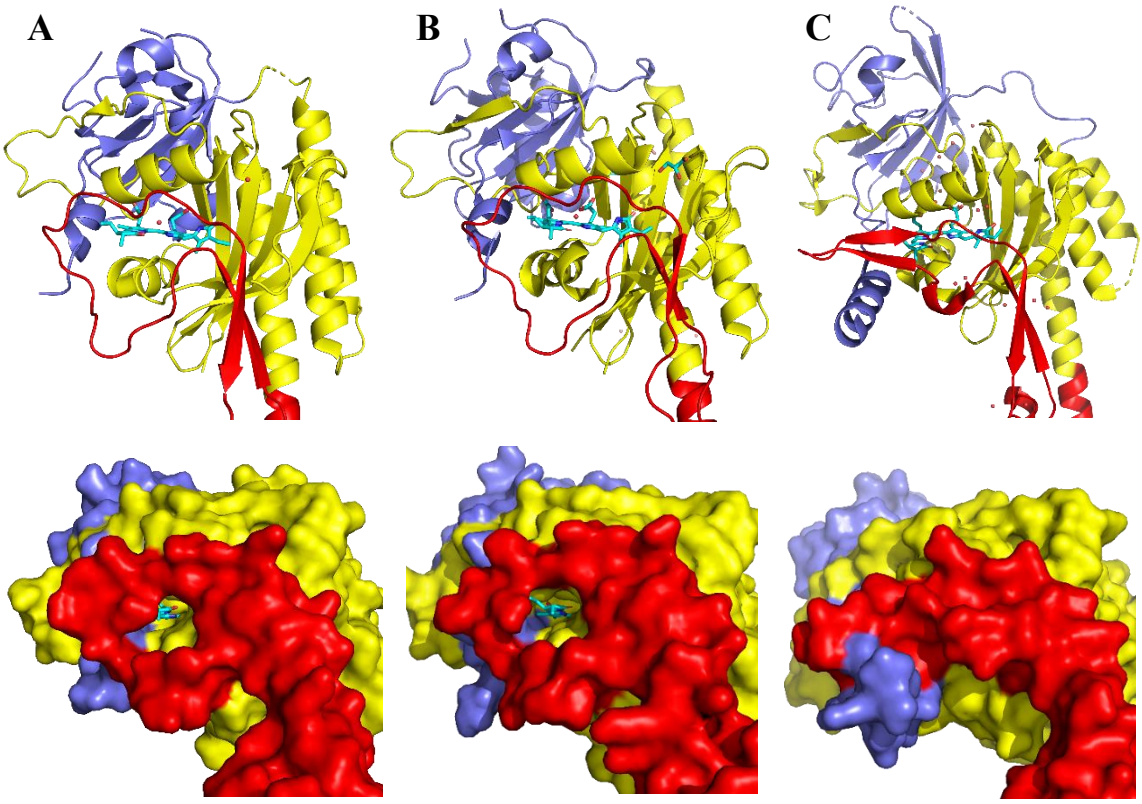


Figure 4.2 Chromophore solvent access in 6TL4 (A), 4OUR (B) and 2VEA (C). Structures are shown as cartoon (upper) and surface (lower).

Indeed, protonation dynamics in phytochromes are correlated with solvent access [108,109]. The NTE, which is absent in both structures, might seal the chromophore binding pocket in Pr, and leave the chromophore exposed to solvent after conformational change to Pfr.

The chromophore has extensive interactions with nearby residues and water molecules. Protein-chromophore interactions in *Gm*phyB(PGP) 6TL4 are closely similar to those of *Arabidopsis* phyB 4OUR (Fig 4.3). Due to the resolution limit, only two and five water molecules were modeled in 6TL4 and 4OUR, respectively. An important water molecule strategically positioned between H354 and the nitrogens of rings *A*, *B* and *C* is seen in all phytochrome structures to date. This so-called pyrrole water, as well as the DIP motif aspartic acid (D303), form hydrogen bonds with the nitrogen atoms of rings *A*, *B* and *C*. Interestingly, a recent XFEL study of *Dr*BphP showed that this water disappears within a picosecond of photon absorption [110].

D303 forms a salt bridge with R575 of the PRxSF motif of the tongue. Studies of bacteriophytochrome showed that the serine of the PRxSF motif replaces the arginine to contact the aspartic acid in Pfr [70]. Mutation of these conserved residues has profound effects on phytochrome photoconversion, Pfr stability, and signalling. For example, D307A in *Arabidopsis* phyB blocks photochromicity (Pfr absorbance formation), whereas R582A does not affect photochromocity but instead inhibits dark reversion [72].

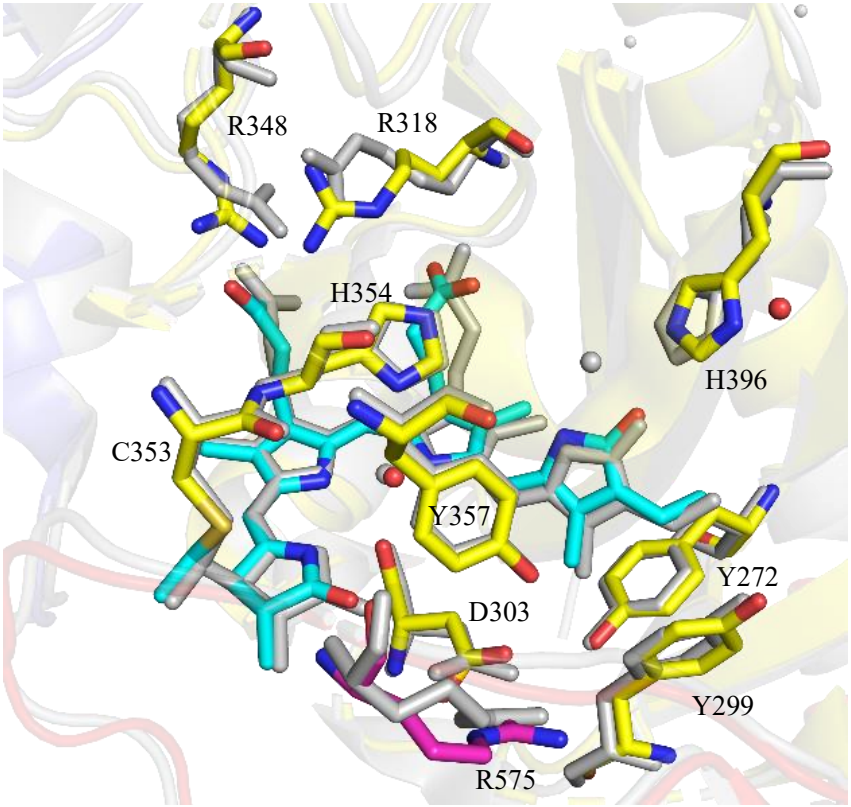


Figure 4.3 Comparison of the chromophore and conserved residues in the binding pocket of *Gm*phyB(PGP) (PCB: cyan, nPAS: slate, GAF: yellow, PHY: red) and *Arabidopsis* phyB 4OUR (grey). Residues are labelled according to *Gm*phyB numbering.

The carbonyl of the ring *D* makes a hydrogen bond with H396, stabilising the chromophore *ZZZssa* configuration in Pr. Undoubtedly, this connection breaks during photoconversion due to the ring *D* photoflip. Three conserved tyrosines around ring *D* are likely to play important roles in phytochrome photoactivation. Crystal structures of bacteriophyothcrome and the MAS NMR structure of oat phyA3 showed that two of them (Y272 and Y299 in *Gm*phyB) undergo side chain movement during photoconversion [70,73]. Remarkably, *At*phyB Y272H seems to mimic the Pfr signalling structure in darkness as the transgenic seedlings are constitutively photomorphogenic (*cop* phenotype) [65]. The homologous mutant in Cph1 leads to strongly enhanced fluorescence [64]. Similarly, the homologous mutant of *Gm*phyA fluoresces too, with a quantum yield of 13.2% for PGP and 10.7% for PG construct (see 3.2.1), slightly lower than Cph1 Y176H (14.4%) [88]. Two arginines close to the propionic side chains of ring *B* and *C* are interesting too as discussed below. The H354 above the chromophore is likely to be critical for protonation [111]. Although these residues are conserved in almost all phytochromes, some position changes are noticeable in the different 3D structures. For example, the propionates of ring *B* and *C* adopt slightly different conformations in 6TL4 and 4OUR, as do the adjacent arginines. R575 of the PRxSF tongue motif is also shifted. In general, therefore, significant differences are seen in and around the chromophore binding pocket.

Prior to this study, little is known about the 3D structure of plant phytochrome A, except for the limited MAS NMR information for oat phyA3(NPGP) both in Pr and Pfr [73]. 6TC7 of *Gm*phyA(PG) represents the first crystal structure for an A-type plant phytochrome. The *Gm*phyA(PG) construct crystallised in many conditions provided by the Morpheus kit (see Table 3.1). We also found that cross-seeding was particularly helpful in finding new crystallisation conditions and improving resolution. Although amino acid side chains could be modelled easily using the 2.1 Å data, the electron density of the chromophore was much weaker and more diffuse (Fig 3.12A), allowing only approximate modelling of the structure. We hypothesized that the chromophore is very mobile in the crystal even in the 100 K cryostream and that this mobility might be reduced at even lower temperatures. However, datasets collected in N2 and He cryostreams (100 K and *ca.* 30 K, respectively) showed similarly clear chromophore density and enabled precise chromophore modelling (Fig 3.12B). The cause of the original problem is not clear. It is perhaps relevant that these crystals, unusually, are able to photoconvert. Thus possibly the original crystals might

have undergone significant photoconversion from the safelight dose during fishing and snap freezing. Indeed, we noticed that a crystal showed a small Pfr-like absorbance peak in icOS, even though the crystal had grown in darkness (Fig 3.23). The safelight photon dosage was minimized carefully in the latter experiments.

Far more water molecules (about 430) could be located in the 1.58 Å structure than at 2.1 Å (176 water molecules). The water molecule on the β-facial side of the chromophore between and hydrogen bonded to the propionate chains of rings *B* and *C* (Fig 4.4) is represented in other phytochrome structures such as *Sorghum* phyB [89]. In Cph1, this water seems to be represented by one on the α-facial side, where it makes additional contact with the conserved R222 [61].

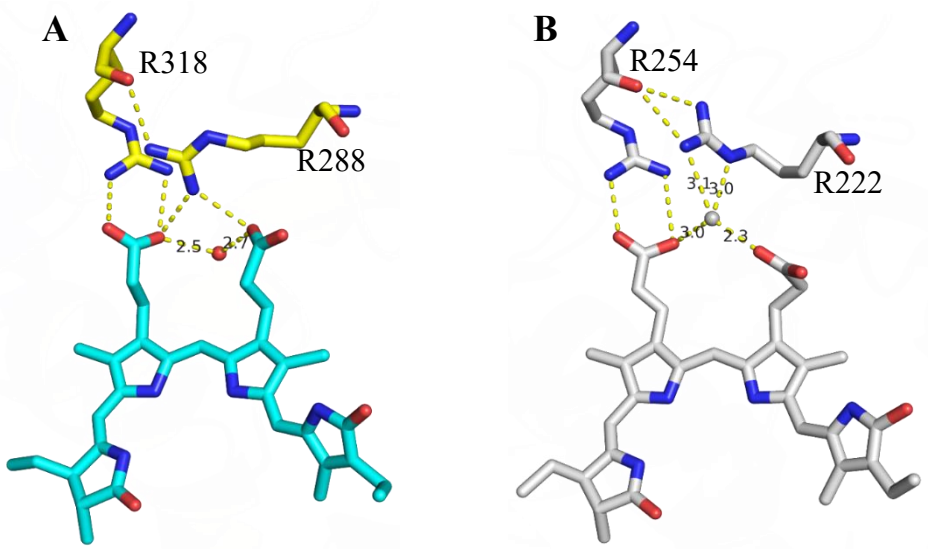


Figure 4.4 Different positions of a water molecule in *Gm*phyA(PG) (PCB: cyan, GAF: yellow) and Cph1 2VEA (grey). Hydrogen bonds are yellow dashed lines with lengths in Å.

Two conserved arginines, R318 and R288 in *Gm*phyA(PG) (R254 and R222 in Cph1), are particularly interesting on account of their interaction with the propionates of rings *B* and *C*. It was suggested that the partnership of the *B*-ring propionate swaps from R254 to R222 upon Pr→Pfr photoconversion in Cph1 and other phytochromes [66]. A similar partner swap was also observed in MAS NMR studies of oat phyA3 [73]. This swap might therefore represent an important component of the photoactivation and/or signalling mechanism. The interactions are, however, rather different in the *Gm*phyB(PGP) and *Gm*phyA(PG) Pr states (Fig 4.5). There, the *B*-ring carboxylate contacts with both arginines, and the *C*-ring propionate is involved too. In *Arabidopsis* phyB R352A

decreases while R322A increases the rate of dark reversion [72], the interactions might therefore have the more mundane function of stabilising Pfr. On the other hand, R352K disrupts the interaction between *Arabidopsis* phyB and PIFs in a co-immunoprecipitation assay [48], implying either a direct role in signalling or an indirect effect via Pfr dark reversion.

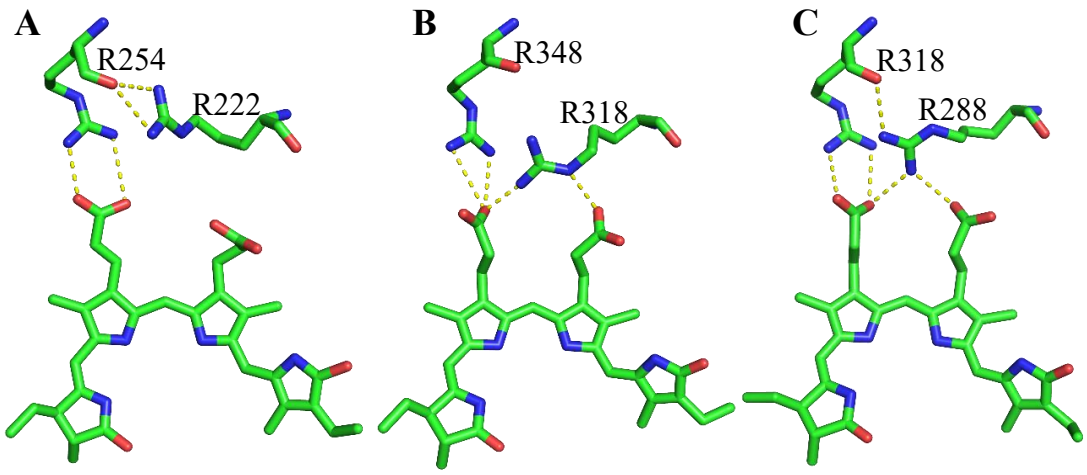


Figure 4.5 Arginine-propionate interaction in Cph1 2VEA (A) *Gm*phyB(PGP) 6TL4 (B) and *Gm*phyA(PG) (C). Hydrogen bonds: yellow dotted lines.

The structure of *Gm*phyA(PG) is closely similar to the MAS NMR model of the chromophore pocket of oat phyA3 Pr (rmsd of 0.517 Å for Cα atoms). Additionally, however, the NMR data provided insight into the equivalent Pfr structure. Upon photoconversion, chromophore and some nearby residues were shown to undergo considerable conformation changes based on bacteriophytochrome structures and NMR structure of oat phyA3 [70,73] (Fig 4.6B). In Pfr the chromophore geometry was *ZZEssa*, as expected, but the *D*-ring was β-facial in contrast to the α-facial disposition in bacteriophytochrome Pfr. The side chain of Y241 (Y242 in *Gm*phyA) shifts radically to make space for the ring *D*, while Y268 (Y269 in *Gm*phyA) only shows a slight movement rather than the backflip seen in bacteriophytochromes. Moreover, instead of being parallel to propionate chain of ring *B* in Pr, the *C* ring propionate bends away to form a hydrogen bond with H372 (H370 in *Gm*phyA) in Pfr. There are three serines (S278, S333, S338) contacting indirectly with PCB (Fig 3.14). In particular, S336 - conserved in prokaryotic and plant phytochromes – when mutated in *Arabidopsis* phyB (S370F) showed abnormal PCB incorporation [47].

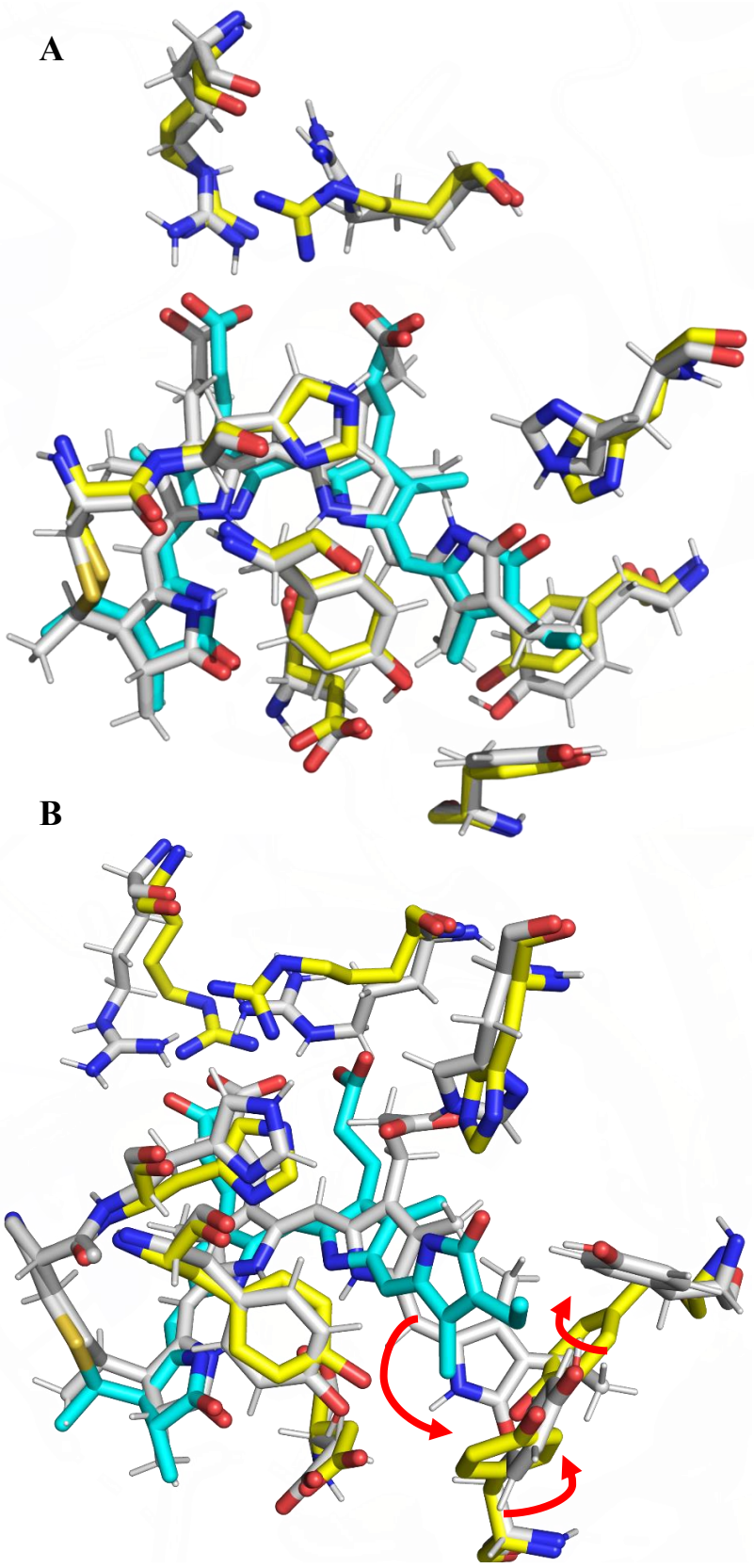


Figure 4.6 Superimposition of *Gm*phyA(PG) (PCB: cyan, GAF: yellow) structure and oat phyA3 MAS NMR model (grey) in Pr (A) and Pfr (B). Red arrows indicate the *D*-ring and tyrosines movements.

Although structures of the photosensory modules of plant phytochromes, especially our high-resolution structures of phyA and phyB of *Glycine max*, provide extensive information on plant phytochrome structure and function, the full-length structure of a plant phytochrome is necessary to understand its signalling mechanism fully. Indeed, two recent cryo-EM studies describe the structures of full-length phyA and phyB from *Arabidopsis* [75,76]. The 17 Å cryo-EM map of *Arabidopsis* phyA was interpreted as a head-to-head dimer [75], however the structural model is unconvincing, not least because of its low resolution, also the model could not interpret the entire electron density, with an additional density close to the photosensory module perhaps representing the NTE. More recently, the cryo-EM structure of full-length *Arabidopsis* phyB was determined at 3.3 Å (7RZW) [76]. Instead of a head-to-head dimer, 7RZW shows that the N-terminal photosensory modules were arranged head-to-tail, and formed a parallelogram-shaped platform together with two PAS2 domains, while C-terminal modules associate head-to-head. The interactions between PAS2 and nPAS-GAF domain, also between PHY domain and the modulator loop (a hairpin extending from PAS2 to the PHY domain) seem to be important for phytochrome dimerization and Pfr stability. The NTE was poorly resolved, while the PAS1 domain was completely missing in the final structure, only appearing when the resolution limit was lowered to 15 Å. It was suggested that the parallelogram-shaped platform formed by photosensory modules and PAS2 domains might mediate interactions between phytochrome and signalling partners. However, it is a very vague point, because both NPGP-GFP-GUS-NLS and NPG-GFP-GUS-NLS constructs function nicely in transgenic plants [26,27], even though they miss the PAS1, PAS2, DHp and CAT domains that make up the 7RZW platform. The *Gm*phyB(PGP) 6TL4 structure closely resembles the corresponding portion of 7RZW in terms of overall structure and details in chromophore binding pocket (Fig 4.7), implying that the structure of the photosensory module is independent of the rest of the complex.

Although the photosensory module signals when it was fused to a dimerization domain and an NLS through direct inhibition on PIFs transactivation activity [30], the C-terminal domain of phyB signals too via mediating PIF3 degradation [31]. Therefore, phyB regulates PIFs via the dual mechanisms [30]: transactivation activity inhibition and degradation of PIF3 are controlled by its N- and C-terminal modules, respectively. The D1040V mutation of *Arabidopsis* phyB disrupts HKM dimerization, bindings to PIF3, photobody

formation and PIF3 degradation [31]. Moreover, the identification of the Quail box within the PAS repeat (see 1.1.3) suggests an indispensable role of the C-terminal module in phytochrome signalling. However, in the context of the full-length structure of *Arabidopsis* phyB (7RZW), the D1040 is buried inside the protein. Therefore, it remains elusive how D1040 mediates dimerization and binding with PIF3. One explanation could be the incorrect folding caused by this mutation. On the other hand, the Quail box region (a 160-residue segment, 624-814 in *Arabidopsis* phyB) is not well-resolved in 7RZW. The residues 622-776 are completely missing, while 777-814 are represented by a so-called modulator loop [76]. The modulator-PHY interaction was suggested to stabilise dimerization while destabilising Pfr [76]. Unquestionably, however, the structure-function correlation requires further studies.

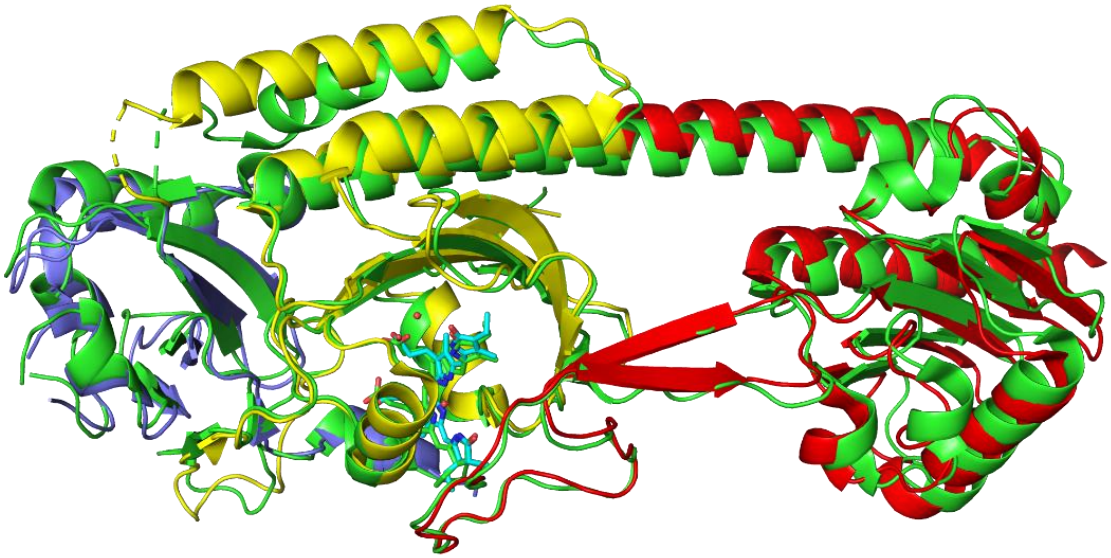


Figure 4.7 Superimposition of the *Gm*phyB(PGP) structure (6TL4: nPAS, slate; GAF, yellow; PHY, red; water: red dots; PCB, cyan.) with the corresponding region of the *At*phyB cryo-EM structure (7RZW, green).

4.3 Crystallisation of plant phytochrome as Pfr

Various Pfr structures of bacteriophytochrome were determined, such as the *Pseudomonas* bathyphytochrome which has Pfr as dark-adapted state [62], the illuminated structure and a Pfr-stabilised mutant structure of *Dr*BphP [68,70]. These studies, together with Pr structures, provide valuable views regarding how phytochrome changes upon light irradiation. For example, the different conformation of the tongue, *D*-ring rotation, and movement of the residues around the chromophore. Whereas plant phytochromes are

physiologically inactive as Pr in darkness, the light-induced Pfr signalling state is a master regulation arguably the most important regulatory molecule in the plant. Its 3D structure is consequently of enormous interest as it would provide invaluable insights into the functional mechanism involved. Currently, however, the only limited structural information on plant phytochrome as Pfr is available, namely from MAS NMR of oat phyA3 [73]. Crystallisation of Pfr is fraught with technical difficulties: first of all, it is impossible to acquire 100% Pfr homogeneous sample after R light irradiation; secondly, plant phytochrome reverts to Pr, a process that is much faster than crystallisation. However, previous studies have yielded *Dr*BphP Pfr structures by using periodic light pulses and/or Pfr-stabilised mutations to maintain a high Pfr proportion [68,70]. We attempted both strategies to obtain 3D structural data for Pfr of plant phytochromes.

In this study, several Pfr-stabilised mutants of both *Gm*phyA and *Gm*phyB were produced and purified for crystallisation trials. Measurements showed that these mutants indeed retarded thermal reversion (Figures 3.17 and 3.20). In addition, mutation studies of one of the tyrosines close to chromophore ring *D* have yielded particularly interesting results. In Cph1 Y176H all but abolished photochromicity, yielding instead a strongly red-fluorescent protein [64]. Even more remarkably, the homologous substitution in *Arabidopsis* phyB induced constitutive photomorphogenesis in transgenic plants [65], implying that it mimics the Pfr conformational signal even though the chromophore remains in the Pr-typical *ZZZssa* conformation. Thus, Pfr-mimic mutants of *Gm*phyA and *Gm*phyB were also produced, characterised and subjected to crystallisation trials (see 3.2.1 and 3.2.2). All the variants could be produced with high yield and good stability, however, as expected, crystallisation and structure derivation proved difficult. Unfortunately, attempts to crystallise *Gm*phyB(NPGP) as Pfr with the help of pulsed irradiation at 590 nm were not successful to date, possibly because of the effects of Pr-Pfr cycling. On the other hand, the R549A and Y242H mutants of *Gm*phyA(PGP) crystallised (Fig 3.22), reproducibility being improved greatly with the help of seeding. However, no useful diffraction could be measured even after extensive work to optimize the conditions. The Y242H crystals failed to diffract even *in situ* at room temperature (see 3.2.2), implying that the cryoprotection / freezing procedures are not the source of the problem. Therefore, further investigations are required to obtain crystals of sufficient diffraction quality to allow the 3D structure to be solved.

Alternatively, it might be possible to obtain structural information on Pfr by photoconversion of Pr crystals. Indeed, *Gm*phyB(PGP) crystals showed a weak absorbance peak in the FR after white light exposure at room temperature (Fig 3.23), suggesting photoconversion could occur in the crystal itself. In this case, however, the crystal lost its diffraction ability as a result of the treatment. Given that phytochromes, especially the PHY domain, might undergo conformation change during photoconversion, it is reasonable to consider that crystal packing might be disturbed by light-induced conformation changes. We were surprised to discover that our *Gm*phyA(PG) construct showed effective photoconversion in solution - despite missing the PHY domain generally required for Pfr production (or stabilisation) in other phytochromes. Remarkably, even crystals of this construct showed photochromicity at room temperature and even retained their diffraction quality (see 3.2.3). Perhaps the PHY domain has a rather different structural significance in phyA, and perhaps its absence is the key to allowing the structural changes involved in photoconversion to take place *in crystallo*. Surprisingly however, although the irradiated *Gm*phyA(PG) crystals diffracted to 2 Å, the dataset revealed little structural changes relative to Pr - other than that electron density in the chromophore region was missing completely. In this context, interestingly, the chromophore showed low electron density in our original 2.1 Å structural solution [89], a problem that was solved when minimal safelight doses were used during crystal mounting (Fig 3.12), implying that irradiation itself causes technical problems. Although this result is disappointing, it does imply that these crystals would be rather well suited to study the associated 3D structural changes kinetically using XFEL and related technologies. Indeed, we were recently granted beamtime at EuXFEL in Hamburg to this end.

Alternatively, *Gm*phyA(PG) was crystallised under orange light pulses. Although some crystallisation conditions overlap with that of Pr crystals, these two crystal types have distinct crystal packing (see Table 7.3 and 7.4). Superimposition of the illuminated state and Pr state identifies conformation changes in the 68-77 residue region (Fig 3.24A). Mass spectroscopy studies have shown that the NTE of plant phyA undergoes conformational changes during photoconversion [112]. Although the NTE is largely missing from the *Gm*phyA(PG) construct, the observed conformational change at the N-terminus of the illuminated structure might reflect the dynamics of the NTE during photoconversion. NMR structure of oat phyA3 in Pfr implied that the NTE is closer to

chromophore than in Pr [73], paradoxically however, the NTE points away from the chromophore in *Gm*phyA(PG) illuminated structure (Fig 4.8). A conserved tyrosine may be a phosphorylation target in *Arabidopsis* phyB, as the phosphomimic mutant (Y104E) is defective in rescuing the *phyB* null mutant phenotypes and in binding with PIF3 [113]. The corresponding tyrosines in oat phyA3 and *Gm*phyA are shown as sticks in Fig 4.8.

The missing density for chromophore in the first structural solution implying chromophore cycling under orange light might be a problem. Indeed, a short period of darkness prior to crystal freezing allowed the chromophore to be resolved (Fig 3.24C). The absorbance of the crystals grown under orange light pulses clearly shows Pfr-like absorbance (around 50%, Figure 3.24B).

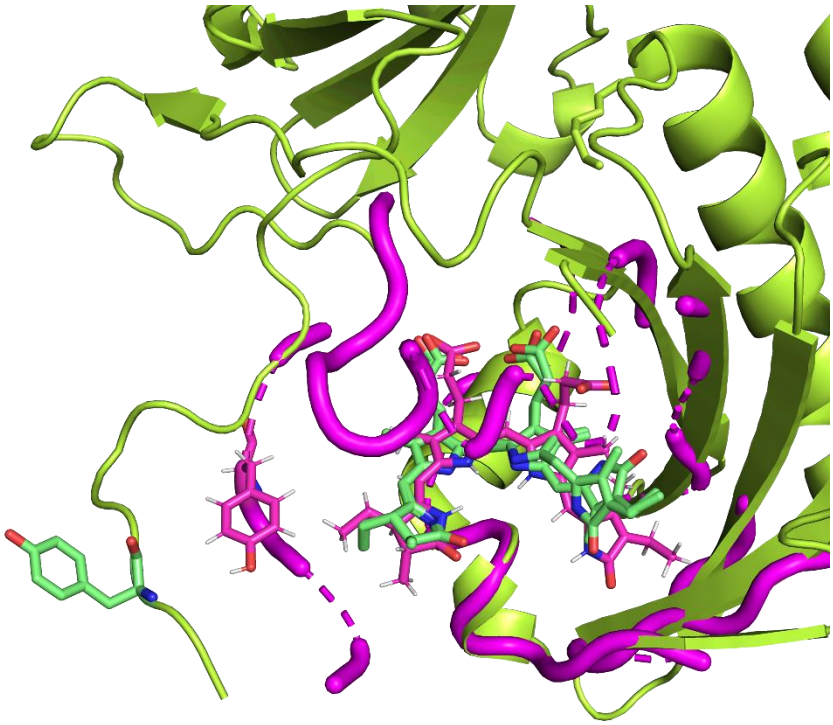


Figure 4.8 Superimposition of oat phyA3 Pfr NMR structure (magenta) and *Gm*phyA illuminated structure (light green). Two tyrosines are shown as sticks.

The supposed photoflip of the Y242 is not observed in the illuminated structure of *Gm*phyA(PG). Interestingly, the homologous tyrosine in *Dr*BphP Pfr structure, which crystallised with R light pulses, remained in the Pr conformation too [68]. CD measurement suggested the *D*-ring of PCB might adapt β-facial configuration in Pfr [15], however, the *D*-ring could not be modelled at β-face in the illuminated structure of *Gm*phA(PG), as it

would then have clashed with the Y242 side chain. Given that the absorbance spectrum of the irradiated crystal clearly shows a mixed state, both *Z* and *E* configurations were modelled in the structure, which is also supported by the electron density (Fig 3.24C).

4.4 Nuclear translocation of plant phytochrome

Plant apophytochromes are synthesized and then assemble with chromophore cofactors exported from the chloroplast to form holophytochrome as Pr. Whereas Pr resides in the cytoplasm in darkness, Pfr formed upon photoconversion is translocated into the nucleus. How state-dependent translocation is brought about is poorly understood - and interestingly, the mechanisms used by phyA and phyB seem to be quite different. phyB nuclear translocation is an enigma. As no NLS has been recognised, it is unlikely that such a function is "unmasked" upon Pfr formation. Although the hypothesis that an NLS function is created upon Pfr formation offers a reasonable alternative [33], none has been found. It has therefore been suggested that Pfr nuclear import of phyB takes place by its attaching itself to PIF transcription factors in the cytoplasm prior to their translocation to the nucleus [34]: like all transcription factors, PIFs carry active NLS motifs. Nuclear translocation of phyA, on the other hand, obligately and specifically requires the function of the transport facilitator FHY1 [35,36,38], which carries both NLS and NES motifs, leading to cyclical nuclear import and export. Indeed, FHY1 is bound specifically by the Pfr state of phyA (and not phyB). The requirement of phyA for FHY1 poses a problem for the PIF-as-carrier model for phyB, however, as phyA Pfr also binds several PIFs, yet it remains trapped in cytoplasm in mutants lacking FHY1 function.

Remarkably, we found a canonical type I NLS [92] (KRKR) in a loop of the GAF domain. Indeed, the K(R/K)K(R/K) motif is conserved in and specific to the phyA lineage (Fig 3.25) yet has remained unnoticed since the cloning of *PHYA*. In this study, we showed that KRKR possesses NLS activity and functions similarly to the classical SV40 NLS to promote nuclear accumulation of YFP in onion epidermal cells, while a single mutation in the motif abolished this activity (Fig 3.26). In one set of experiments with transfected onion epidermal cells, we noticed that phyA:GFP bearing a KRKR→AAAA mutation showed cytoplasmic foci similar to those seen earlier in a *fhy1*/*fhl* double mutant [96], we did not see this in a second experiment, however. If the KRKR motif is indeed a Pfr-functional NLS, what then would explain the absolute necessity for FHY1? It is important

to recognise, however, that the molecular mechanism of FHY1-phyA binding and nuclear translocation is largely unknown and might involve the putative NLS in phyA. This is an interesting question to be answered. A simple explanation would therefore be that the KRKR motif in phyA is somehow unmasked upon Pfr formation (see above for phyB). Indeed, its position at the base of a flexible loop would lend itself to an idea like this. Thus, KRKR activity in phyA would be affected by light but still require FHY1 for translocation. We recall that the mechanism of Pfr-specific binding to FHY1 is not known. Moreover, instead of being a straightforward process, protein nuclear translocation could be regulated at multiple levels, for example, by phosphorylation. Phosphorylation of two residues (S39 and T61) around the NLS in FHY1 reduces its nuclear translocation [114]. Perhaps this occurs around KRKR in phyA too? Light-dependent import systems might be more complex than we think: for example, KRKR acts as an NLS in PAS domain-containing histidine kinases 1 and 2 in *Physcomitrium*, red light irradiation triggering interactions in the cytoplasm [94]. Further investigations are required to decipher the integration of KRKR and FHY1.

4.5 Phytochrome-PIF interactions

Once in the nucleus, Pfr interacts with various signalling partners, including the PIF family, members of the bHLH transcription factor superfamily that regulate expression of numerous genes. Pfr-PIF interaction is a crucial event in the biology of the plant, but how it happens and its consequences at the molecular level are far from being fully understood.

PIF3 was identified in a yeast two-hybrid screen using the C-terminal PAS repeat and HKM module of phyB as bait, clearly suggesting that the binding site for PIF3 is in this region [44]. Paradoxically, however, it was shown later that the binding domain is in the N-terminal region of phyB [45]. Mutagenic studies identified four residues (R110, G111, G348, R352) located in the knot region that are important for phyB N-terminal interaction with PIF3 [47,48]. On the other hand, mapping of the phytochrome binding site in PIFs identified the APA and APB motifs that bind to active phyA and active phyB, respectively. APB exists in all eight PIFs in *Arabidopsis*, while APA is only represented in PIF1 and PIF3. The APB motif comprises only about 15 residues, but is quite conserved in PIFs (E31, L32, G37, and Q38 (*At*PIF3 positions) in particular). Missense mutation of these four residues disrupts PIF interaction with phyB [99].

It was shown by SEC that the *Arabidopsis* PIF6 (1-100 aa) which includes the APB motif interacts with the photosensory module of phyB [97,98]. We investigated the interaction between *At*PIF6 and various *Gm*phyB constructs with a view to finding a protein complex appropriate for structural studies. We showed that *At*PIF6 interacts with *Gm*phyB(NPGP) in a Pfr-dependent manner (Fig 3.29), suggesting that phyB-PIF6 interaction is conserved. The Y272H mutant of *Gm*phyB(NPGP) interacted with PIF6 even in darkness (Fig 3.30), implying that Y272H mutant probably mimics the Pfr signalling conformation constitutively. Interestingly, interaction could not be detected when the NTE was deleted, whereas it was still apparent in the absence of the PHY domain (Fig 3.30). Therefore, the NTE seems to be essential and the PHY domain is dispensable for the interaction between *Gm*phyB(NPGP) and AtPIF6. To confirm this important conclusion, we investigated the effect of *At*PIF6 on the dark reversion of the NPGP and PGP constructs. Indeed, PIF6 slowed dark reversion of *Gm*phyB(NPGP) but not the PGP construct (Fig 3.31), supporting the notion that PIF6 interaction involves the NTE. These data are consistent with studies *in planta* which showed that both NPGP and NPG constructs triggered normal photomorphogenesis in transgenic plants when dimerized and localised in the nucleus [26,27]. As PIFs are the major signalling partners regulating light-responsive gene expression, it is quite unlikely that the NPGP and NPG constructs could transduce light signals without interacting with PIFs.

Following structural studies on bacteriophytochromes [68], tongue refolding is considered to be one of the major structural changes during photoconversion and important for signal transduction. However, according to our SEC studies, the PHY domain is dispensable for phyB-PIF6 interaction (Fig 3.30). In fact, this is very surprising, given that Pfr production is abolished in phyB in the absence of the PHY domain! Perhaps Pfr-like activity is generated transiently, indeed, as in transgenic plants the NPG construct provides for responses to continuous R light but not R light pulses [27].

The NTE seems to be critical to plant phytochrome signalling. First, the NTE contributes to the Pfr thermal stability, its deletion accelerating dark reversion in phyA and phyB [72,115]. Second, phosphorylation of some residues in the seine/threonine-rich NTE attenuates phyB signalling by enhancing dark reversion [13,14]. Third, phytochromes missing the NTE displayed impaired biological activities compared to the WT in transgenic plants [14,115]. Moreover, Our results emphasize the importance of the NTE in phyB(NPGP)-PIF

interaction (Fig 3.30), its conformational changes might be crucial for this interaction. However, it would be ambiguous in the context of the full-length phytochrome, as both N- and C-terminal module interacts with PIFs separately. A complex structure would contribute to revealing this mystery.

5. References

1 Sullivan, J. A., Shirasu, K. & Deng, X. W. The diverse roles of ubiquitin and the 26S proteasome in the life of plants. *Nat. Rev. Genet.* **4**, 948-958, doi:10.1038/nrg1228 (2003).

2 Tepperman, J. M., Zhu, T., Chang, H. S., Wang, X. & Quail, P. H. Multiple transcription-factor genes are early targets of phytochrome A signaling. *Proc. Natl. Acad. Sci. U.S.A.* **98**, 9437-9442 (2001).

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

4 Hughes, J. *et al.* A prokaryotic phytochrome. *Nature* **386**, 663-663 (1997).

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

6 Blumenstein, A. *et al.* The Aspergillus nidulans phytochrome FphA represses sexual development in red light. *Curr. Biol.* **15**, 1833-1838, doi:10.1016/j.cub.2005.08.061 (2005).

7 Froehlich, A. C., Noh, B., Vierstra, R. D., Loros, J. & Dunlap, J. C. Genetic and molecular analysis of phytochromes from the filamentous fungus Neurospora crassa. *Eukaryot. Cell* **4**, 2140-2152, doi:10.1128/EC.4.12.2140-2152.2005 (2005).

8 Clack, T., Mathews, S. & Sharrock, R. A. The phytochrome apoprotein family in Arabidopsis is encoded by five genes: the sequences and expression of PHYD and PHYE. *Plant Mol. Biol.* **25**, 413-427 (1994).

9 Sharrock, R. A. & Quail, P. H. Novel phytochrome sequences in *Arabidopsis thaliana*: structure, evolution, and differential expression of a plant regulatory photoreceptor family. *Genes Dev.* **3**, 1745-1757 (1989).

10 Hirschfeld, M., Tepperman, J. M., Clack, T., Quail, P. H. & Sharrock, R. A. Coordination of phytochrome levels in phyB mutants of arabidopsis as revealed by apoprotein-specific monoclonal antibodies. *Genetics* **149**, 523-535 (1998).

11 Sharrock, R. A. & Clack, T. Patterns of expression and normalized levels of the five Arabidopsis phytochromes. *Plant Physiol.* **130**, 442-456, doi:10.1104/pp.005389 (2002).

12 Hughes, J. Phytochrome three-dimensional structures and functions. *Biochem. Soc. Trans.* **38**, 710-716, doi:10.1042/BST0380710 (2010).

13 Medzihradszky, M. *et al.* Phosphorylation of phytochrome B inhibits light-induced signaling via accelerated dark reversion in Arabidopsis. *Plant Cell* **25**, 535-544, doi:10.1105/tpc.112.106898 (2013).

14 Viczian, A. *et al.* Differential phosphorylation of the N-terminal extension regulates phytochrome B signaling. *New Phytol.*, doi:10.1111/nph.16243 (2019).

15 Rockwell, N. C., Shang, L., Martin, S. S. & Lagarias, J. C. Distinct classes of red/far-red photochemistry within the phytochrome superfamily. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 6123-6127, doi:10.1073/pnas.0902370106 (2009).

16 Tanaka, R. & Tanaka, A. Tetrapyrrole biosynthesis in higher plants. *Annu. Rev. Plant Biol.* **58**, 321-346, doi:10.1146/annurev.arplant.57.032905.105448 (2007).

17 Davis, S. J., Kurepa, J. & Vierstra, R. D. The *Arabidopsis thaliana HY1* locus, required for phytochrome- chromophore biosynthesis, encodes a protein related to heme oxygenases. *Proc. Natl. Acad. Sci. U.S.A.* **96**, 6541-6546 (1999).

18 Kohchi, T. *et al.* The *Arabidopsis HY2* gene encodes phytochromobilin synthase, a ferredoxin-dependent biliverdin reductase. *Plant Cell* **13**, 425-436 (2001).

19 Wahleithner, J. A., Li, L. M. & Lagarias, J. C. Expression and assembly of spectrally active recombinant holophytochrome. *Proc. Natl. Acad. Sci. U.S.A.* **88**, 10387-10391 (1991).

20 Rockwell, N. C., Su, Y. S. & Lagarias, J. C. Phytochrome structure and signaling mechanisms. *Annu. Rev. Plant Biol.* **57**, 837-858, doi:10.1146/annurev.arplant.56.032604.144208 (2006).

21 Wagner, J. R. *et al.* Mutational analysis of *Deinococcus radiodurans* bacteriophytochrome reveals key amino acids necessary for the photochromicity and proton exchange cycle of phytochromes. *J. Biol. Chem.* **283**, 12212-12226, doi:10.1074/jbc.M709355200 (2008).

22 Takala, H., Edlund, P., Ihalainen, J. A. & Westenhoff, S. Tips and turns of bacteriophytochrome photoactivation. *Photochemical & photobiological sciences : Official journal of the European Photochemistry Association and the European Society for Photobiology* **19**, 1488-1510, doi:10.1039/d0pp00117a (2020).

23 van Thor, J. J. *et al.* Light-induced proton release and proton uptake reactions in the cyanobacterial phytochrome Cph1. *Biochemistry* **40**, 11460-11471 (2001).
24 Chen, E., Lapko, V. N., Lewis, J. W., Song, P. S. & Kliger, D. S. Mechanism of native oat phytochrome photoreversion: a time- resolved absorption investigation. *Biochemistry* **35**, 843-850 (1996).
25 Wagner, D., Fairchild, C. D., Kuhn, R. M. & Quail, P. H. Chromophore-bearing NH2-terminal domains of phytochromes A and B determine their photosensory specificity and differential light lability. *Proc. Natl. Acad. Sci. U.S.A.* **93**, 4011-4015 (1996).
26 Matsushita, T., Mochizuki, N. & Nagatani, A. Dimers of the N-terminal domain of phytochrome B are functional in the nucleus. *Nature* **424**, 571-574 (2003).
27 Oka, Y. *et al.* Functional analysis of a 450-amino acid N-terminal fragment of phytochrome B in *Arabidopsis*. *Plant Cell* **16**, 2104-2116 (2004).
28 Mateos, J. L. *et al.* Functional and biochemical analysis of the N-terminal domain of phytochrome A. *J. Biol. Chem.* **281**, 34421-34429 (2006).
29 Viczian, A. *et al.* A short amino-terminal part of Arabidopsis phytochrome A induces constitutive photomorphogenic response. *Mol Plant* **5**, 629-641, doi:10.1093/mp/sss035 (2012).
30 Yoo, C. Y. *et al.* Direct photoresponsive inhibition of a p53-like transcription activation domain in PIF3 by Arabidopsis phytochrome B. *Nat Commun* **12**, 5614, doi:10.1038/s41467-021-25909-5 (2021).
31 Qiu, Y. *et al.* Mechanism of early light signaling by the carboxy-terminal output module of Arabidopsis phytochrome B. *Nat Commun* **8**, 1905, doi:10.1038/s41467-017-02062-6 (2017).
32 Quail, P. H. *et al.* Phytochromes: Photosensory perception and signal transduction. *Science* **268**, 675-680 (1995).
33 Chen, M., Tao, Y., Lim, J., Shaw, A. & Chory, J. Regulation of phytochrome B nuclear localization through light-dependent unmasking of nuclear-localization signals. *Curr. Biol.* **15**, 637-642, doi:10.1016/j.cub.2005.02.028 (2005).
34 Pfeiffer, A. *et al.* Interaction with plant transcription factors can mediate nuclear import of phytochrome B. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 5892-5897, doi:10.1073/pnas.1120764109 (2012).
35 Hiltbrunner, A. *et al.* Nuclear accumulation of the phytochrome A photoreceptor requires FHY1. *Curr. Biol.* **15**, 2125-2130, doi:10.1016/j.cub.2005.10.042 (2005).
36 Hiltbrunner, A. *et al.* FHY1 and FHL act together to mediate nuclear accumulation of the phytochrome A photoreceptor. *Plant Cell Physiol.* **47**, 1023-1034, doi:10.1093/pcp/pcj087 (2006).
37 Zeidler, M., Zhou, Q., Sarda, X., Yau, C. P. & Chua, N. H. The nuclear localization signal and the C-terminal region of FHY1 are required for transmission of phytochrome A signals. *Plant J.* **40**, 355-365 (2004).
38 Zhou, Q. *et al.* FHL is required for full phytochrome A signaling and shares overlapping functions with FHY1. *Plant J.* **43**, 356-370 (2005).
39 Helizon, H. *et al.* Arabidopsis phytochrome A nuclear translocation is mediated by a far-red elongated hypocotyl 1-importin complex. *Plant J.* **96**, 1255-1268, doi:10.1111/tpj.14107 (2018).
40 Genoud, T. *et al.* FHY1 mediates nuclear import of the light-activated phytochrome A photoreceptor. *PLoS Genet.* **4**, e1000143, doi:10.1371/journal.pgen.1000143 (2008).
41 Huq, E. & Quail, P. H. PIF4, a phytochrome-interacting bHLH factor, functions as a negative regulator of phytochrome B signaling in Arabidopsis. *EMBO J.* **21**, 2441-2450 (2002).
42 Leivar, P. *et al.* The Arabidopsis Phytochrome-Interacting Factor PIF7, Together with PIF3 and PIF4, Regulates Responses to Prolonged Red Light by Modulating phyB Levels. *Plant Cell* **20**, 337-352 (2008).
43 Martinez-Garcia, J. F., Huq, E. & Quail, P. H. Direct targeting of light signals to a promoter element-bound transcription factor. *Science* **288**, 859-863, doi:10.1126/science.288.5467.859 (2000).
44 Ni, M., Tepperman, J. M. & Quail, P. H. PIF3, a phytochrome-interacting factor necessary for normal photoinduced signal transduction, is a novel basic helix-loop- helix protein. *Cell* **95**, 657-667 (1998).
45 Ni, M., Tepperman, J. M. & Quail, P. H. Binding of phytochrome B to its nuclear signalling partner PIF3 is reversibly induced by light. *Nature* **400**, 781-784 (1999).
46 Leivar, P. & Quail, P. H. PIFs: pivotal components in a cellular signaling hub. *Trends Plant Sci.* **16**, 19-28, doi:10.1016/j.tplants.2010.08.003 (2011).

47 Oka, Y., Matsushita, T., Mochizuki, N., Quail, P. H. & Nagatani, A. Mutant screen distinguishes between residues necessary for light-signal perception and signal transfer by phytochrome B. *PLoS Genet.* **4**, e1000158, doi:10.1371/journal.pgen.1000158 (2008).
48 Kikis, E. A., Oka, Y., Hudson, M. E., Nagatani, A. & Quail, P. H. Residues clustered in the light-sensing knot of phytochrome B are necessary for conformer-specific binding to signaling partner PIF3. *PLoS Genet.* **5**, e1000352, doi:10.1371/journal.pgen.1000352 (2009).
49 Shin, A. Y. *et al.* Evidence that phytochrome functions as a protein kinase in plant light signalling. *Nat Commun* **7**, 11545, doi:10.1038/ncomms11545 (2016).
50 Ni, W. *et al.* PPKs mediate direct signal transfer from phytochrome photoreceptors to transcription factor PIF3. *Nat Commun* **8**, 15236, doi:10.1038/ncomms15236 (2017).
51 Pham, V. N., Kathare, P. K. & Huq, E. Phytochromes and Phytochrome Interacting Factors. *Plant Physiol.* **176**, 1025-1038, doi:10.1104/pp.17.01384 (2018).
52 Ni, W. *et al.* A mutually assured destruction mechanism attenuates light signaling in Arabidopsis. *Science* **344**, 1160-1164, doi:10.1126/science.1250778 (2014).
53 Zhu, L. *et al.* CUL4 forms an E3 ligase with COP1 and SPA to promote light-induced degradation of PIF1. *Nat Commun* **6**, 7245, doi:10.1038/ncomms8245 (2015).
54 Li, J., Li, G., Wang, H. & Wang Deng, X. Phytochrome signaling mechanisms. *Arabidopsis Book* **9**, e0148, doi:10.1199/tab.0148 (2011).
55 Legris, M., Ince, Y. C. & Fankhauser, C. Molecular mechanisms underlying phytochrome-controlled morphogenesis in plants. *Nat Commun* **10**, 5219, doi:10.1038/s41467-019-13045-0 (2019).
56 Valadon, L. R. G., Osman, M. & Mummery, R. S. Phytochrome mediated carotenoid synthesis in the fungus *Verticillium agaricinum*. *Photochem. Photobiol.* **29**, 605-607 (1979).
57 Tu, S. L., Gunn, A., Toney, M. D., Britt, R. D. & Lagarias, J. C. Biliverdin reduction by cyanobacterial phycocyanobilin:ferredoxin oxidoreductase (PcyA) proceeds via linear tetrapyrrole radical intermediates. *J. Am. Chem. Soc.* **126**, 8682-8693, doi:10.1021/ja049280z (2004).
58 Psakis, G. *et al.* Signaling kinetics of cyanobacterial phytochrome Cph1, a light regulated histidine kinase. *Biochemistry* **50**, 6178-6188, doi:10.1021/bi200612d (2011).
59 Giraud, E. *et al.* Bacteriophytochrome controls photosystem synthesis in anoxygenic bacteria. *Nature* **417**, 202-205 (2002).
60 Wagner, J. R., Brunzelle, J. S., Forest, K. T. & Vierstra, R. D. A light-sensing knot revealed by the structure of the chromophore-binding domain of phytochrome. *Nature* **438**, 325-331, doi:10.1038/nature04118 (2005).
61 Essen, L.-O., Mailliet, J. & Hughes, J. The structure of a complete phytochrome sensory module in the Pr ground state. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 14709-14714, doi:10.1073/pnas.0806477105 (2008).
62 Yang, X., Kuk, J. & Moffat, K. Crystal structure of *Pseudomonas aeruginosa* bacteriophytochrome: Photoconversion and signal transduction. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 14715-14720, doi:10.1073/pnas.0806718105 (2008).
63 Yang, X., Ren, Z., Kuk, J. & Moffat, K. Temperature-scan cryocrystallography reveals reaction intermediates in bacteriophytochrome. *Nature* **479**, 428-432, doi:10.1038/nature10506 (2011).
64 Fischer, A. J. & Lagarias, J. C. Harnessing phytochrome's glowing potential. *Proc. Natl. Acad. Sci. U.S.A.* **101**, 17334-17339, doi:10.1073/pnas.0407645101 (2004).
65 Su, Y. S. & Lagarias, J. C. Light-independent phytochrome signaling mediated by dominant GAF domain tyrosine mutants of *Arabidopsis* phytochromes in transgenic plants. *Plant Cell* **19**, 2124-2139, doi:10.1105/tpc.107.051516 (2007).
66 Song, C. *et al.* Two ground state isoforms and a chromophore *D*-ring photoflip triggering extensive intramolecular changes in a canonical phytochrome. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 3842-3847, doi:10.1073/pnas.1013377108 (2011).
67 Anders, K., Daminelli-Widany, G., Mroginski, M. A., von Stetten, D. & Essen, L. O. Structure of the cyanobacterial phytochrome 2 photosensor implies a tryptophan switch for phytochrome signaling. *J. Biol. Chem.* **288**, 35714-35725, doi:10.1074/jbc.M113.510461 (2013).
68 Takala, H. *et al.* Signal amplification and transduction in phytochrome photosensors. *Nature* **509**, 245-248, doi:10.1038/nature13310 (2014).
69 Kacprzak, S. *et al.* Intersubunit distances in full-length, dimeric, bacterial phytochrome Agp1, as measured by pulsed electron-electron double resonance (PELDOR) between different spin label positions, remain unchanged upon photoconversion. *J. Biol. Chem.* **292**, 7598-7606, doi:10.1074/jbc.M116.761882 (2017).

70 Burgie, E. S., Zhang, J. & Vierstra, R. D. Crystal structure of *Deinococcus* phytochrome in the photoactivated state reveals a cascade of structural rearrangements during photoconversion. *Structure* **24**, 448-457, doi:10.1016/j.str.2016.01.001 (2016).
71 Burgie, E. S. *et al.* Crystallographic and electron microscopic analyses of a bacterial phytochrome reveal local and global rearrangements during photoconversion. *J. Biol. Chem.*, doi:10.1074/jbc.M114.571661 (2014).
72 Burgie, E. S., Bussell, A. N., Walker, J. M., Dubiel, K. & Vierstra, R. D. 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, doi:10.1073/pnas.1403096111 (2014).
73 Song, C. *et al.* 3D structures of plant phytochrome A as Pr and Pfr from solid-state NMR: Implications for molecular function. *Frontiers in plant science* **9**, 498, doi:10.3389/fpls.2018.00498 (2018).
74 Bartesaghi, A. *et al.* Atomic Resolution Cryo-EM Structure of beta-Galactosidase. *Structure* **26**, 848-856 e843, doi:10.1016/j.str.2018.04.004 (2018).
75 Wahlgren, W. Y., Golonka, D., Westenhoff, S. & Moglich, A. Cryo-Electron Microscopy of Arabidopsis thaliana Phytochrome A in Its Pr State Reveals Head-to-Head Homodimeric Architecture. *Frontiers in plant science* **12**, 663751, doi:10.3389/fpls.2021.663751 (2021).
76 Li, H., Burgie, E. S., Gannam, Z. T. K., Li, H. L. & Vierstra, R. D. Plant phytochrome B is an asymmetric dimer with unique signalling potential. *Nature*, doi:10.1038/s41586-022-04529-z (2022).
77 Landgraf, F. T., Forreiter, C., Hurtado, P. A., Lamparter, T. & Hughes, J. Recombinant holophytochrome in *Escherichia coli*. *FEBS Lett.* **508**, 459-462 (2001).
78 Beyer, H. M. *et al.* Generic and reversible opto-trapping of biomolecules. *Acta Biomater.* **79**, 276-282, doi:10.1016/j.actbio.2018.08.032 (2018).
79 Heras, B. & Martin, J. L. Post-crystallization treatments for improving diffraction quality of protein crystals. *Acta Crystallogr D Biol Crystallogr* **61**, 1173-1180, doi:10.1107/S0907444905019451 (2005).
80 Kabsch, W. Xds. *Acta Crystallogr D Biol Crystallogr* **66**, 125-132, doi:10.1107/S0907444909047337 (2010).
81 Evans, P. R. & Murshudov, G. N. How good are my data and what is the resolution? *Acta Crystallogr D Biol Crystallogr* **69**, 1204-1214, doi:10.1107/S0907444913000061 (2013).
82 Winn, M. D. *et al.* Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr* **67**, 235-242, doi:10.1107/S0907444910045749 (2011).
83 McCoy, A. J. *et al.* Phaser crystallographic software. *J Appl Crystallogr* **40**, 658-674, doi:10.1107/S0021889807021206 (2007).
84 Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* **66**, 486-501, doi:10.1107/S0907444910007493 (2010).
85 Murshudov, G. N. *et al.* REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr D Biol Crystallogr* **67**, 355-367, doi:10.1107/S0907444911001314 (2011).
86 Jaedicke, K., Rösler, J., Gans, T. & Hughes, J. Bellis perennis: a useful tool for protein localization studies. *Planta* **234**, 759-768, doi:10.1007/s00425-011-1443-7 (2011).
87 Mailliet, J. *et al.* Spectroscopy and a high-resolution crystal structure of Tyr263 mutants of cyanobacterial phytochrome Cph1. *J. Mol. Biol.* **413**, 115-127, doi:10.1016/j.jmb.2011.08.023 (2011).
88 Nagano, S. *et al.* Improved fluorescent phytochromes for in situ imaging. *Sci Rep* **12**, 5587, doi:10.1038/s41598-022-09169-x (2022).
89 Nagano, S. *et al.* Structural insights into photoactivation and signalling in plant phytochromes. *Nat Plants* **6**, 581-588, doi:10.1038/s41477-020-0638-y (2020).
90 Fischer, A. J. *et al.* Multiple roles of a conserved GAF domain tyrosine residue in cyanobacterial and plant phytochromes. *Biochemistry* **44**, 15203-15215 (2005).
91 Hershey, H. P., Colbert, J. T., Lissemore, J. L., Barker, R. F. & Quail, P. H. Molecular cloning of cDNA for *Avena* phytochrome. *Proc. Natl. Acad. Sci. U.S.A.* **81**, 2332-2336 (1984).
92 Kosugi, S. *et al.* Six classes of nuclear localization signals specific to different binding grooves of importin alpha. *J. Biol. Chem.* **284**, 478-485, doi:10.1074/jbc.M807017200 (2009).
93 Do, H. J. *et al.* Identification of multiple nuclear localization signals in murine Elf3, an ETS transcription factor. *FEBS Lett.* **580**, 1865-1871, doi:10.1016/j.febslet.2006.02.049 (2006).

94 Anami, S. *et al.* Red light-regulated interaction of Per-Arnt-Sim histidine kinases with partner histidine-containing phosphotransfer proteins in Physcomitrium patens. *Genes Cells* **26**, 698-713, doi:10.1111/gtc.12878 (2021).
95 Jaedicke, K., Lichtenthaler, A. L., Meyberg, R., Zeidler, M. & Hughes, J. A phytochrome-phototropin light signaling complex at the plasma membrane. *Proc Natl Acad Sci U S A* **109**, 12231-12236, doi:10.1073/pnas.1120203109 (2012).
96 Rösler, J., Klein, I. & Zeidler, M. Arabidopsis fhl/fhy1 double mutant reveals a distinct cytoplasmic action of phytochrome A. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 10737-10742, doi:10.1073/pnas.0703855104 (2007).
97 Horner, M. *et al.* Production of Phytochromes by High-Cell-Density E. coli Fermentation. *ACS Synth Biol* **8**, 2442-2450, doi:10.1021/acssynbio.9b00267 (2019).
98 Golonka, D. *et al.* Deconstructing and repurposing the light-regulated interplay between Arabidopsis phytochromes and interacting factors. *Commun Biol* **2**, 448, doi:10.1038/s42003-019-0687-9 (2019).
99 Khanna, R. *et al.* A novel molecular recognition motif necessary for targeting photoactivated phytochrome signaling to specific basic helix-loop-helix transcription factors. *Plant Cell* **16**, 3033-3044 (2004).
100 Smith, R. W. *et al.* Interactions Between phyB and PIF Proteins Alter Thermal Reversion Reactions in vitro. *Photochem. Photobiol.* **93**, 1525-1531, doi:10.1111/php.12793 (2017).
101 Vierstra, R. & Quail, P. H. Native phytochrome: inhibition of proteolysis yields a homogeneous monomer of 124 kilodaltons from #Avena#. *Proc. Natl. Acad. Sci. U.S.A.* **79**, 5272-5276 (1982).
102 Vierstra, R. D., Cordonnier, M. M. l., Pratt, L. H. & Quail, P. H. Native phytochrome: immunoblot analysis of relative molecular mass and in-vitro proteolytic degradation for several plant species. *Planta* **160**, 521-528 (1984).
103 Ulijasz, A. T. *et al.* Characterization of two thermostable cyanobacterial phytochromes reveals global movements in the chromophore-binding domain during photoconversion. *J. Biol. Chem.* **283**, 21251-21266, doi:10.1074/jbc.M801592200 (2008).
104 Song, C. *et al.* The D-ring, not the A-ring, rotates in Synechococcus OS-B' phytochrome. *J. Biol. Chem.* **289**, 2552-2562, doi:10.1074/jbc.M113.520031 (2014).
105 Auldridge, M. E., Satyshur, K. A., Anstrom, D. M. & Forest, K. T. Structure-guided engineering enhances a phytochrome-based infrared fluorescent protein. *J. Biol. Chem.* **287**, 7000-7009, doi:10.1074/jbc.M111.295121 (2012).
106 Park, C. M., Bhoo, S. H. & Song, P. S. Inter-domain crosstalk in the phytochrome molecules. *Semin. Cell Dev. Biol.* **11**, 449-456, doi:10.1006/scdb.2000.0200 (2000).
107 Hahn, T. R., Kang, S. S. & Song, P. S. Difference in the degree of exposure of chromophores in the Pr and Pfr forms of phytochrome. *Biochem.Biophys.Res.Commun.* **97**, 1317-1323 (1980).
108 Rumfeldt, J., Kurttila, M., Takala, H. & Ihalainen, J. A. The hairpin extension controls solvent access to the chromophore binding pocket in a bacterial phytochrome: a UV-vis absorption spectroscopy study. *Photochemical & photobiological sciences : Official journal of the European Photochemistry Association and the European Society for Photobiology* **20**, 1173-1181, doi:10.1007/s43630-021-00090-2 (2021).
109 Borucki, B. *et al.* Light-induced proton release of phytochrome is coupled to the transient deprotonation of the tetrapyrrole chromophore. *J. Biol. Chem.* **280**, 34358-34364, doi:10.1074/jbc.M505493200 (2005).
110 Claesson, E. *et al.* The primary structural photoresponse of phytochrome proteins captured by a femtosecond X-ray laser. *eLife* **9**, doi:10.7554/eLife.53514 (2020).
111 Hahn, J. *et al.* Probing protein-chromophore interactions in Cph1 phytochrome by mutagenesis. *FEBS J.* **273**, 1415-1429, doi:10.1111/j.1742-4658.2006.05164.x (2006).
112 von Horsten, S. *et al.* Mapping light-driven conformational changes within the photosensory module of plant phytochrome B. *Sci Rep* **6**, 34366, doi:10.1038/srep34366 (2016).
113 Nito, K., Wong, C. C., Yates, J. R., 3rd & Chory, J. Tyrosine phosphorylation regulates the activity of phytochrome photoreceptors. *Cell reports* **3**, 1970-1979, doi:10.1016/j.celrep.2013.05.006 (2013).
114 Chen, F. *et al.* Phosphorylation of FAR-RED ELONGATED HYPOCOTYL1 is a key mechanism defining signaling dynamics of phytochrome A under red and far-red light in Arabidopsis. *Plant Cell* **24**, 1907-1920, doi:10.1105/tpc.112.097733 (2012).
115 Cherry, J. R., Hondred, D., Walker, J. M. & Vierstra, R. D. Phytochrome requires the 6-kDa N-terminal domain for full biological activity. *Proc. Natl. Acad. Sci. U.S.A.* **89**, 5039-5043 (1992).

6.Table of figures

7. Appendices

Table 7.1 Summary statistics of *Gm*phyB(PGP) crystal structure as Pr

Data collection	
Space group	P $3_1$21
Cell dimensions	102.27, 102.27, 222.92 Å; 90, 90, 120°
Resolution (Å)	47.17-2.90 (3.06-2.90) *
R_{merge}	0.238 (5.838)
Mean I/sigma(I)	12.1 (0.8)
Completeness (%)	99.9(99.9)
Refinement	
Resolution (Å)	47.17-2.90 (3.06-2.90)
No. of reflections	30475
R_{work} / R_{free}	0.2570 / 0.2824
Average B-factor	121.8
rmsds	
Bond lengths (Å)	0.009
Bond angles (°)	1.392
Ramachandran (%)	
favored	91.2
outliers	0.4
PDB	6TL4

* Highest resolution shell is shown in parentheses.

Table 7.2 Summary statistics of 2.1 Å *Gm*phyA(PG) crystal structure as Pr.

Data collection	
Space group	P 2_1
Cell dimensions	55.42, 113.23, 68.46 Å; 90, 93.52, 90°
Resolution (Å)	44.35-2.13 (2.19-2.13) *
R_{merge}	0.127 (1.950)
Mean I/sigma(I)	10.0 (0.9)

Completeness (%)	99.5 (99.8)
Refinement	
Resolution (Å)	44.35-2.13 (2.19-2.13)
No. of reflections	44964
R_{work} / R_{free}	0.189 / 0.242
Average B-factor	48.6
rmsd values	
Bond lengths (Å)	0.02
Bond angles (°)	2.469
Ramachandran (%)	
favored	96.2
outliers	0.6
PDB	6TC7

* Highest resolution shell is shown in parentheses.

Table 7.3 Summary statistics of 1.58 Å *Gm*phyA(PG) structure as Pr

Data collection	
Space group	P 2_1
Cell dimensions	55.29, 111.81, 68.27 Å; 90, 92.31, 90°
Resolution (Å)	43.24-1.58 (1.61-1.58) *
R_{merge}	0.06 (1.243)
Mean I/sigma(I)	8.4 (0.7)
Completeness (%)	99.7 (99.7)
Refinement	
Resolution (Å)	43.24-1.58 (1.61-1.58)
No. of reflections	107044
R_{work} / R_{free}	0.174 / 0.197
Average B-factor	30.229
rmsd values	
Bond lengths (Å)	0.019
Bond angles (°)	2.264

Ramachandran (%)	
favored	97.73
outliers	0

* Highest resolution shell is shown in parentheses.

Table 7.4 Summary statistics of 2.4 Å *Gm*phyA(PG) illuminated structure.

Data collection	
Space group	P 4_3 2 2
Cell dimensions	55, 55, 266Å/90, 90, 90°
Resolution (Å)	46.77-2.42 (2.50-2.42) *
R_{merge}	0.079 (0.82)
Mean I/sigma(I)	19.6 (3.3)
Completeness (%)	100 (100)
Refinement	
Resolution (Å)	46.77-2.42 (2.50-2.42)
No. of reflections	15801
R_{work} / R_{free}	0.246 / 0.288
Average B-factor	65.1
rmsd values	
Bond lengths (Å)	0.015
Bond angles (°)	1.683
Ramachandran (%)	
favored	93.5
outliers	0.65

* Highest resolution shell is shown in parentheses.

Protein sequence:

Cloning site, affinity purification tag, site-directed mutation were shown in red, blue and orange, respectively.

*Gm*phyA(NPGP): MA 1-590aa 6His

MAMSTSRPSQSSSNSGRSRRSARAMALATVDAKLHATFEESGSSFDYSSSVRISGTADGVNQ
PRHDKVTTAYLHHMQKGKMIQPFGCLLALDEKTCKVIAYSENAPEMLTMVSHAVPSVGDHP

ALGIGTDIKTLFTAPSASALQKALGFAEVLLLNPVLIHCKTSGKPFYAIIHRVTGSMIIDFEPVK
PYEVPMTAAGALQSYKLAAKAITRLQSLPSGSMERLCDTMVQEVFELTGYDRVMAYKFHE
DDHGEVIAEITKPGLEPYLGLHYPATDIPQASRFLFMKNKVRMIVDCHAKHVRVLQDEKLPF
DLTLCGSTLRAPHSCHAQYMANMDSIASLVMAVVVNDNEEDGDTDAIQPQKRKRLWGLVV
CHNTTPRFVPFPLRYACEFLAQVFAIHVNKEIELEYQIIEKNILRTQTLLCDLVMRDAPLGIVS
ESPNIMDLVKCDGAALIYKNKVWRLGVTPSESQIREIAFWLSEYHMDSTGFSTDSLSDAGFP
SALSLGDVVCGMAAVRVTAKDVVFWFRSHTAAEIRWGGAKHEAGEKDDGRRMHPRSSFK
VFLDVVKARSLPWKEYEIDAMHSLQLILRNAFKDTESHHHHHH

*Gm*phyB(NPGP): 1-615aa 6His

MASASGAENSSVPPSPLPPPPPPQIHTSRTKLSHHHHNNNNNNNNNIDSTSKAIAQYTEDARL
HAVFEQSGESGRSFDYSQSIRVTSESVPEQQITAYLLKIQRGGFIQPFGSMIAVDEPSFRILAYS
DNARDMLGITPQSVPSLDDKNDAAFALGTDIRTLFTHSSAVLLEKAFSAREISLMNPIWIHSR
TSGKPFYGILHRIDVGIVIDLEPARTEDPALSIAGAVQSQKLAVRAISQLQSLPGGDVKLLCDT
VVESVRELTGYDRVMVYRFHEDEHGEVVAETKRPDLEPYIGLHYPATDIPQASRFLFKQNRV
RMIVDCHASAVRVVQDEALVQPLCLVGSTLRAPHGCHAQYMANMGSTASLVMAVIINGND
EEGVGGRTSMRLWGLVVCHHTSARCIPFPLRYACEFLMQAFGLQLNMELQLAAQSLEKRVL
RTQTLLCDMLLRDSPTGIVTQSPSIMDLVKCDGAALYYQGNYYPLGVTPTEAQIRDIIEWLL
AFHRDSTGLSTDSLADAGYPGAASLGDAVCGMAVAYITEKDFLFWFRSHTAKEIKWGGAKH
HPEDKDDGQRMHPRSSFKAFLEVVKSRSLPWENAEMDAIHSLQLILRDSFKDAEHHHHHH

*Gm*phyA(PGP): M 51-590aa 6His

MISGTADGVNQPRHDKVTTAYLHHMQKGKMIQPFGCLLALDEKTCKVIAYSENAPEMLTM
VSHAVPSVGDHPALGIGTDIKTLFTAPSASALQKALGFAEVLLLNPVLIHCKTSGKPFYAIIHR
VTGSMIIDFEPVKPYEVPMTAAGALQSYKLAAKAITRLQSLPSGSMERLCDTMVQEVFELTG
YDRVMAYKFHEDDHGEVIAEITKPGLEPYLGLHYPATDIPQASRFLFMKNKVRMIVDCHAK
HVRVLQDEKLPFDLTLCGSTLRAPHSCHAQYMANMDSIASLVMAVVVNDNEEDGDTDAIQ
PQKRKRLWGLVVCHNTTPRFVPFPLRYACEFLAQVFAIHVNKEIELEYQIIEKNILRTQTLLCD
LVMRDAPLGIVSESPNIMDLVKCDGAALIYKNKVWRLGVTPSESQIREIAFWLSEYHMDSTG
FSTDSLSDAGFPSALSLGDVVCGMAAVRVTAKDVVFWFRSHTAAEIRWGGAKHEAGEKDD
GRRMHPRSSFKVFLDVVKARSLPWKEYEIDAMHSLQLILRNAFKDTESHHHHHH

*Gm*phyB(PGP): M 85-615aa 6His

MVTSESVPEQQITAYLLKIQRGGFIQPFGSMIAVDEPSFRILAYSDNARDMLGITPQSVPSLDD
KNDAAFALGTDIRTLFTHSSAVLLEKAFSAREISLMNPIWIHSRTSGKPFYGILHRIDVGIVIDL
EPARTEDPALSIAGAVQSQKLAVRAISQLQSLPGGDVKLLCDTVVESVRELTGYDRVMVYRF
HEDEHGEVVAETKRPDLEPYIGLHYPATDIPQASRFLFKQNRVRMIVDCHASAVRVVQDEAL
VQPLCLVGSTLRAPHGCHAQYMANMGSTASLVMAVIINGNDEEGVGGRTSMRLWGLVVCH
HTSARCIPFPLRYACEFLMQAFGLQLNMELQLAAQSLEKRVLRTQTLLCDMLLRDSPTGIVT
QSPSIMDLVKCDGAALYYQGNYYPLGVTPTEAQIRDIIEWLLAFHRDSTGLSTDSLADAGYP
GAASLGDAVCGMAVAYITEKDFLFWFRSHTAKEIKWGGAKHHPEDKDDGQRMHPRSSFKA
FLEVVKSRSLPWENAEMDAIHSLQLILRDSFKDAEHHHHHH

*Gm*phyA(PG): M 51-402aa 6His

MISGTADGVNQPRHDKVTTAYLHHMQKGKMIQPFGCLLALDEKTCKVIAYSENAPEMLTM
VSHAVPSVGDHPALGIGTDIKTLFTAPSASALQKALGFAEVLLLNPVLIHCKTSGKPFYAIIHR
VTGSMIIDFEPVKPYEVPMTAAGALQSYKLAAKAITRLQSLPSGSMERLCDTMVQEVFELTG
YDRVMAYKFHEDDHGEVIAEITKPGLEPYLGLHYPATDIPQASRFLFMKNKVRMIVDCHAK
HVRVLQDEKLPFDLTLCGSTLRAPHSCHAQYMANMDSIASLVMAVVVNDNEEDGDTDAIQ
PQKRKRLWGLVVCHNTTPRFVPFPLRYACEFLAQVFAIHVNKEIELHHHHHH

*Gm*phyB(NPG): 1-428aa 6His

MASASGAENSSVPPSPLPPPPPPQIHTSRTKLSHHHHNNNNNNNNNIDSTSKAIAQYTEDARL
HAVFEQSGESGRSFDYSQSIRVTSESVPEQQITAYLLKIQRGGFIQPFGSMIAVDEPSFRILAYS
DNARDMLGITPQSVPSLDDKNDAAFALGTDIRTLFTHSSAVLLEKAFSAREISLMNPIWIHSR
TSGKPFYGILHRIDVGIVIDLEPARTEDPALSIAGAVQSQKLAVRAISQLQSLPGGDVKLLCDT
VVESVRELTGYDRVMVYRFHEDEHGEVVAETKRPDLEPYIGLHYPATDIPQASRFLFKQNRV
RMIVDCHASAVRVVQDEALVQPLCLVGSTLRAPHGCHAQYMANMGSTASLVMAVIINGND
EEGVGGRTSMRLWGLVVCHHTSARCIPFPLRYACEFLMQAFGLQLNMELQLHHHHHH

*Gm*phyB(NPG)-Y272H

MASASGAENSSVPPSPLPPPPPPQIHTSRTKLSHHHHNNNNNNNNNIDSTSKAIAQYTEDARL
HAVFEQSGESGRSFDYSQSIRVTSESVPEQQITAYLLKIQRGGFIQPFGSMIAVDEPSFRILAYS
DNARDMLGITPQSVPSLDDKNDAAFALGTDIRTLFTHSSAVLLEKAFSAREISLMNPIWIHSR
TSGKPFYGILHRIDVGIVIDLEPARTEDPALSIAGAVQSQKLAVRAISQLQSLPGGDVKLLCDT
VVESVRELTGYDRVMVHRFHEDEHGEVVAETKRPDLEPYIGLHYPATDIPQASRFLFKQNRV
RMIVDCHASAVRVVQDEALVQPLCLVGSTLRAPHGCHAQYMANMGSTASLVMAVIINGND
EEGVGGRTSMRLWGLVVCHHTSARCIPFPLRYACEFLMQAFGLQLNMELQLHHHHHH

*Gm*phyA(PGP)-R549A:

MISGTADGVNQPRHDKVTTAYLHHMQKGKMIQPFGCLLALDEKTCKVIAYSENAPEMLTM
VSHAVPSVGDHPALGIGTDIKTLFTAPSASALQKALGFAEVLLLNPVLIHCKTSGKPFYAIIHR
VTGSMIIDFEPVKPYEVPMTAAGALQSYKLAAKAITRLQSLPSGSMERLCDTMVQEVFELTG
YDRVMAYKFHEDDHGEVIAEITKPGLEPYLGLHYPATDIPQASRFLFMKNKVRMIVDCHAK
HVRVLQDEKLPFDLTLCGSTLRAPHSCHAQYMANMDSIASLVMAVVVNDNEEDGDTDAIQ
PQKRKRLWGLVVCHNTTPRFVPFPLRYACEFLAQVFAIHVNKEIELEYQIIEKNILRTQTLLCD
LVMRDAPLGIVSESPNIMDLVKCDGAALIYKNKVWRLGVTPSESQIREIAFWLSEYHMDSTG
FSTDSLSDAGFPSALSLGDVVCGMAAVRVTAKDVVFWFRSHTAAEIRWGGAKHEAGEKDD
GRRMHPASSFKVFLDVVKARSLPWKEYEIDAMHSLQLILRNAFKDTESHHHHHH

*Gm*phyA(PGP)-Y242H

MISGTADGVNQPRHDKVTTAYLHHMQKGKMIQPFGCLLALDEKTCKVIAYSENAPEMLTM
VSHAVPSVGDHPALGIGTDIKTLFTAPSASALQKALGFAEVLLLNPVLIHCKTSGKPFYAIIHR
VTGSMIIDFEPVKPYEVPMTAAGALQSYKLAAKAITRLQSLPSGSMERLCDTMVQEVFELTG
YDRVMAHKFHEDDHGEVIAEITKPGLEPYLGLHYPATDIPQASRFLFMKNKVRMIVDCHAK
HVRVLQDEKLPFDLTLCGSTLRAPHSCHAQYMANMDSIASLVMAVVVNDNEEDGDTDAIQ
PQKRKRLWGLVVCHNTTPRFVPFPLRYACEFLAQVFAIHVNKEIELEYQIIEKNILRTQTLLCD
LVMRDAPLGIVSESPNIMDLVKCDGAALIYKNKVWRLGVTPSESQIREIAFWLSEYHMDSTG
FSTDSLSDAGFPSALSLGDVVCGMAAVRVTAKDVVFWFRSHTAAEIRWGGAKHEAGEKDD
GRRMHPRSSFKVFLDVVKARSLPWKEYEIDAMHSLQLILRNAFKDTESHHHHHH

*Gm*phyB(NPGP)-R575A

MASASGAENSSVPPSPLPPPPPPQIHTSRTKLSHHHHNNNNNNNNNIDSTSKAIAQYTEDARL
HAVFEQSGESGRSFDYSQSIRVTSESVPEQQITAYLLKIQRGGFIQPFGSMIAVDEPSFRILAYS
DNARDMLGITPQSVPSLDDKNDAAFALGTDIRTLFTHSSAVLLEKAFSAREISLMNPIWIHSR
TSGKPFYGILHRIDVGIVIDLEPARTEDPALSIAGAVQSQKLAVRAISQLQSLPGGDVKLLCDT
VVESVRELTGYDRVMVYRFHEDEHGEVVAETKRPDLEPYIGLHYPATDIPQASRFLFKQNRV
RMIVDCHASAVRVVQDEALVQPLCLVGSTLRAPHGCHAQYMANMGSTASLVMAVIINGND
EEGVGGRTSMRLWGLVVCHHTSARCIPFPLRYACEFLMQAFGLQLNMELQLAAQSLEKRVL
RTQTLLCDMLLRDSPTGIVTQSPSIMDLVKCDGAALYYQGNYYPLGVTPTEAQIRDIIEWLL

AFHRDSTGLSTDSLADAGYPGAASLGDAVCGMAVAYITEKDFLFWFRSHTAKEIKWGGAKH
HPEDKDDGQRMHPASSFKAFLEVVKSRSLPWENAEMDAIHSLQLILRDSFKDAEHHHHHH

*Gm*phyB(NPGP)-F578W

MASASGAENSSVPPSPLPPPPPPQIHTSRTKLSHHHHNNNNNNNNNIDSTSKAIAQYTEDARL
HAVFEQSGESGRSFDYSQSIRVTSESVPEQQITAYLLKIQRGGFIQPFGSMIAVDEPSFRILAYS
DNARDMLGITPQSVPSLDDKNDAAFALGTDIRTLFTHSSAVLLEKAFSAREISLMNPIWIHSR
TSGKPFYGILHRIDVGIVIDLEPARTEDPALSIAGAVQSQKLAVRAISQLQSLPGGDVKLLCDT
VVESVRELTGYDRVMVYRFHEDEHGEVVAETKRPDLEPYIGLHYPATDIPQASRFLFKQNRV
RMIVDCHASAVRVVQDEALVQPLCLVGSTLRAPHGCHAQYMANMGSTASLVMAVIINGND
EEGVGGRTSMRLWGLVVCHHTSARCIPFPLRYACEFLMQAFGLQLNMELQLAAQSLEKRVL
RTQTLLCDMLLRDSPTGIVTQSPSIMDLVKCDGAALYYQGNYYPLGVTPTEAQIRDIIEWLL
AFHRDSTGLSTDSLADAGYPGAASLGDAVCGMAVAYITEKDFLFWFRSHTAKEIKWGGAKH
HPEDKDDGQRMHPRSSWKAFLEVVKSRSLPWENAEMDAIHSLQLILRDSFKDAEHHHHHH

*Gm*phyB(NPGP)-G557E:

MASASGAENSSVPPSPLPPPPPPQIHTSRTKLSHHHHNNNNNNNNNIDSTSKAIAQYTEDARL
HAVFEQSGESGRSFDYSQSIRVTSESVPEQQITAYLLKIQRGGFIQPFGSMIAVDEPSFRILAYS
DNARDMLGITPQSVPSLDDKNDAAFALGTDIRTLFTHSSAVLLEKAFSAREISLMNPIWIHSR
TSGKPFYGILHRIDVGIVIDLEPARTEDPALSIAGAVQSQKLAVRAISQLQSLPGGDVKLLCDT
VVESVRELTGYDRVMVYRFHEDEHGEVVAETKRPDLEPYIGLHYPATDIPQASRFLFKQNRV
RMIVDCHASAVRVVQDEALVQPLCLVGSTLRAPHGCHAQYMANMGSTASLVMAVIINGND
EEGVGGRTSMRLWGLVVCHHTSARCIPFPLRYACEFLMQAFGLQLNMELQLAAQSLEKRVL
RTQTLLCDMLLRDSPTGIVTQSPSIMDLVKCDGAALYYQGNYYPLGVTPTEAQIRDIIEWLL
AFHRDSTGLSTDSLADAGYPGAASLGDAVCGMAVAYITEKDFLFWFRSHTAKEIKWEGAKH
HPEDKDDGQRMHPRSSFKAFLEVVKSRSLPWENAEMDAIHSLQLILRDSFKDAEHHHHHH

*Gm*phyB(NPGP)-Y272H

MASASGAENSSVPPSPLPPPPPPQIHTSRTKLSHHHHNNNNNNNNNIDSTSKAIAQYTEDARL
HAVFEQSGESGRSFDYSQSIRVTSESVPEQQITAYLLKIQRGGFIQPFGSMIAVDEPSFRILAYS
DNARDMLGITPQSVPSLDDKNDAAFALGTDIRTLFTHSSAVLLEKAFSAREISLMNPIWIHSR
TSGKPFYGILHRIDVGIVIDLEPARTEDPALSIAGAVQSQKLAVRAISQLQSLPGGDVKLLCDT
VVESVRELTGYDRVMVHRFHEDEHGEVVAETKRPDLEPYIGLHYPATDIPQASRFLFKQNRV
RMIVDCHASAVRVVQDEALVQPLCLVGSTLRAPHGCHAQYMANMGSTASLVMAVIINGND
EEGVGGRTSMRLWGLVVCHHTSARCIPFPLRYACEFLMQAFGLQLNMELQLAAQSLEKRVL
RTQTLLCDMLLRDSPTGIVTQSPSIMDLVKCDGAALYYQGNYYPLGVTPTEAQIRDIIEWLL
AFHRDSTGLSTDSLADAGYPGAASLGDAVCGMAVAYITEKDFLFWFRSHTAKEIKWGGAKH
HPEDKDDGQRMHPRSSFKAFLEVVKSRSLPWENAEMDAIHSLQLILRDSFKDAEHHHHHH

GFP-PIIF6(1-100aa) in pHB111

MVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTLV
TTLTYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNR
IELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLADHYQ
QNTPIGDGPVLLPDNHYLSTQSKLSKDPNEKRDHMVLLEFVTAAGITLGMDELYKGSAGSA
GMMFLPTDYCCRLSDQEYMELVFENGQILAKGQRSNVSLHNQRTKSIMDLYEAEYNEDFM
KSIIHGGGGAITNLGDTQVVPQSHVAAAHETNMLESNKHVDHHHHHH

8. Abbreviations

Å	Ångstrom
AMS	Ammonium sulfate
APA	Active phyA binding
APB	Active phyB binding
APS	Ammonium persulfate
At	*Arabidopsis thaliana*
a	*anti* conformation of a bridge single bond in a bilin
BESSY	*Berliner Elektronenspeicherring-Gesellschaft für Synchrotronstrahlung* (Berlin electron storage ring company for synchrotron radiation)
bHLH	Basic helix-loop-helix
BV	Biliverdin
CAT	Catalytic ATPase-like
CD	Circular dichroism
COP1	Constitutively photomorphogenic1
CRL	Cullin ring ubiquitin ligase
Cryo-EM	Cryogenic electron microscopy
DHP	Dimerization and histidine phosphoreceptor
*Dr*BphP	*Deinococcus radiodurans* bacteriophytochrome photoreceptor
DTT	Dithiothreitol
E	*entgegengesetzt* (*trans* configuration of a bridge double bond in a bilin)
EDTA	Ethylenediaminetetraacetic acid
EtBr	Ethidium bromide
FHL	FHY1 like
FHY1	Far-red elongated hypocotyl
FR	Far-red light
GAF	cGMP phosphodiesterase/adenylate cyclase/FhlA

GFP	Green fluorescent protein
Gm	*Glycine max*
HKM	Histidine kinase-like module
HO	Heme oxygenase
HPGP	HELIX-PAS-GAF-PHY
HY1	*Arabidopsis* heme oxygenase
HY2	*Arabidopsis* phytochromobilin synthase
HY5	Elongated hypocotyl 5 transcription factor
IPTG	Isopropyl β- d-1-thiogalactopyranoside
kDa	Kilodalton
LB	Lysogeny broth
MLKs	MUT9-like Ser/Thr kinases
NES	Nuclear export signal
NLS	Nuclear localization signal
NMR	Nuclear magnetic resonance
NPGP	NTE-nPAS-GAF-PHY
NPG	NTE-nPAS-GAF
NTE	N-terminal extension
Φ_F	Fluorescence quantum yield
*Pa*BphP	*Pseudomonas aeruginosa* bacteriophytochrome photoreceptor
PAS	Period/ARNT/Singleminded
PCB	Phycocyanobilin
PΦB	Phytochromobilin
PCR	Polymerase chain reaction
PcyA	Ferredoxin oxidoreductase
PDB	Protein data bank
PEG	Polyethylene glycol
PGA	Polyglutamic acid

PG	PAS-GAF
PGP	PAS-GAF-PHY
PHY	Phytochrome-specific
PIF	Phytochrome interacting factor
PPK	Photoregulatory protein kinase
Pfr	Far-red absorbing
Pr	R absorbing
s	*syn* conformation of a bridge single bond in a bilin
Sb	*Sorghum bicolor*
SDS-PAGE	Sodium dodecyl sulfate–polyacrylamide gel electrophoresis
SEC	Size-exclusion chromatography
SPA	Suppressor of phyA
SV 40	Simian virus 40
TB	Terrific broth
TBE	Tris Borate EDTA buffer
TEMED	Tetramethylethylenediamine
YFP	Yellow fluorescent protein
Z	*zusammen* (*cis* configuration of a bridge double bond in a bilin)

9. Acknowledgements

I would like to express my deep gratitude to Prof. Jon Hughes for providing me with the excellent platform that enabled me to carry out my research project, and also for his in-depth, enthusiastic supervision during my Ph.D. His cautious attitude and critical thinking have profoundly affected my way of conducting a research project.

I would like to thank Dr. Soshichiro Nagano for his continuous guidance, support and cooperation. He is always very helpful and has always answered my numerous questions concerning phytochrome and X-ray crystallography. I am very grateful for the knowledge he has given me.

My sincere thanks also go to Prof. Ulrike Alexiev for her supervision and cooperation from the SFB 1078. Her opinions and suggestions during our annual meetings were always friendly, constructive and helpful. I also thank Prof. Mathias Zeidler who gave me a lot of constructive suggestions at our group seminars.

I would like to thank the BESSY II MX team for their assistance and cooperation. I also thank my thesis committee members: Prof. Jon Hughes, Prof. Ulrike Alexiev, Prof. Kai Thormann and Prof. Peter Friedhoff for spending their valuable time on my thesis.

I am very grateful for the support from all other members of the Plant Physiology institute (Peng-Yuan Chen, Silvia Trogu, Nicole Bazant, Anette Mündelein, Christina Lang, Sonja Graf, Tanja Gans and Roland Kürschner) for their valuable help.

Finally, I would like to thank my family, especially my wife, for their unconditional love and support.

10. Publications

Parts of this work were published in:

Nagano, S., **Guan, K.,** Shenkutie, S.M., Feiler, C., Weiss, M., Kraskov, A., Buhrke, D., Hildebrandt, P & Hughes, J. Structural insights into photoactivation and signalling in plant phytochromes. Nat. Plants 6, 581–588 (2020). https://doi.org/10.1038/s41477-020-0638-y

nature plants

ARTICLES

https://doi.org/10.1038/s41477-020-0638-y

Check for updates

Structural insights into photoactivation and signalling in plant phytochromes

Soshichiro Nagano[1], Kaoling Guan[1], Sintayehu Manaye Shenkutie[1], Christian Feiler[2], Manfred Weiss[2], Anastasia Kraskov[3], David Buhrke[3], Peter Hildebrandt[3] and Jon Hughes[1]✉

Plant phytochromes are red/far-red photochromic photoreceptors that act as master regulators of development, controlling the expression of thousands of genes. Here, we describe the crystal structures of four plant phytochrome sensory modules, three at about 2 Å resolution or better, including the first of an A-type phytochrome. Together with extensive spectral data, these structures provide detailed insight into the structure and function of plant phytochromes. In the Pr state, the substitution of phycocyanobilin and phytochromobilin cofactors has no structural effect, nor does the amino-terminal extension play a significant functional role. Our data suggest that the chromophore propionates and especially the phytochrome-specific domain tongue act differently in plant and prokaryotic phytochromes. We find that the photoproduct in period–ARNT–single-minded (PAS)–cGMP-specific phosphodiesterase–adenylyl cyclase–FhlA (GAF) bidomains might represent a novel intermediate between MetaRc and Pfr. We also discuss the possible role of a likely nuclear localization signal specific to and conserved in the phytochrome A lineage.

Canonical phytochromes detect light via their amino-terminal photosensory module comprising N-terminal extension (NTE), period–ARNT–single-minded (PAS), cGMP-specific phosphodiesterase–adenylyl cyclase–FhlA (GAF) and phytochrome-specific (PHY) domains with the help of a bilin chromophore. The red-light-absorbing Pr state can be photoconverted to the far-red-light-absorbing Pfr state, which itself can slowly revert in darkness or be photoconverted back to Pr. Photon absorption by the chromophore in Pr somehow leads to a $Z \rightarrow E$ isomerization followed by thermal reaction steps including protonation changes, a huge ~70 nm shift in the absorbance maximum and major structural changes in the protein, none of whose origins is more than vaguely understood.

Whereas prokaryotic representatives generally function as photon-de/activated histidine kinases signalling as Pr from a carboxy-terminal transmitter module, plant phytochromes (1) have lost their histidine kinase function, (2) have acquired a PAS repeat between the sensor and transmitter-like modules and (3) signal as Pfr. In darkness, the inactive Pr state is cytoplasmic, light-absorption-generating Pfr that is then translocated to the nucleus where it forms a signalling complex that binds, inactivates and finally leads to the destruction of the PIF transcription factors responsible for the typical morphology of dark-grown seedlings, thereby inducing photomorphogenesis. In prokaryotic phytochromes, the tongue—a long hairpin extension of the PHY domain that makes intimate contact with the GAF domain—undergoes radical refolding on photoconversion, possibly shifting the PHY domain and associated structures to initiate signalling[1,2]. In plant phytochromes, however, the PAS–GAF bidomains of the photosensory module can signal effectively[3–6], showing that in this case the PHY domain is dispensable. Nevertheless, solid-state NMR studies imply that tongue refolding also occurs during photoconversion in plant phytochromes[7].

In plants, A-type phytochrome (phyA) predominates in darkness, providing an extraordinarily sensitive photon sensor for seeds and seedlings. In light, phyA is all but lost, and B-type phytochrome (phyB) takes over as the principal sensory photoreceptor, in particular as a day length and shade sensor. The two types show less than 50% sequence identity[8], and their physiologies only partly overlap. To understand the mechanism of action in each case, it is necessary to elucidate the structure of Pr, the changes that occur during Pfr formation and how they are associated with the physical interactions necessary for Pfr translocation to and signalling in the nucleus. Here, we present four plant phytochrome Pr crystal structures, including the first for phyA, as well as extensive spectral data for Pr and the photoproduct, providing new insights to these ends.

Sample characterization

We designed various constructs of *Sorghum bicolor* (Sb) and *Glycine max* (Gm) phyA and phyB sensory modules for expression as phycocyanobilin (PCB) or phytochromobilin (PΦB) holoprotein adducts in *Escherichia coli* (Supplementary Table 1). Following purification (Supplementary Fig. 1), we characterized the Pr states and photoproducts by ultraviolet–visible (UV–vis) absorbance and circular dichroism (CD) spectroscopy (Fig. 1) as well as vibrational spectroscopy (Extended Data Fig. 1), as described later. The UV–vis data show the expected red/far-red photochromicity for the constructs including the PHY domain, although interestingly the Gm.phyA(PG) showed a blue-shifted Pfr-like photoproduct even without the PHY domain. We derived fairly uniform extinction coefficients of ~110 $mM^{-1}\,cm^{-1}$ for PCB adducts in the Pr state. In all cases, the negative CD signals for Pr were inverted in the case of the photoproducts, in contrast to the situation in bacteriophytochromes, as discussed in a previous study in the case of Cph1 (ref. [9]).

The Pr state

We crystallized and solved the three-dimensional structures of Sb.phyB(PG)–PCB, Sb.phyB(PG)–PΦB, Gm.phyB(PGP)–PCB and Gm.phyA(PG) as Pr (Protein Data Bank (PDB) entries 6TBY, 6TC5, 6TL4 and 6TC7 at resolutions of 1.8, 2.1, 2.9 and 2.1 Å, respectively;

[1]Institut für Pflanzenphysiologie, Justus-Liebig-Universität, Gießen, Germany. [2]BESSY II, Helmholtz-Zentrum Berlin für Materialien und Energie, Berlin, Germany. [3]Institut für Chemie, Sekr. PC14, Technische Universität, Berlin, Germany. ✉e-mail: jon.hughes@uni-giessen.de

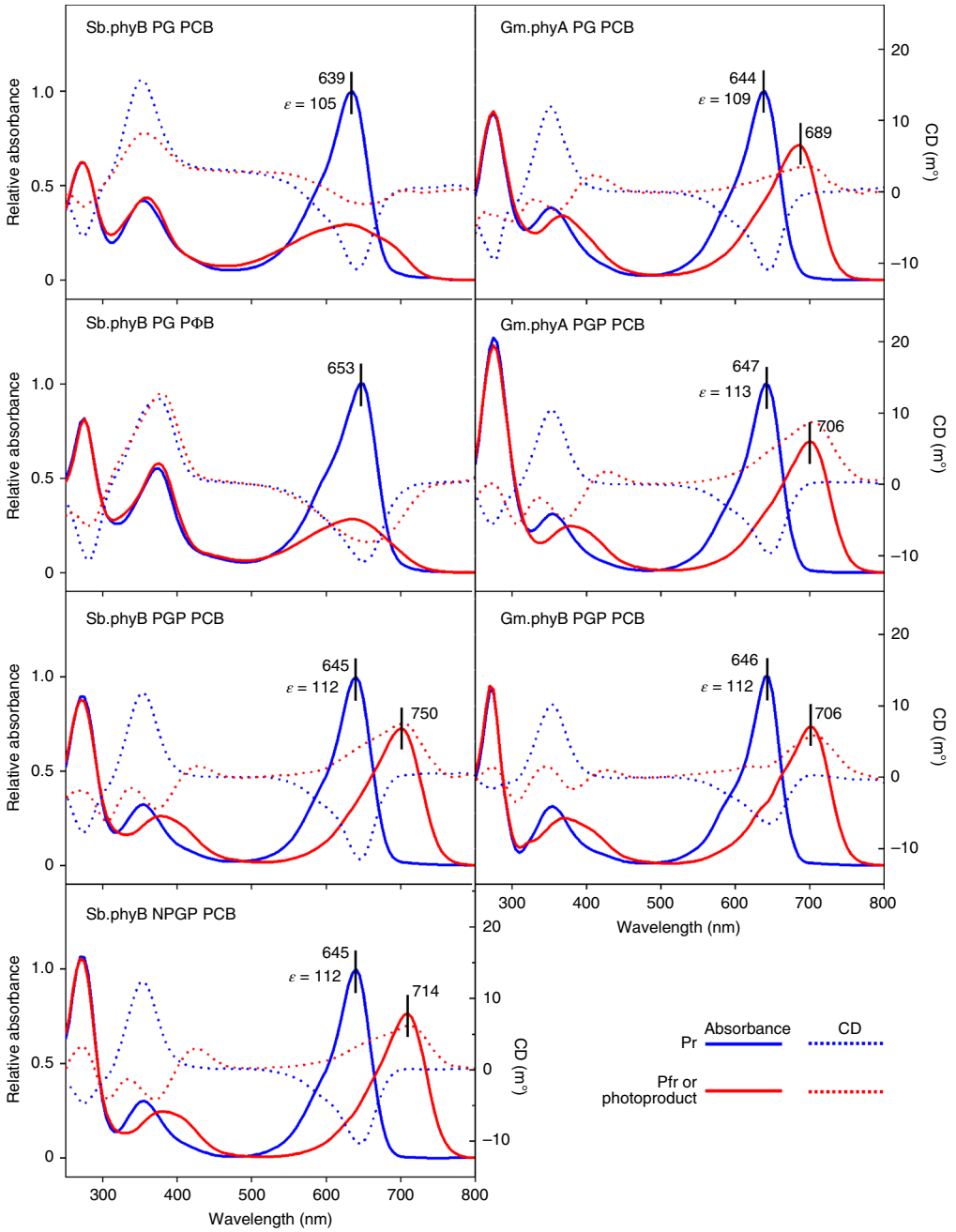


Fig. 1 | UV–vis absorbance and CD spectra of the seven plant phytochrome constructs investigated. Left, Sb.phyB. Right, Gm.phyA and Gm.phyB. λ_{max} for Pr and Pfr as well as the extinction coefficients for Pr are shown (nm and $mM^{-1}\,cm^{-1}$, respectively). The Pfr spectrum at full occupancy was derived by subtracting the appropriate proportion of Pr and rescaling to unity. Representative spectra based on at least two samples are shown. The spectra for each sample were measured several times. The CD spectra are based on five averaged scans with saturating actinic irradiation between each scan. ε, molar absorption coefficient.

see Methods and Supplementary Table 2). Particularly useful comparisons are with the structures of the cyanobacterial phytochrome Cph1 (refs. [10–12]) and *Arabidopsis* phyB (ref. [13]) (2VEA, 3ZQ5 and 4OUR at 2.2, 1.9 and 3.4 Å, respectively).

The Sb.phyB(PG)–PCB 1.8 Å structure (6TBY, Fig. 2) reveals details of the chromophore-binding lobe, including all waters, at near-atomic resolution. The overall root mean square (RMS) shift of the Cα atoms relative to the Cph1 structure 2VEA (ref. [12]) is 1.2 and 1.8 Å for the PAS and GAF domains, respectively. The principal structural difference relative to Cph1 is the A240–A262 region, which is shifted towards the chromophore and almost perfectly helical in phyB. As seen in prokaryotic phytochrome structures, the peptide chain forms a figure-eight knot, a total of 26 N-terminal residues upstream of I125 passing through a GAF domain loop. Notably, in this construct the first 14 of these residues are heterologous, including the 6His tag that is not resolved. Two plant-phytochrome-specific regions, S156–P169 and S394–A409 (the 150s and 380s loops), showed diffuse electron density and were not modelled. Regarding the chromophore (Fig. 2b,c), the tilt angles between the *A*–*B*, *B*–*C* and *C*–*D* rings are 18°, 8° and 49°, respectively, with the C15–C16 double bond showing a strong twist (dihedral angle of 21°). The ethyl group of the *D* ring

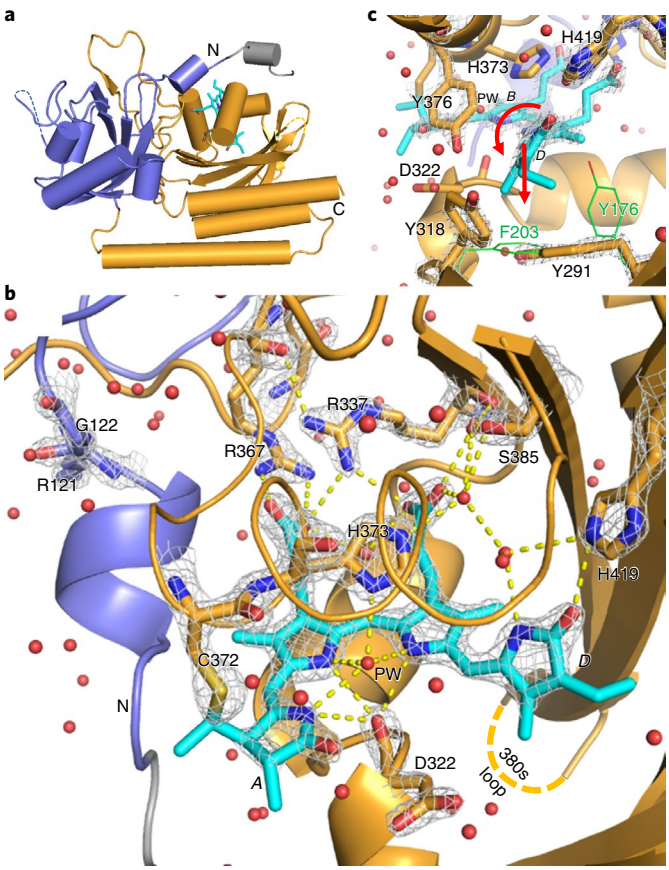


Fig. 2 | Sb.phyB(PG)–PCB 1.8 Å structure (PDB 6TBY). a, Overall view. **b**, Chromophore pocket from above with hydrogen bonds (dashed yellow lines). The helix in the foreground is shown as a ribbon to improve the visibility of the features behind. **c**, Chromophore *D*-ring environment showing the expected anticlockwise rotation and β-facial slump on photoconversion (red arrows). The positions of the Y318 and Y291 homologues F203 and Y176 in the bacteriophytochrome Pfr structure (5C5K) are shown (green). The heterologous N-terminal residues (grey), PAS (slate), GAF (gold), PCB (cyan), waters (red spheres) and electron density at 3 σ (grey mesh) are shown. The N and C termini, pyrrole water (PW) and rings as well as important amino acids are labelled. The missing 150s and 380s loops are shown as dashed lines.

is rotated 66° to the β-face, the staggered position probably being slightly favourable energetically. None of the rings is puckered. Similar to prokaryotic phytochromes, the N–H groups of rings *A*, *B* and *C* are hydrogen-bonded to both the main chain oxygen of D322 below and the pyrrole water above in an almost perfectly symmetrical pyramid. However, the carboxyl groups of both propionic side chains (Ps*B* and Ps*C*) point outwards instead of inwards as in prokaryotic phytochromes. Ps*B* forms a salt bridge to R367, whereas, unlike prokaryotic phytochromes, the paired R337 is directly hydrogen-bonded to both Ps*B* and Ps*C* (Fig. 2b). Despite its orientation, the latter is connected by hydrogen bonds to H372 and thence the N–H groups of rings *A*, *B* and *C* by bridging waters, as in prokaryotic phytochromes. Similarly, H419 is hydrogen-bonded to the *D*-ring carbonyl group, stabilizing the *Za* geometry typical of Pr. Other critical residues such as the tyrosine trio 291, 318 and 376 around the *D* ring (Fig. 2c) are positioned similarly to their homologues Y176, Y203 and Y263 in Cph1.

We addressed possible structural differences between Sb.phyB(PG)–PCB (described above) and the native PΦB adduct by solving its crystal structure to a resolution of 2.1 Å (6TC5). For this, we needed to create a new PΦB model (PDB cofactor O6E, Extended Data Fig. 2) because the 2VO cofactor used in the 4OUR structure[13] is wrong: most prominently, the C10 and C15 methine bridges are saturated in 2VO, a condition that would not only allow free rotation of both the *C* and *D* rings but also lead to an enormous hypsochromic absorbance shift. Relative to Sb.phyB(PG)–PCB, the overall RMS shift of the Cα atoms in the PΦB adduct is 0.50 Å, the only major differences being in the superficial R234–D236 region and near the unresolved loops (Extended Data Fig. 2). Ignoring such outliers, the RMS shift falls below 0.2 Å even when side chains are included. The *A*–*B*, *B*–*C* and *C*–*D* ring tilts are 14°, 5° and 46°, respectively, similar to those in the PCB adduct. On account of the double bond, the vinyl group is co-planar with the *D* ring, whereas the ethyl in the PCB adduct is able to rotate. However, none of the neighbouring protein atoms shows any change in position between the two adducts. These findings are in line with the resonance Raman (RR) spectra (Extended Data Fig. 1), in which the few differences can readily be attributed to the different ring *D* substituent[14,15]. Thus, any structural differences between the PCB and PΦB adducts in the Pr state lie below the resolution levels of our study.

We were eager to solve the structure of Sb.phyB(PGP)–PCB to see the precise effects of the interaction between the PHY domain tongue and the GAF domain, but none of the crystals obtained diffracted sufficiently well. We can, however, rule out that the PHY domain causes even subtle structural differences at least for the chromophore and its immediate protein interactions, as the RR spectra of the Sb.phyB(PG)–PCB and Sb.phyB(PGP)–PCB constructs are nearly identical (Extended Data Fig. 1). We succeeded, however, in solving the equivalent structure of Gm.phyB(PGP)–PCB to 2.9 Å (6TL4, Fig. 3). Not surprisingly, this structure generally resembles those of *Arabidopsis* phyB at 3.4 Å resolution (PDB 4OUR)[13] except for a slight shift in the position of the PHY domain and the fact that the protomers crystallize as a staggered complex rather than as the crystallographic 4OUR-A/B heterodimer (Fig. 3a). The PAS–GAF region is quite similar to that of 4OUR (with RMS Cα differences of 1.3 and 2.0 Å for the A and B subunits, respectively), although the region L219–Q229 is highly deviant in the 4OUR-B subunit (Fig. 3b). All but one of the five sections of the PHY domain missing in 4OUR-A or eight in 4OUR-B are resolved in our structure (Fig. 3c). As the protomers in our case are identical despite different crystal contacts, we expect the structure to include minimal crystallographic artefacts. Indeed, considering the evolutionary divergence, the Gm.phyB(PGP)–PCB structure is remarkably similar to that of Sb.phyB(PG)–PCB with an overall RMS shift of only 1.2 Å for the equivalent Cα atoms. The chromophore *A*–*B*, *B*–*C* and *C*–*D* ring tilts are 11°, 3° and 51°, respectively, also very similar to the smaller *Sorghum* construct. The RR spectra (Extended Data Fig. 1) show that the strongest band—originating from the overlap of the *C*–*D* and *A*–*B* stretching modes[16]—is at a somewhat lower frequency in Gm.phyB(PGP)–PCB (1,641 cm^{-1}) than in the *Sorghum* PG and PGP equivalents (1,646 cm^{-1}). The side chain position of the DIP motif aspartate below the chromophore is hardly changed, despite the availability of the arginine salt-bridge partner (PRxSF motif) when the tongue is present. Although one of its nitrogens is hydrogen-bonded to D303, the R575 finger is crooked so that the other nitrogen binds to the main-chain oxygen of G557 of the WGG motif of the tongue rather than to the aspartate. Indeed, although the structure at the base of the tongue is conserved, the tip of the tongue in phyB is quite different from that of Cph1 and bacteriophytochromes: instead of hydrogen-bonding to form an antiparallel sheet with the other side of the hairpin to seal the chromophore cavity, the incoming strand K560–H573 of phyB takes a quite different course to form a lasso that leaves the chromophore ring *A* accessible to the medium via an extended tunnel (Fig. 3d). These differences with respect to prokaryotic phytochromes are also apparent in the 4OUR phyB structures.

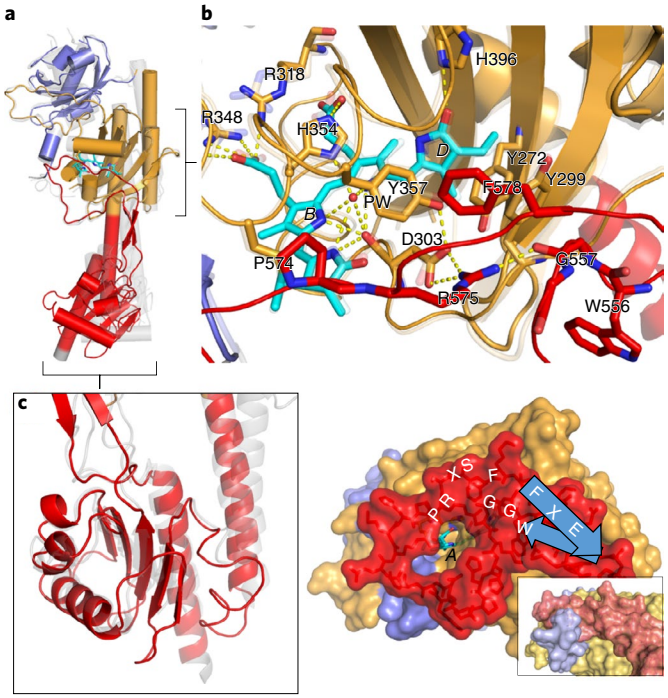


Fig. 3 | Gm.phyB(PGP)–PCB 2.9 Å structure (PDB 6TL4). a, Overall view with the 4OUR-A and -B subunits (grey) superimposed. **b**, PAS–GAF bidomain with the Sb.phyB(PG)–PCB (pale) homologue superimposed. Chromophore rings *B* and *D*, crucial residues with hydrogen bonds (yellow dashed lines) and the pyrrole water are labelled. **c**, PHY domain with the 4OUR-A and -B subunits (grey) superimposed. **d**, Surface rendition of the tongue providing access to the chromophore (ring *A*, labelled). The conserved WGG, PRxSF and W/FxE motifs and the antiparallel sheets are labelled. The inset shows the equivalent view of Cph1 (PDB 2VEA). The region K560–R571 in phyB forms a loop, whereas P454–L469 in Cph1 makes close contact with the returning strand (which is folded similarly in both). Note that the NTE is missing in the plant structure. The PAS (slate), GAF (gold), PHY (red) and PCB (cyan) structures are shown.

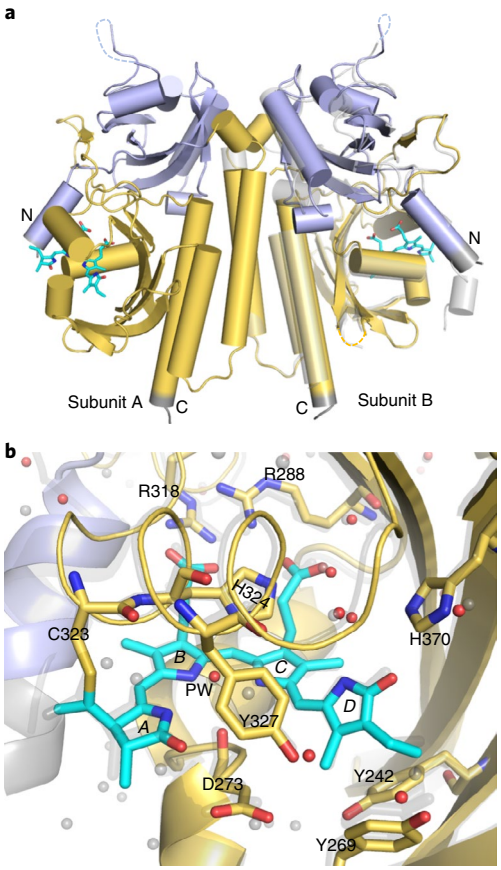


Fig. 4 | Gm.phyA(PG)–PCB 2.1 Å structure (PDB 6TC7). a, Overall view of dimer subunits A and B. **b**, Detail of the chromophore pocket. The Sb.phyB(PG)–PCB (grey) homologue is superimposed on subunit B in both cases. The features shown include PAS (slate), GAF (gold), PCB (cyan) and waters (red spheres). The missing 150s and 380s loops are shown as dashed lines.

We also crystallized Gm.phyA(PG)–PCB to 2.1 Å (6TC7, Fig. 4 and Extended Data Fig. 3), providing the first three-dimensional structure for an A-type plant phytochrome. The crystallographic protomer (Fig. 4a) is a head-to-head dimer, with slight RMS Cα atom position differences of 0.5 Å between the A and B subunits, arising in particular from the superficial T175–G176 and E304–D309 loops. The chromophore (Fig. 4b) position and geometry are almost identical in the two monomers, with the ring tilts showing 17°, 6° and 46° for subunit A and 15°, 8° and 43° for subunit B, respectively. With respect to Sb.phyB(PG)–PCB, the Cα atom positions show an RMS difference of ~1.2 Å for both subunits, although several helices superimpose almost perfectly (Fig. 4a,b). The chromophore region of Gm.phyA(PG)–PCB is closely similar to that in phyB (Fig. 4b). Accordingly, compared with Sb.phyB(PGP)–PCB, the RR spectra show only small differences of the *C*–*D* methine bridge C–C stretching and N–H in-plane bending modes that point to subtle differences in the *C*–*D* twist angle and hydrogen-bonding interactions of the ring *B* and *C* N–H groups (Extended Data Fig. 1; the hydrogen-bond network around the chromophore is shown in Extended Data Fig. 4). Conversely, the regions L145–l153 and V185–M192 show conserved sequence differences between A- and B-type phytochromes, and indeed the folds are dissimilar. The T108–A122 150s and N344–K361 380s loop sequences are also phyA/phyB specific but could be only partially modelled because of their high mobility, as reflected by large B factors. The 60 atoms near the chromophore in the quantum-mechanics/molecular-mechanics (QM/MM)-optimized solid-state NMR model for oat phyA3 (ref. [7]) have an RMS shift of only 0.6 Å from their positions in the crystal structure, the main deviant being H370 on account of its uncertain protonation (Supplementary Fig. 2).

The four structures together with the spectroscopic data reveal novel details of plant phyA and phyB in the Pr state. First, we see very similar structures for the various homologous phyB constructs from *Sorghum*, *Glycine* and at least the 4OUR-A *Arabidopsis* subunit[13]. The close similarity is remarkable in view of the distant evolutionary relationships. Moreover, the novel phyA structure is closely similar to that of phyB, especially in the chromophore region. Second, they allow us to clarify various structural and functional aspects of the cofactor. Substituting PCB with PΦB in phyB has a minimal effect on the structure (Extended Data Fig. 2). The C18–C18^1=C18^2 group is coplanar with the *D* ring, leading to an extension of the delocalized π-electron system that is responsible for a 19 cm^{-1} downshift of the *C*–*D* stretching mode at 1,646 cm^{-1} (Extended Data Fig. 1) and the ~14 nm maximum wavelength (λ_{max}) red shifts in the absorbance and CD spectra (Fig. 1). Although PΦB is the native chromophore in plant phytochromes, PCB adducts function similarly in planta but with the corresponding 10–15 nm blue-shifted action maxima[17,18]. It is also notable that the *C*–*D* ring tilts in all our structures are close to 50°, much larger than in Cph1 but similar to that predicted for oat phyA3 on the basis of RR and solid-state NMR data[7,16,19] (Supplementary Fig. 2). This angle

is significant because, assuming a $\cos^2$-dependence of the π-bond overlap, the coupling of the *D* ring to the rest of the system would start to decline rapidly beyond 50°: the *D* ring is, as it were, poised. Third, as in the case of prokaryotic phytochromes, the PAS–GAF bidomain structure is scarcely affected by the presence of the tongue, despite interactions that seem to be functionally important. Intriguingly, the tongue in Gm.phyB(PGP)–PCB is broad and pierced, conspicuously providing solvent access to the chromophore rather than sealing the pocket as in prokaryotic phytochromes (Fig. 3d). Fourth, although we failed to grow well-diffracting crystals of a sensory module complete with its likely mobile NTE, the spectroscopic data imply that it too has a minimal effect at least on the chromophore and its surroundings in the Pr state. As in other known phytochrome structures, the peptide chain is knotted in all four constructs we crystallized, the N-terminal residues passing through an extended GAF-domain loop. This is particularly intriguing in the case of Sb.phyB(PG)–PCB, in which most of these residues are heterologous: it is quite unclear what features induce knot formation.

In conjunction with the UV–vis absorption, CD and RR spectroscopic findings, we conclude that the high-resolution PAS–GAF Pr structures we present are probably valid in detail for the full-length photoreceptors in planta. The extensive structural similarities seem paradoxical because the Pr state of plant phytochromes has no known physiological function and thus would not itself be under selection pressure. Its structural conservation relates not only to its ability to form functional Pfr but also, as pointed out in review, to its preventing physiological signals being generated in darkness. The similar UV–vis absorbance spectra and extinction coefficients of Pr (Fig. 1) suggest that the primary $S_0 \rightarrow S_1$ photochemistry of the chromophore is conserved, with the CD and RR spectra (Fig. 1 and Extended Data Fig. 1) confirming that the chromophores have similar geometries and interactions with the pocket.

Photoproducts

The Pfr photoproducts of the *Sorghum* and *Glycine* phyB(PGP)–PCB constructs show almost identical absorbance and RR spectra (Fig. 1 and Extended Data Fig. 1) characteristic of a fully protonated chromophore in the *ZZEssa* configuration[20]. The infrared (IR) difference spectra relative to Pr are very similar too, with amide I band changes reflecting a sheet-to-helix transition, supporting the notion of tongue refolding during photoconversion (Extended Data Fig. 1). We note a signal pair at 1,711(+)/1,730(−) cm^{-1} in the NPGP and PGP constructs corresponding to a change of the hydrogen-bonding interactions of a carboxyl side chain on photoconversion, possibly associated with the DIP–PRxSF salt bridge to the tongue in the Pr state. The additional NTE in NPGP constructs leads to a slightly higher helicity increase than in the PGP constructs, either because tongue refolding is enhanced by the NTE or because the NTE itself adopts a more helical structure, as was proposed[21]. The NTE has a small but detectable effect on the RR spectrum of the chromophore in Pfr, specifically the *C*–*D* stretching and hydrogen-out-of-plane modes localized at the *C*–*D* methine bridge. In contrast to bacteriophytochromes, the CD signal switches from negative in Pr to positive values for Pfr in both phyA and phyB (Fig. 1)—as reported earlier for Cph1—indicating that the chirality inverts on photoconversion[9,22]. If the *D* ring rotates anticlockwise, a clash between the *C*- and *D*-ring methyl groups might force the *D* ring to slump to the β-face of the chromophore in Pfr (ref. [9]), as supported by solid-state NMR data for both Cph1 and oat phyA3[7,11,19]. Unexpectedly, however, recent time-resolved X-ray free-electron laser (XFEL) work[23] implies that the *D* ring rotates anticlockwise in bacteriophytochromes too, raising the question as to why the ring remains α-facial in that case.

The photoproducts of the PG constructs provide interesting insights into the role of the PHY domain and the tongue in particular. The associated IR difference spectra indicate small, light-induced changes above 1,700 cm^{-1} due to the different hydrogen bonding of carboxyl amino acid side chains and the carbonyl group of ring *D* (Extended Data Fig. 1). In contrast to the PGP constructs, no signals attributable to a secondary structure transition are apparent, consistent with the loss of the tongue. By RR and denaturation in acidic urea (Extended Data Fig. 1 and Supplementary Fig. 3), we show for all constructs not only that $Z \rightarrow E$ photoisomerization of the *D* ring nevertheless occurs but also that the chromophore is protonated. In the case of phyB, this is quite remarkable, because the photoproducts of the PG constructs display a bleached absorption (Fig. 1). Such bleaching is common in prokaryotic phytochromes and plant phyB when the PHY domain is deleted or the tongue damaged,[12,13,24] yet it has generally been ascribed to chromophore deprotonation. Similarly, the bleached character of MetaRc (the spectral intermediate prior to Pfr known to date[25–28]) is generally considered to result from deprotonation. It seems that this idea must be revised. However, we were also surprised to observe that, in contrast, the photoproduct of the phyA(PG) construct shows a strong Pfr-like absorbance, albeit with a 15 nm blue shift (Fig. 1). We know of only two other phytochromes that show strong red/far-red photochromicity in the absence of the PHY domain, namely SyB.Cph2 (whose photoproduct also shows a 15 nm hypsochromic shift relative to the wild type)[29,30] and a triple mutant of Dr.BphP (with no spectral shift)[31]. It therefore seems that neither the tongue nor chromophore protonation has a simple causal connection to the strong absorbance of the photoproduct. Whereas a protonated *ZZEssa* chromophore is seen in the earlier MetaRa intermediate, its RR spectrum differs radically from that of the PG photoproduct[26] (Extended Data Fig. 1). However, differences are apparent between the RR spectra of the PG and PGP constructs. In the case of phyB, the C–C stretching and hydrogen-out-of-plane modes of the *C*–*D* methine bridge—which give rise to particularly strong and sharp bands in Pfr—are substantially broadened in the PG photoproducts: the doublet fine structure of the hydrogen-out-of-plane mode at 805 and 813 cm^{-1} that reflects conformational substates of Pfr is lost in the PG constructs, implying higher *D*-ring mobility. Since the frequencies of the respective modes are also different for the PG and PGP constructs, we conclude that the PHY domain is required for the final structural adjustment of the chromophore to form Pfr and that the PG photoproduct might represent an intermediate between MetaRc and Pfr. We thus suggest calling it MetaRd. Although the RR spectrum of the phyA(PG) photoproduct resembles that of phyB(PG), the distinctly narrower band widths point to a more rigid fixation of the chromophore in the protein environment, perhaps explaining the strong, Pfr-like absorbance properties. In the case of most other phytochromes, the tongue might be necessary to fix the chromophore in a similar way.

Regardless of the subtle structural differences between the Pfr-like states of phyA(PG) and phyB(PG), the present data demonstrate that the refolding of the tongue is not required for chromophore reprotonation. Under the likely assumption that MetaRd directly precedes Pfr in the photoinduced reaction cascade of the intact photosensor, we consider that proton transfer to the chromophore in MetaRc might trigger tongue refolding. Time-resolved IR difference data for the bacteriophytochrome Dr.BphP suggest that tongue refolding happens in the last step of the phototransformation, although the salt bridge to the tongue probably dissolves, and other backbone refolding occurs much earlier in $Pr \rightarrow Pfr$ photoconversion[32]. In the recent bacteriophytochrome XFEL study, interactions between the chromophore and the tongue begin to change within picoseconds of photoactivation[23]. It was proposed that sheet-to-helix refolding of the tongue plays a central role in signalling by shifting the relative positions of the PHY domains and associated transmitter module in the functional dimer of prokaryotic phytochromes[1,2]. This is far from clear, however, as the expected

PHY domain shifts were not apparent in full-length Agp1[33]. Moreover, once in the nucleus, phyA and phyB PAS–GAF dimers can supress PIF3 function and thereby induce photomorphogenesis autonomously[3,4,6,34], showing that neither the tongue nor the PHY domain nor indeed the rest of the C-terminal region (the PAS repeat and the transmitter-like module) is essential for signalling, and that MetaRd is itself a signalling state. From work on *Arabidopsis* phyB, R110, G111, G363 and R352 are thought to be involved in Pfr binding to PIF3[35,36]. The C-terminal region, however, can mediate the destruction of PIF3 and induce photomorphogenesis autonomously (and thus independently of light)[5], suggesting that its action in the intact photoreceptor is somehow gated by the photosensory module, perhaps via the PHY domain.

Even the light-activated events within the photosensory module are clearly different in plant phytochromes from those in bacteriophytochromes. The salt bridge formed between the aspartate and arginine of the DIP and PRxSF motifs, respectively, is a conspicuous feature of the tongue interaction with the GAF domain in prokaryotic phytochromes, yet in phyB the arginine side chain turns away to hydrogen-bond also to the main-chain oxygen of G557 of the WGG motif on the other side of the tongue (Fig. 3b). Interestingly, mutation of that residue (WGG → WEG) in *Arabidopsis* phyB stabilizes Pfr substantially, as do PRxSF → PAxSF or PRxSW mutations[13,37]. Whether FxE → WxE does too, as in the *Deinococcus* bacteriophytochrome[38], is not known. Similarly, on the other side of the chromophore, the state-dependent Ps*B* partner swap between paired arginines in prokaryotic phytochromes is a less convincing basis for Pfr-dependent PIF3 binding at the PAS–GAF cleft in phyB[36] because in that case—and in phyA too—both arginines interact with Ps*B* and Ps*C* in the Pr state (Figs. 2b and 4b). A serine residue replacing the valine in the VCHH motif conserved in plant phytochromes (V417 in *Sorghum* phyB) stabilizes Pfr, seemingly by forming an alternative hydrogen bond with Ps*C*[13,39].

Solid-state NMR studies of oat phyA3 imply that the NTE is within 5 Å of ring *A* in Pfr but not Pr (ref. [7]), suggesting that the NTE might block solvent access to the chromophore in Pfr. This would contrast with the behaviour of bacteriophytochromes in which the tongue closes and opens access to the chromophore pocket in Pr and Pfr, respectively. Lagarias's suggestion[7] that photoconversion might involve waters moving in and out of the chromophore pocket is interesting in this context, noting that a recent XFEL study[23] reported the disappearance of the pyrrole water within a picosecond of photoactivation, and thus simultaneously with or even preceding the *D*-ring photoflip. Solvent access might alternatively be associated with proton release or reuptake during Pr → Pfr photoconversion[40,41]. The NTE is also significant in Pfr → Pr thermal reversion[20,42–44], perhaps regulated in phyB by the state-dependent phosphorylation and binding of the ARR4 response regulator[45].

Finally, the 380s loop in phyA might be of particular interest because it includes the strongly basic sequence KRKR (residues 359–362), representing a putative Class I nuclear localization signal (NLS)[46]. Positioned at the surface of the GAF domain, the consensus K(R/K)K(R/K) is conserved in and exclusive to the phyA lineage (Extended Data Fig. 5)[47]. It is ironic that this sequence has escaped notice despite high scores from various algorithms including NLS-mapper, PSORT, Euk-mPLoc 2.0 and INSP (nls-mapper.iab.keio.ac.jp, wolfpsort.hgc.jp and www.csbio.sjtu.edu.cn/bioinf, respectively), whereas a hypothetical NLS in B-type phytochromes[48] was sought in vain. The function of the K(R/K)K(R/K) motif in phyA is quite unclear, however. As Pr is cytoplasmic in all wild-type phytochromes, it is clearly ineffective as an NLS in that state. Native phyA Pfr, on the other hand, is translocated efficiently into the nucleus—but the process is not autonomous, requiring the function of FHY1 and its NLS in particular[49]. Although we might therefore conclude that the putative NLS in phyA has no physiological function, Oka et al. showed that when the phyA GAF domain is introduced into phyB, the chimeric holoproteins localize to the nucleus in darkness[50]. Moreover, the effect is lost if the phyA C-terminal region is present too. We propose, therefore, that the putative NLS at the surface of the phyA GAF domain is held inactive by the phyA C-terminal region in Pr and that the suppression is released in Pfr. It nevertheless remains unclear how this integrates with FHY1 to control Pfr-dependent nuclear translocation. We are investigating this intriguing matter in planta using transgenic *Arabidopsis* mutants.

Discussion

In this study, we show that different A- and B-type plant phytochromes show close spectral and structural similarities in the Pr state. Subtle but perhaps important differences in relation to prokaryotic phytochromes include salt bridges (1) between the chromophore propionates and a conserved pair of Arg residues and (2) between conserved Asp and Arg residues associated with the chromophore and the tongue, respectively. The former has been correlated with the Pfr-dependent binding of phyB to PIF3, a crucial event in photomorphogenesis, but our data suggest that the switch might be different in plant phytochromes. The role of the tongue too seems to differ from the proposal that sheet-to-helix refolding leads to PHY domain shifts and de/activation of the histidine kinase output module, especially as plant phytochromes can signal in the absence of the PHY domain and are not histidine kinases. Moreover, if the tongue is missing, the final photoconversion state, Pfr, is either not reached (phyB) or significantly blue-shifted (phyA), although the chromophore is reprotonated in both cases; we therefore suspect that this proton transfer might be necessary for tongue refolding and thus bona fide Pfr formation. Interestingly, in phyB at least, the tongue provides a tunnel between the chromophore and the medium in Pr, in contrast to prokaryotic phytochromes, where the tongue seals the chromophore pocket. It is surprising that Pfr-like absorbance properties require neither the tongue nor the protonation of the chromophore, the crucial factor perhaps being chromophore mobility. Solid-state NMR indeed revealed that Pfr is much more rigid than Pr, which seems to exist in at least two substates[7,11,19,51,52]. The resonances in Pfr that were missing in Pr were interpreted similarly in a recent liquid NMR study of Dr.BphP[53]. A further remarkable discovery is that the tongue in a Y263F mutant of Dr.BphP in the Pr state is helical (Pfr-like), indicating that the photochromic state can be uncoupled from tongue refolding[54]. The homologous mutant in Cph1 did not show this, however[10].

We also noticed a covert NLS conserved in and specific to the phyA lineage. As phytochromes are normally cytoplasmic in darkness, the NLS must be inactive in Pr, but it might have an intriguing role in phyA Pfr, whose translocation to the nucleus is not autonomous but requires the FHY1 transport facilitator. This is only one of the major physiological differences between phyA and phyB. For example, once the Pfr state is formed in light, phyA is rapidly degraded and lost from the cell, whereas phyB is relatively stable. Moreover, the interaction site of PIF3 with Pfr is thought to be different in the two types[55,56]. Given the similarities that we show between the Pr states of phyA and phyB, it will be of great interest to see how their structures and functions differ as Pfr.

Methods

Sample production and spectroscopy. Various domain constructs of *S. bicolor* and *G. max* phyA and phyB sensory modules together with His6 tags (Fig. 1 and Supplementary Table 1) were overproduced using pCDFDuet in *E. coli* engineered to provide bilin cofactors in vivo essentially as described in ref. [57]. The plasmids used are available and will be provided by the authors on request. The complete protein sequences in each case are given in Supplementary Table 1. The holoproteins were purified by Ni^{2+}-affinity and size-exclusion chromatography and characterized by UV–vis absorbance and CD spectroscopy (Agilent 8453, Hitachi U-3010 and Jasco J-715) in their lowest-energy Pr state following saturating far-red (710–750 nm) and red (645–665 nm or, in the case of PΦB adducts, 655–675 nm) irradiation (Fig. 1) as described in ref. [10]. Chromophore configurations around the C15–C16 double bond in holophytochromes following red or far-red irradiation

were probed by denaturation in acidic urea, adding one volume of sample to five volumes of 8 M urea with 5% (v/v) of 36% (w/v) HCl[30,58–60]. This procedure also allowed us to estimate the molar extinction coefficients of the native holoproteins at λ_{max} based on the absorbance change associated with denaturation and protonation and the molar extinction coefficient of free PCB in acid urea (37.6 mM^{-1} cm^{-1} at 680 nm) as described in ref. [61]. The vibrational spectra were measured essentially as described in ref. [62].

Crystallography. All steps were carried out in dim blue-green (490 nm) safelight or using infrared visualization equipment[63].

Sb.phyB(PG)–PCB and –PΦB: 2 µl of Pr holoprotein (at 20 mg ml^{-1} in 50 mM NaCl, 20 mM HEPES pH 7.8, 1 mM EDTA and 1 mM β-mercaptoethanol) was added to 2 µl of precipitant (9% (w/v) PEG4000, 0.1 M Tris/HCl pH 8.5 and 0.5 M NaCl) with n-Octyl-β-D-glucopyranoside (0.25% w/v final concentration) as a hanging drop and equilibrated at 10 °C with 500 µl of precipitant for 1 week in darkness. The crystals were cryoprotected in a solution consisting of reservoir (70% v/v) and sucrose (70% w/v) solutions.

Gm.phyB(PGP)–PCB: 1 µl of holoprotein (at 15 mg ml^{-1} in 20 mM HEPES pH 7.8, 100 mM NaCl, 1 mM EDTA, 1 mM β-mercaptoethanol and 5% v/v glycerol) was added to 1 µl of precipitant (200 mM MES pH 6.5, 200 mM KSCN, 200 mM $LiCl_2$, 1% w/v γ-PGA and 3% w/v PEG3350) as a hanging drop and equilibrated at 10 °C with 500 µl of precipitant for 1 week in darkness. The crystals were cryoprotected in reservoir solution with 25% (v/v) CryoMix 1 (12.5 % v/v diethylene glycol, 12.5 % v/v MPD, 37.5 % v/v 1,2-propanediol and 12.5 % v/v dimethyl sulfoxide (Molecular Dimensions)).

Gm.phyA(PG)–PCB: 400 nl of holoprotein (at 15 mg ml^{-1} in 20 mM HEPES pH 7.8, 200 mM NaCl, 1 mM EDTA, 1 mM β-mercaptoethanol and 5% v/v glycerol) was added to 400 nl of precipitant (0.1 M amino acids (1 M amino acids: 0.2 M DL-glutamic acid monohydrate, 0.2 M DL-alanine, 0.2 M glycine, 0.2 M DL-lysine monohydrochloride and 0.2 M DL-serine), 0.1 M buffer system 1 pH 6.5 (1 M buffer system 1 pH 6.5: 0.5 M imidazole and 0.5 M MES monohydrate (acid)), 30 % v/v precipitant mix 3 (40% v/v glycerol and 20% w/v PEG 4000)) as a sitting drop and equilibrated at 10 °C with 60 µl of precipitant for 2 weeks in darkness. The crystals were cryoprotected with the mother liquor.

The cryoprotected crystals were vitrified and stored in liquid nitrogen before data collection at 100 K in N_2 cryostreams. The datasets for Sb.phyB(PG)–PCB, Sb.phyB(PG)–PΦB and Gm.phyA(PG)–PCB crystals were collected at 0.9184 Å at beamline 14.1 at BESSY II and for Gm.phyB(PGP)–PCB at 0.9677 Å at beamline ID30A-3 at ESRF. The data were processed using XDS (build 20190315)[64] and scaled using AIMLESS (0.7.4)[65] from the CCP4 package (7.0.072)[66]. The crystallographic phases were determined by molecular replacement using PHASER (5.8.0257)[67] and the 2VEA Cph1 structure[12]. The structures were refined by iterative rounds of manual modelling using COOT (Wincoot 0.8.9.2)[68] and automatic refinement with REFMAC5 (5.8.0257)[69] from CCP4 (7.0.072) and PHENIX (1.17-3644)[70] and validated using MolProbity (v.4.5)[71]. Refinement yielded models with the following percentages of residues in Ramachandran favoured or outlier regions: Sb.phyB(PG)–PCB (96.5, 0.0), Sb.phyB(PG)–PΦB (94.4, 0.0), Gm.phyB(PGP)–PCB (91.2, 0.4) and Gm.phyA(PG)–PCB (96.2, 0.6). The summary statistics are given in Supplementary Table 2. The molecular images were created using PyMol 1.7 (Schrödinger): the atom position differences were calculated using SUPER without excluding outliers, and the electron density maps exported from COOT were visualized using ISOMESH. The cofactor ring tilts were calculated using OLEX2 (1.2.10)[72] software.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

The three-dimensional structural data that support the findings of this study have been deposited in wwPDB with the accession codes 6TBY, 6TC5, 6TL4 and 6TC7. The authors declare that all other data supporting the findings of this study are available within the paper and its Supplementary Information files.

Received: 17 December 2019; Accepted: 16 March 2020;
Published online: 4 May 2020

References

1. Anders, K., Daminelli-Widany, G., Mroginski, M. A., von Stetten, D. & Essen, L. O. Structure of the cyanobacterial phytochrome 2 photosensor implies a tryptophan switch for phytochrome signaling. *J. Biol. Chem.* **288**, 35714–35725 (2013).
2. Takala, H. et al. Signal amplification and transduction in phytochrome photosensors. *Nature* **509**, 245–248 (2014).
3. Oka, Y. et al. Functional analysis of a 450-amino acid N-terminal fragment of phytochrome B in *Arabidopsis*. *Plant Cell* **16**, 2104–2116 (2004).
4. Matsushita, T., Mochizuki, N. & Nagatani, A. Dimers of the N-terminal domain of phytochrome B are functional in the nucleus. *Nature* **424**, 571–574 (2003).
5. Qiu, Y. et al. Mechanism of early light signaling by the carboxy-terminal output module of *Arabidopsis* phytochrome B. *Nat. Commun.* **8**, 1905 (2017).
6. Viczian, A. et al. A short amino-terminal part of *Arabidopsis* phytochrome A induces constitutive photomorphogenic response. *Mol. Plant* **5**, 629–641 (2012).
7. Song, C. et al. 3D structures of plant phytochrome A as Pr and Pfr from solid-state NMR: implications for molecular function. *Front. Plant Sci.* **9**, 498 (2018).
8. Sharrock, R. A. & Quail, P. H. Novel phytochrome sequences in *Arabidopsis thaliana*: structure, evolution, and differential expression of a plant regulatory photoreceptor family. *Genes Dev.* **3**, 1745–1757 (1989).
9. Rockwell, N. C., Shang, L., Martin, S. S. & Lagarias, J. C. Distinct classes of red/far-red photochemistry within the phytochrome superfamily. *Proc. Natl Acad. Sci. USA* **106**, 6123–6127 (2009).
10. Mailliet, J. et al. Spectroscopy and a high-resolution crystal structure of Tyr263 mutants of cyanobacterial phytochrome Cph1. *J. Mol. Biol.* **413**, 115–127 (2011).
11. Song, C. et al. Two ground state isoforms and a chromophore D-ring photoflip triggering extensive intramolecular changes in a canonical phytochrome. *Proc. Natl Acad. Sci. USA* **108**, 3842–3847 (2011).
12. Essen, L.-O., Mailliet, J. & Hughes, J. The structure of a complete phytochrome sensory module in the Pr ground state. *Proc. Natl Acad. Sci. USA* **105**, 14709–14714 (2008).
13. Burgie, E. S., Bussell, A. N., Walker, J. M., Dubiel, K. & Vierstra, R. D. Crystal structure of the photosensing module from a red/far-red light-absorbing plant phytochrome. *Proc. Natl Acad. Sci. USA* **111**, 10179–10184 (2014).
14. Kneip, C. et al. Effect of chromophore exchange on the resonance Raman spectra of recombinant phytochromes. *FEBS* **414**, 23–26 (1997).
15. Remberg, A. et al. Raman spectroscopic and light-induced-kinetic characterization of a recombinant phytochrome of the cyanobacterium *Synechocystis*. *Biochemistry* **36**, 13389–13395 (1997).
16. Mroginski, M. A. et al. Structure of the chromophore binding pocket in the Pr state of plant phytochrome phyA. *J. Phys. Chem. B* **115**, 1220–1231 (2011).
17. Wahleithner, J. A., Li, L. M. & Lagarias, J. C. Expression and assembly of spectrally active recombinant holophytochrome. *Proc. Natl Acad. Sci. USA* **88**, 10387–10391 (1991).
18. Elich, T. D. & Lagarias, J. C. Formation of a photoreversible phycocyanobilin-apophytochrome adduct in vitro. *J. Biol. Chem.* **264**, 12902–12908 (1989).
19. Song, C., Essen, L. O., Gärtner, W., Hughes, J. & Matysik, J. Solid-state NMR spectroscopic study of chromophore–protein interactions in the Pr ground state of plant phytochrome A. *Mol. Plant* **5**, 698–715 (2012).
20. Velazquez Escobar, F. et al. Structural communication between the chromophore-binding pocket and the N-terminal extension in plant phytochrome phyB. *FEBS Lett.* **591**, 1258–1265 (2017).
21. Parker, W., Partis, M. & Song, P. S. N-terminal domain of *Avena* phytochrome: interactions with sodium dodecyl sulfate micelles and N-terminal chain truncated phytochrome. *Biochemistry* **31**, 9413–9420 (1992).
22. Litts, J. C., Kelly, J. M. & Lagarias, J. C. Structure–function studies on phytochrome: preliminary characterization of highly purified phytochrome from *Avena sativa* enriched in the 124-kilodalton species. *J. Biol. Chem.* **258**, 11025–11031 (1983).
23. Claesson, E. et al. The primary structural photoresponse of phytochrome proteins captured by a femtosecond X-ray laser. *eLife* **9**, e53514 (2020).
24. Wagner, J. R., Brunzelle, J. S., Forest, K. T. & Vierstra, R. D. A light-sensing knot revealed by the structure of the chromophore-binding domain of phytochrome. *Nature* **438**, 325–331 (2005).
25. Eilfeld, P. & Rüdiger, W. Absorption spectra of phytochrome intermediates. *Z. Naturforsch.* **40c**, 109–114 (1985).
26. Kneip, C. et al. Protonation state and structural changes of the tetrapyrrole chromophore during the Pr→Pfr phototransformation of phytochrome: a resonance Raman spectroscopic study. *Biochemistry* **38**, 15185–15192 (1999).
27. Mizutani, Y., Tokutomi, S. & Kitagawa, T. Resonance Raman spectra of the intermediates in phototransformation of large phytochrome: deprotonation of the chromophore in the bleached intermediate. *Biochemistry* **33**, 153–158 (1994).
28. Mroginski, M. A. et al. Elucidating photoinduced structural changes in phytochromes by the combined application of resonance Raman spectroscopy and theoretical methods. *J. Mol. Struct.* **993**, 15–25 (2011).
29. Ulijasz, A. T. et al. Characterization of two thermostable cyanobacterial phytochromes reveals global movements in the chromophore-binding domain during photoconversion. *J. Biol. Chem.* **283**, 21251–21266 (2008).
30. Song, C. et al. The D-ring, not the A-ring, rotates in *Synechococcus* OS-B' phytochrome. *J. Biol. Chem.* **289**, 2552–2562 (2014).
31. Auldridge, M. E., Satyshur, K. A., Anstrom, D. M. & Forest, K. T. Structure-guided engineering enhances a phytochrome-based infrared fluorescent protein. *J. Biol. Chem.* **287**, 7000–7009 (2012).

32. Ihalainen, J. A. et al. Chromophore–protein interplay during the phytochrome photocycle revealed by step-scan FTIR spectroscopy. *J. Am. Chem. Soc.* **140**, 12396–12404 (2018).
33. Kacprzak, S. et al. Intersubunit distances in full-length, dimeric, bacterial phytochrome Agp1, as measured by pulsed electron-electron double resonance (PELDOR) between different spin label positions, remain unchanged upon photoconversion. *J. Biol. Chem.* **292**, 7598–7606 (2017).
34. Park, E. et al. Phytochrome B inhibits binding of phytochrome-interacting factors to their target promoters. *Plant J.* **72**, 537–546 (2012).
35. Oka, Y., Matsushita, T., Mochizuki, N., Quail, P. H. & Nagatani, A. Mutant screen distinguishes between residues necessary for light-signal perception and signal transfer by phytochrome B. *PLoS Genet.* **4**, e1000158 (2008).
36. Kikis, E. A., Oka, Y., Hudson, M. E., Nagatani, A. & Quail, P. H. Residues clustered in the light-sensing knot of phytochrome B are necessary for conformer-specific binding to signaling partner PIF3. *PLoS Genet.* **5**, e1000352 (2009).
37. Viczian, A., Klose, C., Adam, E. & Nagy, F. New insights of red light-induced development. *Plant Cell Environ.* **40**, 2457–2468 (2017).
38. Burgie, E. S., Zhang, J. & Vierstra, R. D. Crystal structure of *Deinococcus* phytochrome in the photoactivated state reveals a cascade of structural rearrangements during photoconversion. *Structure* **24**, 448–457 (2016).
39. Yang, X., Kuk, J. & Moffat, K. Conformational differences between the Pfr and Pr states in *Pseudomonas aeruginosa* bacteriophytochrome. *Proc. Natl Acad. Sci. USA* **106**, 15639–15644 (2009).
40. van Thor, J. J. et al. Light-induced proton release and proton uptake reactions in the cyanobacterial phytochrome Cph1. *Biochemistry* **40**, 11460–11471 (2001).
41. Borucki, B. et al. Light-induced proton release of phytochrome is coupled to the transient deprotonation of the tetrapyrrole chromophore. *J. Biol. Chem.* **280**, 34358–34364 (2005).
42. Viczian, A. et al. Differential phosphorylation of the N-terminal extension regulates phytochrome B signaling. *New Phytol.* **225**, 1635–1650 (2020).
43. Remberg, A., Ruddat, A., Braslavsky, S. E., Gärtner, W. & Schaffner, K. Chromophore incorporation, Pr to Pfr kinetics, and Pfr thermal reversion of recombinant N-terminal fragments of phytochrome A and B chromoproteins. *Biochemistry* **37**, 9983–9990 (1998).
44. Cherry, J. R., Hondred, D., Walker, J. M. & Vierstra, R. D. Phytochrome requires the 6-kDa N-terminal domain for full biological activity. *Proc. Natl Acad. Sci. USA* **89**, 5039–5043 (1992).
45. Sweere, U. et al. Interaction of the response regulator ARR4 with phytochrome B in modulating red light signaling. *Science* **294**, 1108–1111 (2001).
46. Kosugi, S. et al. Six classes of nuclear localization signals specific to different binding grooves of importin alpha. *J. Biol. Chem.* **284**, 478–485 (2009).
47. Mathews, S., Lavin, M. & Sharrock, R. A. Evolution of the phytochrome gene family and its utility for phylogenetic analyses of angiosperms. *Ann. Missouri Bot. Gard.* **82**, 296–321 (1995).
48. Chen, M. & Chory, J. Phytochrome signaling mechanisms and the control of plant development. *Trends Cell Biol.* **21**, 664–671 (2011).
49. Zeidler, M., Zhou, Q., Sarda, X., Yau, C. P. & Chua, N. H. The nuclear localization signal and the C-terminal region of FHY1 are required for transmission of phytochrome A signals. *Plant J.* **40**, 355–365 (2004).
50. Oka, Y. et al. Arabidopsis phytochrome A is modularly structured to integrate the multiple features that are required for a highly sensitized phytochrome. *Plant Cell* **24**, 2949–2962 (2012).
51. Song, C. et al. On the collective nature of phytochrome photoactivation. *Biochemistry* **50**, 10987–10989 (2011).
52. Kim, P. W. et al. Unraveling the primary isomerization dynamics in cyanobacterial phytochrome Cph1 with multi-pulse manipulations. *J. Phys. Chem. Lett.* **4**, 2605–2609 (2013).
53. Gustavsson, E. et al. Modulation of structural heterogeneity controls phytochrome photoswitching. *Biophys. J.* **118**, 415–421 (2020).
54. Takala, H. et al. On the (un)coupling of the chromophore, tongue interactions, and overall conformation in a bacterial phytochrome. *J. Biol. Chem.* **293**, 8161–8172 (2018).
55. Al-Sady, B., Ni, W. M., Kircher, S., Schäfer, E. & Quail, P. H. Photoactivated phytochrome induces rapid PIF3 phosphorylation prior to proteasorne-mediated degradation. *Mol. Cell* **23**, 439–446 (2006).
56. Khanna, R. et al. A novel molecular recognition motif necessary for targeting photoactivated phytochrome signaling to specific basic helix-loop-helix transcription factors. *Plant Cell* **16**, 3033–3044 (2004).
57. Landgraf, F. T., Forreiter, C., Hurtado, P. A., Lamparter, T. & Hughes, J. Recombinant holophytochrome in *Escherichia coli*. *FEBS Lett.* **508**, 459–462 (2001).
58. Hirose, Y. et al. Green/red cyanobacteriochromes regulate complementary chromatic acclimation via a protochromic photocycle. *Proc. Natl Acad. Sci. USA* **110**, 4974–4979 (2013).
59. Zhao, K. H. & Scheer, H. Type I and type II reversible photochemistry of phycoerythrocyanin α-subunit from *Mastigocladus laminosus* both involve Z–E isomerization of phycoviolobilin chromophore and are controlled by sulfhydryls in apoprotein. *Biochim. Biophys. Acta Bioenerg.* **1228**, 244–253 (1995).
60. Ishizuka, T., Narikawa, R., Kohchi, T., Katayama, M. & Ikeuchi, M. Cyanobacteriochrome TePixJ of *Thermosynechococcus elongatus* harbors phycoviolobilin as a chromophore. *Plant Cell Physiol.* **48**, 1385–1390 (2007).
61. Brandlmeier, T., Scheer, H. & Rudiger, W. Chromophore content and molar absorptivity of phytochrome in the Pr form. *Z. Naturforsch. C* **36**, 431–439 (1981).
62. Fernandez Lopez, M. et al. Role of the propionic side chains for the photoconversion of bacterial phytochromes. *Biochemistry* **58**, 3504–3519 (2019).
63. Mailliet, J. et al. Dwelling in the dark: procedures for the crystallography of phytochromes and other photochromic proteins. *Acta Crystallogr. D* **65**, 1232–1235 (2009).
64. Kabsch, W. XDS. *Acta Crystallogr. D* **66**, 125–132 (2010).
65. Evans, P. R. & Murshudov, G. N. How good are my data and what is the resolution? *Acta Crystallogr. D* **69**, 1204–1214 (2013).
66. Winn, M. D. et al. Overview of the CCP4 suite and current developments. *Acta Crystallogr. D* **67**, 235–242 (2011).
67. McCoy, A. J. et al. Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).
68. Emsley, P. B., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D* **66**, 486–501 (2010).
69. Murshudov, G. N. et al. REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. D* **67**, 355–367 (2011).
70. Liebschner, D. et al. Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallogr. D* **75**, 861–877 (2019).
71. Williams, C. J. et al. MolProbity: more and better reference data for improved all-atom structure validation. *Protein Sci.* **27**, 293–315 (2018).
72. Dolomanov, O. V., Bourhis, L. J., Gildea, R. J., Howard, J. A. K. & Puschmann, H. OLEX2: a complete structure solution, refinement and analysis program. *J. Appl. Crystallogr.* **42**, 339–341 (2009).

Acknowledgements

We thank SFB 1078 for funding (subprojects B6 to P.H. and B8 to J.H.). The X-ray diffraction measurements were carried out at BL14.1 at BESSY II (Helmholtz-Zentrum Berlin für Materialien und Energie (HZB)) and at ID30-A3 at ESRF with support from CALIPSOplus (Grant Agreement 730872 of the EU Framework Programme HORIZON 2020) and HZB. We thank L.-O. Essen (University of Marburg) for collaboration in the early stages of the project, W. Wende (Giessen) for cooperation in measuring the CD spectra and C. Lang (Giessen) for expert technical assistance. We dedicate this paper to the memory of Winslow Briggs.

Author contributions

S.N., G.K. and S.M.S. designed, cloned, expressed, purified and crystallized the holophytochrome constructs. S.N. solved the structures with suggestions from C.F. and M.W. S.N. and J.H. interpreted the structures. A.K., D.B. and P.H. measured and interpreted the vibrational spectra. J.H. wrote the manuscript with the participation of P.H. and in discussion with all authors. J.H. devised and co-ordinated the project.

Competing interests

The authors declare no competing interests.

Additional information

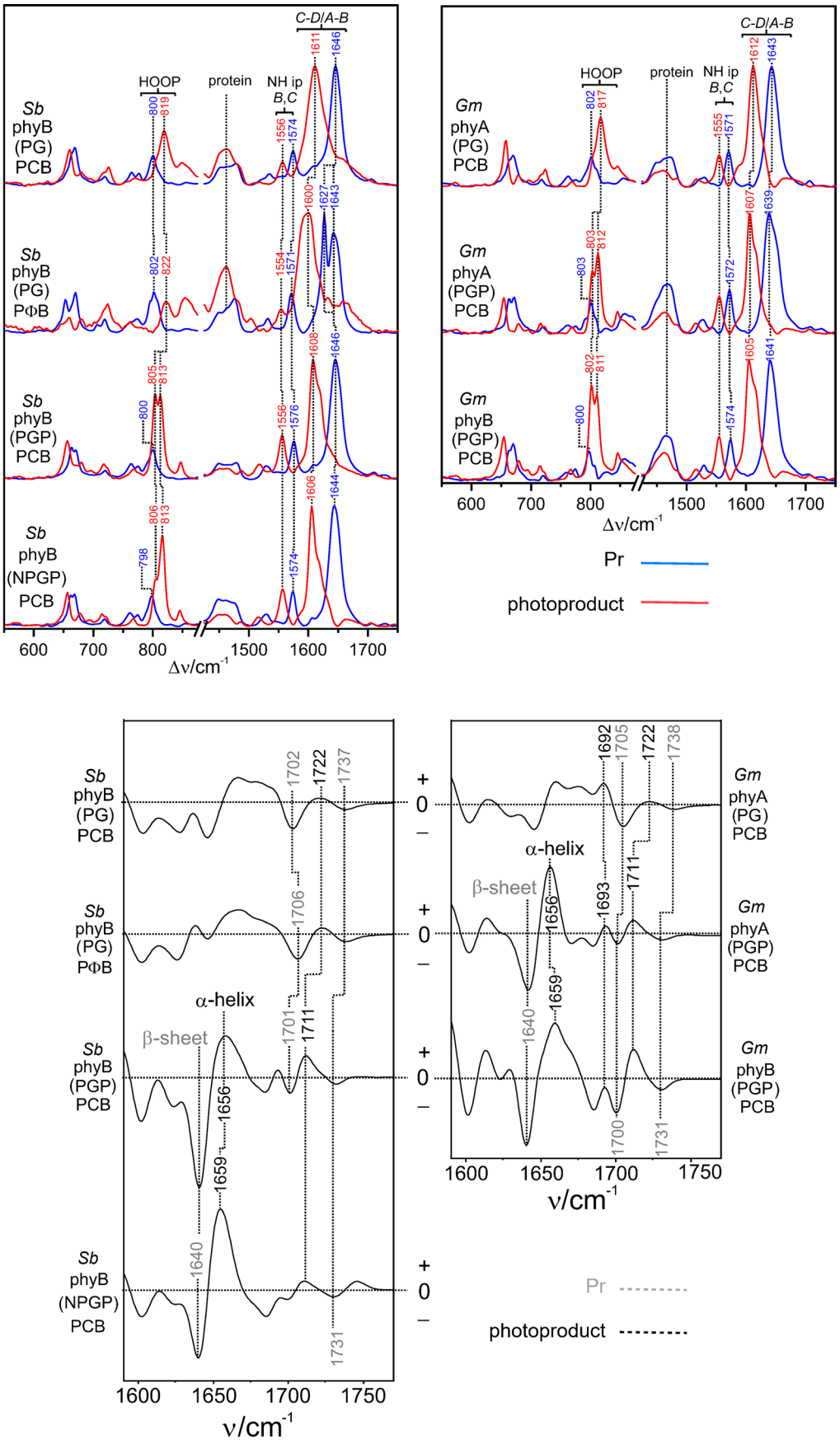
Extended data is available for this paper at https://doi.org/10.1038/s41477-020-0638-y.

Supplementary information is available for this paper at https://doi.org/10.1038/s41477-020-0638-y.

Correspondence and requests for materials should be addressed to J.H.

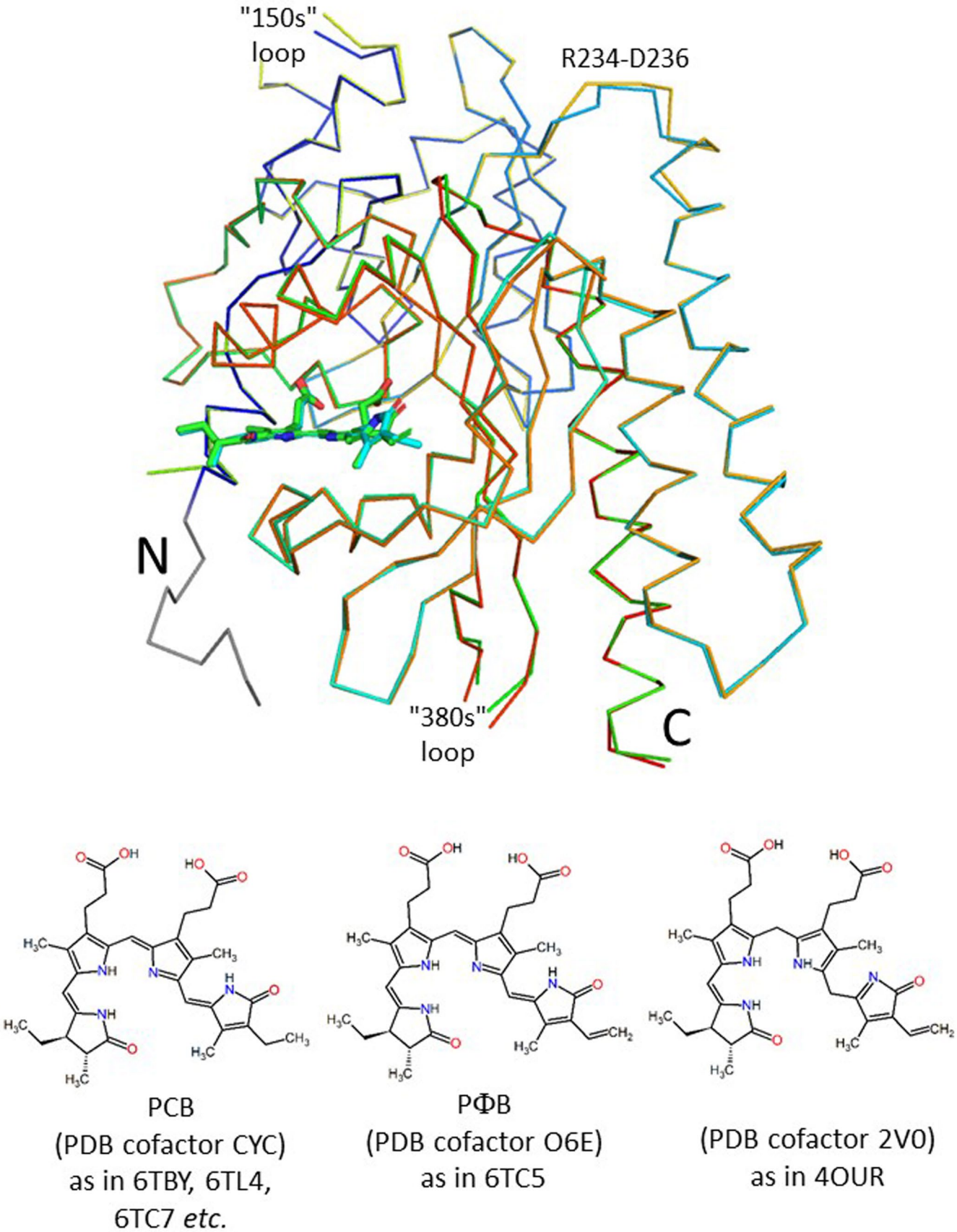
Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

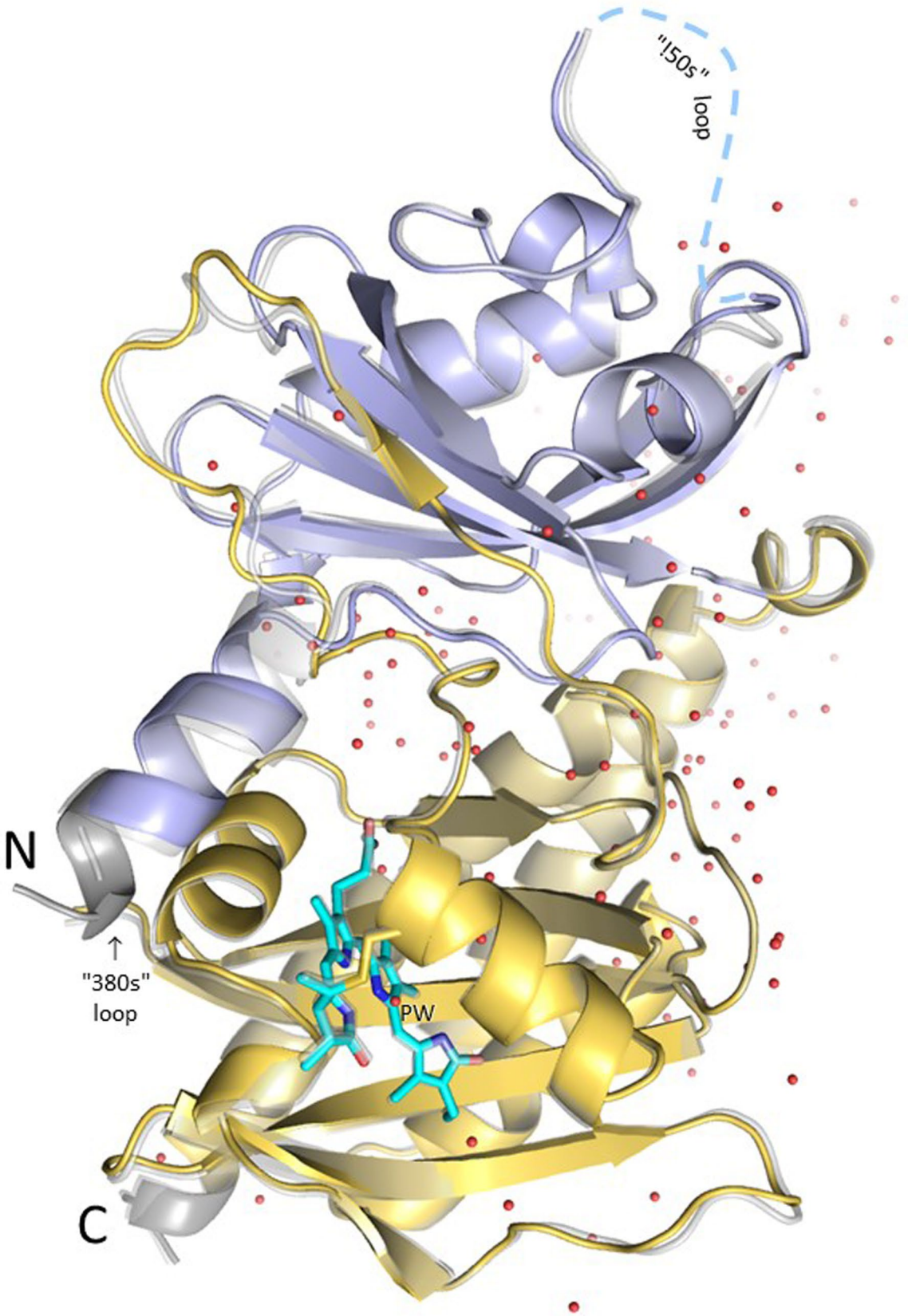


Extended Data Fig. 1 | See next page for caption.

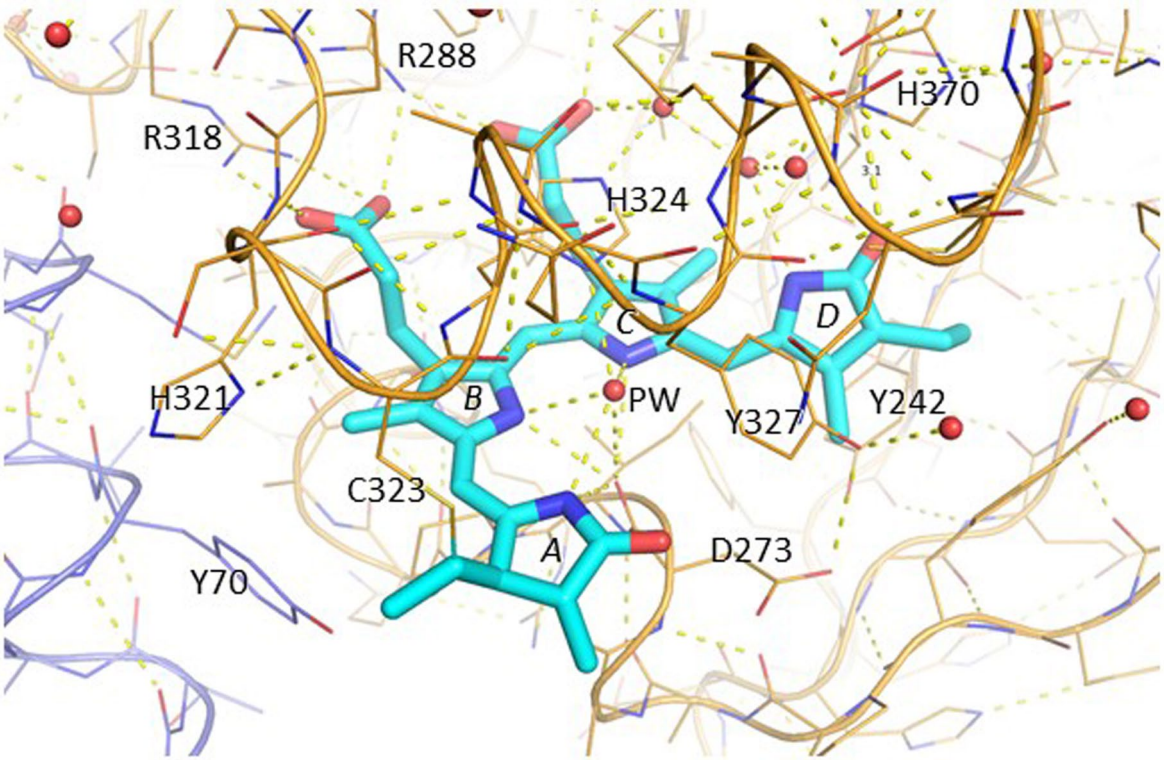
Extended Data Fig. 1 | RR and IR difference spectra of phyA and phyB constructs. Left, *Sorghum bicolor*; Right *Glycine max*. **Above**: RR spectra of the Pr states (blue traces) and their photoconversion products (red traces) obtained upon 670 and 750 nm irradiation at ambient temperature. All spectra were measured at 90 K with 1064 nm excitation. The spectral regions labelled are indicative of (i) the methine bridge configurations and conformations (C=C stretching modes of the *A-B* and *C-D* methine bridges at *ca*. 1600–1650 cm^{-1}), (ii) pyrrole nitrogen protonation state (N-H in-plane bending modes of rings *B* and *C* at *ca*. 1550 – 1580 cm^{-1}) and (iii) the *C-D* methine bridge torsion (hydrogen-out-of-plane [HOOP] mode at *ca*. 795 - 825 cm^{-1}). The broad feature at *ca*. 1460 cm^{-1} is largely due to non-resonant Raman bands of the protein. The high intensity of this feature relative to the RR bands of the chromophore indicates that the latter experience a low resonance enhancement. **Below**: IR "photoproduct minus Pr" difference spectra obtained upon irradiation with 670 and 750 nm at ambient temperature. The positive signals indicated by black lines and labels refer to the photoproduct, whereas the grey lines and labels mark the signals of the Pr state. Representative spectra based on at least two samples are shown. Spectra for each sample were measured several times. Each spectrum is based on 1000 separate FT scans.



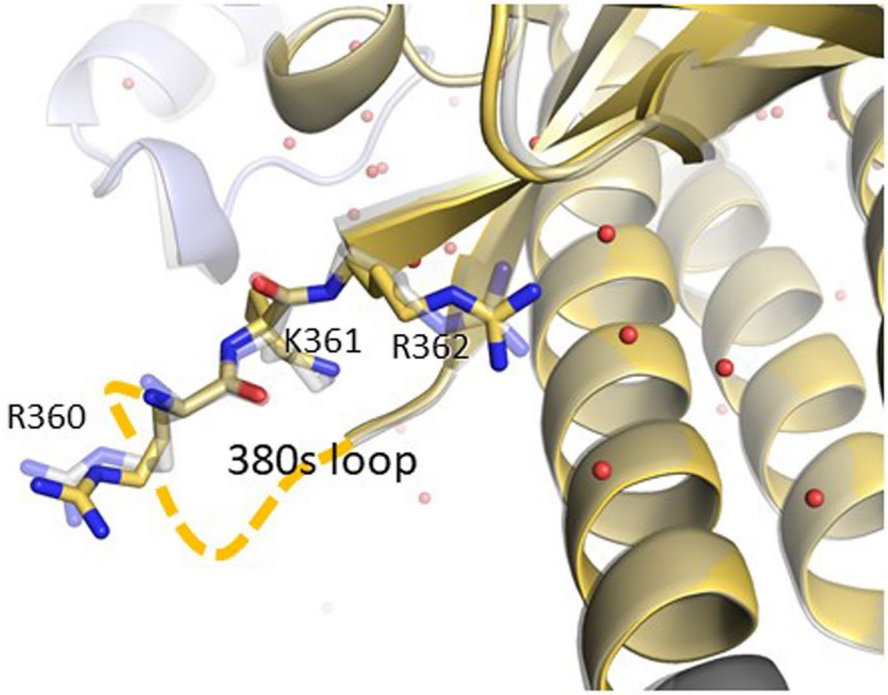
Extended Data Fig. 2 | Sb.phyB(PG)-PCB and -PΦB structures are almost identical. Above: Superimposition of peptide chains with chromophores. The N- and C-termini, the two unresolved loops and the somewhat deviant R234-D236 regions are labelled. Inset: superimposition of the PCB (cyan) and P(B (green) *D* rings. **Below**: Chemical structures of PCB, PΦB, and the incorrect model used in 4OUR (PDB cofactor codes CYC, O6E and 2VO, respectively). Note that the uncharged structures are shown, whereas in both Pr and Pfr holoprotein states all four cofactor nitrogens are protonated.



Extended Data Fig. 3 | Gm.phyA(PG)-PCB 2.1 Å structure (PDB 6TC7) of subunit B. PCB (cyan), PAS (slate), GAF (gold), PCB (cyan), waters (red spheres). PW, pyrrole water. Subunit A (grey) is superimposed.



Extended Data Fig. 4 | Gm.phyA(PG)-PCB 2.1 Å structure (PDB 6TC7) of subunit B including side chains and hydrogen bonding network. Note that the weak (3.1 Å) *D*-ring carbonyl – H370 hydrogen bond in subunit B is somewhat stronger (3.0 Å) in subunit A. PCB (cyan), PAS (slate), GAF (gold), PCB (cyan), waters (red spheres). PW, pyrrole water.




```
sm   -PSGGGGGGG QHKGRRLWGL VVCHHTSPRS
cp   -----YSRGA IQRGRKLWGL VVCQHTSPRT
ac   ------SAGH SSQGIKLWGL VVCHHTSPRY
atA  GD-APDATTQ PQKRKRLWGL VVCHNTTPRF
cpA  NE---GPALQ QQKRKRLWGL VVCHNSSPRF
psA  GD--SADAVL PQKKKRLWGL VVCHNTTPRF
stA  GE--SSDSSQ SQKRKRLWGL VVSHNTTPRF
asA  DEAESEQPAQ QQKKKKLWGL LVCHHESPRY
osA  DEVGADQPAQ QQKRKKLWGL LVCHHESPRY
zmA  DEPEPEQPPQ QQKKKRLWGL IVCHHESPRY
atB  G----SNVAS GRSSMRLWGL VVCHHTSSRC
atD  G---NGVNTG GRNSMRLWGL VVCHHTSARC
atE  ---------- ---SSKLWGL VVGHHCSPRY
stB  ------AVGG GRNSMRLWGL VVGHHTSVRS
osB  D--HNIARGS IPSAMKLWGL VVCHHTSPRC
atC  E-----MNRD LQTGRHLWGL VVCHHASPRF
```

Extended Data Fig. 5 | Putative Class I NLS specific to A-type plant phytochromes. Above. Superimposition of Gm.phyA(PG)-PCB subunit B (GAF domain, gold) with subunit A (grey) superimposed. Although the more mobile N-terminal section of the "380s" loop is missing (gold dashes), the final triad of the putative NLS (R360-R362) is resolved. **Below**. Alignment of the "380s" loop region in plant phytochromes (from Mathews *et al.* [47]). sm, *Selaginella martensii*; cp, *Ceratodon purpureus*; ac, *Adiantum caperis-veneris*; atA-D, *Arabidopsis thaliana* PHYA-D; cpA, *Curcurbita pepo* PHYA; psA, *Pisum sativum* PHYA; stA/B, *Solanum tuberosum* PHYA/B, asA-D, *Avena sativa* PHYA; osA/B, *Oryza sativa* PHYA/B; zmA, *Zea mais* PHYA. The K(R/K)K(R/K) consensus is boxed red.

nature research

Corresponding author(s): Hughes, J.

Last updated by author(s): Mar 6, 2020

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed
☒	☐ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☒	☐ The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐ A description of all covariates tested
☒	☐ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☒	☐ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☒	☐ For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐ Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	Standard firmware was used to collect spectral data. Crystallographic data was processed using XDS (build 20190315) and scaled using AIMLESS 0.7.4 from the CCP4 package (7.0.072).
Data analysis	CCP4 package (7.0.072): COOT (Wincoot 0.8.9.2) and REFMAC5 (5.8.0257); PHENIX (1.17-3644); PHASER (5.8.0257), MolProbity (4.5), OLEX2 (1.2.10); PyMOL 1.7.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The co-ordinates have been deposited in PDB as 6TBY, 6TC5, 6TL4 and 6TC7. They will be released prior to publication.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	For each of the 4 constructs whose structures are reported: Numerous crystals were screened for X-ray diffraction, datasets were collected from the best in each case. The final structural solution derived from the best dataset from the best crystal. Data collected in this manner is sufficient to provide a definitive structure within the limits described in SI Figure 2.
Data exclusions	Initial datasets were replaced by better ones as determined by the distribution of reflections.
Replication	The structures from individual constructs crystalized under similar conditions were essentially identical.
Randomization	This question is inappropiate for biomolecular structural analysis.
Blinding	This question is inappropiate for biomolecular structural analysis.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology
☒	☐ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

I declare that I have completed this dissertation single-handedly without the unauthorized help of a second party and only with the assistance acknowledged therein. I have appropriately acknowledged and cited all text passages that are derived verbatim from or are based on the content of published work of others, and all information relating to verbal communications. I consent to the use of an anti-plagiarism software to check my thesis. I have abided by the principles of good scientific conduct laid down in the charter of the Justus Liebig University Giessen "Satzung der Justus-Liebig-Universität Gießen zur Sicherung guter wissenschaftlicher Praxis" in carrying out the investigations described in the dissertation.

Place and Date: Giessen
Jan 12, 2023

Signature: