Biophysical characterisation of atypical phytochromes for development of novel optogenetic tools.

A thesis submitted to the University of Manchester for the degree of
Doctor of Philosophy in the
Faculty of Science and Engineering

2019

Uzma Choudry

School of Chemistry

Contents

Word count: 42,994

List of Abbreviations

ADoB12 - Adenosylcobalamin

ATP - Adenosine Triphosphate

BLUF - Sensor of blue light using FAD

Bphs - Bacterial phytochromes

BV - Biliverdin

CBCRs - Cyanobacteriochromes

CCA - Complementary Chromatic Adaption

CCT - C terminal domain

COP1 - CONSTITUTIVELY PHOTOMORPHOGENIC 1

EADS - Evolution Associated Difference Spectra

EMR - Electromagnetic Radiation

Fd - Ferredoxin

FRET- Fluorescence Re

FTIR - Fourier transform infra-red

GAD - GAL4 transactivation domain

GAF - cGMP phosphodiesterase/ Adenylate cyclase Fha1

GBD - GAL4 DNA-binding-domain

GEFs - Guanine nucleotide exchange factors

HAMP - Histidine kinases, adenylyl cyclases, methyl binding proteins, phosphatases

HK - Histidine Kinase

HO1 - Heme-oxygenase

IR - Infrared

LB - Luria Broth

LED - Light emitting diode

LOV - Light-oxygen-voltage sensing

MCP - Methyl-accepting chemotaxis protein

NLS - Nuclear localization signal

OPA - Optical parametric amplifier

OPO - Optical parametric oscillator

PAS - Per-ARNT-Sim

PBS - CpcG2-phycobilisome supercomplex

PCB - Phycocyanobilin

PCM - Photosensory Core Module

PcyA - Phycocyanobilin:Ferredoxin Oxidoreductase

PHR- Photolyase homology region

PHY - Phytochrome specific

Phy - Phytochrome

PIF3 - Phytochrome interaction factor 3

PIFs - Phytochrome interacting factors

PMT - Photomultiplier tube

PVB - Phycoviobilin

PVDF - Polyvinylidene fluoride

PYP - Photoactive yellow protein

PΦB - Phytochromobilin

RBSs - Ribosome binding sites

RcaE - Regulator of chromatic adaptation

RR - Response Regulator

S/WAXS - Small/ Wide Angle X-Ray Scattering

SAXS - Small Angle X-Ray Scattering

TA - Transient Absorption

TCS - Two Component Signaling

Abstract

Optogenetics refers to the use of light for the control of gene expression. This is achieved via light sensing domains, which are coupled to regulatory proteins that control gene expression. Starting in 2002 with channel rhodopsins the field has undergone rapid development and since then, rhodopsins, BLUF domains and phytochromes have now all been exploited for this purpose. Optogenetics has become an indispensible tool for neuroscientists to study neural networks and brain function. It is clear that optogenetics also has the potential to revolutionise the field of synthetic biology by allowing light-dependent control of gene expression in biosynthetic pathways.

The phytochromes have been the subject of detailed photochemical and structural studies. Although phytochromes have proven to be an attractive target for new applications in the field of optogenetics, their use has been limited due to their narrow wavelength sensitivity. However, several atypical phytochromes that exhibit broad spectral properties have been discovered recently. Here, we have expressed and purified two of these atypical phytochromes, one from cyanobacteria and one from algae.. Several biophysical techniques, including time-resolved visible and infrared spectroscopy, cryo-trapping and time-resolved X-Ray scattering, were used to understand the photocycles and characterise the mechanism of photoreactions of these atypical phytochromes in more detail. These findings are discussed in relation to the photocycles of the well-studied typical red/far-red phytochromes. In order to exploit the broad spectral diversity, we used these atypical phytochromes to engineer an optogenetic toolkit that allows light-induced modulation of gene expression. The modular nature of phytochromes makes them an ideal target for coupling different photosensory core modules with selected effector domains in a 'mix and match' manner. We demonstrated that this approach can be used to make functional chimaeras, which showed both *in-vitro* and *in-vivo* activity. This work was done in parallel with refactoring of a published optogenetic system onto the linalool biosynthesis pathway, which will eventually be used as a benchmark to test the chimaeric systems against. Ultimately, this work has set the foundations to allow an optogenetic toolkit to

be established for the light-dependent control of biosynthetic pathways for biotechnology applications.

Declaration

No portion of the work referred to in the thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning

Copyright Statement

i.   The author of this thesis (including any appendices and/or schedules to this thesis) owns certain copyright or related rights in it (the "Copyright") and s/he has given The University of Manchester certain rights to use such Copyright, including for administrative purposes.

ii. Copies of this thesis, either in full or in extracts and whether in hard or electronic copy, may be made only in accordance with the Copyright, Designs and Patents Act 1988 (as amended) and regulations issued under it or, where appropriate, in accordance with licensing agreements which the University has from time to time. This page must form part of any such copies made.

iii. The ownership of certain Copyright, patents, designs, trademarks and other intellectual property (the "Intellectual Property") and any reproductions of copyright works in the thesis, for example graphs and tables ("Reproductions"), which may be described in this thesis, may not be owned by the author and may be owned by third parties. Such Intellectual Property and Reproductions cannot and must not be made available for use without the prior written permission of the owner(s) of the relevant Intellectual Property and/or Reproductions.

iv. Further information on the conditions under which disclosure, publication and commercialisation of this thesis, the Copyright and any Intellectual Property and/or Reproductions described in it may take place is available in the University IP Policy (see http://documents.manchester.ac.uk/DocuInfo.aspx?DocID=2442 0),   in any relevant Thesis restriction declarations deposited in the University Library, The University Library's regulations (see

Acknowledgements

This is very last minute, and I don't know where to begin.

I want to start by thanking my exceptionally supportive family, who have been my champions and my strength. They have pushed me to work hard and challenge myself but also been there to provide TLC when required. Your love means everything in the world!

The five years I spent at the MIB not only gave me an opportunity to shape a career but most importantly build some precious relationships. I was welcomed into the team with open arms and received an unparalleled level of support and encouragement. Maybe this is the reason I stayed for so long. Like for many other Munros and Scruttons – the lab became a second home. The boys (y'all know who you are) enjoyed pulling my leg, and that still hasn't changed, but Shirley, Michi and Marina often came to my rescue – you ladies have a special place in my heart. I met incredibly kind and humble people along the way, who have each inspired me in some way. Like for many of us, this time contributed tremendously to shaping who I am today, and I am incredibly lucky to have had such a brilliant experience.

Helen, Dave Mansell, Shirley and John Waller initially trained me in the lab and have been my champions ever-since, you are the best! Thanks to Sarah, George, John and Siobhan, I delved into the world of start-ups and spin-outs. Who knew that would lead me to a career in Investments and Venture Capital? Special thanks to some of my dearest friends: Sarah, Max, Ravina, Tiziana, Saba, Beth and Mark for being my strength in times of weakness, and to continue being there for me as I grapple with the first world problems of grown-up life and dating. I love you guys.

Mark, Johnny, Lewis, Charlie, Robin, Toby and Alex – I miss making your lives ever so fun (read: miserable) haha, you guys made my time at the MIB unimaginably more fun. Hanan - thank you for being ever so sweet, considerate

and warm. Your advice is highly appreciated. Derren and Samantha, without your help and support I don't think I would've known what to do with lasers, cryostats, phytochromes ...... life (other than blind myself and others). Thank you ever so much Samantha, Queen of the lasers, for encouraging me over the past year to finish off the remaining 1/3rd of my thesis. All your encouragement made a huge difference. I also just loved how amused you were about all things biology. I very much appreciated your advice on my move to London. Derren, I don't have words and don't know what to say other than that if I could go back, I would still do this Ph.D., not for my undying love of phytochromes and disco lights (definitely the latter) but because I thoroughly enjoyed working with you. I don't think anyone else would have the patience to listen to me moan about lab, make-up and gym problems! You are just an exceptional human being! John thank you so much for helping me work safely with radioactive reagents. I tried very hard to walk out of the radiation facility a transformed hulk – we can all dream!

Obviously, Nigel – this would not have been possible unless you had offered me a summer internship (Lister Studentship) and then a Ph.D. I hope many other students benefit from such internships. I appreciated the freedom I had under your supervision to delve into different parts of the project (continuously being reminded to narrow down). Your backing to take on other interests whether that was Biotech YES or UMIP fellowship helped me develop beyond a good scientist. I will cherish the memories from my time at the MIB and miss baking for you guys!

UMIP deserves mention. I enjoyed the UMIP Fellowship – a brilliant learning opportunity that opened so many doors for me. Siobhan D, you have been a great mentor. I appreciate all of the help and advice you and Tony provided me, which helped me start my career in Venture Capital at Octopus Ventures. Finally, this wouldn't be complete without a mention to my first manager, Simon King, who helped me manage my transition out of academia into the corporate world. I am grateful to you for offering me time off to write up my thesis and generally being mindful, reminding me to not to work when away on holiday.

What I have found myself summing up here is that success is only meaningful, enjoyable and possible when surrounded by a great bunch of people that root out for you and help you grow – and I count myself incredibly lucky to have been in that privileged position.

We live on an island surrounded by a sea of ignorance. As our island of knowledge grows, so does the shore of our ignorance. - John Archibald Wheeler

"It is better to deserve honors and not have them than to have them and not to deserve them." - Mark Twain

1 Introduction

1.1 Outline

The work carried out during this Ph.D. is presented as an alternative format thesis consisting of three manuscripts.

The studies carried out during this PhD involved collaborations with several researchers from the University of Manchester, Central Laser Facility (STFC, Harwell) and Institut de Structurale Biologie (Grenoble).

Manuscript titles, publication status and roles of authors:

Manuscript 1, Chapter 2: "Initial characterisation of novel atypical phytochromes."
Uzma Choudry, Derren J Heyes, Samantha J O Hardman, Nigel S Scrutton.
This manuscript is pending submission.
U.C. wrote the paper with input from D.J.H. and S.J.O.H; U.C prepared samples and performed basic characterization; U.C. and D.J.H performed cryogenic and laser flash photolysis measurements; S.J.O.H performed visible pump probe spectroscopy; N.S.S. and D.J.H conceived of and designed the work.

Manuscript 2, Chapter 3: "Photochemical mechanism of an atypical algal phytochrome."
Uzma Choudry, Derren J. Heyes, Samantha J. O. Hardman, Michiyo Sakuma, Igor V. Sazanovich, Joyce Woodhouse, Eugenio De La Mora, Martin N. Pedersen, Michael Wulff, Martin Weik, Giorgio Schirò, Nigel S. Scrutton.
ChemBioChem 2018, **19**, 1036 – 1043
U.C., D.J.H., S.J.O.H., and G. S. wrote the paper with input from all authors; U.C. and M.S. prepared and samples; U.C. performed basic characterization; U.C. and D.J.H performed laser flash photolysis measurements; S.J.O.H performed visible pump probe spectroscopy; D.J.H, S.J.O.H., U.C. and I.S. performed infra-red time resolved measurements; G.S., D.J.H, S.J.O.H., J.W., E.D.L.M., M.P., M.W. and M.W., performed WAXS measurements; S.J.O.H.

analyzed the visible and infra-red spectroscopy data; G.S analyzed the WAXS data.; N.S.S. and D.J.H conceived of and designed the work.

Manuscript 3: Chapter 4: "Design of new chimaeric phytochromes for optogenetic control of gene expression."

Uzma Choudry, Derren J Heyes, John M X Hughes, Samantha J O Hardman, Nigel S Scrutton.

This manuscript is pending submission.

U.C. wrote the paper with input from D.J.H. and S.J.O.H; U.C designed and assembled constructs, prepared samples and performed basic characterization; U.C. and D.J.H performed laser flash photolysis measurements; S.J.O.H performed visible pump probe spectroscopy; U.C. performed radiolabelling measurements with assistance from J.M.X.H; N.S.S. and D.J.H conceived of and designed the work.

Given the broad range of topics covered in this thesis, an introduction to photochemistry, photoreceptors and optogenetics is outlined below.

1.2 Photochemistry

Light is the main source of energy for all life. Nature has evolved to find mechanisms in which light signals drive essential metabolic processes, allowing organisms to adapt to the changing environment. Almost all processes within plants are controlled by light in some way. A key example is photosynthesis, a photochemical process fundamental to all life on Earth, which occurs in plants, photosynthetic bacteria and algae. Chlorophyll absorbs sunlight and uses the energy from the photon to split water into oxygen while making glucose, which is converted to pyruvate and releases adenosine triphosphate (ATP) by cellular respiration. Another example is vitamin D synthesis in the skin from cholesterol through a chemical reaction that is dependent on sunlight, UVB radiation.[7-10]

Photoreceptors are light activated proteins that receive and respond to light signals. The received light signal leads to chemical changes within the protein that transduce specific output signals leading to a physiological change. Photoreceptors can be found widely in nature, from the human eye to plants and microbes, where they fulfil a myriad of functions. In plants, these photoreceptors control many crucial stages of plant development, such as seed germination, flowering and shade avoidance, through metabolic entrainment to adapt to the environment. Vision and circadian rhythm are also dependent on photoreceptors.[11-15]

Light is a versatile signal, carrying information in the form of wavelength (energy), direction, polarisation and intensity. Photoreceptors receive their signal through the absorption of a photon. Absorbance is a process whereby an electron within a molecule is excited to a higher, excited, state, S_n, from a ground state, S_0, by the energy provided by a photon in a process that occurs on the femtoseconds timescale. The energy of a photon is reciprocally related to the wavelength of light as shown in the equation: $E = \frac{hc}{\lambda}$. According to the

Stark-Einstein law, a molecule absorbing a single photon will undergo a single electronic transition. Thus, absorbance of multiple photons to bring about a single transition is forbidden except in very specific experimental conditions.

The chemical and physical properties of an excited state molecule are often drastically different from the ground state molecule. The molecule can only remain in the excited state for a short period of time before it undergoes a series of events to return to the ground state or form a new species that will reach its own ground state. The absorbance process and the subsequent photo-physical processes are shown in the form of a Jablonski diagram as shown in Figure 1.1.[16-18]
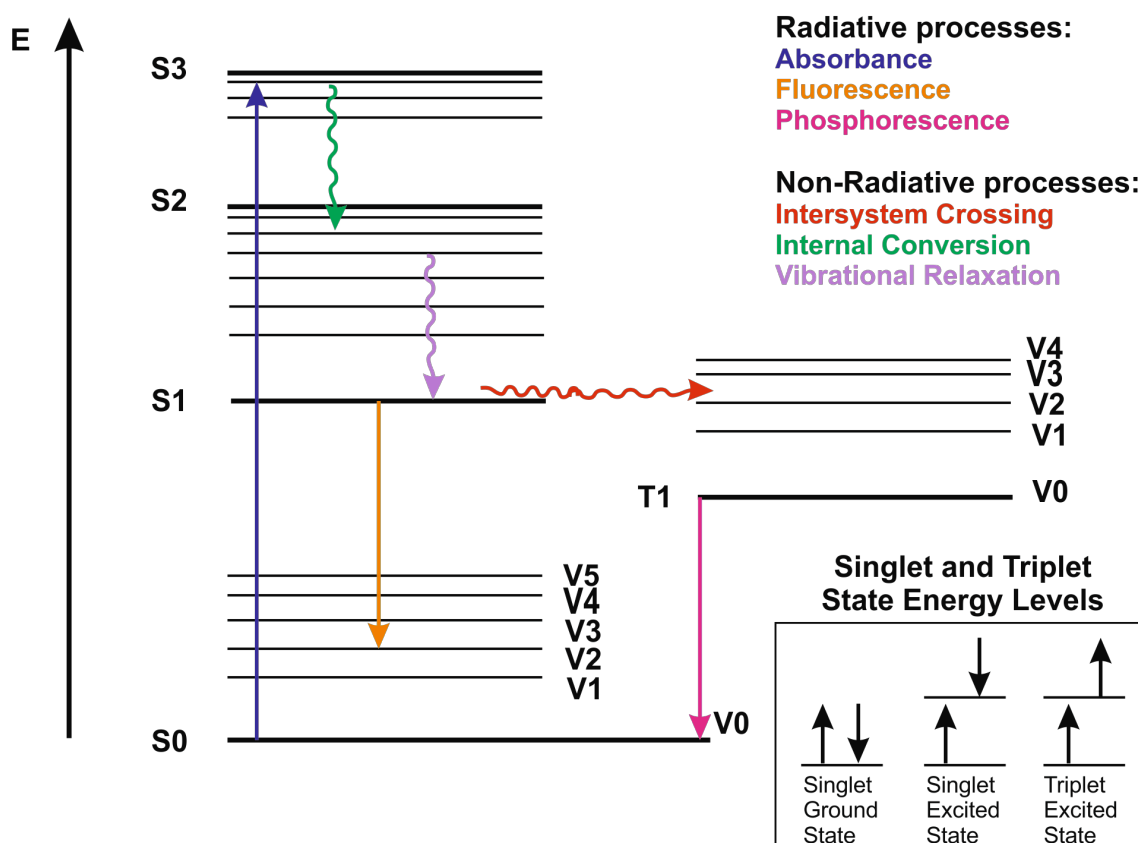
Figure 1.1. Jablonski diagram – energy diagram that demonstrates the different transitions that a molecule may undergo upon absorption of a photon. E denotes the amount of energy. Each column represents a specific spin multiplicity. Bold horizontal lines define the limit of each electronic state. Each electronic energy state has multiple vibronic levels within it.

Each electronic energy state has within it many vibronic energy levels as shown in Figure 1.1. Only a fraction of the vibrational energy levels are shown because there are a large number of possible vibrations in every molecule. Each vibrational energy state is divided further into rotational energy levels (not shown). As the electronic energy increases, the difference in energy between levels continually reduces, and the overlap between vibrational energy level increases. Once an electron is excited energy is dissipated via several routes

which are discussed below and shown in Figure 1.1.[16,17] The faster the transition the more likely it is to happen as it will outcompete the slower processes. The typical timescales at which the absorbance and deactivation processes occur are shown in Table 1.1.

Table 1.1. Timescales of the radiative and non-radiative transition processes.

Transition process	Timescale (s)
Absorbance	10^{-15}
Internal Conversion	10^{-14} to 10^{-11}
Vibrational Relaxation	10^{-14} to 10^{-11}
Fluorescence	10^{-9} to 10^{-7}
Intersystem Crossing	10^{-8} to 10^{-3}
Phosphorescence	10^{-4} to 10^{-1}

1.2.1 Internal Conversion

As discussed above, at increased energy levels the electronic states are closer together and this results in the overlap between vibrational levels. If such overlap of vibrational levels and electronic states occurs then the excited electron can transition from an electronic state in one vibrational level to another in a lower electronic state. Internal conversion is a form of vibrational relaxation so it is likely to occur immediately after absorption, however, it is unlikely to be the final transition because at lower energies - energies approaching the ground state - there is a bigger energy difference between the electronic states, thus there is reduced overlap between vibrational levels. The slow return to the ground state means that other transitive processes outcompete internal conversion for an electron in its first excited state.[17]

1.2.2 Vibrational Relaxation

Vibrational relaxation allows the dissipation of energy via transferring the energy provided by the photon to other vibrational modes as kinetic energy. Depending on the phase of the solvent, this kinetic energy can either be transferred to surrounding molecules or it can stay within the molecule. As the relaxation occurs between vibrational levels, there is no change in the electronic

state. This process usually takes 10^{-14} to 10^{-11} seconds. As this transition is very fast, this is likely to occur immediately after absorption.[17]

1.2.3 Fluorescence

Another energy dissipation route is fluorescence, whereby the energy from the absorbed photon is released in the form of a photon. According to Kasha's rule fluorescence is only likely to occur between the lowest excited state and the ground state. The released photon is almost always of a lower energy than the photon absorbed because some of the energy is lost in internal conversion and vibrational relaxation. Fluorescence occurs on the timescale of picoseconds to nanoseconds.[17]

1.2.4 Intersystem Crossing and Phosphorescence

The slowest horizontal transition in the Jablonski diagram is intersystem crossing (Figure 1.1). In this process the electron changes multiplicity between an excited singlet state to an excited triplet state. Based on electronic rules this transition is classically forbidden. However, combining the electronic energy with vibrational energy makes intersystem crossing weakly possible as it can outcompete fluorescence. Intersystem crossing leads to energy dissipation through several different routes that occur on microseconds to seconds timescale. One such process is phosphorescence, through which the electron returns from the excited triplet state to the ground state in a radiative process.

1.2.5 Other Relaxation Processes

There are other non-emitting processes, such as loss of energy through molecular collisions. Overlap of absorption and emission spectra can result in energy transfer. These non-emitting processes can compete with fluorescence.[17] The competing energy dissipation processes occur over a broad range of timescales shown in Table 1.1, each covering a certain time frame. Thus, these processes can be probed using time-resolved laser spectroscopy, which allows the reaction to be triggered with a laser and provides excellent time resolution.

The efficiency of a certain process can be measured quantitatively by calculating the quantum yield (Φ) of the process. The quantum yield of a de-excitation process is the ratio of molecules de-exciting via this transition process compared to those that entered the excited state. The equation is shown below:

$$\Phi = \frac{Yield\ of\ photochemical\ product}{Number\ of\ photons\ absorbed}$$

The quantum yield of processes that are optimised to harness light is maximised for generating a signal through minimising the quantum yield of competing, non-productive processes such as thermal de-excitation.

Photoexcitation can also result in processes such as luminescence, chemiluminescence, photoisomerisation, photodissociation, photosensitsation and photorearrangements. When near an electronically excited molecule, a second molecule can be excited through space in a process known as photosensitisation, or fluorescence resonance energy transfer (FRET). Two major classes of photoisomerisation are *cis-trans* isomerisation and open-closed ring transition. The former isomerisation involves rotation around a double bond, whereas the latter involves bond cleavage and formation.[19]

1.3 Spectroscopy

Spectroscopy has been an indispensible tool for the study of any interaction of electromagnetic radiation (EMR) with matter, including biological molecules. The results of these measurements are often presented in the form of a spectrum, which is a plot of the measured parameter (*e.g.* absorption or emission) against wavelength or frequency. When an electromagnetic radiation beam is passed through a sample, the photons may be refracted, reflected or absorbed to affect the chemical bonds and electrons in the sample. This leads to the photochemical processes covered in section 1.2.[20-22] Visible and Infrared absorption and X-ray scattering techniques will be covered in this section, as these are the techniques employed in this thesis to study the photochemistry of photoreceptors that are targeted for optogenetic applications.

On a very basic level, in a spectroscopy experiment light emitted from a light source is passed through a sample to a detector, which records the properties of the received light in the form of a spectrum (Figure 1.2).

Light Source   Sample Cell   Detector
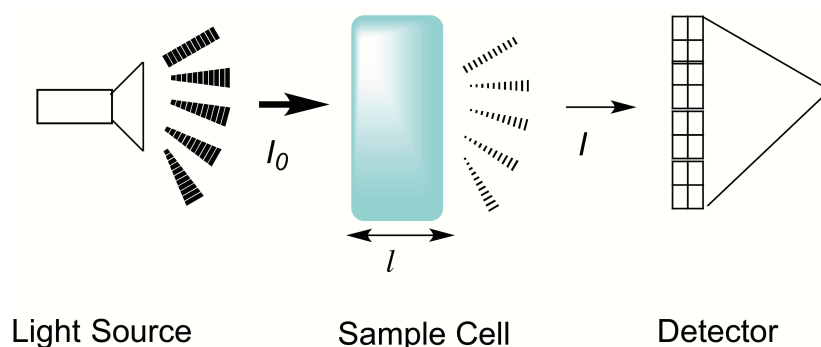
Figure 1.2. Schematic illustration of basic absorption spectroscopy. Radiation from the light source, I_0, is absorbed by the sample in sample cell of length, l, and the fraction of light that passes through the sample, I, is quantified by the detector.

The amount of light transmitted through the sample depends on the path length of the sample; it's extinction coefficient, and concentration. This is defined by the Beer-Lambert Law as

$$T = I/I_0 = 10^{-\varepsilon cl}$$

where T is transmittance and equals to light passing through the sample I divided by and incident light I_0. [23] Transmittance can be converted to absorbance as

$$A = -\log (I/I_0) = \varepsilon cl$$

1.3.1 Infrared Spectroscopy

Covalent bonds in organic molecules behave like springs and are not rigid. Organic molecules are always in motion as their bonds bend, twist, rock and stretch at room temperature, as indicated in Figure 1.3. Each of these modes has an energy, which is usually in the 2-20 μm wavelength or infrared (IR)

region and will depend on the bond mode, the atoms involved, and the surrounding environment.

| Symmetric Stretching | Assymmetric Stretching | Scissoring | Rocking |
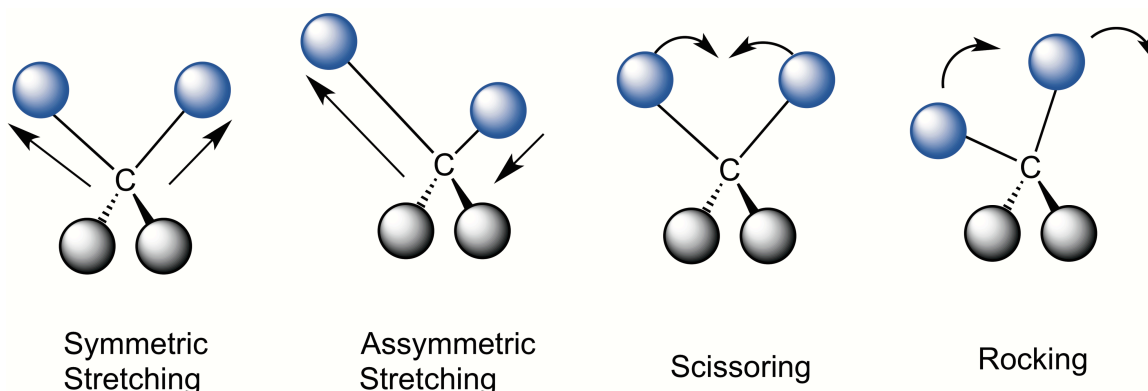
Figure 1.3. Some examples of vibrational modes within organic molecules at room temperature.

The strength of IR spectroscopy is that different functional groups (*e.g.* carbonyl, amine and carbon-hydrogen bonds) have different characteristic absorption frequencies (Figure 1.4). Whether bonds absorb IR light depends on whether the absorption results in a periodic change in the dipole moment of the molecule. The greater the polarity of the bond, the stronger its IR absorption. The carbonyl bond is very polar and therefore absorbs very strongly, whereas a carbon-carbon bond is non-polar, unless it has un-identical sp2 carbons, and so it does not absorb and is often therefore IR-inactive.[20,21,24-27]
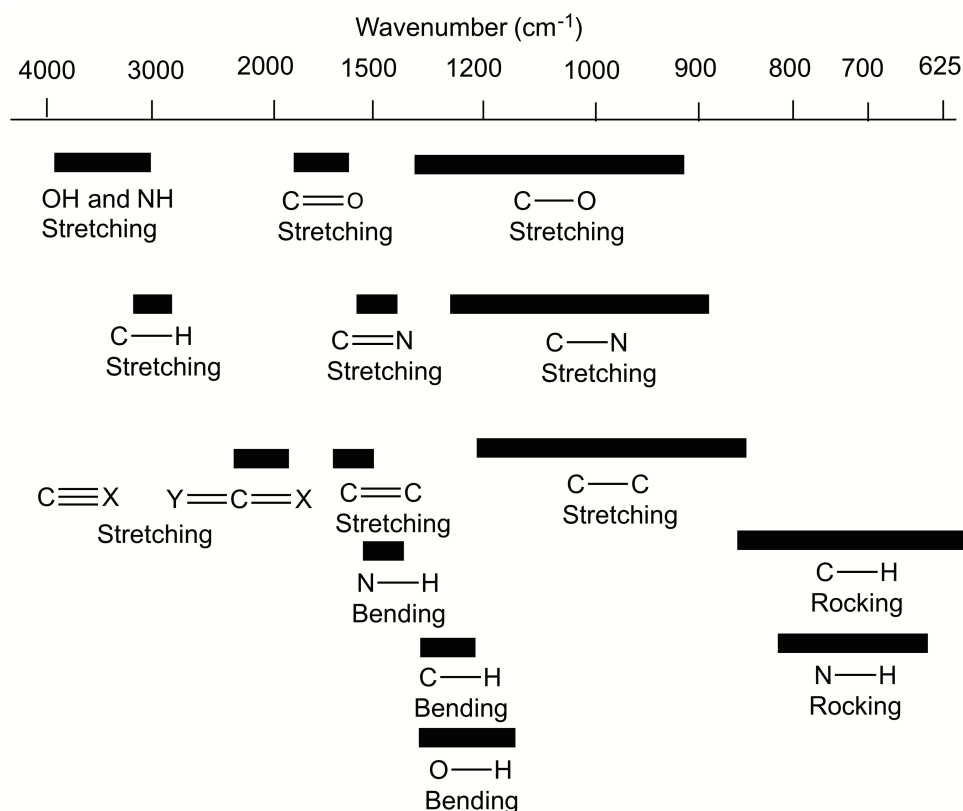
Figure 1.4. Vibrational modes and energies of some common types of molecular bonds.

Although IR does not allow you to decipher complete structure, even for simple molecules, it is a valuable tool when used other analytical methods to monitor functional groups. It is therefore good for monitoring reactions or, in our case, photochemical processes. Another technique for measuring vibrational modes is Raman spectroscopy, however this technique is not used in this project and therefore not covered here.

1.3.2 UV/Visible Spectroscopy

Unlike IR spectroscopy, absorption of higher energy, shorter wavelength UV/ visible light very often results in electronic transitions as discussed in section 1.2. The setup of a UV/ Visible absorption spectrometer includes a tungsten filament or another light source to produce visible white light and a deuterium arc lamp to produces UV light. A diffraction grating within a monochromator is used to separate the different wavelengths of light so that light of a single

wavelength reaches the detector at one time. The detector is usually a photomultiplier tube, a photodiode array or a photodiode, which filter the light.

1.3.3 Time-resolved Laser Spectroscopy

Many chemical processes that occur within proteins occur over a wide range of timescales (Figure 1.5). No single instrument can cover this large temporal range and so different methods are often required to study these processes. Time-resolved spectroscopy can provide information on both the lifetimes and the spectral properties of photoinduced intermediates.
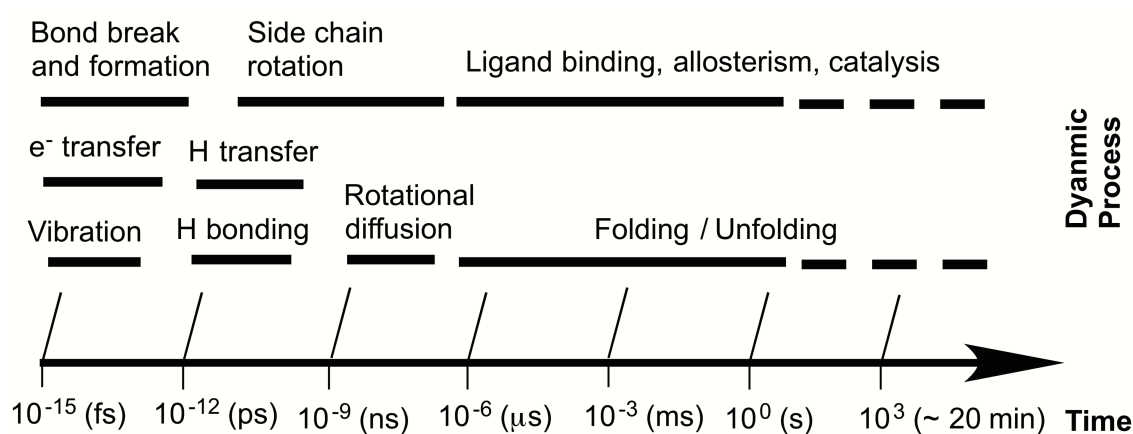
Figure 1.5. Protein dynamics with associated timescales.

To investigate fast processes, light properties need to be controlled in a precise manner. Invented in 1960, lasers provide unprecedented control of light parameters, such as intensity, wavelength, direction, duration of pulse, polarisation and phase. Narrow and intense beams of light can be generated that have enough energy to populate intermediate states to a detectable concentration. As such, transient, short-lived species, such as excited states and reactive intermediates can be detected and followed in real time, allowing dynamics of photochemical reactions to be characterised in unprecedented detail.[20,21,28,29] This type of measurement, broadly known as pump-probe spectroscopy, usually involves a laser pulse (the 'pump') to trigger the photochemical reaction and various types of spectroscopic probe (*e.g.* UV/vis or IR absorbance, X-ray scattering, etc) to monitor the reaction progress.

To monitor processes occurring on ultrafast, fs – ns, timescales ultrafast transient absorption spectroscopy can be used. These experiments use ultrafast laser systems, which typically produce a train of sub-ps duration pulses at kHz repetition rates. The photoreaction is initiated with a pump pulse and another, synchronised, probe pulse, is used to follow absorbance changes at a time delay, t. An optical chopper is used to halve the frequency of the pump beam, which allows for "pump on" and "pump off" readings that are used to generate difference spectra at a higher sensitivity than simply observing the absolute level of transmitted light. One of the beams passes *via* a delay stage, the position of which is varied to vary the time delay. At these timescales, photophysical processes such as photochemical conversions, vibrational relaxation, internal conversion and intersystem crossing can occur.[24,27-31]

To monitor slower, ns – ms, kinetics, laser flash photolysis can be used. This technique utilises a single pump pulse to initiate a photoreaction and then a series of probe measurements monitor changes in absorbance over a range if timescales. The source of pump pulse can be a flash lamp or a light-emitting diode (LED). The minimum time resolution of the flash photolysis set up is restricted to nanoseconds, which is determined by the speed of the detector response. The probe light is typically generated by a Xe lamp and can be either monochromatic and detected at a single wavelength with a photomultiplier tube (PMT) to monitor kinetics, or detected as a full spectrum on a photodiode array, with lower sensitivity. At these timescales, the observed signals are assumed to show formation of triplet states, intermediates and photoproducts of the reaction rather than photophysical loss processes.[32]

1.3.4 X-ray scattering

Scattering techniques produce deflection patterns when a sample is illuminated with X-rays, which can be used to deduce the nano- and micro-scale molecular structure and dynamics of non-crystalline materials including biological macromolecules. Diffraction and scattering occur due to interference of X-rays elastically scattered from electrons within the sample. Solution scattering techniques, such as small angle X-ray scattering (SAXS) and wide angle X-ray scattering (WAXS), are powerful approaches to provide structural insights of
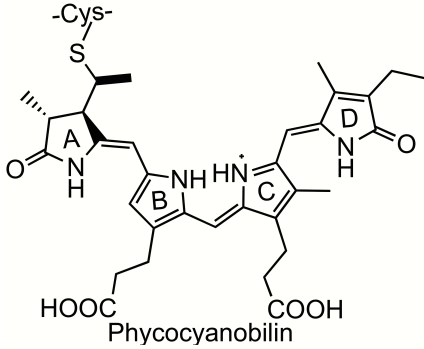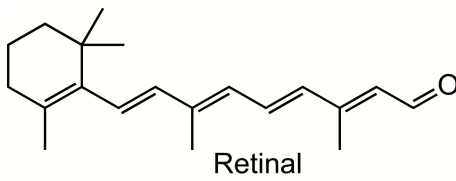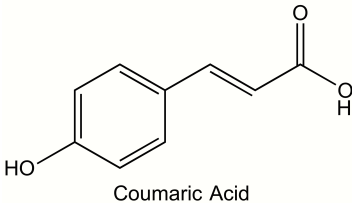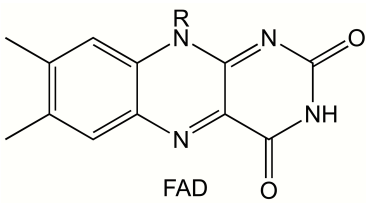
proteins in solution. These techniques can be applied to samples in solution and allow large and small structural dynamics and conformational changes to be probed. With this approach, rotationally symmetric scattering images are produced from randomly oriented proteins. The integration of these scattering images result in one-dimensional scattering profiles that give structural insights such as mass, flexibility, gyration, compactness and overall shape.[33] By applying the time-resolved approach to this technique you can build a picture of conformational and structural changes of a molecule overtime when triggered or activated.
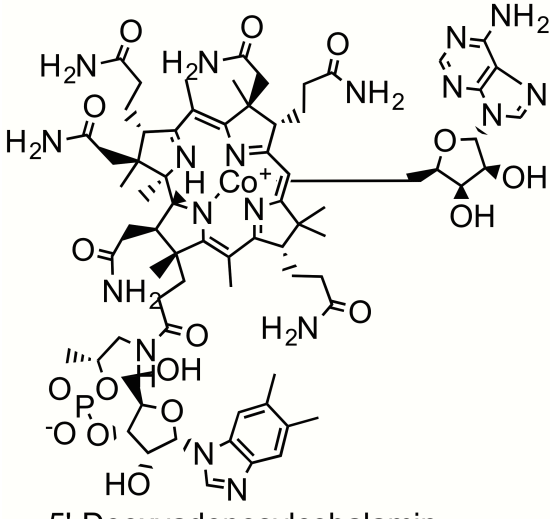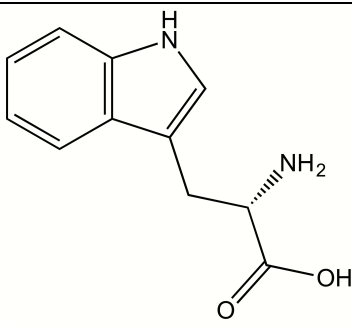
1.4 Photoreceptors

The light-sensing moiety within photoreceptors is known as a chromophore, usually a small molecule that absorbs in or near the visible spectral region. Light-induced changes in the configuration of the chromophore (*e.g.* through an isomerisation step or cysteinyl–adduct formation) initiate the formation of a signalling state, which communicates with the downstream signal transducer to elicit a specific physiological response [34].Photoreceptors are ideal candidates for studying the role of dynamic changes within the protein in relation to their function. This is because the reactions can be triggered with fast laser illumination and the resulting photochemical changes can be followed using various spectroscopy techniques, with excellent time resolution.[10,35]

Photoreceptors were originally separated into six main classes (rhodopsins, phytochromes, xanthopsins, cryptochromes, BLUF and phototropism) based upon the structure of their chromophore, as shown in Table 1.2. The recent discovery of vitamin B_{12} containing photoreceptors, and the UV-sensitive photoreceptor UVR8 has further expanded the known number of photoreceptor families.

Table 1.2: Chromophore-based classification of photoreceptors. Chromophore structure and details of the primary photochemical event are included. Adapted from: [10]

Class of Photoreceptor based on the Chromophore	Photosensor family	Photochemistry
Bilins Phycocyanobilin	Phytochromes, Cyanobacterial phytochromes. Cyanobacterio-chromes, Algal Phytochromes	*E/Z* isomerisation
Retinal	Rhodopsin	*E/Z* Isomerisation
Coumaric Acid	Xanthopsin	*E/Z* isomerisation
Flavins FAD	BLUF Proteins	Proton transfer
	Cryptochromes	Electron transfer

	Phototropin	Cysteinyl adduct formation
\n5'-Deoxyadenosylcobalamin	B_{12}	Photolysis of C-Co bond
\nTryptophan	UVR-8	Electron and proton transfer; salt bridge disruption

In phytochromes, rhodopsins and xanthopsins, photon absorption results in an initial *E/Z* isomerisation step. However, Blue Light Utilising FAD (BLUF), cryptochromes and phototropins all contain a flavin chromophore and their initial photochemistry is driven by photon transfer, electron transfer and cysteinyl-adduct formation, respectively. The B_{12} containing photoreceptors have a 5′-deoxyadenosylcobalamin chromophore, which transmits the light signal via adduct formation, ultimately leading to a change in oligomerisation state of the photoreceptor.[36,37] A novel tryptophan-based cofactor acts as the UVR-8 chromophore, which senses UVB light. UVB irradiation leads to the disruption of tryptophan and arginine mediated salt bridges, thus leading to monomerisation of the protein.[38]

The initial photoreaction results in local changes in the protein structure. In many cases this forms a signalling state that transduces the photon absorption

signal to a signal transduction partner to elicit a physiological response. All classes of photoreceptor are water soluble except rhodopsins, which are membrane proteins. The majority of the work in this thesis has focussed on the bilin-containing photoreceptors, namely the cyanobacteriochromes and the phytochromes, so these will be discussed in more detail in the next section. A brief description of the other main classes of photoreceptor is given below.

1.4.1 Rhodopsins

Rhodopsins are membrane photoreceptors that contain a retinal chromophore. The retinal is covalently linked to a lysine residue in the protein as a Schiff base. Though well known as photoreceptors in the eyes, rhodopsins can also be found in microorganisms and algae.[39-41] Channelrhodopsins are responsible for phototaxis in flagellate algae.[42,43] Algal rhodopsins that act as light driven proton pumps are found in photosynthetic organisms, yet their physiological role remain unknown. Microbial rhodopsins known as bacteriorhodopsins pump protons across a membrane in response to light to generate a proton gradient that is then converted to chemical energy.[43] Enzymerhodopsins are made up of sensor domains that are linked to effector domains such as kinases and adenylate cyclase. The enzymatic activity allows adaptation of behavioural responses.[44]

Microbial rhodopsins are seven-helix transmembrane proteins.[45] The helices traverse the membrane forming a barrel-like bundle. The retinal is covalently attached to the middle of the bundle.[46] The light induced conformational changes occurring during the photocycle have been revealed through crystallography and cryo-electron microscopy.[47,48] However, this photocycle may not stand true for the algal rhodopsins. Although there is a lack of high-resolution plant rhodopsin structures, plant rhodopsins are expected to be structurally similar to homologous microbial rhodopsins.[47]

Studies of the retinal cofactor in rhodopsin have elucidated the photochemistry of the reaction. Animal rhodopsins are optimised for 11-*cis* to all-*trans* light driven isomerisation, whereas microbial rhodopsins are optimised for all-*trans* to 13-*cis* isomerisation. In bacterial rhodopsins the isomerisation results in a series

of structural transitions with proton translocations as shown in Figure 1.6.[49] The all-*trans* form is the active form in bacterial rhodopsin. Within 5 ps of absorbing a 570 nm photon, the bacterial rhodopsin retinal undergoes a *trans* to *cis* isomerisation around bond 13 converting the BR 570 state to the K 590 state.[50] The ultrafast isomerisation is followed by the subsequent formation of several intermediates (L, K, M, N, O), which are shown with respective lifetimes in the photocycle in Figure 1.6. Within tens of milliseconds of the photoexcitation the *cis*-intermediate relaxes back to the BR ground state.[43,49-54]
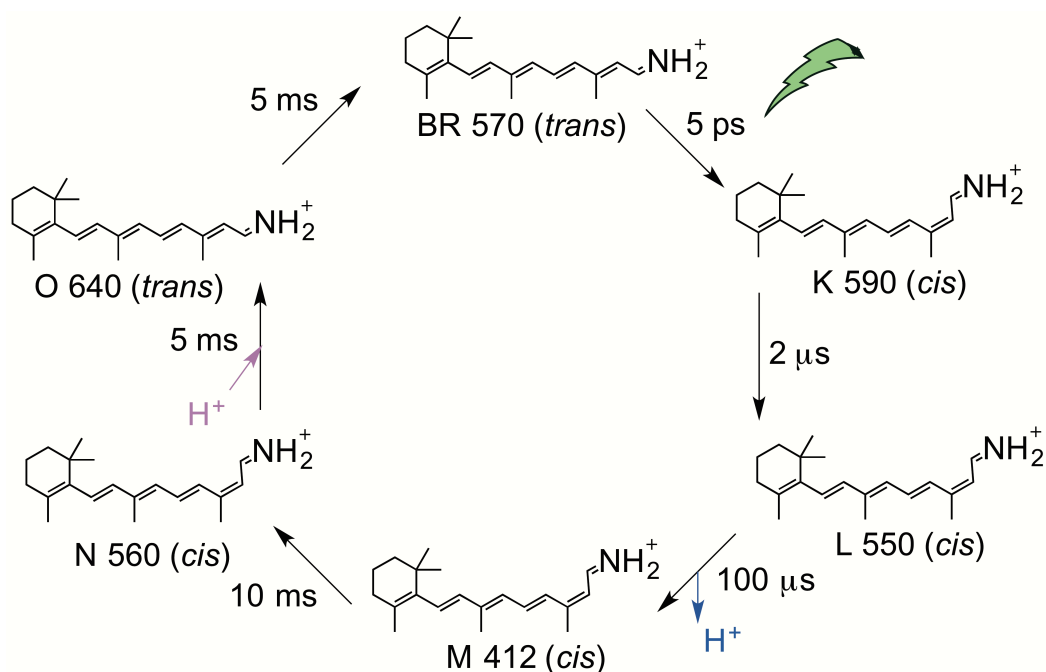
Figure 1.6. Bacterial rhodopsin retinal isomerisation upon photon absorption. The intermediates with conventional names: K, L, M, N, O are shown with their respective lifetimes.

1.4.2 Xanthopsin

The xanthopsin family of photoreceptors bind a cinnamyl chromophore (*e.g.* coumaric acid) *via* a thioester linkage to a cysteine buried deep within the protein. The prototypical xanthopsin is a small photoactive yellow protein (PYP). It was first isolated from the phototrophic bacterium *Ectothiorhodospira (Halorhodospira) halophile* [55] and later found in other purple bacteria. Xanthopsins have different roles in different organisms, including chalcone synthesis in *Rhodospirillum centenum* and initiation of a photophobic tactile response in *H. halophile*.[56,57] As yet, no interacting partner proteins have been

found and its physiological mechanism remains elusive.[58] One function suggested for PYP is the regulation of cell buoyancy as the PYP genes are associated with those for gas vesicle formation in *Rhodobacter* species.[59-61]

Compared to the photosensory domains of other photoreceptors, the photocycle of PYP has been studied in detail at the atomic level. Green photon absorption leads to the primary photochemical event, which is a *trans* to *cis* isomerisation around the sole double bond in the chromophore, as shown in Figure 1.7. The isomerisation occurs within a few nanoseconds to form a 465 nm-absorbing I_1 state. Two different intermediates, I_0 and $I_0^{\#}$, have been identified on sub-nanosecond timescales before formation of I_1. Within 500 µs I_1 is protonated concomitant with protein unfolding to form the second intermediate, I_2. The *cis*-intermediate I_2 thermally reverts to the *trans* state within 150 ms of the isomerisation. I_1 and I_2 are also referred to as pR and pB, respectively.[53,62]
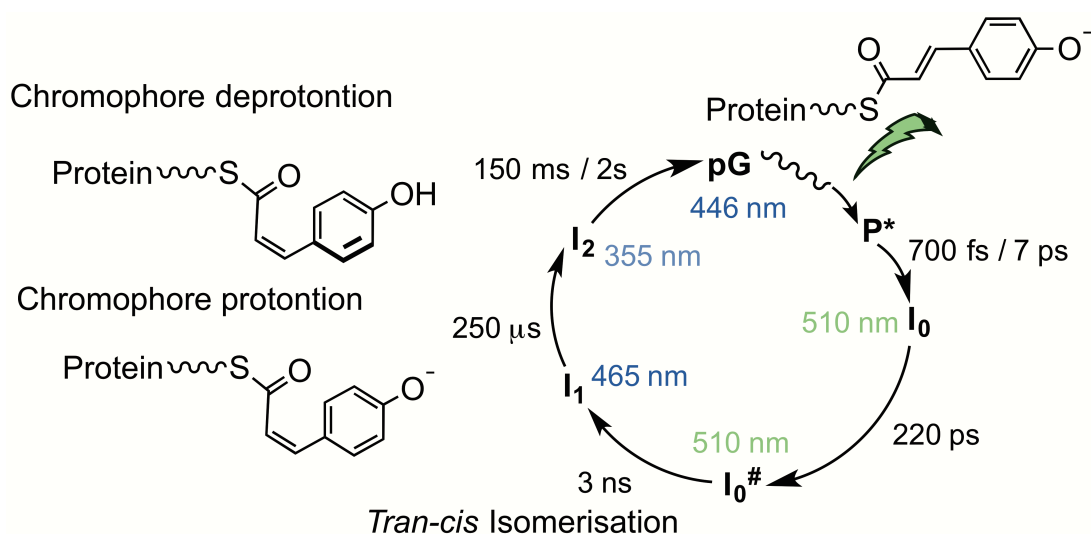
Figure 1.7. Photocycle of PYP. Intermediates are shown with the respective lifetimes and λ_{max}.

In the ground state the protein is constrained because the phenolate anion of the chromophore is stabilised by hydrogen bonds and the carbonyl end is pinned by the covalent thioester bond (Figure 1.8).[53,63-68]
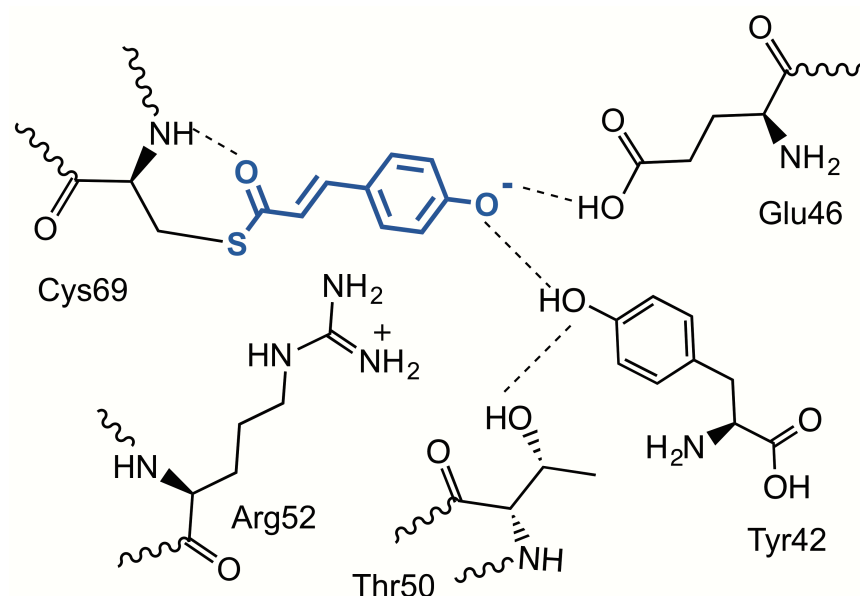
Figure 1.8. The structure of the PYP chromophore within its local environment.

Each intermediate formed upon relaxation from the nanoseconds to second timescale has been studied using time-resolved crystallography.[58,68,69] This revealed that the isomerisation event causes strain in the chromophore and protein. Relieving the strain leads to several changes in the tertiary structure, which start with the breakage of hydrogen bonds that stabilise the phenolate.

1.4.3 BLUF Proteins

BLUF domains contain the blue-absorbing FAD chromophore. BLUF proteins are usually dimeric and can be found with or without an effector domain, which are usually involved in cyclic nucleotide metabolism.[70,71] Absorption of blue light by the chromophore of the P_B state leads to conformational changes in the FAD, forming the signalling state on the sub-nanosecond timescale.[72] When the ground state absorbs a photon of blue light, an electron is transferred from a conserved tyrosine residue to the FAD within 7 ps to form a short-lived radical pair (Figure 1.9), involving FAD$^{\bullet-}$ and Tyr$^{\bullet+}$. Within 6 ps the anionic flavin radical is protonated to form the neutral semiquinone radical FADH$^{\bullet}$. A side chain glutamine then rotates, followed by the electron being returned to the tyrosine residue to form the slightly red-shifted signalling state, FAD$_{red\text{-}shifted}$ (P_R), within 65 ps (Figure 1.9). This state is slightly red-shifted as it differs in the hydrogen bonds it forms in the ground state.[73] The quantum yield of P_R formation has

been measured as 0.25 to 0.4 and it reverts back to the ground state within 5 seconds.[63,71-73]
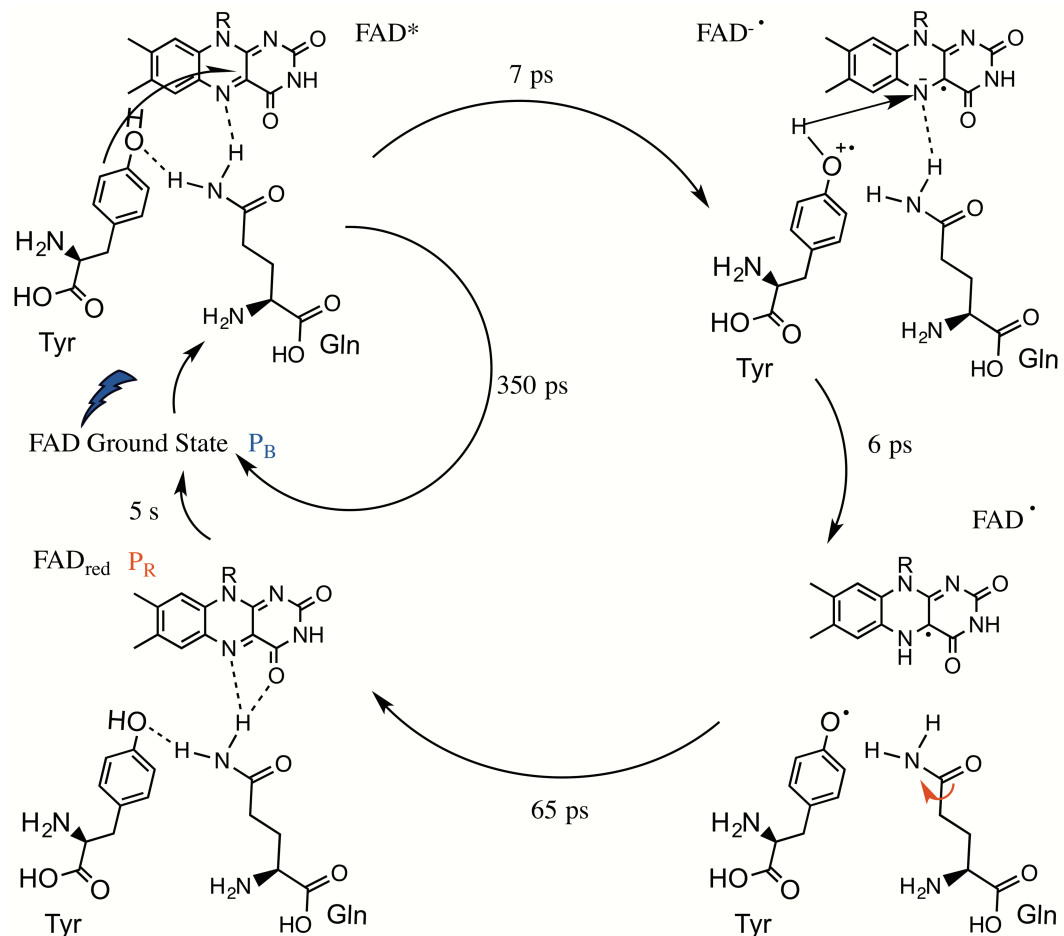
Figure 1.9. Blue light triggered transformation of FAD during BLUF photocycle as mediated by the conserved tyrosine and glutamine residues, forming the P$_R$ state from the P$_B$ state. Adapted from[2]

1.4.4 Cryptochromes

Cryptochromes also usually contain the blue-absorbing FAD chromophore and are known to modulate several processes within plants and animals, including entrainment by light of the circadian clock, circadian oscillators in animal brains and photomorphogenesis.[12,74-76] The photosensory domains of *Arabidopsis thaliana* cryptochromes, AtCry1 and AtCry2, are related to DNA photolyase. It is therefore called the photolyase homology region (PHR), but it has no DNA repair activity.[77,78] Unlike photolyases, cryptochromes contain a C-terminal extension (CCT) that may serve the function of effector domain.

Cryptochromes are thought to play a key role in magnetoreception within birds. The cryptochrome photoreceptor is proposed to process magnetic information for bird migratory orientation.[79] The magnetic compass is reliant on 373 nm, 424 nm, 502 nm, and 565 nm light for orientation in birds. 590 nm, 635 nm and 645 nm light illumination result in disorientation.[80-85] The radical-pair based magnetoreception mechanism proposes that the concentration of signalling-active radical, FADH•, formed during the cryptochrome photocycle (as shown in Figure 1.10) is altered by an applied magnetic field. This magnetic sensitivity results from the interconversion of the singlet and triplet state of the radical pairs.[86-88]
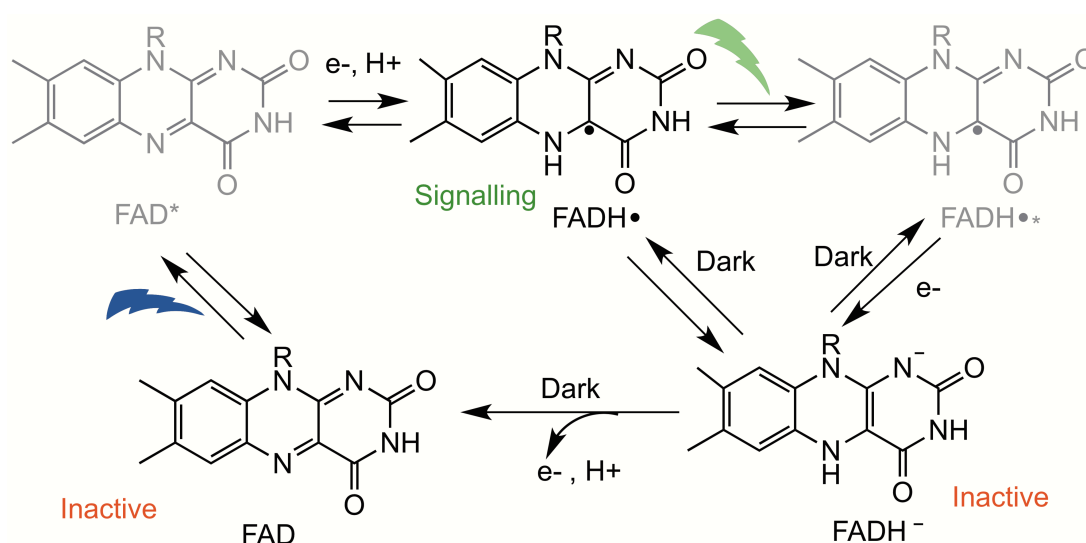
Figure 1.10. The formation of the cryptochrome signalling state, FADH•, and the signal quenching state, FADH-.

The FAD chromophore is non-covalently bound in the helical domain of cryptochrome's 16 helices bundle. Absorption of a photon of blue light by the fully oxidised form of FAD leads to the formation of a semi-reduced semiquinone intermediate. The semiquinone accumulates in the activated signalling form and is further fully reduced to FADH-, which ceases the signalling. Through dark reversion, the FADH- re-oxidises to FAD (Figure 1.10). Green light absorption by the FADH• radical shifts the equilibrium to the fully reduced FADH- form, which is the inactive form.[80,86,88-91]

Signal transduction *via* cryptochromes is known to modulate nuclear gene expression. However, plenty remains to be elucidated about the downstream

signal transduction of cryptochromes, such as the kinases, phosphatases and ubiquitin ligases required for light regulation of cryptochromes. A recent study has revealed that CRYII does not bind FAD *in-vitro* and is postulated to be a vestigial flavoprotein. CRYII does not have the structural features required to bind photoactive chromophore.[92] How this impacts on *in vivo* function however remains to be determined.

1.4.5 Phototropins

Phototropins contain LOV (light-oxygen-voltage sensing) domains, which also bind the blue-light absorbing FAD chromophore in a cleft lined with conserved residues. LOV domains are not only found in plants but also fungal and bacterial proteins. In plants, three classes of proteins utilise LOV domains. Phototropins 1 and 2 are the first family which modulate some of the fast responses, such as stomata closure and chloroplast movements,[93,94] whereas the Zeitlupe family control some of the slower responses.[95,96] The third family is the Aureochromes, which modulate photomorphogenesis in photosynthetic organisms.[97]

Absorption of blue light by the 'dark' LOV1-445/ D_{447} state of the FAD leads to formation of two triplet state intermediates, LOV1-715a and LOV1-715b, *via* intersystem crossing.[98] This process occurs on the picosecond timescale. These two intermediates decay with time constants of 800 ns and 4 µs respectively to a blue-shifted, LOV1-390/ S_{390} state through the formation of a thioether bond between the C4α and a cysteine residue, *via* a radical pair intermediate (Figure 1.11).[99-101] S_{390} is formed with a quantum yield of ~0.3 in phototropins[102] and thermally reverts to the dark state over hundreds of seconds.[103] The S_{390} is a signalling state that activates a kinase.[104] The S_{390} state can be photolysed with near-UV light to generate D_{447}.[105] The quantum yields and kinetics of different LOV domains vary in their photocycle[102] due to differences in the chromophore environment within the photoreceptor.
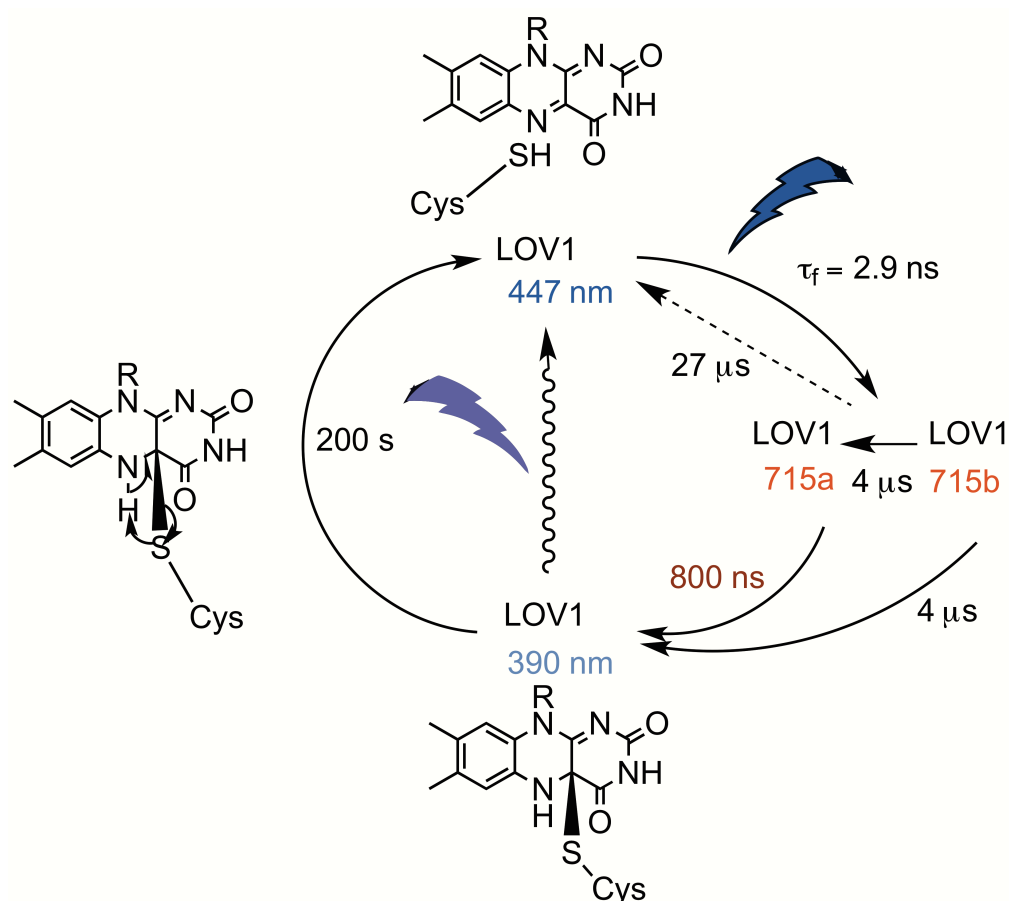
Figure 1.11. The photocycle of phototropin LOV domain; all intermediates shown with their respective lifetimes. Adapted from [106]

Phototropins contain two types of LOV domain, LOV1 transduces the photon-generated signal to a serine/threonine kinase and leads to the autophosphorylation of serine and threonine residues,[94,107] whereas LOV2 acts to attenuate the LOV1 signal.[108]

1.4.6 Vitamin-B$_{12}$ Containing Photoreceptors

Cobalamins are well-studied enzyme co-factors, which usually act as methyl- or adenosyl- donors. B$_{12}$-binding domains have also been found in many non-enzymatic proteins and 5′-deoxyadenosylcobalamin (AdoB$_{12}$) has recently been discovered to be essential for light sensing in the CarH gene repressor,[109-111] which controls expression of carotenoids in *Myxobacteria*. In the dark AdoB$_{12}$ binding leads to CarH tetramer formation, which binds to the operator and results in the repression of gene expression. Exposure to light of certain wavelengths leads to photolysis of the upper 5'deoxyadenosyl-Co ligand, which

leaves a large cavity and triggers reorientation of a helix relative to the B_{12} binding domain. This leads to tetramer disassembly and results in the activation of carotenoid gene expression. [112,113] The full photocycle of *Thermus thermophiles* CarH has been unravelled from femtoseconds to seconds timescale.[36] Upon photoexcitation of CarH two intermediates are formed, with tens of femtoseconds lifetime. These dissociate into radical pairs within 100 ps before recombining to form the ground state within 10 ns.

B_{12}-binding domains with no obvious output domains have also been discovered to control gene expression upon light absorption.[114] Many B_{12}-binding domains without output domains remain uncharacterised. Further studies of these domains will likely lead to expansion of the B_{12}-binding photoreceptor family.

1.4.7 UVR8

Due to its high energy, UVB light (280-315 nm) has damaging effects on biological molecules, such as DNA. However, plants have evolved to withstand the UVB induced stress. UVR8 is the only known plant photoreceptor that perceives UVB radiation and it's sensitivity to these wavelengths is due to a number of tryptophan residues. Upon UVB absorbance the UVR8 dimer dissociates into signalling-active monomers that interact with signalling components to modulate gene expression. UVB absorbance leads to electron transfer from tryptophan-233 to tryptophan-285 and then to arginine-338. The electron transfer steps occur concomitant with proton transfer from tryptophan-233 to aspartic acid-129, which leaves arginine-338 as a neutral radical. This disrupts salt-bridges involving these residues as well as other key salt-bridges. This leads to a decrease in the energy of protein-protein interactions and dissociation of the dimer into its monomer units,[38,115,116] which then interact with signalling components to modulate gene expression. Nuclear accumulation of the monomer also occurs, which interact with COP1 (Constitutively Photomorphogenic 1) through its C-terminal extension.[117-121] COP1 is a negative regulator of photomorphogenesis in *Arabidopsis thaliana* and it acts as an E3 ubiquitin ligase, targeting proteins for proteasomal degradation.[122-124]

1.5 Phytochromes

Phytochromes are a red and far-red light sensing family of proteins.[125] They play a major role in regulation of plant photomorphogenesis, flowering, seed germination and shade avoidance, among other physiological roles such as phototaxis and circadian rhythm.[126-129] As well as plants, phytochromes are also found in bacteria, denoted as bacteriophytochromes (Bphs), algae, fungi and cyanobacteria.[130,131] Although phytochromes have been predominantly found to act as photosensors that enable organisms to adapt to changing conditions, there are examples of phytochromes that serve a different function, such as a redox sensor.[132,133] A linear tetrapyrrole molecule, usually a variant of bilin, is the chromophore responsible for light sensing in these proteins. Heme oxygenases convert heme to biliverdin IXα, which is the precursor to the different variants of bilins (Figure 1.12).[134] The bilins in different phytochrome subfamilies can differ in their oxygenation states.
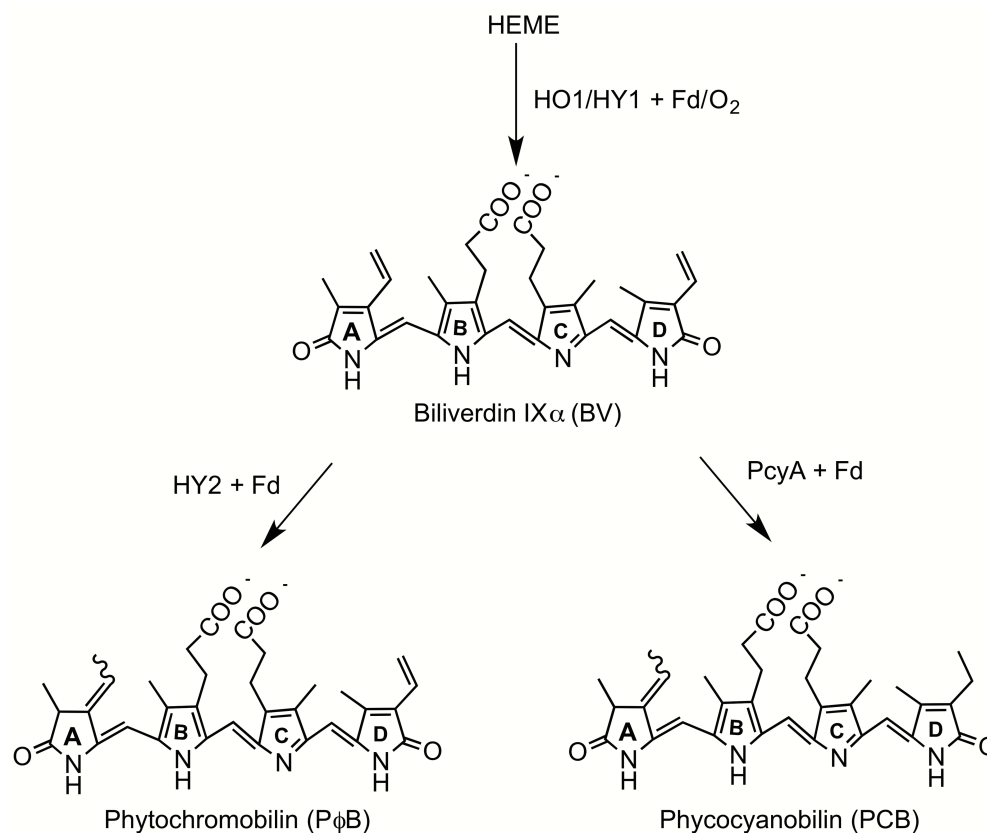
Figure 1.12. Synthesis of biliverdin IXα from heme by heme oxygenase. The reduction of bilin to phytochromobilin and phycocyanobilin is shown. Ferredoxin (Fd)-dependent bilin reductase (HY2) reduces biliverdin to phytochromobilin. Phycocyanobilin:Ferredoxin Oxidoreductase (PcyA) catalyses the four electron reduction of biliverdin to phycocyanobilin.

Phytochromes are modular proteins that are composed of a N-terminal photosensory core module (PCM) and a C-terminal effector domain.[135,136] The PCM region generally consists of 3 domains, termed PAS, GAF and PHY domains, although the Cph2 phytochrome from cyanobacteria is a PAS-domain-less phytochrome that maintains the GAF and PHY domains.[35,137-139] Cyanobacteria also have cyanobacteriochromes (CBCRs), a distant relative of phytochromes that only have GAF domains as their PCM and respond to a broad spectrum of light (see section 1.6).[140] More recently, phytochrome proteins have been found in eukaryotic algal genomes and were found to absorb green, red, orange, blue and violet light (Figure 1.13), presumably due

to changes in the GAF domains. As red and far-red light is attenuated at certain depths, phytochromes with evolved spectral tuning allow algae to sense light signals and to adapt to the changing environment.[131]
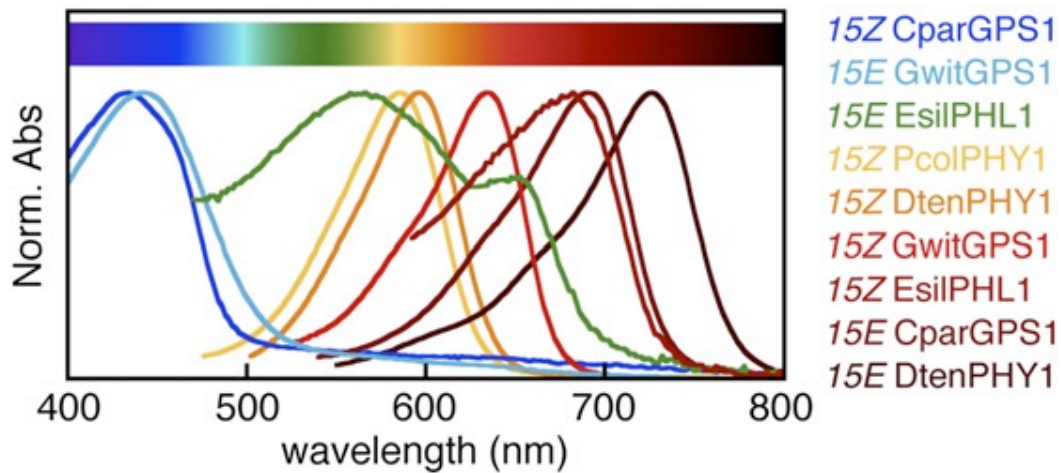
Figure 1.13. Spectra of the recently identified algal phytochromes. These phytochromes have broad spectral tuning that spans the entire visible spectrum. Spectra source: Rockwell *et al.* (2014).[141]

The effector domain in most phytochromes is a histidine kinase (HK), which autophosphorylates upon signal detection, although phytochromes with different HK domains have been discovered.[142,143] The HK is usually coupled to a response regulator (RR) to which the phosphate group is relayed (Figure 1.14). The RR is usually a transcription factor that is activated or inactivated by the phosphorylation. In its active form the RR binds to a promoter or operator to induce or repress gene expression. This modulation is referred to as a two-component signalling (TCS) pathway (Figure 1.14).
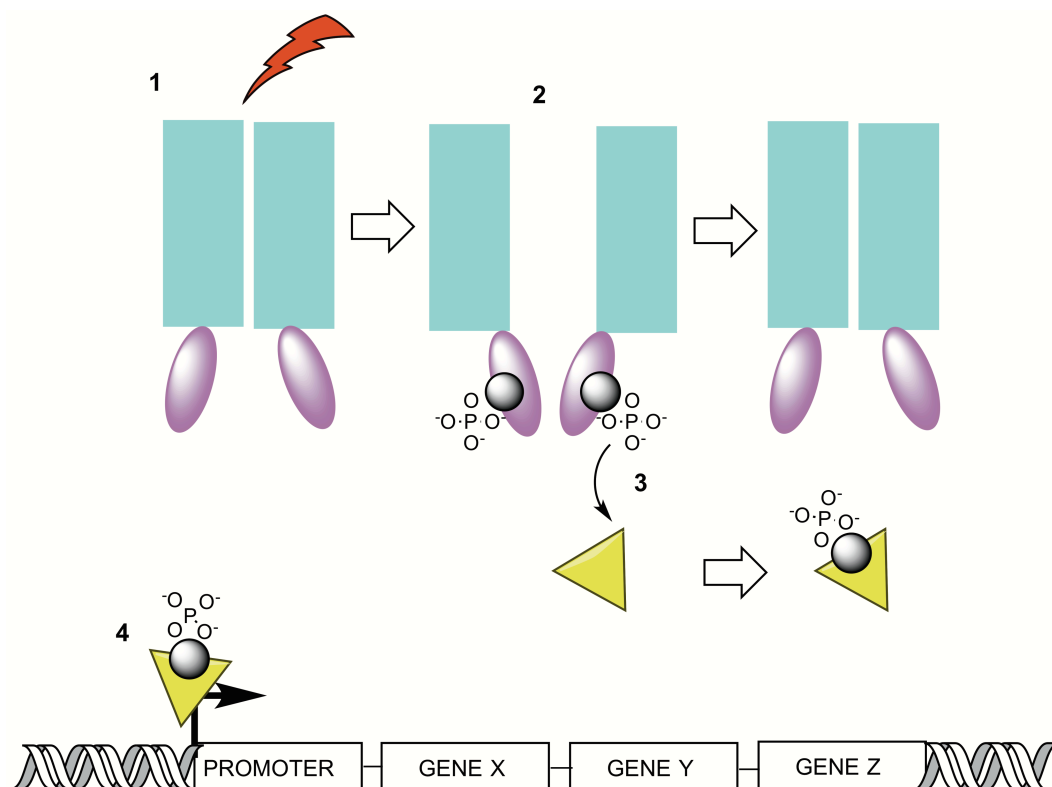
Figure 1.14. Schematic summary of phytochrome signalling. Signalling is initiated with light activated (as shown in step 1) autophosphorylation of effector domain (2), followed by phosphotransfer onto the response regulator (3), which is responsible for modulating gene expression (4). PCM units are represented by blue rectangles, effector domain represented by violet ovals, response regulator represented by mustard triangle.

1.5.1 Structure

Phytochromes from plants and cyanobacteria differ to Bphs in the substituents found on the pyrrole ring of the bilin and the mode of attachment of the tetrapyrrole ring to the protein. Bphs contain biliverdin IXα (BV) whereas cyanobacteria and plant phytochromes consist of phycocyanobilin (PCB) and phytochromobilin (PΦB), respectively (Figure 1.12). The linkage of the chromophore is made via an intrinsic bilin lyase activity of the GAF domain.[144] The bilin forms a thioether bond with a conserved cysteine residue within the GAF domain, which is found within the PCM.

The PCM of typical/plant-like phytochromes is composed of PAS (Per-ARNT-Sim), GAF (cGMP phosphodiesterase/ Adenylate cyclase Fha1) and PHY (phytochrome specific) domains.[58,135,145] The PAS domain is a sensor domain that senses oxygen, redox potential, light and small molecules. Some

phytochrome proteins, mainly from plants, may contain multiple PAS domains within the same protein, as shown in Figure 1.15. GAF domains are structurally similar to PAS domains but contain three additional helices that are at the dimer interface.[145] The C-terminal domain of these proteins is an effector domain, usually a HK, which has ATPase activity, although other types of output domains are also found in some phytochromes (Figure 1.15).
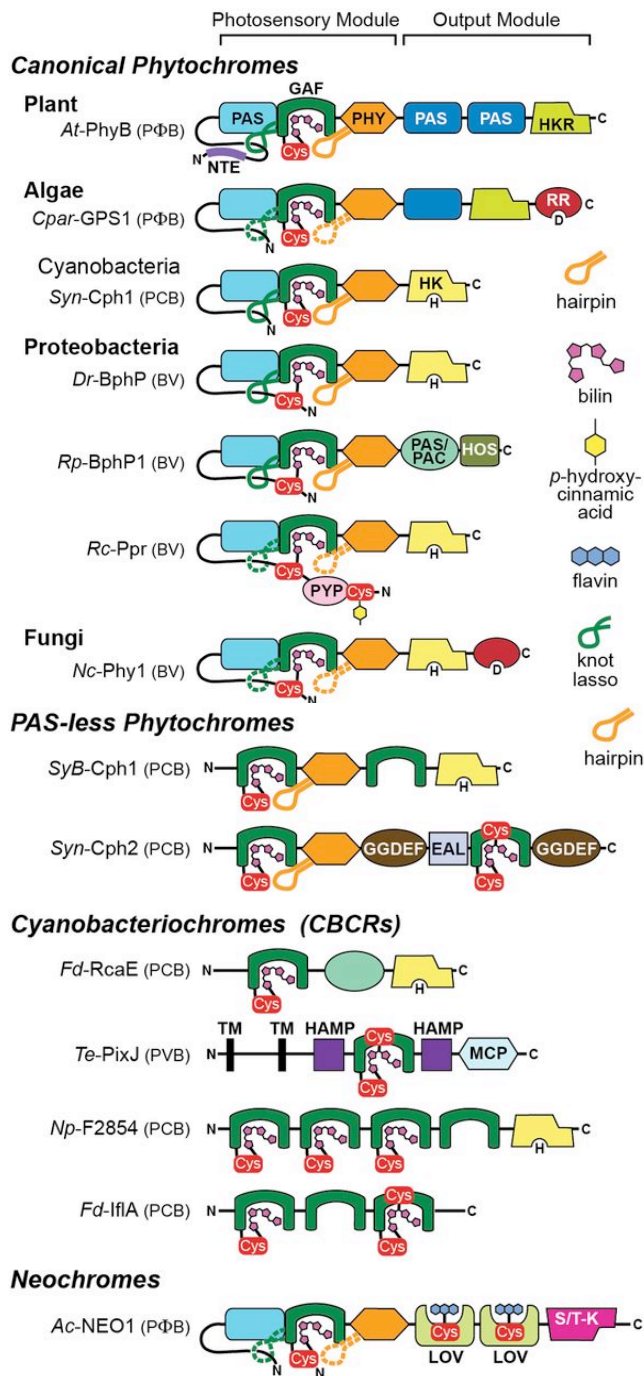
Figure 1.15. Domain architecture of plant, bacterial, algal and fungal, phytochromes together with cyanobacteriochromes and neochromes. The photosensory module consisting of PAS, GAF, PHY domains are shown along with domains that make up the effector module. Taken from [146]

1.5.1.1 Photosensory Core Module Structure

The first crystal structure of a classic red/far-red phytochrome was from non-photosynthetic *Deincoccus radiodurans*.[147] A detailed structure of the full PCM was obtained more recently from *Synechocystis sp.*, Cph1,[139] (Figure 1.16) and *Pseudomonas aeruginosa*, PaBphP.[148] These structures helped advance our understanding of the Pr to Pfr conversion.[149] [137,139,147,148,150-165] The *Deincoccus radiodurans* structural studies are reviewed comprehensively by Burgie *et al.* 2016.[150]
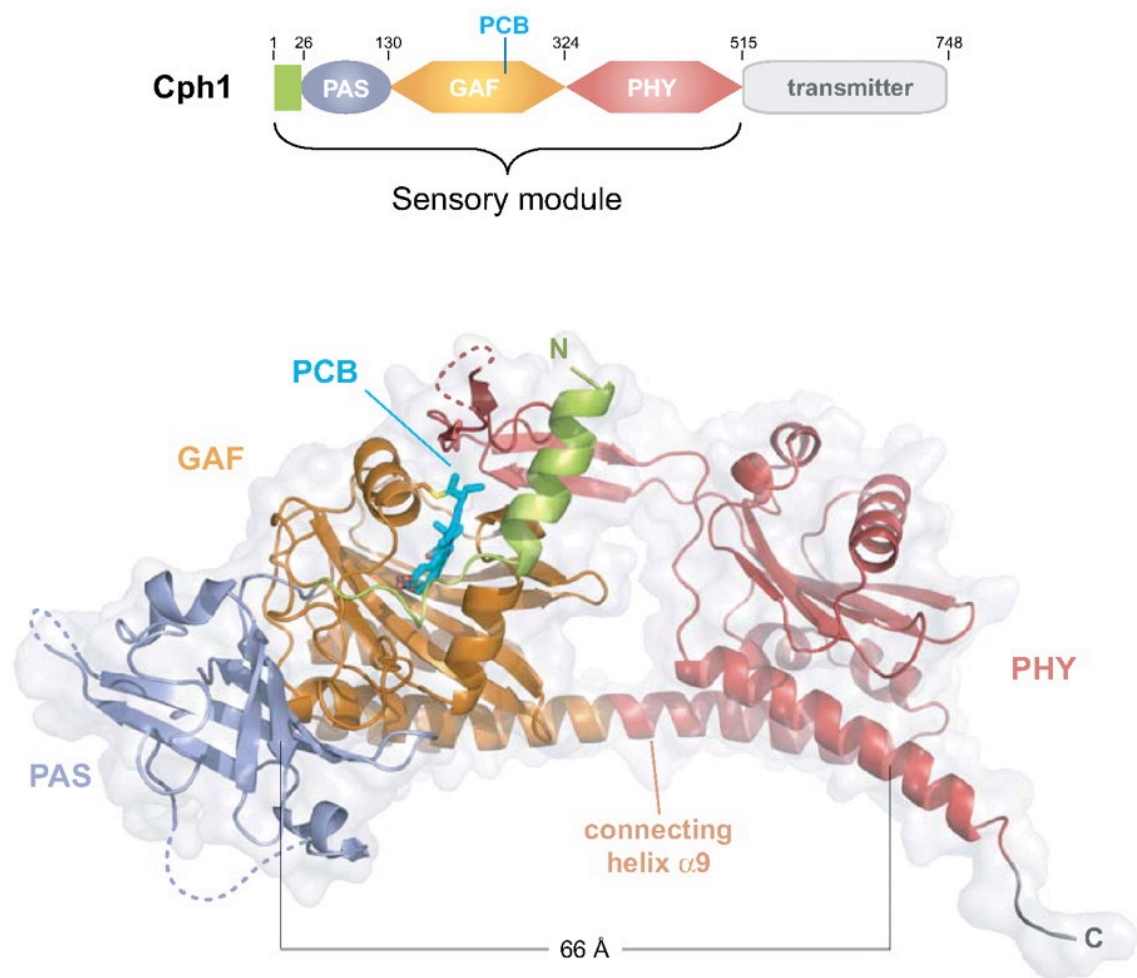
Figure 1.16. The colour coded domain architecture and structure of Cph1 from *Synechocystis 6803*. Disordered loop-regions are shown by dotted lines. Taken from Essen *et al.* (2008).

The PAS, GAF and PHY domains have the common fold of five anti-parallel β-sheets. The arrangement of these domains is linear yet the PAS and GAF

Page | 45

domains are integrated via an unusual Figure-of-eight knot structure that is conserved in all PAS-containing phytochromes, along with the extended arms formed by both the PAS and PHY domain.[139,147,148] The function of the knot structure is unknown and the knot is missing in PAS-less phytochromes. However, Cph2 PAS-less phytochromes contain the sequence of the GAF domain that forms the lasso portion of the knot.[138] The structures of the PCM have allowed insight into the interactions of the PHY and GAF domains and the similarities between these two domains. The αE-helix of the GAF domain is continuous with the αA helix of the PHY domain, thus a long continuous helix, 66 Å long, spans the entire PCM (Figure 1.16). The extended PHY arm serves to protect the bilin from solvent.[136,139,148]

The structure of the PCM of Cph1 from *Synechocystis 6803* consists of PAS-GAF and PHY domains. It has the conserved knot structure and a 46-residue protruding tongue structure between β16 and α15 of the PHY domain. The tongue extends back as a long β-hairpin from the PHY domain to the GAF domain, making intimate contact with the knot in addition to the GAF surface (Figure 1.17). This tongue structure is also conserved in all classes of phytochromes and it contains several highly conserved residues that are involved in interaction with the GAF domain. The tongue can vary in its length, but only at the tips. The tips only weakly interact with the GAF domain as it is partly disordered. The shortest tongues are found in the non-canonical PAS-less Cph2. The tongue covers the chromophore-binding pocket, which along with PAS and GAF isolates the chromophore from solvent. The conserved F475 residue of the tongue closes the chromophore pocket, thus shielding it from the solvent (Figure 1.17). The tongue also serves to stabilise the Figure of eight knot formed by the PAS and GAF domains.[139]
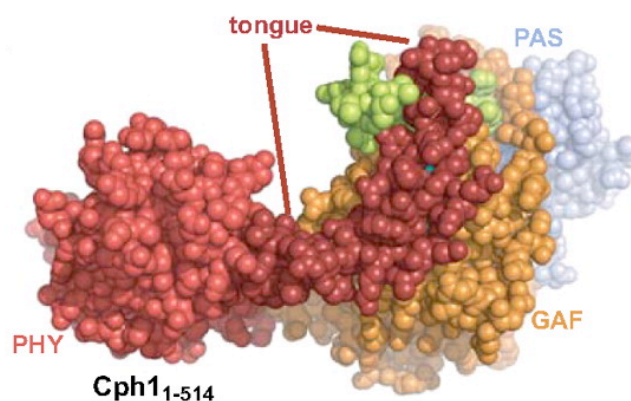
Figure 1.17 – Spacefill model of the Cph1 tongue structure (dark-red) and PCM from Essen *et al*. (2008). The chromophore (cyan) is hidden from the solvent by the tongue.

Photoexcitation of the red-light absorbing state (P_R) leads to isomerisation across the C15-C16 double bond (see section 1.5.3), which occurs concomitantly with flipping of ring D. Two Tyr residues, one above ring D and one at the side of ring D (Phe in Bphs), help to stabilise the far-red-light absorbing state (P_{FR}) conformation upon ring flipping by forming a hydrogen bond network with the carbonyl group and an Asp residue, which disturbs a salt bridge between the Asp and an Arg residue, shown in Figure 1.19.[163,166-168] It is believed that these rearrangements upon red light illumination, lead to a change in the fold of the tongue from a β-hairpin to an α-helix and cause the tongue to shorten as the distance between the GAF and PHY domain is reduced by 2.5 Å (Figure 1.18). The shortening of the tongue leads to dimer opening which allows autophosphorylation.[139,161,163,169] Takala *et al.'*s (2014) solution and crystal structures of the P_R and P_{FR} forms show that the PCM dimer opens up by several nanometres upon illumination of the P_R form.[163] Mutation of a conserved Arg residue (R472) to an Ala residue does not have a profound effect on the absorbance properties of Cph1 phytochrome but this mutant fails to dimerise in the P_R form. This indicates that the tongue plays an essential role in transmitting the light signal to the protein surface. This arginine residue hydrogen bonds to a main chain oxygen between two Gly residues (G451 and G452). R472 is thought to couple photoisomerisation led changes of ring D to the tongue conformation. Another Arg residue, R254, on the other side of the bilin forms a salt bridge with the ring B propionate and is also involved in mediating the signal transduction to the protein surface.[139,170]
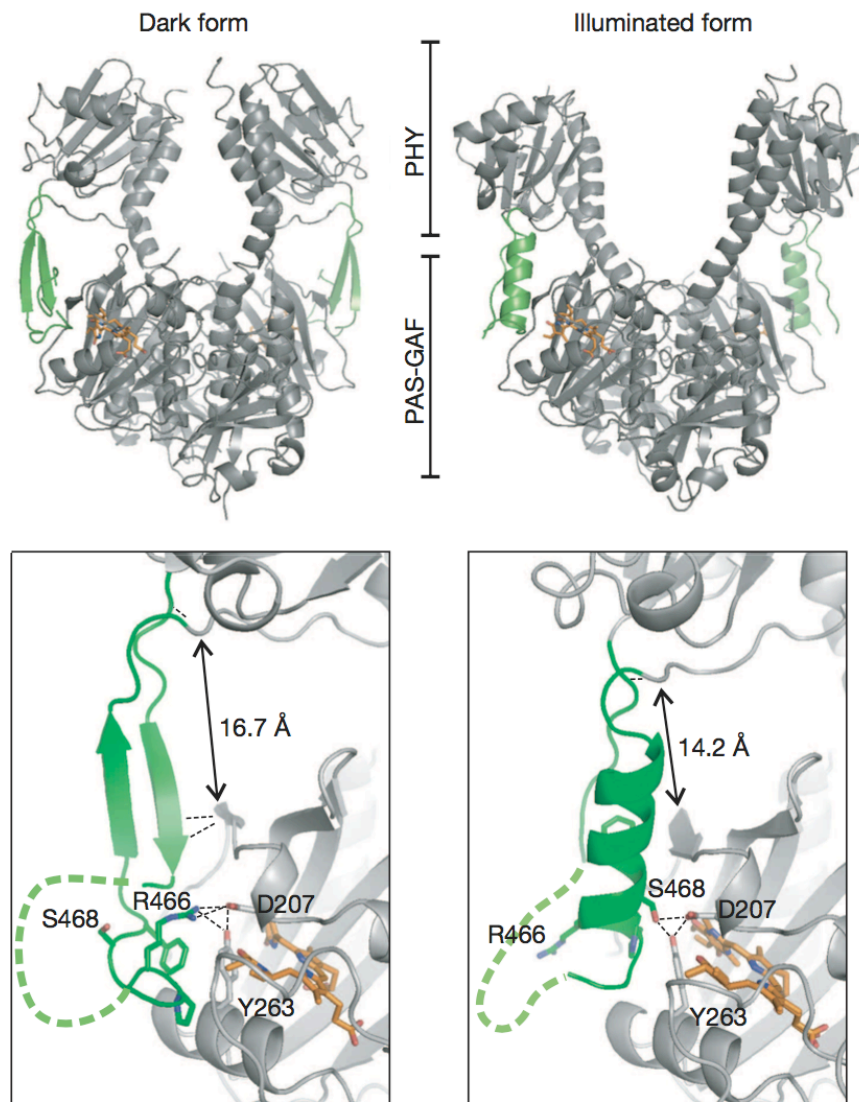
Figure 1.18. Crystal structures of *Deincoccus radiodurans* in dark (P_R) and illuminated state (P_{FR}). The PHY tongue (green) changes fold from a β-hairpin in the dark form to a shorter α-helix in the illuminated form. The interactions of the tongue structure are shown. Asp 207 and Arg 466 salt bridge visible in the beta hairpin conformation is disrupted in the α-helix conformation. The tongue shortens by 2.5Å upon fold-change. The chromophore (orange) is protected from the solvent by the PHY-tongue. Structure diagrams taken from Takala *et al*. (2014).[163]

The proposed structural mechanism relies on three factors:

1) The PAS, GAF and PHY domains, which are connected by the tongue are very rigid and cannot deform to accommodate for the variation in tongue length upon illumination.

2) The tongue and its junction must be rigid to maintain the GAF and PHY a fixed distance away from each other in the P_R form.

3) The long helix connecting PAS, GAF and PHY is important to direct the shortening of the tongue, which allows the dimer to open (Figure 1.19).

These requirements have been fulfilled by the structures revealed thus far.[139,163]

1.5.1.2 Full-Length Models

Several methods have been used to try to find the structure of full-length phytochromes, although up until recently, it has remained elusive. A recent low resolution structure of a full-length phytochrome from *Xanthomonas campestris*, which has a PAS9 effector domain has been published.[171] This structure shows a head to head dimer, in which the output module contributes to the dimer interface helices. Evans *et al*. (2008) proposed a small-angle-X-ray-scattering (SAXS) model by fitting the X-ray crystal structure of DrBphP-CBD and a HK domain into RpBphP2 envelope. SAXS analysis of full length RpBphP2 revealed that it has a Y-letter shaped structure. According to the SAXS model, each CBD/PCM is contained within one arm of the Y and the PHY domain and HK-braid twisting over one another and continue through the base of the arms. The ATP binding site of the HK opposes the phosphoacceptor His residue.[172] *Yang et al.* (2008) have taken the PaBphP PCM structure and the HK domain used in the SAXS model to propose a structure that takes less of a Y-letter shape. Hence, there is slight variation between the two different models. A recent electron microscopy study revealed two-fold symmetry of the full length DrBphP in addition to the parallel dimerisation interface. However, the full HK domain is not visible which can be attributed to the high flexibility of the HK domain. The twisted structure of HK at the neck of the full-length model also fits this data.[148]

Recent time-resolved X-ray scattering studies on full-length phytochromes have revealed that light induced signal transduction to the HK domains leads to rotational motion (Figure 1.19), which is postulated to lead to a degree of uncoiling of the HK monomers that are otherwise coiled together.[164] The coil is reminiscent of the α-helix linker, the angular orientation of which has been demonstrated to be essential for the HK functionality.[173] Structural changes prior to formation of the Meta-Rc intermediate state were not detectable by

time-resolved X-ray scattering suggesting the early re-arrangements are small and below the detection limit.[164] The study also suggests that the millisecond step within the PAS-GAF moiety controls the larger transformation of the entire molecule as this step is associated with the Meta-Rc state, which is itself associated with the deprotonation step.[174,175] Bjorling *et al.* 2016 suggest that this deprotonation controls the ultimate biochemical activity of the phytochrome. An additional final process is detected, however, the nature of this step remains unclear.[164]
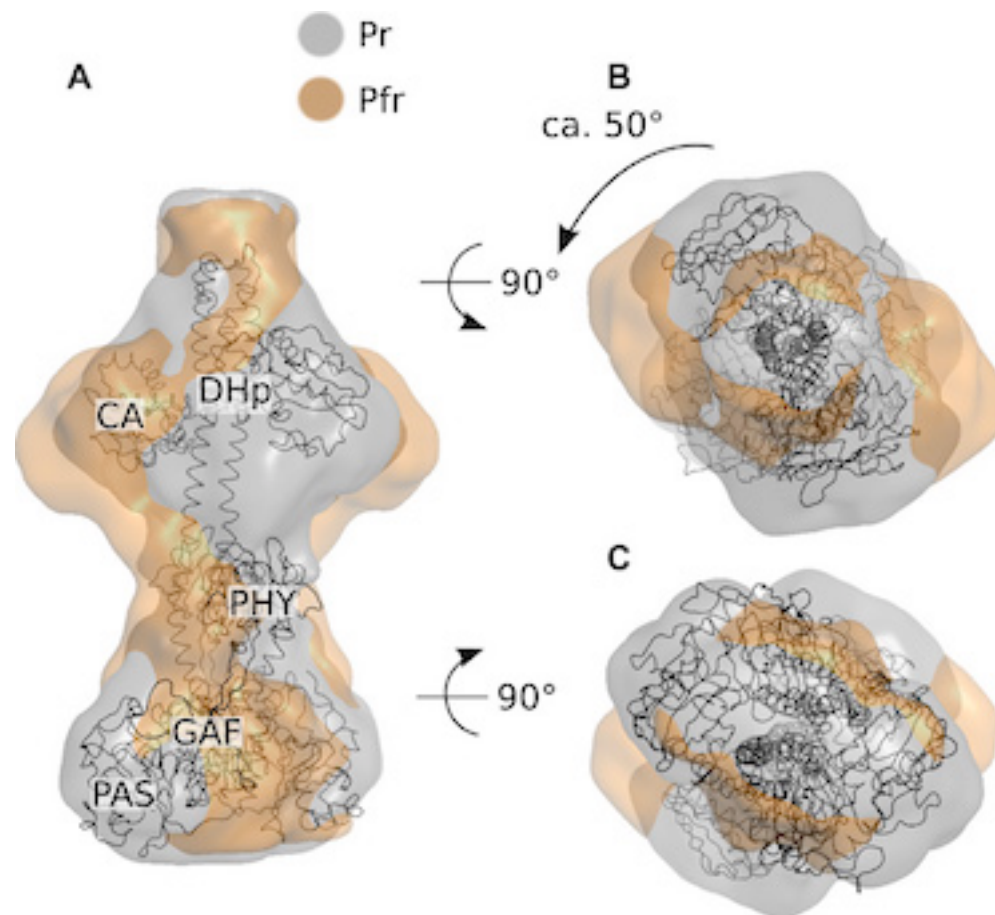
Figure 1.19. Cartoon representation of HK rotation relative to the PCM upon formation of P$_{FR}$ (gold) state from P$_R$ (silver) state as suggested by Bjorling *et al.* (2016). A) Side view, **B)** Top view and **C)** Bottom view of DAMMIN generated ab-initio and homology model of full length phytochrome, DrBphP, from *D. radiodurans*. Taken from [164]

1.5.1.3 Dimerisation

Phytochromes were initially thought to dimerise at the HK domain, however this was disproved when the crystal structure of the DrBphP PCM revealed that the PCM alone could form dimers.[176,177] Crystal structures have shown that PaBphP,[153] DrBphP[153] ,RpBphP3[151] and most bacterial phytochrome[148] PCMs form a head to head parallel dimer stabilised by a six-helix bundle and three helices from each GAF (Figure 1.20). Opposing Glu and Arg residues form a salt bridge to support dimerisation.[148] This salt bridge is broken upon ring D rotation when the phytochrome is illuminated. A recent study looking at dimer interfaces within the PCM of a *D. radiodurans* revealed two dimer interfaces, one formed due to the hydrophobic interactions at the GAF - GAF interface and the other between the HK domains. The interactions between the HK domains serve to stabilise the dimer.[178] Crystal structures of *P. aeruginosa* phytochrome suggest that the PHY domains approach each other.[148,149] However, it is likely that there may be differences in dimer interface interactions among different phytochromes.
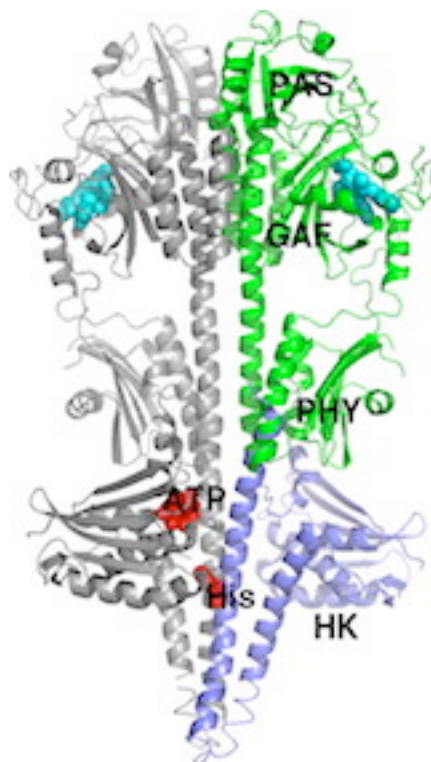
Figure 1.20. A model of full length *PaBphP* phytochrome showing the PCM and HK domains. Taken from [148]

1.5.2 Chromophore

The bilin chromophore ligates to apophytochromes autocatalytically, although in light of Essen *et al.'s* Cph1 structure, in which the chromophore-binding pocket is closed, it is unclear how the chromophore enters the pocket.[139] Data suggests that both Cph1 and Bph phytochrome chromophores adopt a *Z* configuration in the P_R state and *E* configuration in the P_{FR} state.[139,147,152] The ring A of the chromophore in plant-type and Cph1 phytochromes is attached to a conserved Cys in the GAF domain via a carbon-thioether linkage.[139] In contrast, the chromophore in Bphs is linked through two carbon atoms of ring A to a cysteine in the PAS domain. The Cph1 chromophore is less twisted than in the Bphs structures. Bound BV in the P_R form has a maximal absorbance at 700 nm, whereas PCB and PΦB bound proteins in the P_R state absorb maximally at 630 nm, which overlaps with the maximal absorbance region of chlorophyll. This potentially allows plants to sense shade from other plants and avoid competition.[179] Switching of BV to PCB and PΦB from bacteria to cyanobacteria and plants, respectively, is likely to be due to evolution of life from anoxygenic to oxygenic conditions, because BV may be susceptible to photodamage.[180,181]

1.5.3 Photochemistry

Phytochrome photocycles have been studied using optical spectroscopy since their discovery in the 1950s.[182-185] Techniques including Fourier transform infra-red (FTIR), UV/Vis absorption and resonance Raman spectroscopy have allowed characterisation of the intermediate states within the photocycle. The initial photochemical events that occur on picosecond time scales have been investigated using ultrafast Vis[186-191] and IR[192,193] transient absorption, and fluorescence spectroscopy.[194-196] Upon absorption of a photon of a specific wavelength of light, the phytochrome chromophore goes through an initial *E/Z* isomerisation and initiates a series of changes within the chromophore and surrounding protein matrix. All four pyrrole nitrogens of the bilin chromophore remain protonated in both the P_R and P_{FR} state.[197,198] There are no counter ions to stabilize the positive charge on these nitrogens, although His290 is

believed to be a proton sink/source, which hydrogen bonds via three waters to the nitrogens of rings A, B and C.[199] Although there is a significant transient proton release and uptake from phytochromes during the photoconversion, there is no obvious proton conductance channel.[139,174]

Time-resolved Raman and optical spectroscopy have been used to probe the mechanism of the photoconversion and the identity of the intermediates.[154,191] Low temperature and nanosecond spectroscopy studies identified three intermediates in the P_{FR} to P_R photocycle of Agp1 from *Agrobacterium tumefacien*s and cyanobacterial Cph1 with lifetimes ranging from nanoseconds to milliseconds (Figure 1.21).[154,164,174,175,181,200,201]
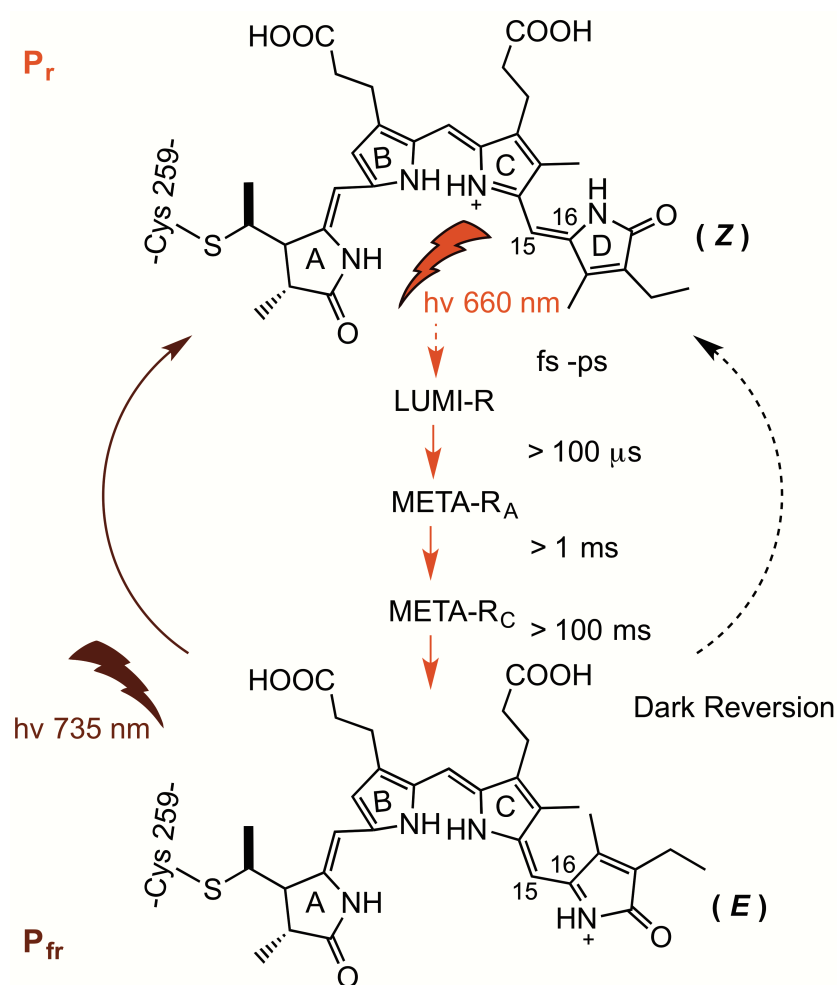
Figure 1.21: Postulated photocycle of Cph1 phytochrome. The timescales of formation of intermediates within the P_R to P_{FR} photoreaction are shown.

The proposed "flip and rotate" model explains the conformational changes within the chromophore and the surrounding protein upon isomerisation. The C15-C16 double bond isomerisation occurs on picosecond timescales concomitantly with the ring D flip which is coupled to rotamer changes of the aromatic residues around ring D [137,174,175,197,200]. Rotation of ring D leads to lengthening of the conjugated system, thus leading to the longer wavelength absorption maximum in the P_{FR} state. After the rotation of ring D, the entire chromophore and surrounding protein matrix rotates around an axis centred at ring A and perpendicular to rings C and B. This leads to changes in the interactions of the propionate substituents on rings B and C with the surrounding protein. Transition to a Meta-R intermediate *via* deprotonation of the D ring, and relaxation to P_{FR} through reprotonation occurs on a timescale of hundreds of microseconds. Observations in Cph1 and Phy A suggest a pronounced conformational change around ring D and the propionate side chains of ring C. However, Ulijasz *et al.'s* (2010) NMR study of SyB-Cph1 suggests pronounced changes around ring A, whereas ring D remains in the same conformation between the P_R and P_{FR} states in a PAS-less phytochrome [158]. Upon absorption of far-red light, after the initial photoisomerisation step, the reverse, P_{FR} to P_R, photoreaction does not proceed by the same reaction pathway as the P_R to P_{FR} photoreaction. P_{FR} can also thermally relax back to the P_R state; this can take several days, therefore P_R is referred to as the 'dark' state [197,200].

There are many notable photochemical differences between the Cph1-like and BphP-like phytochromes, which suggest that the initial change in direction of ring D is different in the two types of phytochromes. PaBphP residues Y163 and Y190 surrounding ring D are conserved in phytochromes (Figure 1.22). Substituting the residues equivalent to PaBphP Y163 in Cph1 and phyB confers a non-photoconvertible, fluorescent phytochrome.[148] However, the same is not observed in PaBphP or DrBphP. Circular dichroism studies confirm that BphP ring D rotates clockwise and Cph1 ring D rotates anti-clockwise.[202] The DrBphP Asp207 mutant is also unable to photoconvert and loses its excitation energy through fluorescence. The same is not observed in the equivalent Cph1 variant. This Asp residue is found in the highly conserved DIP motif, which is located below the chromophore.[169]
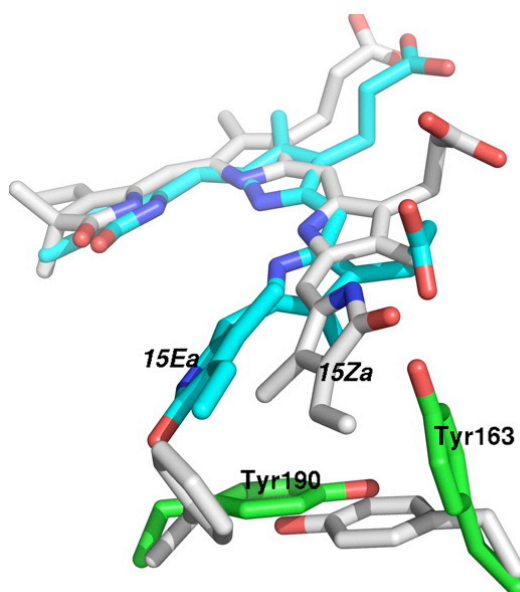
Figure 1.22. Conformation of key tyrosine residues (green) surrounding the chromophore within *PaBphP* phytochrome. Taken from [149]

1.5.4 Signalling Mechanism

The P_{FR} and P_R states of phytochromes interact differently with partner proteins and thus activate/inactivate different signalling cascades.[34] Phytochromes that have a dark P_{FR} state, rather than the P_R state, have been identified in nitrogen fixing symbionts and are called bathyphytochromes because they absorb longer wavelength light in their ground state. As far-red light is able to penetrate through soil, the dark adapted P_{FR} state responds well in the symbiont.[203]

Due to limited knowledge of the structure of a full-length phytochrome, any details of the signal transduction through phytochromes are inferred indirectly. HK is the most widespread regulatory domain among phytochromes, which suggests the evolution of phytochromes from a two-component system sensor with a tetrapyrrole-binding pocket. As the PHY domain is close to the chromophore this makes it a good potential signal transducer.[148] The distance between the kinase active site and the phosphate accepting histidine residue suggests an in-*trans* or *cis* autophosphorylation and this postulation has also been supported by electron microscopy data.[148,204-206]

The signal from the GAF is transduced to the PHY via the central helix bundle, this bundle continues through to the HK and activates a typical two-component phosophorelay system, which begins with an ATP phosphate transfer to a HK His residue. Studies on Cph1 have shown that both the length and the sequence of the α-helix linker between the PCM and HK domains is important for signal transduction.[207-210] The phosphate is then transferred to an Asp residue on a response regulator (RR). The phosphorylation affects the binding of the RR to its DNA or protein partner, thus eliciting a response. Some phytochromes have the RR on the same polypeptide as the HK, whereas other RRs exist as separate proteins. Partner RRs are not known for most phytochromes, although it is established that the signal transduction to the RR is via two component signalling.[137]

The RR coupled to Cph1 is Rcp1 (response regulator for cyanobacterial phytochrome 1). His538 is essential for autophosphorylation of the Cph1 HK domain and the phosphotransfer to Rcp1. The Cph1 P_R form has increased autophosphorylation and phosphotransferase activity towards Rcp1 compared to the P_{FR} form. However, this does not apply to all phytochromes. PrBphP2, 3, and 6, Cph A, Cph B and Agp1 show increased P_R dependent enzyme activity, whereas, Agp2 and PsBph show higher P_{FR} enzyme activity.[211-213] There is also variation between species. Rcp1 dephosphorylation may occur through phosphotransfer to another regulatory molecule.[214] The function of Rcp1 is currently unknown, as its target promoters are unknown.

Although Bphs and Cphs transduce signal through a phosophorelay mechanism, plant phytochromes have complex signalling mechanisms, which involve nuclear translocation and accumulation to form speckles within nuclear bodies, as shown in Figure 1.23.[140,214-216] This nuclear localisation signal (NLS) is at the C-terminal PAS domains and it can be deduced that that the regulatory domain serves to both maintain the homodimer in phytochrome signalling and for the nuclear localisation of the phytochrome.[217,218] It is hypothesised that P_R to P_{FR} conversion upon red light absorption leads to conformational changes that lead to autophosphorylation of a bound anchoring molecule (X), which leads to P_R: X-P dissociation and exposure of the NLS within the PAS domain. With NLS exposure, P_{FR} migrates to the nucleus where it interacts with

transcription factors and regulates gene expression. At least one of the dimers must be in the P_{FR} state for nuclear localisation. Once in the nucleus the $P_{FR}P_{FR}$ homodimer is more likely to compartmentalise to nuclear bodies (Figure 1.23), the function of which is yet unknown.[219]
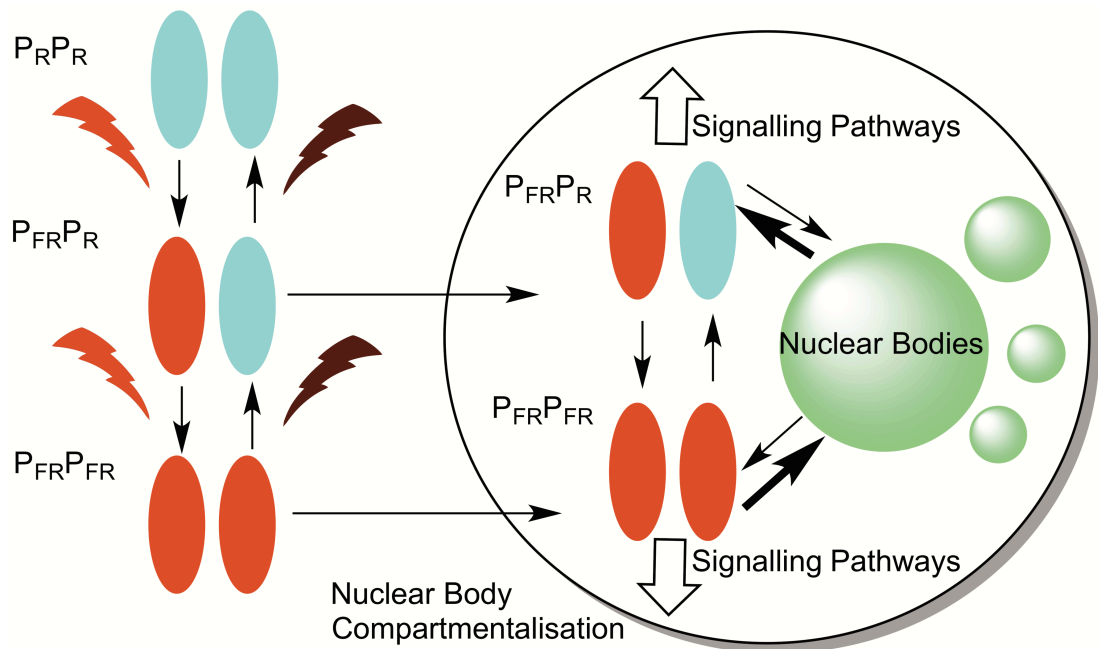
Figure 1.23. Schematic diagram illustrates nuclear localisation of phytochromes upon red light activation, using PhyB as a model. Heterodimers of $P_{FR}P_R$ and $P_{FR}P_{FR}$ homodimers translocate to the nucleus, where these regulate signalling functions through regulating gene transcription. Bold arrows illustrate that the $P_{FR}P_{FR}$ homodimers are more likely to form nuclear bodies.

Phytochromes with output domains other than HK have also been identified. The regulatory domain of these phytochromes is replaced with another catalytic or regulatory domains during the course of evolution.[132,143] As few such phytochromes exist it can be assumed that such chimeras don't remain functional. Neochromes have arisen from such domain exchange, combining the plant PCM with phototropin.[220] BphP1 from *Rhodopsuedomonas palustris* and *Bradyrhizobium ORS278* contain a second PAS domain instead of a HK and, BrBphP3 consists only of PAS-GAF-PHY domains.[221] These transduce the light signal via interactions with other targets that are currently unidentified. Phytochromes with MCP (methyl-accepting chemotaxis protein) domains and HAMP (histidine kinases, adenylyl cyclases, methyl binding proteins,

phosphatases) domains have also been discovered.[140,206,213,222-226] GGDEF domain, or GGDEF and EAL domain containing microbial phytochromes have also been identified. These domains have diguanylate cyclase and phosphodiesterase activity and regulate cyclic-dimeric-GMP levels.[213,227,228]

In *Arabidopsis*, phytochromes (PhyA, PhyB, PhyC, PhyD, PhyE) control the activity of phytochrome interacting factors (PIFs), which play largely negative roles in light responsiveness, such as repressing seed germination and promoting shade avoidance. In the red light absorbing state phytochromes post-translationally control PIFs either through degradation of PIFs or inhibition of PIF-DNA binding.[229-231]

Little is known about the physiological roles of many bacterial and cyanobacterial phytochromes. However, due to bacterial mRNA's polycistronic organisation, the partner proteins or RRs are encoded downstream on the same operon as the phytochrome. This genomic organisation has allowed identification of cognate RRs and downstream physiological function.[232] For example, DrBphP has been implicated in up-regulation of carotenoid synthesis in response to red light absorption,[232] regulator of chromatic adaptation (RcaE) is involved in complementary chromatic adaptation (CCA),[233] and BrBphP1 is involved in regulation of photosystem expression, which is involved in N_2-fixing in *Bradyrhizobium*.[142]

1.6 Cyanobacteriochromes

1.6.1 Structure and Photocycle

Cyanobacteriochromes (CBCRs) are distant relatives of phytochromes that mediate phototactic and photochromatic responses in cyanobacteria.[140,234] Although CBCRs use the same photochemical mechanism, namely *E/Z* isomerisation around the C15-C16 double bond concomitant with ring D rotation,[147,148,170] the absorption properties of the *E* and *Z* states, output domains and signal transduction pathways of CBCRs are different to those in phytochromes.

Phytochromes require a PAS-GAF-PHY PCM for their photochemistry whereas the CBCR GAF domain alone is sufficient to achieve reversible photochemistry. As CBCRs don't contain PAS and PHY domains, rings A and B of the bilin are solvent exposed. CBCRs can also contain multiple GAF domains, allowing integration of multiple light signals. Key protein-chromophore interactions are also conserved between CBCRs and phytochromes including a conserved tyrosine around ring D that forms a hydrogen bond with the carbonyl oxygen in the Z isomer state,[139,147,206,235] and an Asp residue hydrogen bonded to the D-ring amide in the 15E state.[148,149,235]

Based on their photocycles and structure, four CBCR families have been identified, known as insert-Cys, red/green, DXCF, and green/red. As CBCRs can integrate multiple light signals, they provide complete coverage of the visible and near-UV spectrum and thus have a diverse range of photocycles.[236-239] For example, the DXCF family has known blue, teal, green, yellow and orange light-absorbing 15E photoproducts and violet, blue and green light-absorbing 15Z dark states. This is because DXCF CBCRs can autocatalytically isomerise the phycocyanobilin chromophore into phycoviobilin.[238,240,241] DXCF CBCRs usually contain a mixed population of PCB and phycoviobilin (PVB). The chromophore is bound in a GAF cleft formed by 6 stranded anti-parallel β sheets and 3 proximal α helices.[225] The photochemistry and photocycles of many CBCRs have been studied in detail.[242-245]

Insert-Cys and DXCF CBCRs use a conserved "second cysteine" residue to form a thioether linkage to the C10 of the chromophore, between rings B and C.[246,247] This allows detection of the near-UV/blue light in the 15Z dark state as the double cysteine linkage alters the conjugation through the bilin molecule.[246] Green/Red CBCRs, such as CcaS and RcaE, have photocycles that involve proton transfer between the chromophore and surrounding protein matrix and this leads to a big change in peak absorption upon photoconversion.[248] Table 1.3 shows properties of selected cyanobacteriochromes.[140]

Table 1.3: Properties of CBCRs. Adapted from Ikeuchi *et al.* 2008.[140]

CBCR	Chromophore	Photoconversion	Physiology
AnPixJ_GAF2	PCB	Pr (648nm, *Z*) Pg (543 nm, *E*)	Phototaxis?
SyCikA	?	P_V (400 nm) P_Y (565 nm)	Circadian Rhythm?
SyPixJ1_GAF2	PVB	Pb (430–5 nm, *Z*) Pg (535 nm, *E*)	Phototaxis
SyScaS_GA	PCB	Pg (535 nm, *Z*) Pr (672 nm, *E*)	Induction of Antenna
TePixJ_GAF	PVB	Pb (433 nm, *Z*) Pg (531 nm, *E*)	Phototaxis

A CBCR that has been studied in detail, due to the crystal structure of the GAF domain being solved, is PixJ (DXCF family) from *Thermosynechoccoccus elongatus* (Te). It is a functional homolog of SyPixJ1, which modulates phototaxis in the unicellular motile *Synechocystis* sp. PCC 6803.[223,249] It photoconverts between a blue light-absorbing P_B (428 nm absorption max.) state and a green light-absorbing P_G (529 nm absorption max.) state.[234,240] TePixJ distal α-helices form a helical bundle to form a parallel dimer. This CBCR transposes the PCB C4=C5 double bond to C2-C3 to form PVB. TePixJ links to PVB via a second thioether bond between the C10 and Cys494 residue as well as the C3 to Cys522 linkage.[224] This leads to a bisected π conjugation system in the bilin, which yields a blue-light absorbing system. Blue light irradiation causes *Z* to *E* isomerisation of the C15-C16 double bond with the concomitant ring-D flip. This is followed by C10-Cys494 linkage rupture and migration of the chromophore within the GAF domain.[224,235,246,250,251] The light signal is transduced to the HK, PixL.[223]

1.6.2 Signalling Mechanism

CBCRs can contain multiple different types of output domains, which are typically involved in signalling systems of chemotaxis. Methyl-accepting chemotaxis protein (MCP) signalling domains and histidine kinase, adenylate

cyclases methyl binding proteins, phosphatases (HAMP) domains are found in TePixJ, SyPixJ1 and AnPixJ, and control flagellar motility.[136,206,213,222,225,227] Other CBCRs use RRs to control different physiological functions, such as complementary chromatic adaptation (CCA). One such photoreceptor that regulates CCA in *Synechocystis* is CcaS, which is a PCB-containing CBCR that absorbs light in the green and red region of the visible spectrum. CcaS is more efficiently phosphorylated in the P_R form than in the P_G form. The phosphate is transferred onto the RR, CcaR, which results in a change in the DNA binding affinity of CcaR. CcaR binds to the operator within the promoter of CpcG2 gene and activates its transcription. CpcG2 is a unique variant of a linker peptide of phycobilisome called CpcG1. CpcG1 containing CpcG2 linker forms a CpcG2-phycobilisome supercomplex (PBS), which transfers light to photosystem I, whereas CpcG1-PBS transfers light to photosystem II. The expression of the *cpcG2* gene triggered by green light suggests that CpcG2-PBS accumulation occurs to compensate for inefficient light-harvesting by photosystem I by the chlorophylls. This is because green light excites PBS more efficiently than chlorophylls.[237]

1.7 Optogenetics

Optogenetics is a fast-developing field focused on exploiting light to control processes occurring within specific cells in a non-invasive manner with fine spatial and temporal control using genetically encoded light-activated proteins. Optogenetics has traditionally been used in the field of neuroscience to modulate the activity of a defined population of neurons with millisecond precision to probe their function. It has been an indispensible tool for neuroscientists to study neural networks and brain function.[252] However, there are certain shortcomings of this technique that need to be addressed and optimised. These shortcomings are all related to targeting, in other words the precise delivery of the optogenetic probe and precise delivery of the light. Microbial opsins were the first photoreceptors to be widely explored as optogenetic tools and have the widest range of applications. Channelrhodopsins produce action potentials upon photon-absorption and this has led to their widespread use in neuroscience to activate and silence specific cells with fine spatial and temporal precision.[253-255]

Not only can optogenetics be used to modulate and probe the activity of specific cells, it can be applied to modulate protein-protein interactions and protein-DNA interactions. Recently, a group from Tokyo engineered a photoactivatable CRISPR-Cas9 system for optically modulated genome editing.[256] Through designing a chassis that allows scientists to "plug and play" phytochromes with varying spectral specificities, the CRISPR-Cas9 can be used to target multiple genomic sequences at the same time. Another study targeted cell signalling by introducing a light activated cAMP synthesis enzyme into sperm cells lacking the endogenous enzyme to restore sperm cell motility, hence reinstating fertility.[256]

1.7.1 Plant phytochromes as optogenetic tools

Levskay*a et al.* (2009) engineered an optically modulated system utilising the plant phytochrome PhyB, which reversibly associates with phytochrome interaction factor 3 (PIF3) upon illumination with 650 nm light. The study

demonstrated that this dimerisation interaction could be used to anchor YFP-tagged PIF to mCherry tagged-PhyB, which is anchored to the plasma membrane. The group constructed a PIF-Tiam-DH-PH chimaera from PIF3 domain and the catalytic domains (DH-PH) of Rac guanine nucleotide exchange factors (GEFs), Tiam, which is involved in spatial regulation of the actin cytoskeleton at the polarised edge of motile cells. When these constructs are co-transfected into fibroblasts, illumination with red light leads to association of membrane anchored PhyB with PIF3 that is fused to Tiam-DH-PH. This translocation of Tiam-DH-PH to the membrane activates it, causing production of lamellipodia within 20 minutes (Figure 1.24). As phytochromes are reversible photoreceptors, illumination with far-red light could disrupt the interaction between PIF3 and PhyB. However, the recruitment is only reversible by 750 nm light in constructs containing tandem PAS repeats.[257]
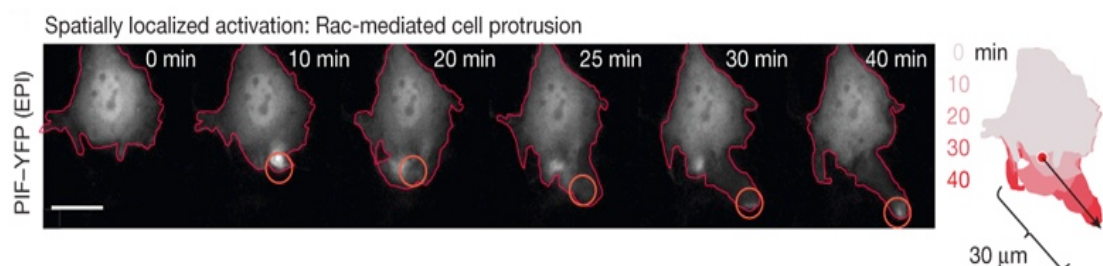
Figure 1.24. Rac-mediated cell protrusion photoactivated by red light. From Levskaya *et al.* (2009).

In a similar study, Levskaya and colleagues used light-dependent association of PhyB and PIF6 to translocate Tiam and intersectin to the plasma membrane of fibroblast cells.[257] Tiam and intersectin activate the GTPases Rac1 and Cdc42, respectively. The activation of these GTPases causes formation of cell protrusion upon red-light illumination. This study built upon an earlier study conducted by Shimizu and colleagues who demonstrated that photoactivated dimerisation of PhyB fused to GAL4 DNA-binding-domain (GBD), and PIF3 fused with GAL4 transactivation domain (GAD), enables transcription of the luciferase reporter gene (Figure 1.25).[258] This is because dimerisation of PIF3 and PhyB pull the GBD and GAD together, which activates LacZ expression. The group showed that the level of luciferase expression can be modulated by titration of the number of photons delivered by each shot.[258]
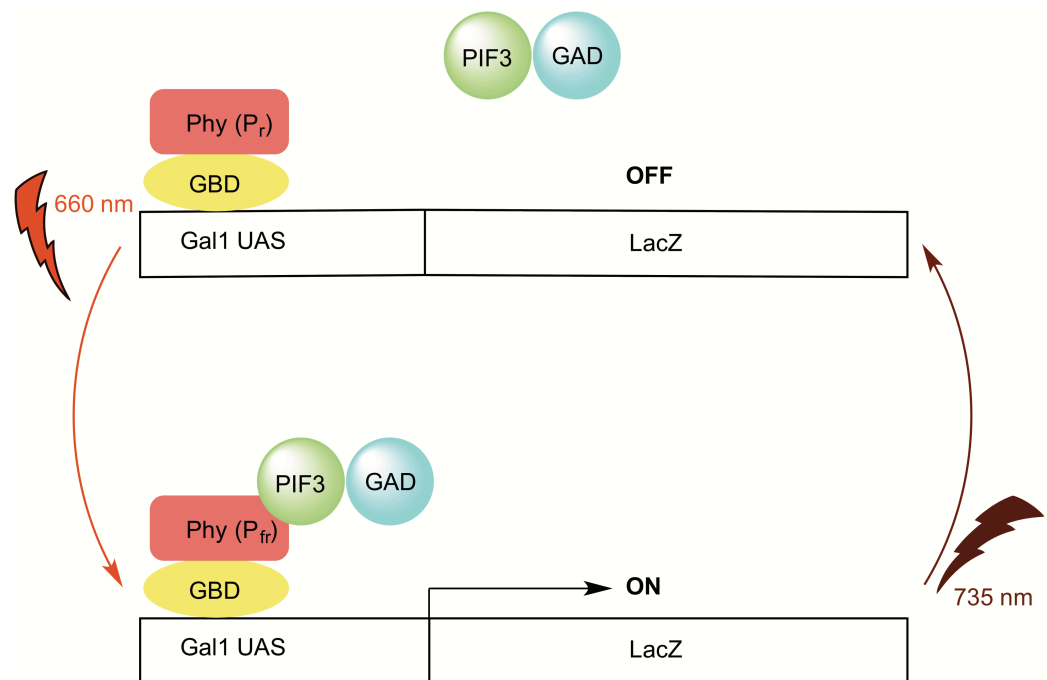
Figure 1.25. Schematic showing red light-activated dimerisation of GBD-fused PhyB with GAD-fused PIF3. This dimerisation is confirmed by activation of LacZ expression. Adapted from Shimizu *et al.* 2002.[258]

Yazawa *et al.* (2009) engineered a light-dependent dimerisation system that activates a transcription factor. In this study, GI was fused to Gal4 DNA binding domain and FKF1 was fused to VP16 transactivation domain. FKF1 and GI are proteins that control flowering in *Arabidopsis thaliana* and FKF1 is a LOV-domain containing photoreceptor. FKF1 and GI bind together upon absorption of a 450 nm photon, which brings together Gal4 and VP16 to activate the transcription of a luciferase reporter gene. The activity of luciferase could be increased 1000-fold upon red light illumination.[259] These examples demonstrate how light can be used to control protein interactions with high spatial and temporal precision.

1.7.2 Cyanobacterial phytochromes as optogenetic tools

Levskaya *et al.* (2005) constructed a red light-dependent gene expression system using Cph1 PCM from *Synechocystis sp*. 6803. The PCM of Cph1 was fused to the HK section of *Escherichia coli (E. coli)* EnvZ through a common

coiled coil linker. The group fused an *E. coli* HK to the photosensor because the RR of photoreceptors do not have known DNA-binding domains. This EnvZ HK is part of the EnvZ-OmpR two-component system, which regulates porin expression in response to osmotic shock. Aligning the sequences of Cph1 and EnvZ identified potential crossover points between Cph1 and EnvZ. The length and composition of the linker peptide that links PCM to the HK can affect signal transduction and therefore, a series of chimaeras with varying linker lengths were constructed. These were transformed into a ΔEnvZ *E. coli* strain. The strain also consisted of a chromosomal fusion between the *ompC* promoter and a luciferase reporter, which produces a black compound. OmpR regulates the *ompC* promoter. As *E. coli* doesn't produce bilin, the genes responsible for PCB production were introduced into the strain. The chimaeras were activated at 37°C using broad-spectrum light. Light repressed the gene expression, leading to the production of a high-contrast replica of the image projected on the biological film.[207] This system has been developed further to engineer a red-green TCS sensor.[260]

1.7.3 Cyanobacteriochromes as optogenetic tools

More recently, Voigt and colleagues reported the development of a green/red photoswitchable TCS from cyanobacteria.[260] Under green light illumination, the TCS induces the expression of a phycobilisome-related gene. It consists of a membrane-associated HK and its regulator is CcaR. Absorbance maxima are at 535 nm in the green region and 672 nm in the red region of the visible spectrum. Absorption of a 535 nm photon leads to CcaS autophosphorylation and phosphotransfer to CcaR, which activates the transcription of the phycobilisome linker protein CpcG2. Absorption of red light quenches the activation.[237] The group co-expressed the CcaS/CcaR system with their previously built system, Cph1-EnvZ, discussed above to achieve multichromatic control of gene expression. As the two systems have different chromatic specificities and functional outputs, they should be able to function alongside. CcaS is inactivated by red light, and red light activates Cph1/EnvZ system, thus light can be differentially applied to specifically induce each system. The group amplified the *ccaS/ccaR/cpcG2* cassette from *Synechocystis PCC 6803*, cloned it into a vector and replaced CpcG2 with LacZ. Both systems were delivered

into *E-coli* and two-colour optical control of gene expression was demonstrated (Figure 1.26).[260]
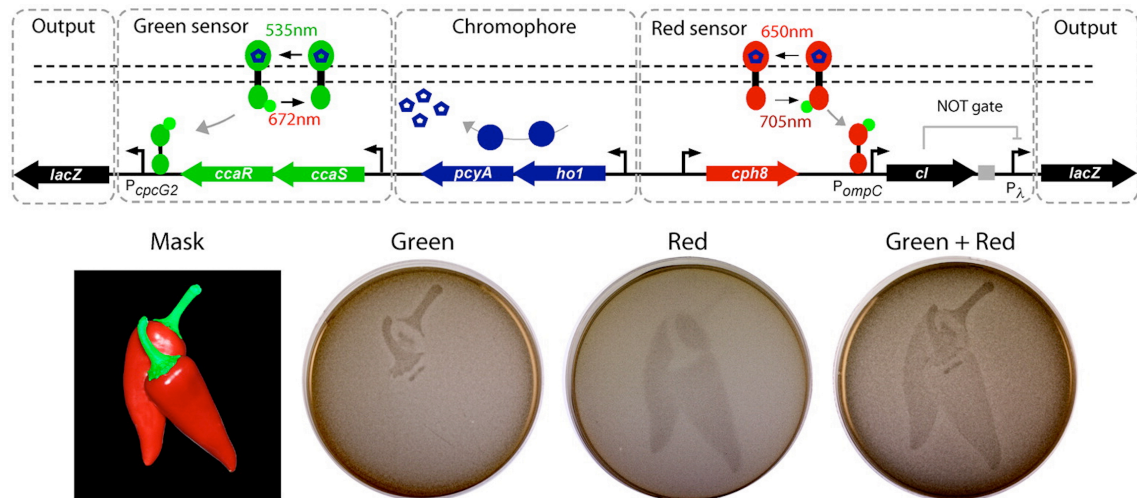
Figure 1.26. Tabor *et al.'s* (2011) two colour bacterial photography. Simplified configuration of the DNA encoding the constructed systems is shown at the top. A composite green-red image was projected onto the three plates containing a lawn of green sensor (CcaS/CcaR) only strain, red sensor (Cph8/EnvZ) only strain and both sensors-containing strain. Taken from Tabor *et al.* 2011 [260].

Tabor and colleagues optimised and refactored these Cph1/EnvZ and CcaS/CcaR TCSs further.[260] Schmidl *et al. (2014)* compressed the four genes of each system onto two plasmids and co-expressed them in the same cell. The independent control of each TCS was retained and resulted in highly accurate and precise quantitative, spatiotemporal control of gene expression. The group engineered a lawn of *E. coli* to perform a challenging image processing algorithm of edge detection by linking Cph8-OmpR to a cell-cell communication system. To move the expression cassette from one plasmid to another, the dosage of the *cph8* gene was reduced. To counteract the consequences of this, the *cph8* gene was put under the control of a constitutive promoter and the ribosome binding sites (RBSs) were replaced with computationally designed synthetic RBSs. Furthermore, several mutants of the *cph8* were monitored and the best performing mutant, *cph8* G772V was carried forward in the study. The promoter of *ompC* was reduced in size to remove unwanted regulatory elements to reduce the DNA footprint of the system. This study demonstrates

Page | 66

the importance of refactoring and optimising when constructing systems for heterologous expression [6]. In addition, the UirS UirR system, which responds to UV and green light, has also been repurposed for regulating expression within gene circuits.[261] It is the only UV-Violet responsive transcription regulatory tool.

Argeitis *et al.* (2016) engineered a fully automated platform for precise long-term expression of media in liquid cultures.[262] The group dynamically regulated the culture growth rate by putting the intracellular supply of methionine under optogenetic control to demonstrate the potential of optogenetic feedback. MetE regulation generates a feedback loop between the global cell physiology and *metE* gene expression. The gene encoding methionine synthase, *metE*, was modulated using the CcaS CcaR system and tailor-made hardware and software. Despite large global perturbations in the system (*e.g.* change in growth media, temperature shift, etc) the feedback operation ensured that the system functions reliably.[262]

1.7.4 Other Examples

Ohlendorf *et al.* constructed and optimised two single-plasmid systems, named pDusk and pDawn, which use blue light to either induce or repress gene expression, respectively.[263] The pDusk plasmid encodes a LOV-containing HK called YF1, which phosphorylates its RR, FixJ, in the absence of blue light. The RR then activates the FixK2 promoter to drive gene expression. The systems were tested with red fluorescent protein DsRed, which was placed in the multiple cloning site of the plasmid containing the pDusk. Increasing blue light intensity resulted in reduced DsRed expression. The signal polarity of the light-regulated gene expression was then inverted by insertion of a gene-inversion cassette into pDusk to obtain pDawn. pDawn triggered the inverted response, whereby increased blue light illumination resulted in increased DsRed expression.[263]

1.8 Outlook

The optogenetics studies discussed here demonstrate how signal transduction can be altered using photoactivatable switches that modulate gene expression in response to light for biological and biotechnological applications. Furthermore, these switches are excellent regulatory tools compared to small molecule chemical inducers, which do not provide spatial resolution as they need to diffuse across cells. Additionally, chemically induced signals are not easily reversible. Chemical groups trapped in photocaged moieties overcome this problem as they allow spatiotemporal control over the release of the inducer but these are expensive and difficult to obtain and are not photoreversible. They remain highly diffusive and are difficult to deliver to cells. Use of photoreceptors for optical control of gene expression avoids all of these problems. Despite all of the exciting developments in the field, some of which are discussed above, optogenetics is still in its early days. Only a limited number of optogenetic tools have been characterised. The discovery of new photoreceptors and photoexcitation mechanisms will expand the optogenetic toolkit.

Refactoring and optimisation of existing tool kits, as reported by Schmidl and colleagues,[6] will allow users to tune the absorbance and signalling efficiency of the systems, as well as making these systems suitable for use in the field of synthetic biology by minimising the footprint of these engineered systems. The PCM-HK coupling in this manner may allow design of signal transduction mechanisms that are not observed in nature. Although there is a good mechanistic understanding of these photoreceptors, there are gaps in our knowledge and thus, a better mechanistic and structural understanding of phytochromes and other photoreceptors will aid the design of these for optogenetics.

The recent discovery of a plethora of new CBCRs and eukaryotic algal phytochromes has provided a wide range of proteins that have yet to be studied. The broad spectral properties of these CBCRs and eukaryotic algal

phytochromes can be useful in engineering of crop plant species for diverse light environments. This will be extremely useful for modern agricultural practices where terrestrial crops are grown at high density, which leads to competition for light. Engineered photoreceptor biobricks that can be effortlessly assembled in a combinatorial manner to pair spectrally tuned photoreceptors with different downstream RRs would also prove useful for control of operons in synthetic biology. The main limitations pertaining to optogenetics is the specific targeting of the light and design of light delivery tools. With continuing developments optogenetics will find use in areas outside of neuroscience.

1.9 Thesis Aims and Objectives

The overall aim of this research project was to develop optogenetic regulators for the modulation of gene expression within DNA constructs built for synthetic biology applications. Due to their modular architecture, phytochromes are a good candidate for protein engineering through domain fusion and this would allow specific tuning of spectral properties and output functions in the design of novel optogenetic tools. However, the different modules of the PCM integrate together in an intricate manner and each fold serves a purpose in mediating the signal received from the photon. This makes designing and engineering chimaeric constructs challenging because care must to be taken to maintain the useful features of naturally occurring phytochromes, such as i). the functionality of the important folds and conserved residues, ii). high photoreaction quantum yields, iii). reversibility of the photoreceptor cycle, iv). dark reversion time, and v). stability.[3,264,265]

Phytochromes and LOV domains have been widely used to construct photosensors, however, only phytochromes will be studied in the present work. Each type of sensor domain may be linked to a range of effector domains and *vice versa*. Existence of these photoreceptors as sensor-effector pairs in nature is evidence against any tertiary-structure-specific interactions that would make the mechanism of signal transduction specific to each photosensor. Thus, the domain fusion approach to engineering constructs may be productive. One approach to alter the absorption spectra of phytochromes would be through changing the substituents of the four-pyrrole rings. However, this approach would require external provision of the bilin rather than relying on endogenous uptake of natural bilin and an extensive screening process. An alternative approach would be to swap GAF domains with others that have different optical properties. However, GAF domains are part of the intricate structural knot fold, and thus swapping GAFs may disrupt this knotted structure and render the phytochrome inactive. This may be a better approach for CBCRs that only have GAF as their PCM because the PCM architecture is not as extensive, thus the disruption may be minimal. Another approach, one that is explored here, is the

coupling of a number of atypical phytochrome PCMs with effector domains of TCS proteins with known partner proteins and target promoters. Through this approach any wavelength of light can be coupled to any output domain to build a 'plug and play' toolkit.

The development of phytochromes as optogenetic tools is hindered because target promoters of phytochrome controlled response regulators are not known. We attempt to tackle this problem by fusing atypical phytochrome PCMs with TCS system effector domains that have known RRs and target promoters. The constructs will be engineered to allow easy insertion and removal of these two units, allowing combinatorial fusion of different light sensitivities with different output signals. As the photocycles of these atypical phytochromes have not been probed yet, very little is known about the photochemical mechanisms and properties of these proteins. Therefore, we also probe the photodynamics and photocycles of these atypical phytochromes using a range of time-resolved approaches, including UV/VIS, IR and X-ray probes. The photocycle studies will inform the design of these phytochromes as optogenetic tools. Ultimately, these optogenetic tools will be indispensible for precise temporal control of gene expression within whole biosynthetic pathways.

1.10 Reference List:

[1]   G. Nagel, D. Ollig, M. Fuhrmann, S. Kateriya, A. M. Musti, E. Bamberg, P. Hegemann, *Science* **2002**, *296*, 2395–2398.

[2]   J. T. M. Kennis, T. Mathes, *Interface Focus* **2013**, *3*, 20130005–20130005.

[3]   A. Möglich, R. A. Ayers, K. Moffat, *J. of Mol. Biol.* **2009**, *385*, 1433–1444.

[4]   X. Yang, A. P.-T. Jost, O. D. Weiner, C. Tang, *Mol. Biol. Cell* **2013**, *24*, 2419–2430.

[5]   F. Zhang, L.-P. Wang, E. S. Boyden, K. Deisseroth, *Nat. Methods* **2006**, *3*, 785–792.

[6]   S. R. Schmidl, R. U. Sheth, A. Wu, J. J. Tabor, *ACS Synth. Biol.* **2014**, *3*, 820–831.

[7]   W. Z. Mostafa, R. A. Hegazy, *J Adv Res* **2015**, *6*, 793–804.

[8]   C. Reinbothe, M. El Bakkouri, F. Buhr, N. Muraki, J. Nomata, G. Kurisu, Y. Fujita, S. Reinbothe, *Trends Plant Sci.* **2010**, *15*, 614–624.

[9]   T. D. Lamb, *Prog Retin Eye Res* **2013**, *36*, 52–119.

[10]   M. A. van der Horst, K. J. Hellingwerf, *Acc. Chem. Res.* **2004**, *37*, 13–20.

[11]   A. R. Cashmore, *Cell* **2003**, *114*, 537–543.

[12]   M. A. Gauger, A. Sancar, *Cancer Res.* **2005**, *65*, 6828–6834.

[13]   K. Palczewski, *J. Biol. Chem.* **2012**, *287*, 1612–1619.

[14]   V. J. Kefalov, *J. Biol. Chem.* **2012**, *287*, 1635–1641.

[15]   J. A. Sullivan, X. W. Deng, *Dev. Biol.* **2003**, *260*, 289–297.

[16]   E. V. Anslyn, D. A. Dougherty, *Modern Physical Organic Chemistry*, University Science Books, **2006**.

[17]   H. H. Jaffe, A. L. Miller, *J. Chem. Educ.* **1966**, *43*, DOI 10.1021/ed043p469.

[18]   E. B. Priestley, *J. Chem. Phys.* **1968**, *49*, 622.

[19]   W. J. A. Woittiez, *Advanced Organic Chemistry. Reactions, Mechanisms, and Structure.*, Recueil Des Travaux Chimiques Des Pays-Bas, Chichester, **2010**.

[20]   R. E. Hester, R. B. Girling, *Spectroscopy of Biological Molecules*, Royal Society of Chemistry, Cambridge, **1991**.

[21]   J. M. Hollas, *Modern Spectroscopy*, Journal of Chemical Education, Chichester, **2003**.

[22]   H. Muller, B. Schweizer, *Biochemical Applications for UV/Vis Spectroscopy: DNA, Protein, and Kinetic Analysis*, Perkin Elmer, Norwalk, **1996**.

[23]   P. Atkins, J. de Paula, in *Atkins Physical Chemistry*, Oxford, **2014**.

[24] M. D. Fayer, *Ultrafast Infrared Vibrational Spectroscopy*, Taylor & Francis Inc, Bosa Roca, **2013**.

[25] P. Myron, S. Siddiquee, S. A. Azad, *Vib. Spectrosc.* **2017**, *89*, 26–36.

[26] H. H. Mantsch, D. Chapman, *Infrared Spectroscopy of Biomolecules*, Wiley-Liss, New York, **1996**.

[27] T. Elsaesser, J. G. Fujimoto, D. A. Wiersma, W. Zinth, *Ultrafast Phenomena XI*, Springer, Garmisch-Partenkirchen, **2012**.

[28] R. L. Fork, B. I. Greene, C. V. Shank, *Applied Physics Letters* **1981**, *38*, 671–672.

[29] V. Sundström, *Annu Rev Phys Chem* **2008**, *59*, 53–77.

[30] W. Demtröder, *Laser Spectroscopy*, Springer Berlin Heidelberg, Berlin, **1981**.

[31] P. M. Dantus, *Femtosecond Laser Pulses: Principles and Experiments*, American Chemical Society, New York, **1999**.

[32] G. Porter, *Science* **1968**, *160*, 1299–1307.

[33] S. P. Meisburger, W. C. Thomas, M. B. Watkins, N. Ando, *Chem. Rev.* **2017**, *117*, 7615–7672.

[34] M. Chen, J. Chory, C. Fankhauser, *Annu. Rev. Genet.* **2004**, *38*, 87–117.

[35] A. Möglich, X. Yang, R. A. Ayers, K. Moffat, *Annu. Rev. Plant Biol.* **2010**, *61*, 21–47.

[36] R. J. Kutta, S. J. O. Hardman, L. O. Johannissen, B. Bellina, H. L. Messiha, J. M. Ortiz-Guerrero, M. Elías-Arnanz, S. Padmanabhan, P. Barran, N. S. Scrutton, et al., *Nat Commun.* **2015**, *6*, 7907.

[37] S. Padmanabhan, M. Jost, C. L. Drennan, M. Elías-Arnanz, *Annu Rev Biochem* **2017**, *86*, 485–514.

[38] J. M. Christie, A. S. Arvai, K. J. Baxter, M. Heilmann, A. J. Pratt, A. O'Hara, S. M. Kelly, M. Hothorn, B. O. Smith, K. Hitomi, et al., *Science* **2012**, *335*, 1492–1496.

[39] M. Fuhrmann, A. Stahlberg, E. Govorunova, S. Rank, P. Hegemann, *J. Cell. Sci.* **2001**, *114*, 3857–3863.

[40] R. A. Bogomolni, J. L. Spudich, *Proc. Natl. Acad. Sci. U.S.A.* **1982**, *79*, 6250–6254.

[41] D. Oesterhelt, W. Stoeckenius, *Proc. Natl. Acad. Sci. U.S.A.* **1973**, *70*, 2853–2857.

[42] O. A. Sineshchekov, E. G. Govorunova, J. L. Spudich, *Photochem. Photobiol.* **2009**, *85*, 556–563.

[43] P. Hegemann, *Annu. Rev. Plant Biol.* **2008**, *59*, 167–189.

[44] S. Kateriya, G. Nagel, E. Bamberg, P. Hegemann, *Physiology* **2004**, *19*, 133–137.

[45]   R. D. Finn, J. Mistry, B. Schuster-Böckler, S. Griffiths-Jones, V. Hollich, T. Lassmann, S. Moxon, M. Marshall, A. Khanna, R. Durbin, et al., *Nucleic Acids Res.* **2006**, *34*, D247–51.

[46]   G. Nagel, T. Szellas, W. Huhn, S. Kateriya, N. Adeishvili, P. Berthold, D. Ollig, P. Hegemann, E. Bamberg, *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 13940–13945.

[47]   R. Henderson, J. M. Baldwin, T. A. Ceska, F. Zemlin, E. Beckmann, K. H. Downing, *J. of Mol. Biol.* **1990**, *213*, 899–929.

[48]   H. Luecke, B. Schobert, H. T. Richter, J. P. Cartailler, J. K. Lanyi, *J. of Mol. Biol.* **1999**, *291*, 899–911.

[49]   F. Tavanti, V. Tozzini, *Molecules* **2014**, *19*, 14961–14978.

[50]   J. Herbst, *Science* **2002**, *297*, 822–825.

[51]   M. A. Lawson, D. N. Zacks, F. Derguini, K. Nakanishi, J. L. Spudich, *Biophys. J.* **1991**, *60*, 1490–1498.

[52]   U. Haupts, J. Tittor, D. Oesterhelt, *Annu Rev Biophys Biomol Struct* **1999**, *28*, 367–399.

[53]   P. Changenet-Barret, A. Espagne, P. Plaza, K. J. Hellingwerf, M. M. Martin, *New J. Chem.* **2005**, *29*, 527–534.

[54]   J. K. Lanyi, G. Váró, *Isr. J. Chem.* **2013**, *35*, 365–385.

[55]   T. E. Meyer, *Biochim. Biophys. Acta* **1985**, *806*, 175–183.

[56]   W. W. Sprenger, W. D. Hoff, J. P. Armitage, K. J. Hellingwerf, *J. Bacteriol.* **1993**, *175*, 3096–3104.

[57]   Z. Jiang, L. R. Swem, B. G. Rushing, S. Devanathan, G. Tollin, C. E. Bauer, *Science* **1999**, *285*, 406–409.

[58]   A. Möglich, R. A. Ayers, K. Moffat, *Structure* **2009**, *17*, 1282–1294.

[59]   J. A. Kyndt, T. E. Meyer, M. A. Cusanovich, J. J. Van Beeumen, *FEBS Lett.* **2002**, *512*, 240–244.

[60]   J. A. Kyndt, T. E. Meyer, M. A. Cusanovich, *Photochem. Photobiol. Sci.* **2004**, *3*, 519–530.

[61]   J. A. Kyndt, J. K. Hurley, B. Devreese, T. E. Meyer, M. A. Cusanovich, G. Tollin, J. J. Van Beeumen, *Biochemistry* **2004**, *43*, 1809–1820.

[62]   Y. O. Jung, J. H. Lee, J. Kim, M. Schmidt, K. Moffat, V. Srajer, H. Ihee, *Nat Chem* **2013**, *5*, 212–220.

[63]   S. Anderson, S. Crosson, K. Moffat, *Acta Crystallogr. D Biol. Crystallogr.* **2004**, *60*, 1008–1016.

[64]   H. Chosrowjan, Y. Shibata, N. Mataga, Y. Imamoto, F. Tokunaga, in *Ultrafast Phenomena XI*, Springer Berlin Heidelberg, Berlin, Heidelberg, **1998**, pp. 684–686.

[65] M. A. van der Horst, W. Laan, S. Yeremenko, A. Wende, P. Palm, D. Oesterhelt, K. J. Hellingwerf, *Photochem. Photobiol. Sci.* **2005**, *4*, 688–693.

[66] W. D. Hoff, I. H. van Stokkum, H. J. van Ramesdonk, M. E. van Brederode, A. M. Brouwer, J. C. Fitch, T. E. Meyer, R. van Grondelle, K. J. Hellingwerf, *Biophys. J.* **1994**, *67*, 1691–1705.

[67] L. Ujj, S. Devanathan, T. E. Meyer, M. A. Cusanovich, G. Tollin, G. H. Atkinson, *Biophys. J.* **1998**, *75*, 406–412.

[68] H. Ihee, S. Rajagopal, V. Srajer, R. Pahl, S. Anderson, M. Schmidt, F. Schotte, P. A. Anfinrud, M. Wulff, K. Moffat, *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 7145–7150.

[69] B. L. Taylor, I. B. Zhulin, *Microbiol. Mol. Biol. Rev.* **1999**, *63*, 479–506.

[70] M. Gomelsky, G. Klug, *Trends Biochem. Sci.* **2002**, *27*, 497–500.

[71] S. Anderson, V. Dragnea, S. Masuda, J. Ybe, K. Moffat, C. Bauer, *Biochemistry* **2005**, *44*, 7998–8005.

[72] M. Gauden, I. H. M. van Stokkum, J. M. Key, D. C. Lührs, R. van Grondelle, P. Hegemann, J. T. M. Kennis, *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 10895–10900.

[73] S. Masuda, K. Hasegawa, A. Ishii, T.-A. Ono, *Biochemistry* **2004**, *43*, 5304–5313.

[74] M. Ahmad, A. R. Cashmore, *Nature* **1993**, *366*, 162–166.

[75] R. Brudler, K. Hitomi, H. Daiyasu, H. Toh, K.-I. Kucho, M. Ishiura, M. Kanehisa, V. A. Roberts, T. Todo, J. A. Tainer, et al., *Mol. Cell* **2003**, *11*, 59–67.

[76] C. Lin, D. Shalitin, *Annu. Rev. Plant Biol.* **2003**, *54*, 469–496.

[77] D. Shalitin, H. Yang, T. C. Mockler, M. Maymon, H. Guo, G. C. Whitelam, C. Lin, *Nature* **2002**, *417*, 763–767.

[78] Q.-H. Li, H.-Q. Yang, *Photochem. Photobiol.* **2007**, *83*, 94–101.

[79] D. A. Kishkinev, N. S. Chernetsov, *Biol. Bull.* **2015**, *5*, 46–62.

[80] R. Wiltschko, S. Denzau, D. Gehring, P. Thalau, W. Wiltschko, *J. Exp. Biol.* **2011**, *214*, 3096–3101.

[81] W. Wiltschko, R. Wiltschko, *J Comp Physiol A* **1995**, *177*, 363–369.

[82] R. Wiltschko, U. Munro, H. Ford, K. Stapput, P. Thalau, W. Wiltschko, *J Comp Physiol A* **2014**, *200*, 399–407.

[83] W. Wiltschko, U. Munro, H. Ford, R. Wiltschko, *Nature* **1993**, *364*, 525–527.

[84] W. Wiltschko, R. Wiltschko, *J Comp Physiol A* **1999**, *184*, 295–299.

[85] M. Liedvogel, K. Maeda, K. Henbest, E. Schleicher, T. Simon, C. R. Timmel, P. J. Hore, H. Mouritsen, *PLoS ONE* **2007**, *2*, e1106.

[86] T. Ritz, S. Adem, K. Schulten, *Biophys. J.* **2000**, *78*, 707–718.

[87]   K. Maeda, A. J. Robinson, K. B. Henbest, H. J. Hogben, T. Biskup, M. Ahmad, E. Schleicher, S. Weber, C. R. Timmel, P. J. Hore, *Proc. Natl. Acad. Sci. U.S.A.* **2012**, *109*, 4774–4779.

[88]   A. A. Lee, J. C. S. Lau, H. J. Hogben, T. Biskup, D. R. Kattnig, P. J. Hore, *J. Royal Soc. Interface* **2014**, *11*, 20131063–20131063.

[89]   J.-P. Bouly, E. Schleicher, M. Dionisio-Sese, F. Vandenbussche, D. Van Der Straeten, N. Bakrim, S. Meier, A. Batschauer, P. Galland, R. Bittl, et al., *J. Biol. Chem.* **2007**, *282*, 9383–9391.

[90]   M. Ahmad, *Curr. Opin. Plant Biol.* **2016**, *33*, 108–115.

[91]   K. S. Ilia A Solov'yov, *J. Phys. Chem. B* **2012**, *116*, 1089–1099.

[92]   R. J. Kutta, N. Archipowa, L. O. Johannissen, A. R. Jones, N. S. Scrutton, *Sci. Rep.* **2017**, *7*, 44906.

[93]   J. M. Christie, *Annu. Rev. Plant Biol.* **2007**, *58*, 21–45.

[94]   J. M. Christie, P. Reymond, G. K. Powell, P. Bernasconi, A. A. Raibekas, E. Liscum, W. R. Briggs, *Science* **1998**, *282*, 1698–1701.

[95]   E. Demarsy, C. Fankhauser, *Curr. Opin. Plant Biol.* **2009**, *12*, 69–74.

[96]   D. C. Nelson, J. Lasswell, L. E. Rogg, M. A. Cohen, B. Bartel, *Cell* **2000**, *101*, 331–340.

[97]   F. Takahashi, D. Yamagata, M. Ishikawa, Y. Fukamatsu, Y. Ogura, M. Kasahara, T. Kiyosue, M. Kikuyama, M. Wada, H. Kataoka, *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 19625–19630.

[98]   T. E. Swartz, S. B. Corchnoy, J. M. Christie, J. W. Lewis, I. Szundi, W. R. Briggs, R. A. Bogomolni, *J. Biol. Chem.* **2001**, *276*, 36493–36500.

[99]   E. Schleicher, R. M. Kowalczyk, C. W. M. Kay, P. Hegemann, A. Bacher, M. Fischer, R. Bittl, G. Richter, S. Weber, *J. Am. Chem. Soc.* **2004**, *126*, 11067–11076.

[100] M. Salomon, W. Eisenreich, H. Dürr, E. Schleicher, E. Knieb, V. Massey, W. Rüdiger, F. Müller, A. Bacher, G. Richter, *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 12357–12361.

[101] D. Matsuoka, T. Iwata, K. Zikihara, H. Kandori, S. Tokutomi, *Photochem. Photobiol.* **2007**, *83*, 122–130.

[102] J. T. M. Kennis, S. Crosson, M. Gauden, I. H. M. van Stokkum, K. Moffat, R. van Grondelle, *Biochemistry* **2003**, *42*, 3385–3392.

[103] M. Kasahara, T. E. Swartz, M. A. Olney, A. Onodera, N. Mochizuki, H. Fukuzawa, E. Asamizu, S. Tabata, H. Kanegae, M. Takano, et al., *Plant Physiol.* **2002**, *129*, 762–773.

[104] K. Okajima, S. Kashojiya, S. Tokutomi, *J. Biol. Chem.* **2012**, *287*, 40972–40981.

[105]   Y. Kawaguchi, Y. Nakasone, K. Zikihara, S. Tokutomi, M. Terazima, *J. Am. Chem. Soc.* **2010**, *132*, 8838–8839.

[106]   T. Kottke, J. Heberle, D. Hehn, B. Dick, P. Hegemann, *Biophys. J.* **2003**, *84*, 1192–1201.

[107]   J. M. Christie, T. E. Swartz, R. A. Bogomolni, W. R. Briggs, *Plant J.* **2002**, *32*, 205–219.

[108]   D. Matsuoka, S. Tokutomi, *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 13337–13342.

[109]   J. E. Johnson Jr, F. E. Reyes, J. T. Polaski, R. T. Batey, *Nature* **2012**, *492*, 133–137.

[110]   A. Peselis, A. Serganov, *Nat Struct Mol Biol* **2012**, *19*, 1182–1184.

[111]   G. H. Reed, *Curr Opin Chem Biol* **2004**, *8*, 477–483.

[112]   J. M. Ortiz-Guerrero, M. C. Polanco, F. J. Murillo, S. Padmanabhan, M. Elias-Arnanz, *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, 7565–7570.

[113]   Z. Cheng, H. Yamamoto, C. E. Bauer, *Trends Biochem. Sci.* **2016**, *41*, 647–650.

[114]   Z. Cheng, K. Li, L. A. Hammad, J. A. Karty, C. E. Bauer, *Mol. Microbiol.* **2014**, *91*, 649–664.

[115]   M. Wu, Å. Strid, L. A. Eriksson, *J. Phys. Chem. B* **2014**, *118*, 951–965.

[116]   Di Wu, Q. Hu, Z. Yan, W. Chen, C. Yan, X. Huang, J. Zhang, P. Yang, H. Deng, J. Wang, et al., *Nature* **2012**, *484*, 214–219.

[117]   M. Heilmann, J. M. Christie, J. T. M. Kennis, G. I. Jenkins, T. Mathes, *Photochem. Photobiol. Sci.* **2015**, *14*, 252–257.

[118]   C. Cloix, E. Kaiserli, M. Heilmann, K. J. Baxter, B. A. Brown, A. O'Hara, B. O. Smith, J. M. Christie, G. I. Jenkins, *Proc. Natl. Acad. Sci. U.S.A.* **2012**, *109*, 16366–16370.

[119]   B. A. Brown, C. Cloix, G. H. Jiang, E. Kaiserli, P. Herzyk, D. J. Kliebenstein, G. I. Jenkins, *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 18225–18230.

[120]   E. Kaiserli, G. I. Jenkins, *Plant Cell* **2007**, *19*, 2662–2673.

[121]   J.-J. Favory, A. Stec, H. Gruber, L. Rizzini, A. Oravecz, M. Funk, A. Albert, C. Cloix, G. I. Jenkins, E. J. Oakeley, et al., *EMBO J.* **2009**, *28*, 591–601.

[122]   A. Oravecz, A. Baumann, Z. Máté, A. Brzezinska, J. Molinier, E. J. Oakeley, E. Adám, E. Schäfer, F. Nagy, R. Ulm, *Plant Cell* **2006**, *18*, 1975–1990.

[123]   Y. Saijo, J. A. Sullivan, H. Wang, J. Yang, Y. Shen, V. Rubio, L. Ma, U. Hoecker, X. W. Deng, *Genes Dev.* **2003**, *17*, 2642–2647.

[124]   M. T. Osterlund, C. S. Hardtke, N. Wei, X. W. Deng, *Nature* **2000**, *405*, 462–466.

[125]   W. L. Butler, K. H. Norris, H. W. Siegelman, S. B. Hendricks, *Proc. Natl.*

Acad. Sci. U.S.A. **1959**, *45*, 1703–1708.

[126]   T. Shinomura, *J. Plant Res.* **1997**, *110*, 151–161.

[127]   J. W. Reed, A. Nagatani, T. D. Elich, M. Fagan, J. Chory, *Plant Physiol.* **1994**, *104*, 1139–1149.

[128]   K. A. Franklin, S. J. Davis, W. M. Stoddart, R. D. Vierstra, G. C. Whitelam, *Plant Cell* **2003**, *15*, 1981–1989.

[129]   E. Johnson, M. Bradley, N. P. Harberd, G. C. Whitelam, *Plant Physiol.* **1994**, *105*, 141–149.

[130]   A. Blumenstein, K. Vienken, R. Tasler, J. Purschwitz, D. Veith, N. Frankenberg-Dinkel, R. Fischer, *Curr. Biol.* **2005**, *15*, 1833–1838.

[131]   N. C. Rockwell, D. Duanmu, S. S. Martin, C. Bachy, D. C. Price, D. Bhattacharya, A. Z. Worden, Lagarias, *Proc. Natl. Acad. Sci. U.S.A.* **2014**, *111*, 3871–3876.

[132]   L. Vuillet, M. Kojadinovic, S. Zappa, M. Jaubert, J.-M. Adriano, J. Fardoux, L. Hannibal, D. Pignol, A. Verméglio, E. Giraud, *EMBO J.* **2007**, *26*, 3322–3331.

[133]   N. B. Ivleva, M. R. Bramlett, P. A. Lindahl, S. S. Golden, *EMBO J.* **2005**, *24*, 1202–1210.

[134]   S. H. Bhoo, S. J. Davis, J. Walker, B. Karniol, R. D. Vierstra, *Nature* **2001**, *414*, 776–779.

[135]   P. H. Quail, *Plant Cell Environ.* **1997**, *20*, 657–665.

[136]   B. L. Montgomery, Lagarias, *Trends Plant Sci.* **2002**, *7*, 357–366.

[137]   N. C. Rockwell, Y.-S. Su, Lagarias, *Annu. Rev. Plant Biol.* **2006**, *57*, 837–858.

[138]   A. T. Ulijasz, G. Cornilescu, D. von Stetten, S. Kaminski, M. A. Mroginski, J. Zhang, D. Bhaya, P. Hildebrandt, R. D. Vierstra, *J. Biol. Chem.* **2008**, *283*, 21251–21266.

[139]   L.-O. Essen, J. Mailliet, J. Hughes, *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 14709–14714.

[140]   M. Ikeuchi, T. Ishizuka, *Photochem. Photobiol. Sci.* **2008**, *7*, 1159–1167.

[141]   N. C. Rockwell, Lagarias, *Chem. Eur. J. of Chem. Phys.* **2010**, *11*, 1172–1180.

[142]   E. Giraud, J. Fardoux, N. Fourrier, L. Hannibal, B. Genty, P. Bouyer, B. Dreyfus, A. Verméglio, *Nature* **2002**, *417*, 202–205.

[143]   B. Karniol, J. R. Wagner, J. M. Walker, R. D. Vierstra, *Biochem. J.* **2005**, *392*, 103–116.

[144]   Lagarias, Lagarias, *Proc. Natl. Acad. Sci. U.S.A.* **1989**, *86*, 5778–5780.

[145]   Y. S. Ho, L. M. Burden, J. H. Hurley, *EMBO J.* **2000**, *19*, 5288–5299.

[146]   E. S. Burgie, R. D. Vierstra, *Plant Cell* **2015**, *26*, 4568–4583.

[147]   J. R. Wagner, J. S. Brunzelle, K. T. Forest, R. D. Vierstra, *Nature* **2005**, *438*, 325–331.

[148]   X. Yang, J. Kuk, K. Moffat, *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 14715–14720.

[149]   X. Yang, J. Kuk, K. Moffat, *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 15639–15644.

[150]   E. S. Burgie, J. Zhang, R. D. Vierstra, *Structure* **2016**, *24*, 448–457.

[151]   X. Yang, E. A. Stojkovic, J. Kuk, K. Moffat, *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 12571–12576.

[152]   J. Hahn, H. M. Strauss, P. Schmieder, *J. Am. Chem. Soc.* **2008**, *130*, 11170–11178.

[153]   J. R. Wagner, J. Zhang, J. S. Brunzelle, R. D. Vierstra, K. T. Forest, *J. Biol. Chem.* **2007**, *282*, 12298–12309.

[154]   F. Andel, Lagarias, R. A. Mathies, *Biochemistry* **1996**, *35*, 15997–16008.

[155]   N. C. Rockwell, Lagarias, *Plant Cell* **2006**, *18*, 4–14.

[156]   J. J. van Thor, M. Mackeen, I. Kuprov, R. A. Dwek, M. R. Wormald, *Biophys. J.* **2006**, *91*, 1811–1822.

[157]   W. Rüdiger, F. Thümmler, E. Cmiel, S. Schneider, *Proc. Natl. Acad. Sci. U.S.A.* **1983**, *80*, 6244–6248.

[158]   A. T. Ulijasz, G. Cornilescu, C. C. Cornilescu, J. Zhang, M. Rivera, J. L. Markley, R. D. Vierstra, *Nature* **2010**, *463*, 250–254.

[159]   A. Nagatani, *Curr. Opin. Plant Biol.* **2010**, *13*, 565–570.

[160]   E. S. Burgie, A. N. Bussell, J. M. Walker, K. Dubiel, R. D. Vierstra, *Proc. Natl. Acad. Sci. U.S.A.* **2014**, *111*, 10179–10184.

[161]   E. S. Burgie, T. Wang, A. N. Bussell, J. M. Walker, H. Li, R. D. Vierstra, *J. Biol. Chem.* **2014**, *289*, 24573–24587.

[162]   S. Nagano, P. Scheerer, K. Zubow, N. Michael, K. Inomata, T. Lamparter, N. Krauß, *J. Biol. Chem.* **2016**, *291*, 20674–20691.

[163]   H. Takala, A. Björling, O. Berntsson, H. Lehtivuori, S. Niebling, M. Hoernke, I. Kosheleva, R. Henning, A. Menzel, J. A. Ihalainen, et al., *Nature* **2014**, *509*, 245–248.

[164]   A. Björling, O. Berntsson, H. Lehtivuori, H. Takala, A. J. Hughes, M. Panman, M. Hoernke, S. Niebling, L. Henry, R. Henning, et al., *Sci Adv* **2016**, *2*, e1600920–e1600920.

[165]   K. Anders, A. Gutt, W. Gärtner, L.-O. Essen, *J. Biol. Chem.* **2014**, *289*, 25590–25600.

[166]   R. E. Hunt, L. H. Pratt, *Biochemistry* **1981**, *20*, 941–945.

[167]   S. Noack, N. Michael, R. Rosen, T. Lamparter, *Biochemistry* **2007**, *46*, 4164–

4176.

[168]   B. Esteban, M. Carrascal, J. Abian, T. Lamparter, *Biochemistry* **2005**, *44*, 450–461.

[169]   J. R. Wagner, J. Zhang, D. von Stetten, M. Günther, D. H. Murgida, M. A. Mroginski, J. M. Walker, K. T. Forest, P. Hildebrandt, R. D. Vierstra, *J. Biol. Chem.* **2008**, *283*, 12212–12226.

[170]   C. Song, G. Psakis, C. Lang, J. Mailliet, W. Gärtner, J. Hughes, J. Matysik, *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, 3842–3847.

[171]   L. H. Otero, S. Klinke, J. Rinaldi, F. Velazquez Escobar, M. A. Mroginski, M. Fernández López, F. Malamud, A. A. Vojnov, P. Hildebrandt, F. A. Goldbaum, et al., *J. of Mol. Biol.* **2016**, *428*, 3702–3720.

[172]   K. Evans, J. G. Grossmann, A. P. Fordham-Skelton, M. Z. Papiz, **2006**, *364*, 655–666.

[173]   R. Ohlendorf, C. H. Schumacher, F. Richter, A. Möglich, *ACS Synth. Biol.* **2016**, *5*, 1117–1126.

[174]   J. J. van Thor, B. Borucki, W. Crielaard, H. Otto, T. Lamparter, J. Hughes, K. J. Hellingwerf, M. P. Heyn, *Biochemistry* **2001**, *40*, 11460–11471.

[175]   B. Borucki, D. von Stetten, S. Seibeck, T. Lamparter, N. Michael, M. A. Mroginski, H. Otto, D. H. Murgida, M. P. Heyn, P. Hildebrandt, *J. Biol. Chem.* **2005**, *280*, 34358–34364.

[176]   J. L. Mateos, J. P. Luppi, O. B. Ogorodnikova, V. A. Sineshchekov, M. J. Yanovsky, S. E. Braslavsky, W. Gärtner, J. J. Casal, *J. Biol. Chem.* **2006**, *281*, 34421–34429.

[177]   T. Matsushita, N. Mochizuki, A. Nagatani, *Nature* **2003**, *424*, 571–574.

[178]   H. Takala, A. Björling, M. Linna, S. Westenhoff, J. A. Ihalainen, *J. Biol. Chem.* **2015**, *290*, 16383–16392.

[179]   K. A. Franklin, G. C. Whitelam, *Ann. Bot.* **2005**, *96*, 169–175.

[180]   J. Hughes, T. Lamparter, F. Mittmann, E. Hartmann, W. Gärtner, A. Wilde, T. Börner, *Nature* **1997**, *386*, 663–663.

[181]   B. Borucki, S. Seibeck, M. P. Heyn, T. Lamparter, *Biochemistry* **2009**, *48*, 6305–6317.

[182]   P. Schmidt, T. Gertsch, A. Remberg, W. Gärtner, S. E. Braslavsky, K. Schaffner, *Photochem. Photobiol.* **1998**, *68*, 754–761.

[183]   H. Linschitz, V. Kasche, *J. Biol. Chem.* **1966**, *241*, 3395–3403.

[184]   R. E. Kendrick, C. J. Spruit, *Photochem. Photobiol.* **1977**, *26*, 201–214.

[185]   P. F. Aramendia, B. P. Ruzsicska, S. E. Braslavsky, K. Schaffner, *Biochemistry* **2002**, *26*, 1418–1422.

[186]   Mark Bischoff, Gudrun Hermann, A. Sabine Rentsch, D. Strehlow,

Biochemistry **2000**, *40*, 181–186.

[187]   K. Heyne, J. Herbst, D. Stehlik, B. Esteban, T. Lamparter, J. Hughes, R. Diller, *Biophys. J.* **2002**, *82*, 1004–1016.

[188]   H. Kandori, K. Yoshihara, S. Tokutomi, *J. Am. Chem. Soc.* **1992**, *114*, 10958–10959.

[189]   M. Bischoff, G. Hermann, S. Rentsch, D. Strehlow, S. Winter, H. Chosrowjan, *J. Phys. Chem. B* **2000**, *104*, 1810–1816.

[190]   S. Rentsch, M. Bischoff, G. Hermann, D. Strehlow, *Appl Phys B* **1998**, *66*, 259–261.

[191]   F. Andel, K. C. Hasson, F. Gai, P. A. Anfinrud, R. A. Mathies, *Biospectroscopy* **1997**, *3*, 421–433.

[192]   J. J. van Thor, K. L. Ronayne, M. Towrie, *J. Am. Chem. Soc.* **2007**, *129*, 126–132.

[193]   C. Schumann, R. Gross, N. Michael, T. Lamparter, R. Diller, *Chem. Eur. J. of Chem. Phys.* **2007**, *8*, 1657–1663.

[194]   A. R. Holzwarth, J. Wendler, B. P. Ruzsicska, S. E. Braslavsky, K. Schaffner, *Biochim. Biophys. Acta* **1984**, *791*, 265–273.

[195]   J. Wendler, A. R. Holzwarth, S. E. Braslavsky, K. Schaffner, *Biochim. Biophys. Acta* **1984**, *786*, 213–221.

[196]   A. R. Holzwarth, E. Venuti, S. E. Braslavsky, K. Schaffner, *Biochim. Biophys. Acta* **1992**, *1140*, 59–68.

[197]   T. Rohmer, C. Lang, C. Bongards, K. B. S. S. Gupta, J. Neugebauer, J. Hughes, W. Gärtner, J. Matysik, *J. Am. Chem. Soc.* **2010**, *132*, 4431–4437.

[198]   H. M. Strauss, J. Hughes, P. Schmieder, *Biochemistry* **2005**, *44*, 8244–8250.

[199]   J. Hahn, H. M. Strauss, F. T. Landgraf, H. F. Gimenèz, G. Lochnit, P. Schmieder, J. Hughes, *FEBS J.* **2006**, *273*, 1415–1429.

[200]   A. Remberg, I. Lindner, T. Lamparter, J. Hughes, C. Kneip, P. Hildebrandt, S. E. Braslavsky, W. Gärtner, K. Schaffner, *Biochemistry* **1997**, *36*, 13389–13395.

[201]   C. Kneip, P. Hildebrandt, W. Schlamann, S. E. Braslavsky, F. Mark, K. Schaffner, *Biochemistry* **1999**, *38*, 15185–15192.

[202]   N. C. Rockwell, L. Shang, S. S. Martin, Lagarias, *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 6123–6127.

[203]   G. Rottwinkel, I. Oberpichler, T. Lamparter, *J. Bacteriol.* **2010**, *192*, 5124–5133.

[204]   P. Casino, V. Rubio, A. Marina, *Cell* **2009**, *139*, 325–336.

[205]   H. Li, J. Zhang, R. D. Vierstra, *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 10872–10877.

[206]   D. Bhaya, *Mol. Microbiol.* **2004**, *53*, 745–754.

[207]   A. Levskaya, A. A. Chevalier, J. J. Tabor, Z. B. Simpson, L. A. Lavery, M. Levy, E. A. Davidson, A. Scouras, A. D. Ellington, E. M. Marcotte, et al., *Nature* **2005**, *438*, 441–442.

[208]   O. Kwon, D. Georgellis, E. C. C. Lin, *J. Biol. Chem.* **2003**, *278*, 13192–13195.

[209]   T. Jin, M. Inouye, *J. of Mol. Biol.* **1994**, *244*, 477–481.

[210]   R. Utsumi, R. E. Brissette, A. Rampersaud, S. A. Forst, K. Oosawa, M. Inouye, *Science* **1989**, *245*, 1246–1249.

[211]   T. Lamparter, N. Michael, F. Mittmann, B. Esteban, *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 11628–11633.

[212]   B. Karniol, R. D. Vierstra, *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 2807–2812.

[213]   M. E. Auldridge, K. T. Forest, *Crit. Rev. Biochem. Mol. Biol.* **2015**, *46*, 67–88.

[214]   K. Yeh, *Science* **1997**, *277*, 1505–1508.

[215]   A. H. West, A. M. Stock, *Trends Biochem. Sci.* **2001**, *26*, 369–376.

[216]   M. Chen, R. Schwab, J. Chory, *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *100*, 14493–14498.

[217]   M. Chen, Y. Tao, J. Lim, A. Shaw, J. Chory, *Curr. Biol.* **2005**, *15*, 637–642.

[218]   K. C. Yeh, Lagarias, *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 13976–13981.

[219]   I. Schepens, P. Duek, C. Fankhauser, *Curr. Opin. Plant Biol.* **2004**, *7*, 564–569.

[220]   N. Suetsugu, F. Mittmann, G. Wagner, J. Hughes, M. Wada, *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 13705–13709.

[221]   M. Jaubert, J. Lavergne, J. Fardoux, L. Hannibal, L. Vuillet, J.-M. Adriano, P. Bouyer, D. Pignol, E. Giraud, A. Verméglio, *J. Biol. Chem.* **2007**, *282*, 7320–7328.

[222]   S. Yoshihara, M. Katayama, X. Geng, M. Ikeuchi, *Plant Cell Physiol.* **2004**, *45*, 1729–1737.

[223]   T. Ishizuka, T. Shimada, K. Okajima, S. Yoshihara, Y. Ochiai, M. Katayama, M. Ikeuchi, *Plant Cell Physiol.* **2006**, *47*, 1251–1261.

[224]   A. T. Ulijasz, G. Cornilescu, D. von Stetten, C. Cornilescu, F. Velazquez Escobar, J. Zhang, R. J. Stankey, M. Rivera, P. Hildebrandt, R. D. Vierstra, *J. Biol. Chem.* **2009**, *284*, 29757–29772.

[225]   R. Narikawa, Y. Fukushima, T. Ishizuka, S. Itoh, M. Ikeuchi, *J. of Mol. Biol.* **2008**, *380*, 844–855.

[226]   Thierry Rohmer, Holger Strauss, Jon Hughes, Huub de Groot, Wolfgang Gärtner, A. Peter Schmieder, Jörg Matysik, *15N MAS NMR Studies of Cph1*

Phytochrome: Chromophore Dynamics and Intramolecular Signal Transduction, American Chemical Society, **2006**.

[227]   T. Schirmer, U. Jenal, *Nat. Rev. Microbiol.* **2009**, *7*, 724–735.

[228]   U. Römling, M. Gomelsky, M. Y. Galperin, *Mol. Microbiol.* **2005**, *57*, 629–639.

[229]   G. Bae, G. Choi, *Annu. Rev. Plant Biol.* **2008**, *59*, 281–311.

[230]   P. Leivar, P. H. Quail, *Trends Plant Sci.* **2011**, *16*, 19–28.

[231]   J. Jeong, G. Choi, *Mol. Cells* **2013**, *35*, 371–380.

[232]   S. J. Davis, A. V. Vener, R. D. Vierstra, *Science* **1999**, *286*, 2517–2520.

[233]   D. M. Kehoe, A. Gutu, *Annu. Rev. Plant Biol.* **2006**, *57*, 127–150.

[234]   N. C. Rockwell, R. Ohlendorf, A. Möglich, *Proc. Natl. Acad. Sci. U.S.A.* **2013**, *110*, 806–807.

[235]   R. Narikawa, T. Ishizuka, N. Muraki, T. Shiba, G. Kurisu, M. Ikeuchi, *Proc. Natl. Acad. Sci. U.S.A.* **2013**, *110*, 918–923.

[236]   N. C. Rockwell, S. S. Martin, K. Feoktistova, Lagarias, *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, 11854–11859.

[237]   Y. Hirose, T. Shimada, R. Narikawa, M. Katayama, M. Ikeuchi, *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 9528–9533.

[238]   N. C. Rockwell, S. S. Martin, A. G. Gulevich, Lagarias, *Biochemistry* **2012**, *51*, 1449–1463.

[239]   N. C. Rockwell, S. S. Martin, Lagarias, *Biochemistry* **2012**, *51*, 9667–9677.

[240]   T. Ishizuka, A. Kamiya, H. Suzuki, R. Narikawa, T. Noguchi, T. Kohchi, K. Inomata, M. Ikeuchi, *Biochemistry* **2011**, *50*, 953–961.

[241]   G. Enomoto, Y. Hirose, R. Narikawa, M. Ikeuchi, *Biochemistry* **2012**, *51*, 3050–3058.

[242]   Y. Fukushima, M. Iwaki, R. Narikawa, M. Ikeuchi, Y. Tomita, S. Itoh, *Biochemistry* **2011**, *50*, 6328–6339.

[243]   Y. Chen, J. Zhang, J. Luo, J.-M. Tu, X.-L. Zeng, J. Xie, M. Zhou, J.-Q. Zhao, H. Scheer, K.-H. Zhao, *FEBS J.* **2012**, *279*, 40–54.

[244]   A. F. E. Hauck, S. J. O. Hardman, R. J. Kutta, G. M. Greetham, D. J. Heyes, N. S. Scrutton, *J. Biol. Chem.* **2014**, *289*, 17747–17757.

[245]   C. Slavov, X. Xu, K.-H. Zhao, W. Gärtner, J. Wachtveitl, *Biochim. Biophys. Acta* **2015**, *1847*, 1335–1344.

[246]   E. S. Burgie, J. M. Walker, G. N. Phillips, R. D. Vierstra, *Structure* **2013**, *21*, 88–97.

[247]   N. C. Rockwell, S. L. Njuguna, L. Roberts, E. Castillo, V. L. Parson, S. Dwojak, Lagarias, S. C. Spiller, *Biochemistry* **2008**, *47*, 7304–7316.

[248]   D. M. Kehoe, A. R. Grossman, *Science* **1996**, *273*, 1409–1412.

[249]   S. Yoshihara, F. Suzuki, H. Fujita, X. X. Geng, M. Ikeuchi, *Plant Cell Physiol.*

2000, *41*, 1299–1304.

[250]   C. C. Cornilescu, G. Cornilescu, E. S. Burgie, J. L. Markley, A. T. Ulijasz, R. D. Vierstra, *J. Biol. Chem.* **2014**, *289*, 3055–3065.

[251]   N. C. Rockwell, S. S. Martin, Lagarias, *Biochemistry* **2012**, *51*, 3576–3585.

[252]   O. Yizhar, L. E. Fenno, T. J. Davidson, M. Mogri, K. Deisseroth, *Neuron* **2011**, *71*, 9–34.

[253]   F. Zhang, J. Vierock, O. Yizhar, L. E. Fenno, S. Tsunoda, A. Kianianmomeni, M. Prigge, A. Berndt, J. Cushman, J. Polle, et al., *Cell* **2011**, *147*, 1446–1457.

[254]   L. L. Looger, *Nat. Methods* **2011**, *9*, 43–44.

[255]   D. M. Shcherbakova, A. A. Shemetov, A. A. Kaberniuk, V. V. Verkhusha, *Annu Rev Biochem* **2015**, *84*, 519–550.

[256]   Y. Nihongaki, F. Kawano, T. Nakajima, M. Sato, *Nat. Biotechnol.* **2015**, *33*, 755–760.

[257]   A. Levskaya, O. D. Weiner, W. A. Lim, C. A. Voigt, *Nature* **2009**, *461*, 997–1001.

[258]   S. Shimizu-Sato, E. Huq, J. M. Tepperman, P. H. Quail, *Nat. Biotechnol.* **2002**, *20*, 1041–1044.

[259]   M. Yazawa, A. M. Sadaghiani, B. Hsueh, R. E. Dolmetsch, *Nat. Biotechnol.* **2009**, *27*, 941–945.

[260]   J. J. Tabor, A. Levskaya, C. A. Voigt, *J. of Mol. Biol.* **2011**, *405*, 315–324.

[261]   P. Ramakrishnan, J. J. Tabor, *ACS Synth. Biol.* **2016**, *5*, 733–740.

[262]   A. Milias-Argeitis, M. Rullan, S. K. Aoki, P. Buchmann, M. Khammash, *Nat Commun.* **2016**, *7*, 12546.

[263]   R. Ohlendorf, R. R. Vidavski, A. Eldar, K. Moffat, A. Möglich, *J. of Mol. Biol.* **2012**, *416*, 534–542.

[264]   A. Möglich, P. Hegemann, *Nature* **2013**, *500*, 406–408.

[265]   J. Duebel, K. Marazova, J.-A. Sahel, *Curr. Opin. Ophthalmo.* **2015**, *26*, 226–232.

Chapter 2 Initial characterisation of novel atypical phytochromes.

Abstract

Phytochromes are typically red and far-red light absorbing photoreceptors that modulate a range of light driven physiological processes in plants, fungi and bacteria. Historically, the relatively limited range of light sensitivity of phytochromes has prevented their use in optogenetics applications. Recently, a number of atypical phytochromes, which have domain architecture similar to the typical phytochromes, have been discovered that absorb across the entire visible spectrum. Here, we expressed and purified a number of these atypical phytochromes for initial biophysical characterisation. We studied the photoreactions of one such protein, namely the phytochrome from *Nostoc punctiforme*, in further detail using time-resolved spectroscopy and cryo-trapping experiments. Consequently, we have identified the intermediates that are formed upon excitation, together with the rates of their formation / decay. This has allowed us to characterise the complete the photocycle of this phytochrome, which will be important for any subsequent optogenetics application in the future.

2.1 Introduction

Phytochromes are tetrapyrrole-containing photoreceptors that are typically known to be responsive to red and far-red region of the visible spectrum.[1] These photoreceptors are found in bacteria, plant and fungi, and modulate a large number of light driven processes, such as photomorphogenesis, leaf opening, seed germination and circadian clocks.[2-5] The exceptional spatial and temporal control of gene expression and physiological processes by these photoreceptors make them an attractive target for applications in the optogenetics field.

Phytochromes are multi-domain proteins composed of a number of PAS, GAF, PHY domains within the photo-sensory region, and HK and ATPase domains within the effector region.[6] The tetrapyrrole chromophore, usually a form of bilin, is found in the GAF domain and there can be multiple GAF domains with different reductive forms of the bilin cofactor. Light captured by the bilin

provides enough energy for an *E/Z* isomerisation of a double bond within the tetrapyrrole. This changes the length of the conjugated electron system and therefore, the absorption properties of the phytochrome.[7] The isomerisation also leads to a helix to β-sheet structural change within a tongue region of the phytochromes, leading to shortening of the tongue that ultimately pulls on the effector domain to facilitate autophosphorylation.[7-11] These localised and large-scale structural changes occur across a wide range of timescales. The isomerisation of the bilin occurs in the excited state on the picoseconds timescale, whereas the secondary structural change take place on the µs to ms timescale and the larger domain changes on ms – s timescale. Time-resolved UV/Vis, IR and Raman spectroscopy have been used to study the photoisomerisation step. [12-18] The tongue refolding triggered by the initial isomerisation has been studied using X-Ray crystallography and time-resolved X-Ray scattering.[19,20] In addition, cryo-trapping has been used to identify the energetic and thermal barriers associated with the formation of different intermediates along the photoreaction for different photoreceptors. [21-23]

The different light-absorbing states of phytochromes either interact with or release from a specific response regulator (RR) to activate and/or inactivate different signalling cascades via phosphotransfer. Due to the lack of a structure of a full-length phytochrome, the details on the signal transduction have only been inferred indirectly.[24,25] The phosphorylation of the RR affects it's binding to its DNA or partner protein, eliciting a response. For some phytochromes the RR can be found on the same polypeptide as the HK but the RRs are not known for all phytochromes. Although significant mechanistic and structural studies have been conducted on phytochromes, there is still a lot unknown about the way these proteins function and how they control the downstream physiological processes. In addition, phytochromes with output domains other than HKs have also been identified, such as another PAS domain, but little is also known about these systems. [26-28]

The recent discovery of phytochromes with unusually diverse absorption properties is particularly important for optogenetic applications as it could potentially increase the range of wavelengths available for these purposes.[29] These include a set of phytochromes from the eukaryotic algae and

cyanobacteria that share similarity in domain architecture with typical phytochromes. The study reported photosensory properties of seven different phytochromes from a diverse range of algae, which showed that the algae have adapted to sense shorter wavelengths in the orange, green and blue region of the visible spectrum at depths where red and far-red light gets attenuated.[29] Herein, we report the expression and purification of a number of atypical phytochromes from *Dolihomastix tenuilepis (Dten)*, *Nostoc punctiforme (Npf)* and *Gloeochaete wittrockiana (Gwit)*. Time resolved spectroscopy and cryo-trapping studies were used to probe the mechanism of Npf phytochrome and identify any intermediates that are formed during the photocycle.

2.2 Results and Discussion

2.2.1 Expression Trials of Dten, Npf and Gwit phytochromes

The genes for the PCMs of Dten, Npf and Gwit phytochrome were sub-cloned into pET15b, pET21a and pET21a-var (a variant of pET21a in which the C-terminal His tag is not encoded) vectors to produce protein with N-terminal and C-terminal His-tags, as well as non-tagged protein. Expression of these phytochromes was optimised and the results of these trials are shown in Figure 2.1. The Dten PCM is 69 kDa, Npf photosensory core module (PCM) is 58 kDa and Gwit PCM is 58 kDa. Bands of these sizes are present in the total cell lysate fractions for all three proteins, but the proteins are largely insoluble. To confirm whether these bands result from the expected proteins, a Western blot against the His-tags was performed as the PCMs in pET15b and pET21a were N-terminally and C-terminally His-tagged, respectively. The blot confirmed that all three proteins were expressed, although the Gwit PCM was only very poorly expressed. pET21a and pET15b plasmids showed the best expression for Dten and Npf PCM, respectively. Previous studies have also suggested that the Gwit phytochrome is a poorly expressing protein.[29,30] Npf has also been reported as a poor-expressing protein, which had previously prevented the study of its photocycle.[30]
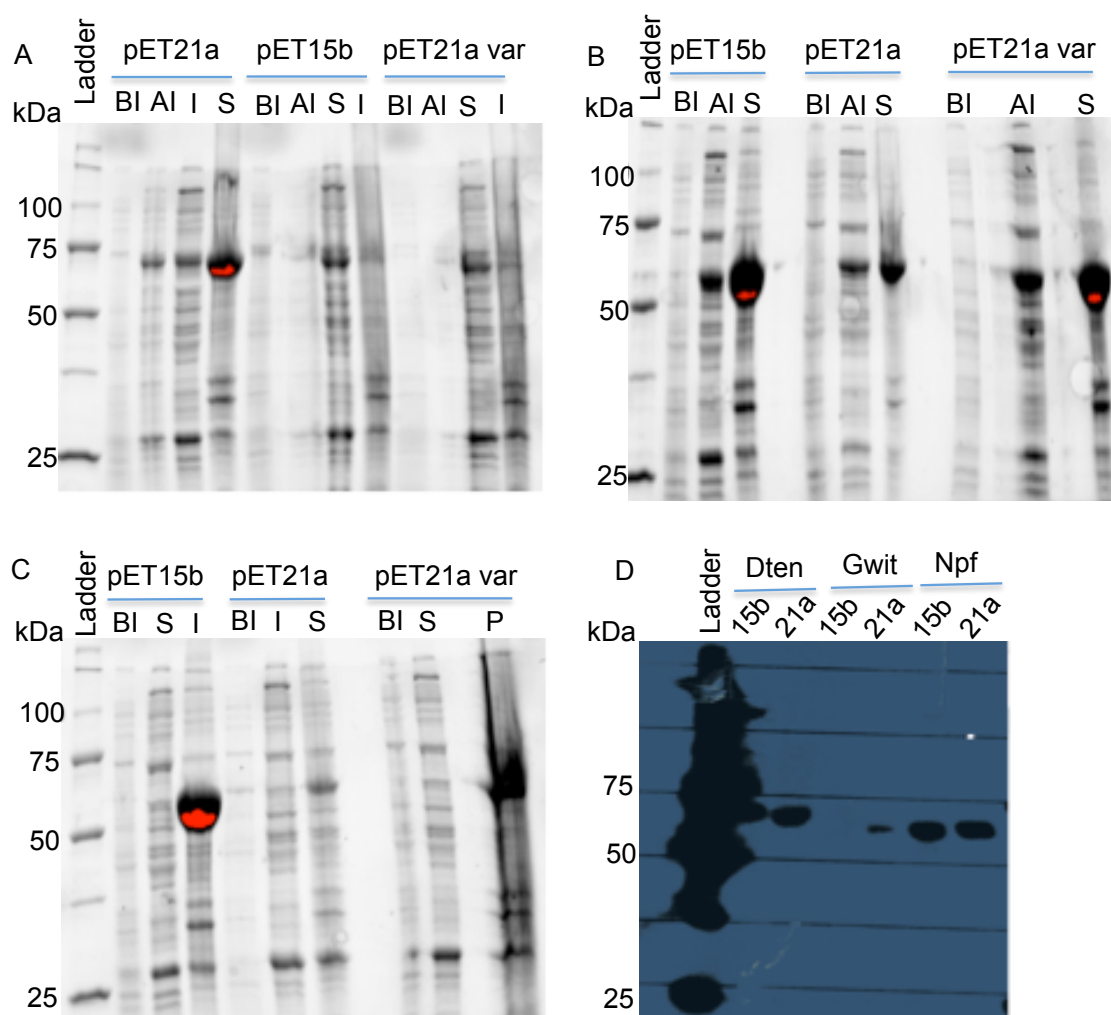
Figure 2.1. SDS-PAGE gels and Western blot showing expression trials of Dten, Npf and Gwit phytochrome in vectors pET15b, pET21a and pET21a var. Lane 1 is a ladder with the molecular weights shown on the left of each Figure. Expression trials for Dten phytochrome (**A**), Npf phytochrome (**B**) and Gwit phytochrome (**C**) expressed in pET21a, pET15b and pET21a var vectors. Before induction (BI) after induction (AI), insoluble (I) and soluble (S) samples were analysed. **D.** Western Blot: Dten phytochrome, Gwit phytochrome and Npf phytochrome expressed in pET15b (15b) and pET21a (21a) vectors.

In an attempt to improve the solubility of these phytochromes, a set of expression trials was carried out using two strains of *E. coli* cells and differing concentrations of benzyl alcohol. SDS-PAGE gels showing the results of these trials are shown in Figure 2.2. One set of trials was conducted using *E.coli* ArcticExpress (DE3) cells, which co-expresses two cold-adapted chaperonins. This cell line was used in the anticipation that the chaperonins might aid correct folding of the phytochromes, resulting in increased solubility of the phytochrome proteins. Both ArcticExpress cells and HMS174 cells were used for protein

expression at varying concentrations of benzyl-alcohol, because addition of benzyl-alcohol prior to induction shocks the cells and results in the production of *E.coli* endogenous chaperones.[31-33] It was anticipated that benzyl-alcohol triggered chaperone production might aid protein folding and hence phytochrome solubilisation. Osmolytes are known to stabilise proteins in their native folded form and assist refolding of the unfolded peptide, therefore expression trails were also conducted with 1 M sorbitol and 2.5 mM betaine.
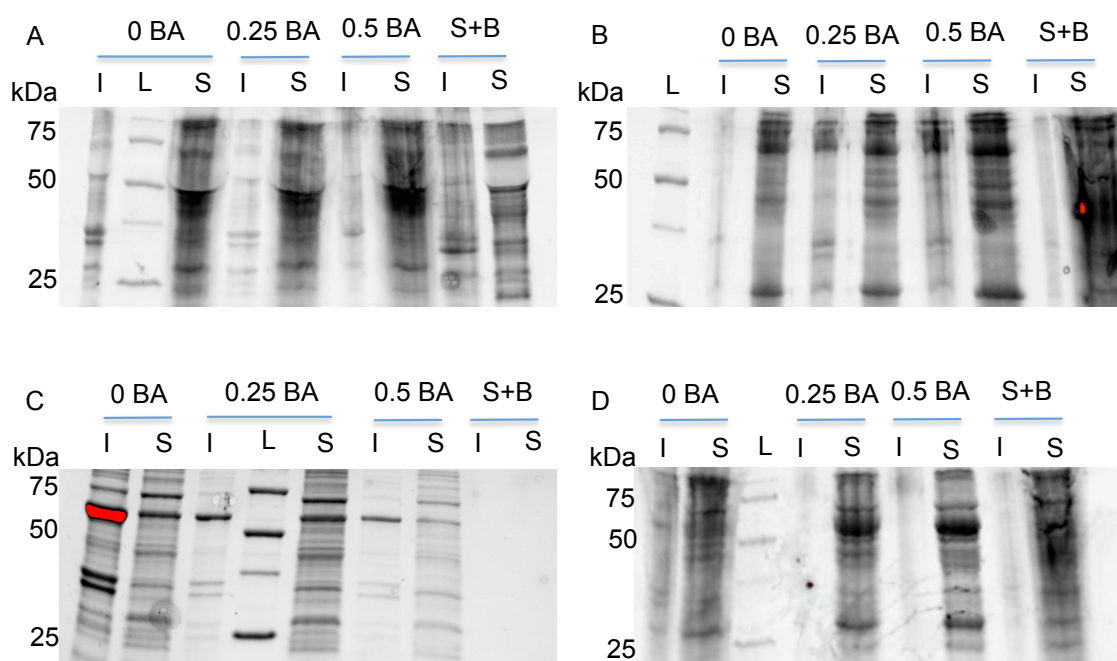
Figure 2.2. SDS-PAGE gels and Western blot showing expression trials for Dten, Npf and Gwit phytochromes under different conditions. A: Dten (67kDa) pET21a in ArcticExpress cells**. B:** Dten pET21a in HMS174 cells. **C:** Npf (58kDa) pET15b in ArcticExpress cells. **D:** Npf pET15b in HMS174 cells. Insoluble (I) and soluble (S) samples analysed from cultures grown in 0 mg/ml Benzyl Alcohol (BA), 0.25 mg/ml BA, 0.5 mg/ml BA and 1M sorbitol with 2.5mM Betaine (S+B)**.** Ladder lane labelled as L.

Although the proteins were more soluble when expressed with the addition of benzyl alcohol, they were no longer able to bind to Ni^{2+} or Co^{2+} resin. The refolding that led to increased solubility of the phytochromes may have caused internalisation of the His-tag and therefore, the phytochromes could no longer bind to the affinity resins. Hence, this approach to enhance solubilisation was not adopted. In an attempt to increase protein yield, starter or inoculation cultures were grown from the transformation plate the morning of protein

expression instead of the night before. It is known that cells can lose plasmids over time and since these cells contained two plasmids, the probability of losing a plasmid is higher. This method resulted in a higher yield of soluble phytochrome protein.

2.2.2 Purification of Dten and Npf phytochromes

The Dten and Npf phytochromes were purified using Ni^{2+} affinity chromatography followed by a gel filtration step (Figure 2.3). The Gwit phytochrome expression levels were not considered high enough to take forward for further studies. This purification protocol resulted in isolation of pure Dten and Npf phytochromes as shown in Figure 2.3 B and C.
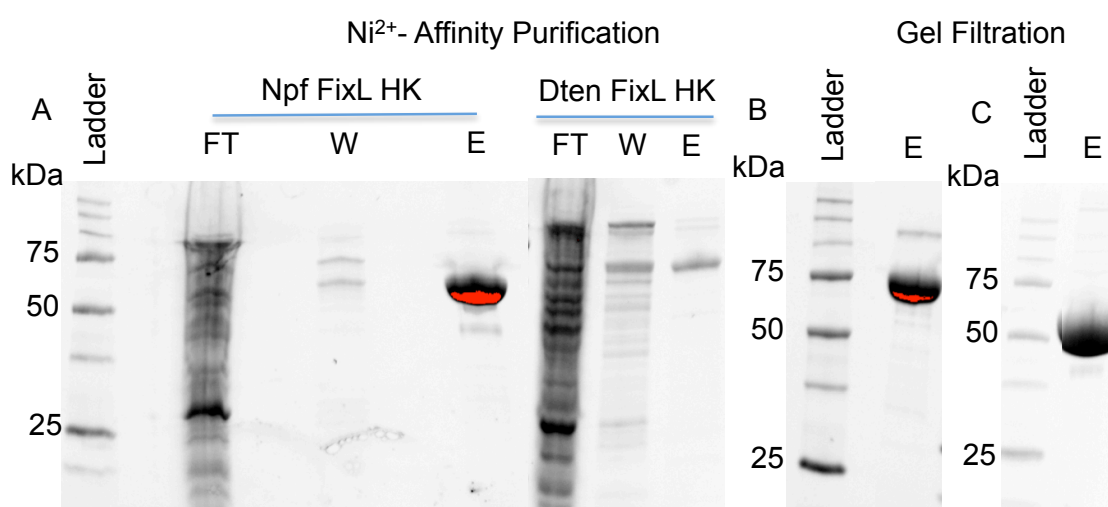
Figure 2.3. **SDS PAGE gels of the 2-step purification of the Dten and Npf phytochromes.** Ladder with the molecular weights shown on the left of each Figure. **A** Nickel-affinity chromatography of Npf PCM (58kDa) and Dten PCM (67kDa), flow through (FT), wash (W) and elute (E) fractions analysed on gel; **B** Dten PCM gel filtration; **C** Npf PCM gel filtration.

2.2.3 Initial Spectral Characterisation of Dten and Npf phytochromes

To assess the concentration of the phytochromes post-purification and to check whether the PCMs were active, steady-state absorbance spectroscopy was used. The Dten phytochrome switches between a 598 nm orange light-absorbing state, with a smaller peak at 354 nm, and a 714 nm far-red light-absorbing state, which has a smaller peak at 373 nm and a shoulder at 614 nm (Figure 2.4A). The protein develops a blue colour when illuminated with 735 nm light and a dark-green colour when illuminated with 595 nm light. The 714 nm absorbing state thermally converts to the 598 nm absorbing state over time over

the time course of about 2 hours. A full spectroscopic characterisation and kinetic characterisation of the entire photocycle of the Dten PCM is subsequently described in chapter 3.
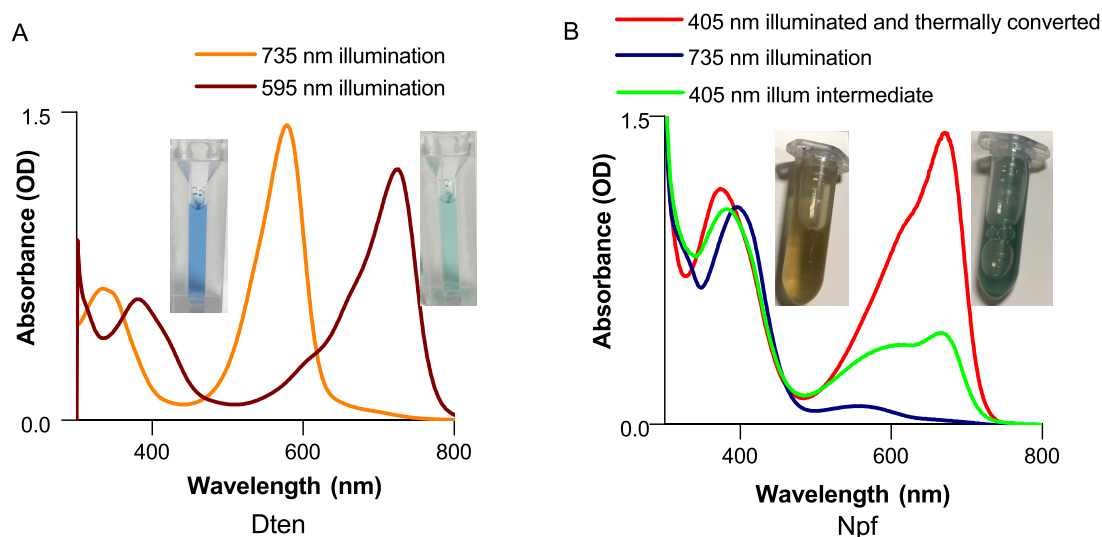
Figure 2.4. Absorption spectra of the different states of Dten and Npf phytochromes. A Dten phytochrome in the orange and far-red light-absorbing states, photographs of actual samples in each state shown. **B** Npf phytochrome in the blue and red light-absorbing states, photographs of actual samples in each states shown.

The Npf PCM forms a blue light-absorbing, Pb, state with a peak at 390 nm, upon illumination with 735 nm light (Figure 2.4B). Upon 405 nm illumination the protein appears to form an initial intermediate state that has absorbance bands at 575 nm and 675 nm. Over about 2.5 hours this state thermally converts to a red light-absorbing state (Figure 2.5), Pr, that shows a marked increase in the absorbance band at 675 nm and 373 nm. This is an unusual photocycle as the published phytochrome photocycles report two signalling states, whereas our data reveals an intermediate state that overtime grows into a third signalling state.
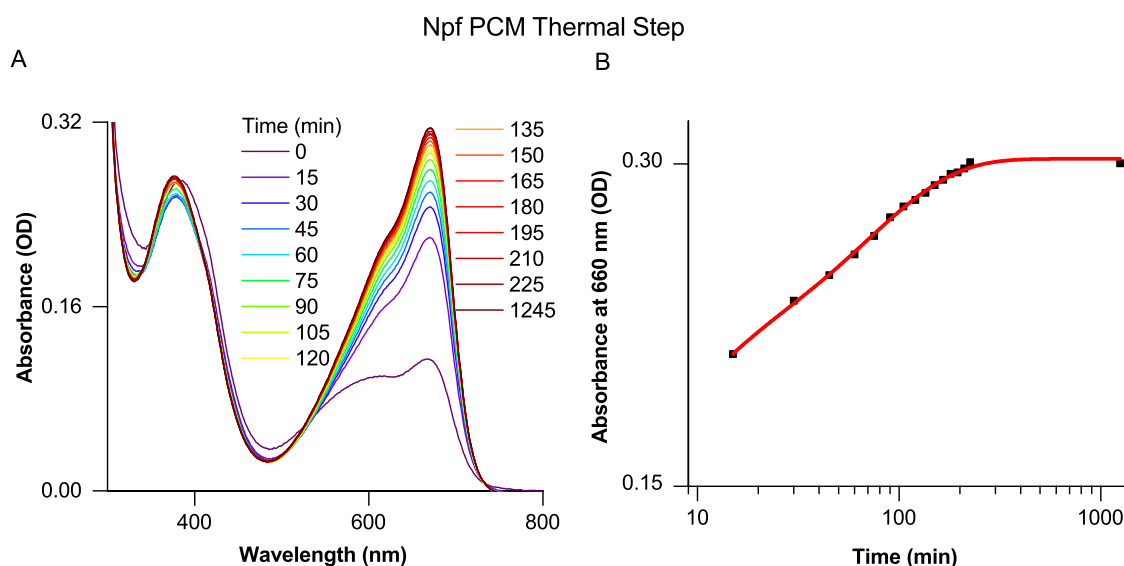
Figure 2.5. Spectra recorded at regular time intervals to follow the thermal conversion step. A. Spectra of Npf phytochrome Pb to Pr photoreaction – dark step – at regular time intervals over a few hours. **B.** Kinetics of formation of the Pr state.

2.2.4 Cryo-trapping of intermediates upon conversion of the Pr to Pb states for Npf Phytochrome

To further investigate the photoreactions of the Npf phytochrome time-resolved spectroscopy and trapping experiments at cryogenic temperatures (cryo-trapping) were conducted to capture the different intermediates formed during the reaction cycle. Absolute and difference absorbance spectra of the fully converted Pb and Pr states at 77 K confirm that both states are stable at cryogenic temperatures (Figure 2.6). In addition, the spectra obtained are very similar in shape to the room-temperature measurements (Figure 2.4B), although some of the smaller peaks have been sharpened up at 77 K.
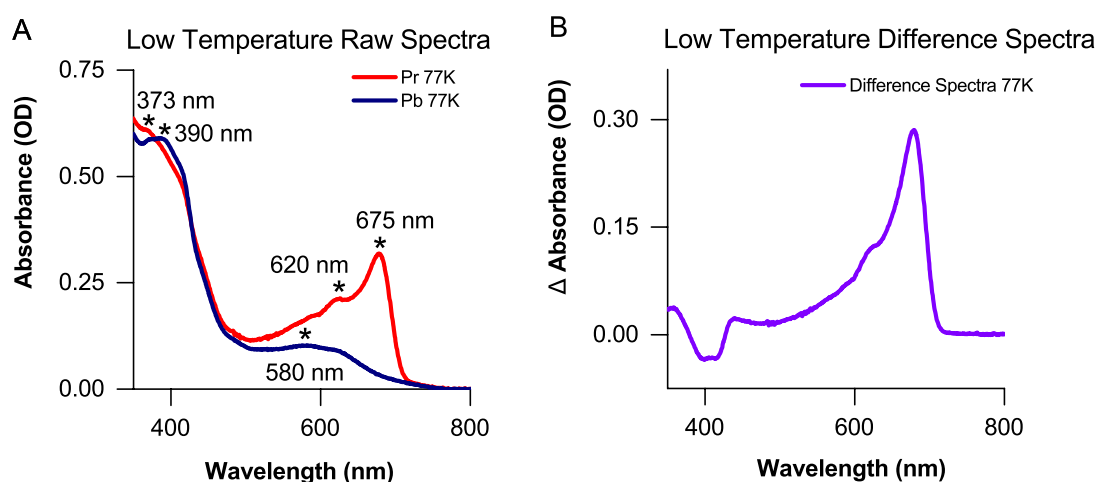
Figure 2.6. Low temperature absorption spectra of the Npf phytochrome. A: Absorption spectra of Pb and Pr states at 77 K. **B**: Difference spectrum of Npf Pr state in relation to Pb state at 77 K.

Cryo-trapping studies can be used to identify the energetic and thermal barriers associated with the formation of different intermediates along the reaction pathway. The reaction was monitored to follow the formation of the Pr and Pb states after illumination with blue and red light respectively, at temperatures ranging from 77 K to 293 K in 10 K increments. The Pr to Pb conversion was monitored by illuminating samples with red light at temperatures ranging from 77 to 210 K, and also by subsequently warming illuminated samples to higher temperatures in the dark from 210 to 293 K (Figure 2.7A). This allowed the reaction to be split into light-dependent and light-independent (dark) steps. The absolute 77 K absorbance spectra for the light-dependent step, recorded after illumination for 15 mins at temperatures between 77 and 210 K, are shown in Figure 2.7B. In order to accurately identify any changes in absorbance difference spectra were calculated by subtraction of the starting Pr ground state spectrum (Figure 2.7C). The data appear to show that two light-dependent steps can be identified upon illumination at different temperatures with multiple spectral changes taking place. A temperature dependence of these light steps was obtained by plotting the absorbance changes at 675 nm, 625 nm and 390 nm against the temperature of illumination (Figure 2.7D). An initial light-dependent reaction, observed at temperatures below 140 K, involves the bleach of the Pr ground state, which absorbs maximally at 675 nm, concomitant with the formation of a species absorbing at 625 nm. Upon illumination at

temperatures above 170 K there is a more significant decrease in the 675 nm band and the positive feature centred at 625 nm shifts towards shorter wavelengths. In addition, this is coupled to a decrease in an absorbance band at 390 nm. Both of these light-dependent steps (midpoint temperatures at ~100 K and 170 K) can occur at temperatures below 200 K, which is generally regarded to be the 'glass transition' temperature of proteins.[21-23] As most protein motions tend to be frozen out below this temperature it implies that the light steps do not require any degree of protein conformation change in order to proceed. Hence, it is most likely that they represent the isomerisation of the bilin cofactor and minor reorganisation of the bilin binding pocket.[10,13-17,34,35]

Figure 2.7. Low temperature absorption spectra to show cryo-trapping of intermediates in the Pr to Pb conversion. A: Absorbance spectra obtained at 77 K after illumination (or dark incubation) of the Pr state of Npf PCM with red light at temperatures between 77 K and 293 K. **B:** Absorbance spectra of the light-dependent step of Pr to Pb conversion obtained at 77 K after illumination with red light at temperatures between 77 K and 210 K. **C:** Difference spectra of the light-dependent step, using the non-illuminated 77 K spectrum as the background. **D:** Change in absorption at 390 nm, 625 nm and 675 nm after illumination at a range of temperatures between 77 K and 210 K. The dotted line identifies the temperatures at which significant changes in absorbance occur.

In order to identify any light-independent or dark steps in the Pr to Pb conversion samples that had been illuminated at 210 K were warmed to progressively higher temperatures in the dark and absolute absorbance spectra recorded at 77 K (Figure 2.8A). Absorbance difference spectra, calculated by subtraction of the spectrum of the state that was illuminated at 210 K, appear to show that there are two additional dark steps to form the final Pb state (Figure

2.8B). At temperatures between 220 K – 240 K there is a decrease in absorbance at 625 nm, which is accompanied by an increase in absorbance at 675 nm. At temperatures above 240 K there is a subsequent increase in the 390 nm absorbing state, concomitant with a decrease in absorbance at 440 nm. The temperature dependencies of these dark steps were obtained by plotting the absorbance changes at 675 nm, 625 nm, 440 nm and 395 nm against the temperature at which the samples were incubated (Figure 2.8C). This shows that both of these light-independent steps can only occur well above 200 K, with midpoint temperatures at ~230 K and ~260 K. Given that proteins generally undergo a glass transition at ~200 K, below which large structural changes are generally frozen out, it is likely that the absorbance changes that accompany these dark steps correlate with large-scale protein dynamics.[21,22,36,37] Therefore, based on these cryo-trapping experiments it appears that the Pr to Pb conversion proceeds via 4 separate steps, two that are light-dependent and two that are light-independent. The species that result from these steps can be represented by the absolute absorbance spectra recorded at 140 K, 210 K, 240 K and 293 K (Figure 2.9A and B).
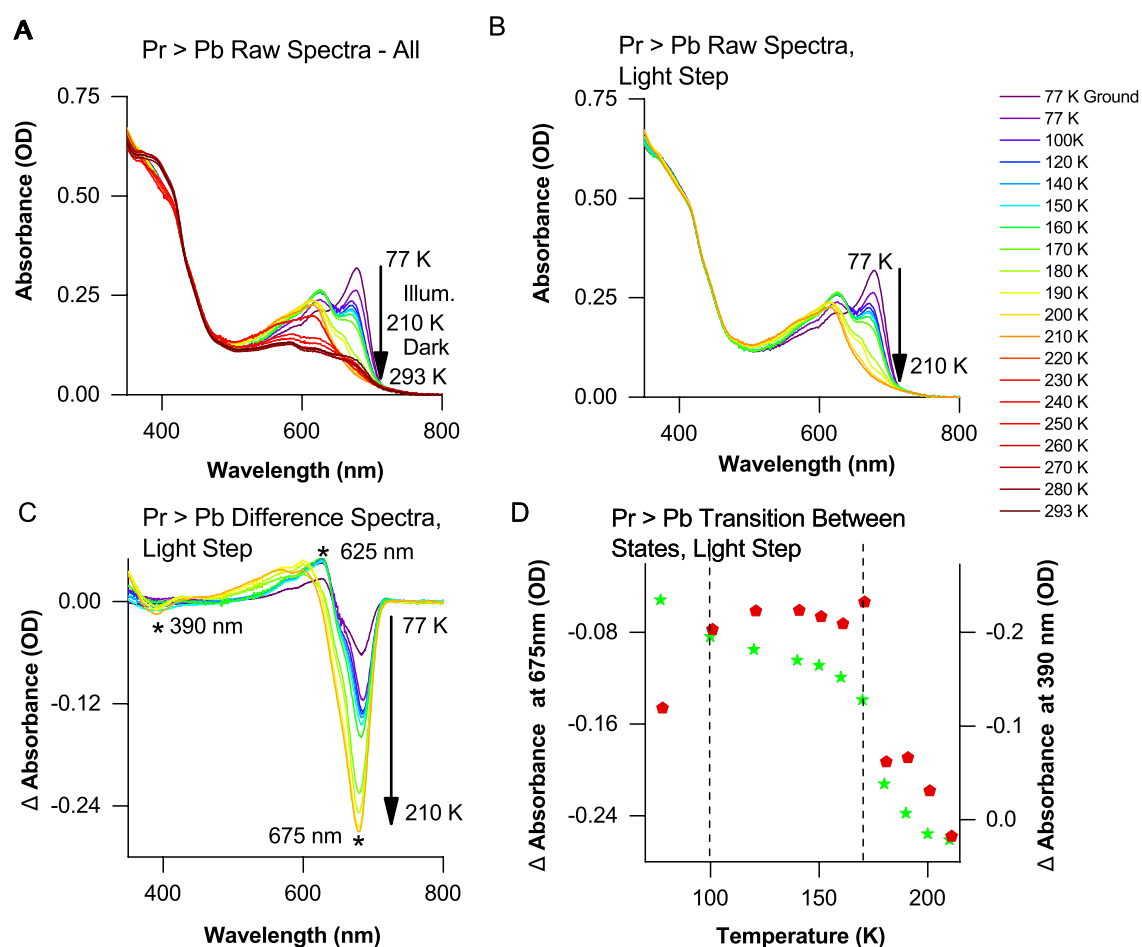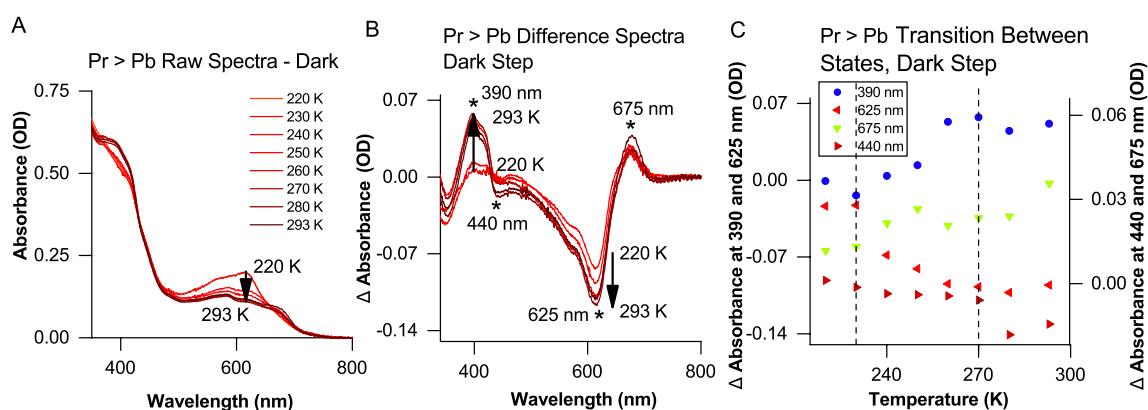

Figure 2.8. Low temperature absorption spectra to show cryo-trapping of intermediates in the Pr to Pb conversion. A. Absorbance spectra of the dark step obtained at 77 K after illumination with red light of the Pr state of Npf PCM at 210 K and subsequent incubation in the dark at a range of temperatures between 220 K and 293 K**. B:** Difference spectra of the dark steps measured at a temperature range between 220 K and 293 K, using the sample illuminated at 210 K as the background. **C:** Change in absorption at 395 nm, 440 nm, 625 nm and 675 nm after illumination at a range of temperatures between 220 K and 293 K. The dotted line identifies the temperature at which noticeable changes in absorbance occur.

2.2.5 Kinetics of Pr to Pb Photoconversion in Npf Phytochrome using Time-Resolved Spectroscopy

Time-resolved spectroscopy experiments were used in order to measure the rates of formation and decay of the intermediate states in the Pr to Pb photoconversion. Time-dependent changes in absorbance were recorded using laser flash photolysis from 0.25 µs to 2.5 s after photoexcitation of the Pr state with 660 nm laser pulse (Figure 2.9C). Global analysis, using a sequential model of the complete dataset, resulted in the Evolution Associated Difference Spectra (EADS) shown in Figure 2.9D. The first EADS, representing the species formed within the 0.25 µs time resolution of the experiment shows a bleach at 675 nm, and a positive feature at 535 nm. The difference spectrum of this intermediate is very similar to the species that was formed after illumination at 210 K, which represents the intermediate formed after isomerisation and minor changes to the bilin binding pocket. The subsequent EADS, formed in 76.2 ms and 2.38 s respectively, are very similar in the red region of the spectrum, with a further decrease in absorbance at 675 nm and a larger decrease at 625 nm. However, in the blue region extra bleach features at 373 and 440 nm appear in EADS3, which represents the final Pb state of the protein. These spectra are comparable to the intermediates trapped at 240 K and 293 K.
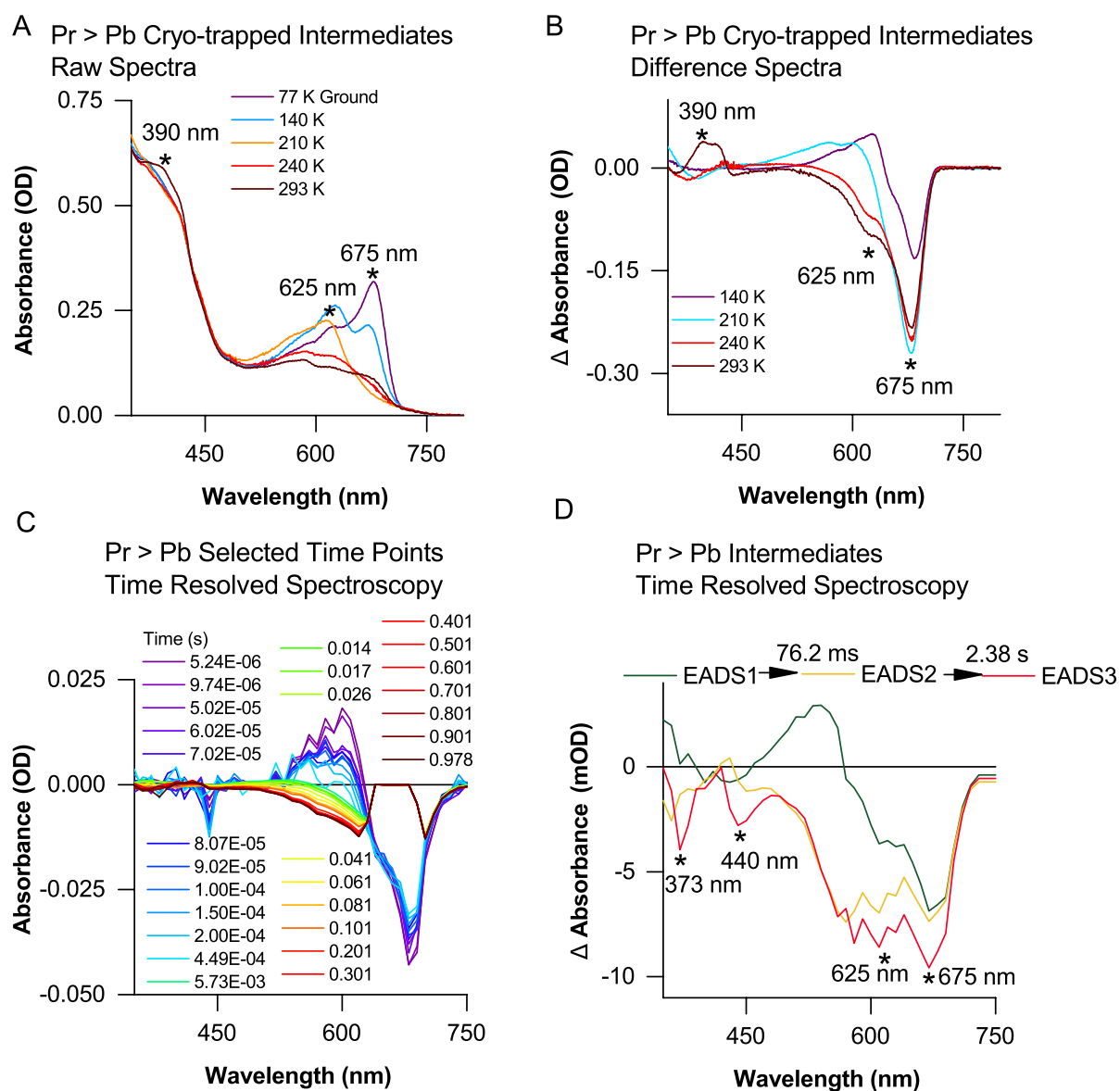
Figure 2.9. Time resolved and cryo-trapped spectral changes observed in the Pr to Pb photoconversion upon excitation at 660 nm. A. Raw absorbance spectra of intermediates cryo-trapped at different temperatures. **B.** Difference spectra of intermediates cryo-trapped at different temperatures. **C.** Absorbance difference spectra of the Pr to Pb photoreaction recorded at regular time intervals after excitation at 660 nm. **D.** EADS obtained from global analysis of the time resolved spectroscopy data recorded from 2.5 μs to 2.5 s after excitation at 660 nm.

2.2.6 Cryo-trapping of intermediates upon conversion of the Pb to Pr states for Npf Phytochrome

The Pb to Pr reaction was monitored by illuminating samples with blue light at temperatures ranging from 77 to 210 K, and also by subsequently warming

illuminated samples to higher temperatures in the dark from 210 to 293 K (Figure 2.10A). Although the spectral changes are much smaller in magnitude than those observed for the Pr to Pb photoconversion, the reaction was again split into light-dependent and light-independent (dark) steps. The absolute 77 K absorbance spectra for the light-dependent step, recorded after illumination for 15 mins at temperatures between 77 and 210 K, are shown in Figure 2.10B. Absorbance difference spectra, calculated by subtraction of the starting Pb ground state spectrum (Figure 2.10C) appear to show that an initial light-dependent reaction can be observed at temperatures below 200 K. This involves the bleach of the Pb ground state, which absorbs maximally at 390 nm, concomitant with the formation of a species that has absorbance peaks at 580 nm, 620 nm and 430 nm. The temperature dependence of this step was obtained by plotting the absorbance change at 350 nm, 390 nm, 430 nm, 580 nm and 675 nm against the temperature (Figure 2.10D). This reveals that the light step (midpoint temperature ~180 K) can occur at temperatures below 200 K and again, suggest that no large-scale protein conformation changes are required at this stage. Hence, it is likely that this light step simply represents the isomerisation of the bilin cofactor.

Figure 2.10 Low temperature absorption spectra to show cryo-trapping of intermediates in the Pb to Pr conversion, Absorbance spectra obtained at 77 K after illumination of the Pr state of Npf PCM with blue light (or dark incubation) at temperatures between 77 K and 293 K. **B:** Absorbance spectra of the light-dependent step of Pb to Pr conversion obtained at 77 K after illumination with blue light at temperatures between 77 K and 210 K. **C:** Difference spectra of the light-dependent step, using the non-illuminated sample as the background. **D:** Change in absorption at 350nm, 395 nm, 430 nm, 580 nm and 675 nm after illumination at a range of temperatures between 77 K and 200 K. The dotted line identifies the temperature at which noticeable changes in absorbance are initiated – 180 K.

In order to identify any light-independent or dark steps in the Pb to Pr conversion samples that had been illuminated at 210 K were warmed to progressively higher temperatures in the dark and absolute absorbance spectra recorded at 77 K (Figure 2.11A). Absorbance difference spectra, calculated by subtraction of the spectrum of the state that was illuminated at 210 K, appear to show that there is at least one additional dark step to form the final Pr state (Figure 2.11B). At temperatures between 220 K – 270 K there is a bleach at

430 nm and 580 nm, concomitant with formation of a 675 nm and 390 nm absorbing state. These features grow in even further after incubation at 293 K, probably reflecting the onset of the slow thermal conversion to the final Pr state observed in the previous room temperature measurements (Figure 2.5). The temperature dependencies of these dark steps were obtained by plotting the absorbance changes at 390 nm, 430 nm, 580 nm and 675 nm against the temperature at which the samples were incubated (Figure 2.11C). This shows that both of the light-independent step(s) can only occur well above 200 K, with midpoint temperatures at ~230-240 K. Again, it is likely that this step to form the state that is the precursor to final thermally converted Pr state involves large-scale protein dynamics.[21,22,36,37] Therefore, based on these cryo-trapping experiments it appears that the Pb to Pr conversion proceeds via at least 2 separate identifiable steps, one that is light-dependent and one that is light-independent. The intermediate species that result from these steps can be represented by the absolute absorbance spectra recorded at 200 K, 270 K and 293 K (Figure 2.12 A and B).
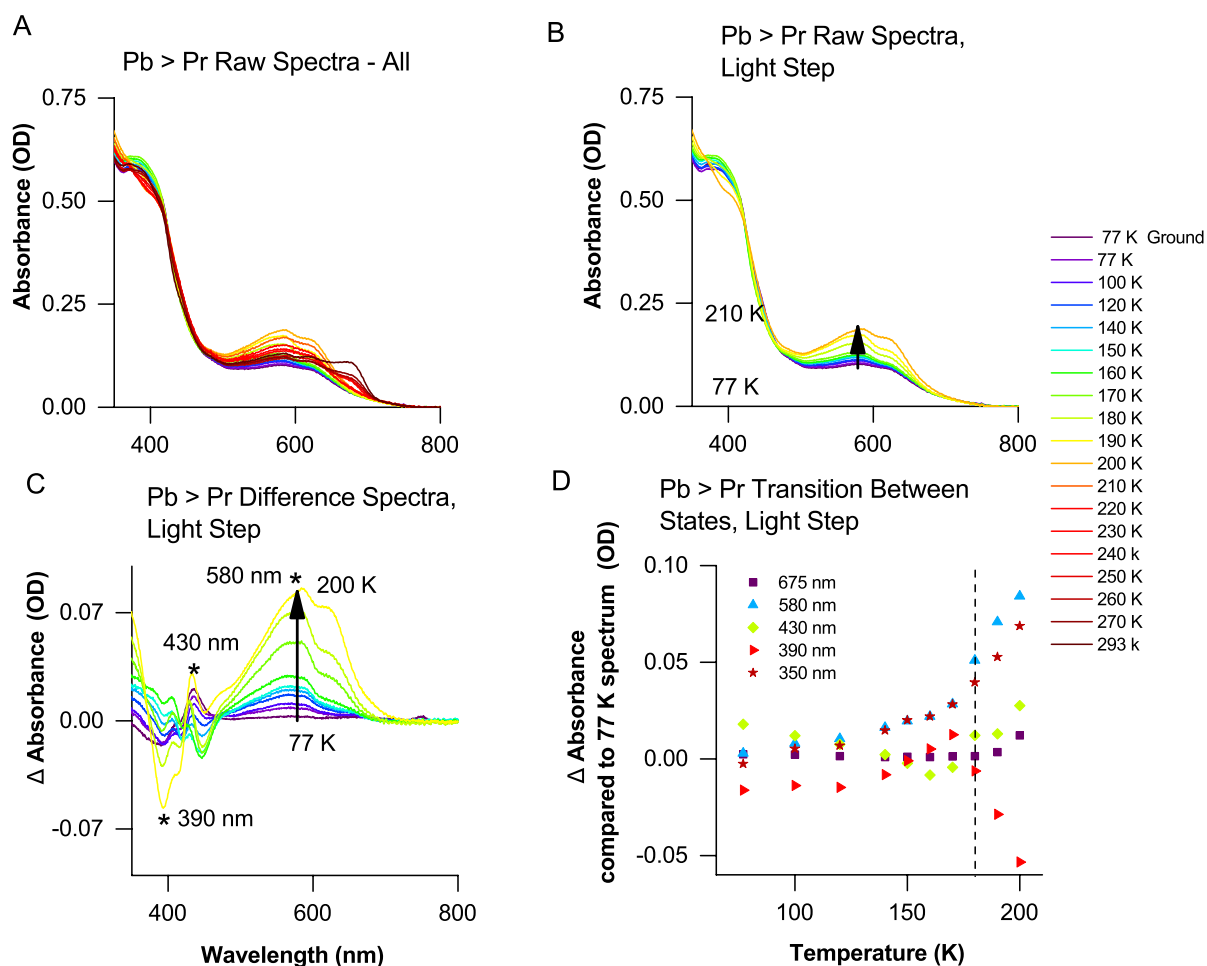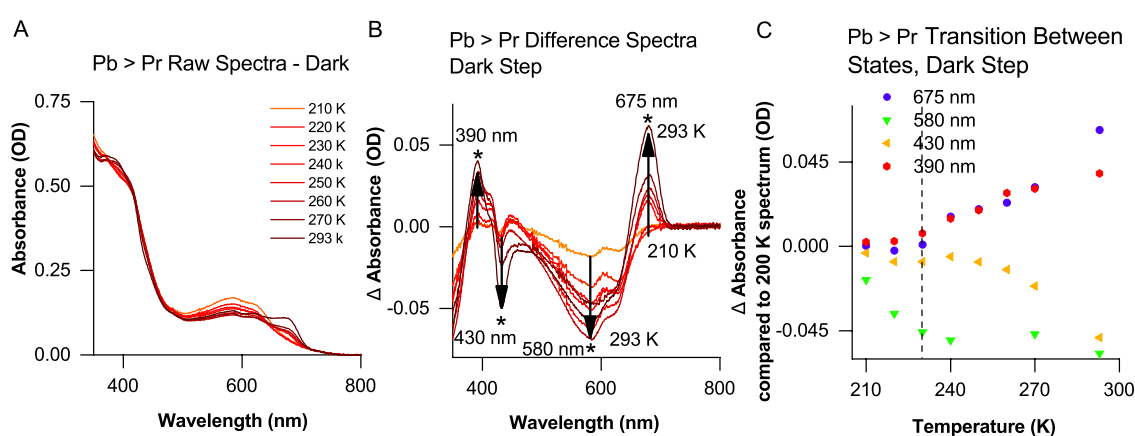

Figure 2.11 Low temperature absorption spectra to show cryo-trapping of intermediates in the Pb to Pr conversion A. Absorbance spectra of the dark step obtained at 77 K after illumination with blue light of the Pb state of Npf PCM at 210 K and subsequent incubation in the dark at a range of temperatures between 210 K and 293 K. **B:** Difference spectra of the dark step using the sample illuminated at 200 K as the background. **C:** Change in absorption at 350 nm, 390nm, 430 nm, 580 nm and 675 nm after incubation at a range of temperatures between 210 K and 293 K. The dotted line identifies the temperature at which noticeable changes in absorbance occur.

2.2.7 Kinetics of Pb to Pr Photoconversion in Npf Phytochrome using Time-Resolved Spectroscopy

Finally, time-resolved spectroscopy studies were used to measure the rates of formation and decay of the intermediate states in the Pb to Pr photoconversion. Time-dependent changes in absorbance were recorded using laser flash photolysis from 0.25 µs to 2.5 s after photoexcitation of the Pr state with a 420 nm laser pulse (Figure 2.12C). It should be noted that the magnitude of the spectral changes observed in the laser flash photolysis experiments are very small for the Pb to Pr photoconversion, resulting in data with a lower signal-to-noise-ratio compared to the Pr to Pr photoconversion. Global analysis using a sequential model of the complete dataset resulted in the EADS shown in Figure 2.12D. The first EADS, representing the species formed within the 0.25 µs time resolution of the experiment shows a broad increase in an absorbance band centred at 430 nm, with minor changes in the red region of the spectrum. The difference spectrum of this intermediate is different to the species that was formed after illumination at 180 K, and hence it is likely that the protein has already undergone photoisomerisation and further conversion within the time resolution of these experiments. EADS2, formed in 19.2 ms, shows a small increase in absorbance at 660 nm and is similar to the intermediate observed at 250 K in the cryo-trapping experiments. However, unlike the cryo-trapping studies, the emergence of the 675 nm absorbing species is not seen in the time-resolved measurements, probably as this state forms over much longer timescales through a slow thermal conversion step, with lifetimes of 6.3 and 70 minutes.

A Pb > Pr Cryotrapped Intermediates Raw Spectra

B Pb > Pr Cryotrapped Intermediates Difference Spectra

C Pb > Pr Selected Time Points Time Resolved Spectroscopy

D Pb > Pr Intermediates Time Resolved Spectroscopy
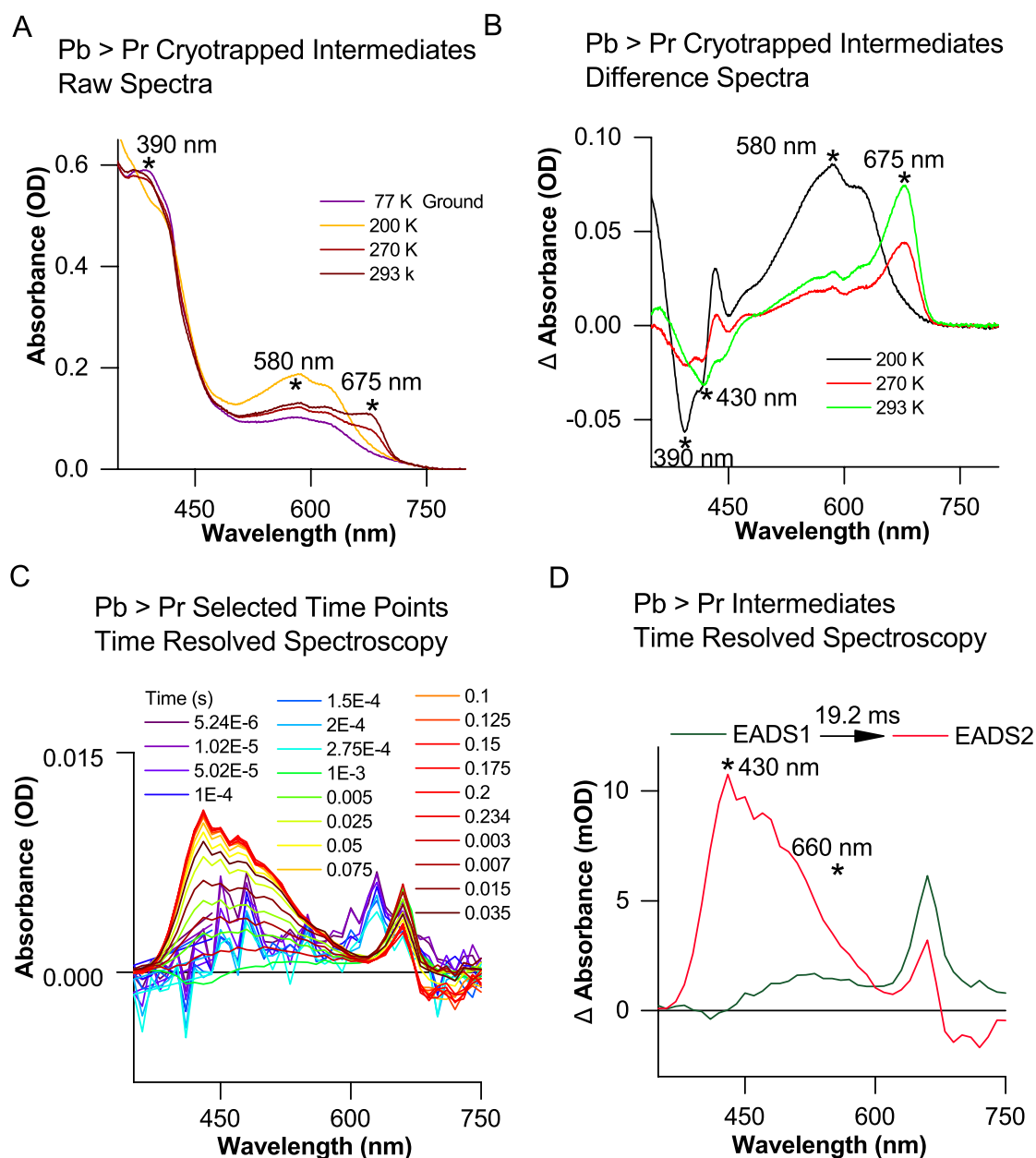
Figure 2.12. **Time resolved spectral changes observed in the Pr to Pb photoconversion upon excitation at 420 nm. A.** Raw spectra of different intermediates cryo-trapped at different temperatures in the formation of the Pr state from the Pb state. **B.** Difference spectra of different intermediates cryo-trapped at different temperatures in the formation of the Pr state from the Pb state. **C.** Absorbance difference spectra of the Pb to Pr photoreaction recorded at regular time intervals after excitation at 660 nm. **D.** Plot of EADS data obtained from global analysis of the time resolved spectroscopy data for upon excitation of the Pb state.

2.3 Conclusion

We have reported the expression, purification and initial characterisation of the photosensory regions of different atypical phytochromes. Flash photolysis and low temperature studies were used to characterise any intermediate states that are formed during the forward and reverse photoreactions of one of these proteins, namely Npf phytochrome. The cryogenic measurements have facilitated the identification of the thermal and energetic barriers for different steps in the photocycle, allowing assignment of these steps to localised changes within the chromophore or more global changes in protein structure. large scale protein dynamics. The time-resolved spectroscopy data have been used to identify the timescales at which many of these intermediates appear, leading to a complete description of the photocycle of this protein (Figure 2.13). Overall, it appears as though the blue / red photocycle of Npf phytochrome is broadly similar to the red / far-red photocycle in typical phytochromes, with an initial photoisomerisation followed by multiple large-scale structural changes in the protein.[38-42] However, it should be noted that the slow thermal conversion that leads to the formation of the final Pr state of the protein is particularly unusual and to our knowledge hasn't been observed in typical red /far-red phytochromes. These studies have provided the first mechanistic understanding of an atypical phytochrome and illustrates the complex photocycles that underpin their function.
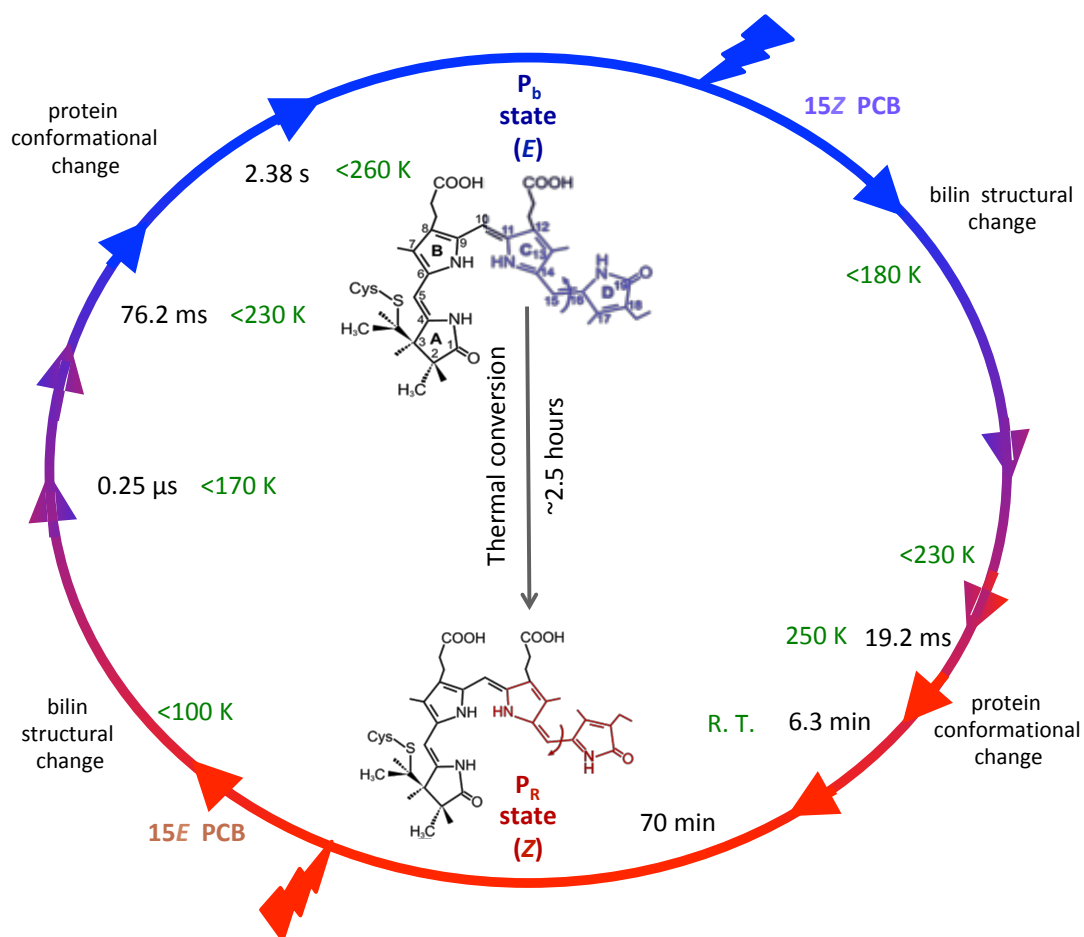
Figure 2.13 – Photocycle of Npf phytochrome displaying the steps by transient absorption and cryo-trapping measurements

2.4 Methodology:

2.4.1 Culture Conditions and Expression Trials

Expression trials were set-up using Luria Broth (LB) media. All expression trials were conducted using 50 ml media and large-scale cultures were on litre scales. The starter culture media was inoculated with HMS174 colonies containing pCola duet and Npf pET15b and incubated at 30°C, 190 rpm. When the cultures reached an OD_{600} of 0.1, the temperature was reduced to 25°C. The cultures were induced with 100 μM IPTG when the cultures reached an OD_{600} of between 0.4 and 0.6, and grown overnight. Trials conducted with benzyl alcohol were grown as above except for the addition of 0 g/L, 0.25 g/L and 0.5 g/L benzyl alcohol 20 minutes before induction. To harvest the cells, the cultures were centrifuged at 4000 rpm for 15-20 minutes and the pellets were stored at 4°C. Trials conducted with ArcticExpress cells were performed according to the supplier protocol and benzyl alcohol was added to the same concentration as outlined above 20 minutes before induction. Trials conducted with sorbitol and betaine involved addition of 1 M sorbitol and 2.5 mM sorbitol to the media.

2.4.2 Cell Lysis

For expression trials, the cell pellets were re-suspended in binding buffer. A Sonics Vibracell sonicator was used to lyse the cells by 4 pulses of sonication at 25% amplitude for 20 seconds with a 50 seconds interval between each pulse. For large-scale grow-ups and purification, the pellets were re-suspended using a 2:3 ratio of cell pellet:binding buffer containing Sigma Aldrich EDTA-free protease inhibitor. The cells were lysed using a cell disruptor (Constant Systems) at 15 kpsi pressure. The lysate was centrifuged at 19,000 rpm for one hour.

2.4.3 Affinity Chromatography Purification

Nickel-IDA resin was used as the first purification step for all phytochromes that were expressed. The resin was equilibrated with 5 column volumes buffer A (50 mM Tris pH 7.5, 300 mM NaCl, 5 mM imidazole, 7.5 % glycerol, 1 % β-

mercaptoethanol) and incubated with the protein for 3 to 4 hours. The protein flow thorough was collected and the resin was washed with 4 column volumes of buffer B (50 mM Tris pH 7.5 300 mM NaCl, 30 mM imidazole, 7.5 % glycerol, 1 % β-mercaptoethanol). The protein was eluted with buffer C (50 mM Tris pH 7.5, 300 mM NaCl, 250 mM imidazole, 7.5% glycerol, 1 % β-mercaptoethanol).

2.4.4 Size Exclusion Purification

After the affinity chromatography purification the proteins were concentrated down to a volume of 5 ml for injection onto HiLoad 26/60 200 (GE Healthcare). Chromatography was performed on an ÄKTA Pure (GE Healthcare). The gel filtration column was equilibrated with two column volumes of gel filtration equilibration buffer D (50 mM Tris pH 7.5, 300 mM NaCl, 7.5 % glycerol, 1% β-mercaptoethanol) before the protein was injected. After injection, the protein was eluted in two-column volume gel filtration equilibration buffer. The relevant fractions were analysed using SDS-PAGE to pool together the fractions containing the most pure protein. The protein was then concentrated down using a Sartorius Stedim Biotech vivaspin column with a 30 kDa cut-off, flash-frozen and stored at -80°C.

2.4.5 SDS-Polyacrylamide Gel Electrophoresis

To assess protein purity aliquots of fractions from protein purifications were resolved on 12% Mini-PROTEAN® TGX™ Stain-Free™ precast gels along with 7 μl of Precision Plus Protein™ Standard Unstained marker (Bio-Rad), using the Laemmli buffer compositions [43]. SDS loading buffer (50 mM Tris-HCl, 100 mM DTT, 2 % SDS [w/v] 0.1 % bromophenol blue [w/v], 10 % glycerol [v/v], 2.5 % 2-mercaptothanol [v/v]) was added to the protein sample in a 1:1 ratio. The gels were run at a constant voltage, 300 V, for 20 to 25 minutes. The gel was imaged on a Gel Doc EX Imager, using Image Lab 3.0.1 Software (Bio-Rad).

2.4.6 Cryo-trapping of intermediates:

For cryo-trapping studies, the chimeric protein was expressed and purified as above and suspended in 80% cryo-buffer (44 % glycerol v/v, 20 % sucrose v/v, 50 mM Tris-HCl). The protein was converted to a single state using LED illumination at room temperature prior to cooling and the temperature was controlled between 77 K and 293 K in a liquid nitrogen cooled cryostat (Optistat

DN). Once the cryostat reached the desired temperature, the protein was equilibrated at the temperature for 15 minutes. Then the reaction was triggered by 10 minutes of illumination with the relevant LED or for the dark steps 10 minutes incubation in dark. The sample was then allowed to warm up in 10-K steps up to 293 K, illuminated Pr state at 660 nm and Pb state at 405 nm or incubated in the dark for the dark steps, for 10 min at each temperature point. Samples were then cooled back down to 77 K prior to record a UV-visible absorbance spectrum in a Cary 50 spectrometer (Agilent Technology). The power outputs of the 405 nm LED and the 660 nm LED were set to 750 µmol $m^2 \, s^{-1}$. Typical protein concentrations used in these experiments were between 5 µM and 10 µM.

2.4.8 UV/Vis Spectroscopy

A Cary 50 UV-Vis Spectrophotometer, Agilent Technologies, was used to obtain the steady-state absorbance spectra of these phytochromes. All spectra were recorded from 350 to 800 nm. Thorlabs mounted high power LEDs of relevant wavelength were used in the dark to convert the phytochromes into their different isomeric states. For phytochromes that undergo dark thermal conversion over time, the cuvette was incubated in the dark within the spectrometer. The data was analysed using OriginPro 2015 Software, Origin Lab Corporation.

2.4.9 Nanosecond laser flash photolysis

The phytochrome photodynamics were studied using a laser system consisting of an optical parameter oscillator (OPO) pumped by a Q-switched Nd: YAG laser (Brilliant B, Quantel). Transient absorption (TA) spectra of the phytochromes were recorded after excitation with pump-beams centered at relevant wavelengths using the OPO pumped by the third harmonic of the laser. A quartz cuvette with a 2 mm path-length for pump-excitation and 10 mm path length for probing of the TA spectra was used. TA spectra were recorded using an LKS-60 Flash-Photolysis instrument (Applied Photophysics Ltd.). Three repeats of 350 nm to 750 nm scans were recorded in 10 nm increments in a selective time window with and without illumination with LEDs of relevant wavelengths. The data was analysed using OriginPro 2015 Software, Origin Lab Corporation. Protein concentration used for these experiments varied between 10 µM and 15 µM of protein.

Global Analysis of the flash photolysis data was performed using the open-source software Glotaran to obtain plots of evolution associated difference spectra (EADS).[44] The pre-excitation data was subtracted for the analysis and both the 'slow' and 'fast' datasets were fitted to a sequential model. According to the sequential model, one species converts to another, which persists for the lifetime of the experiment.

2.5 Reference List:

[1]   W. L. Butler, K. H. Norris, H. W. Siegelman, S. B. Hendricks, *Proc. Natl. Acad. Sci. U.S.A.* **1959**, *45*, 1703–1708.

[2]   T. Shinomura, *J. Plant Res.* **1997**, *110*, 151–161.

[3]   J. W. Reed, A. Nagatani, T. D. Elich, M. Fagan, J. Chory, *Plant Physiol.* **1994**, *104*, 1139–1149.

[4]   K. A. Franklin, S. J. Davis, W. M. Stoddart, R. D. Vierstra, G. C. Whitelam, *Plant Cell* **2003**, *15*, 1981–1989.

[5]   E. Johnson, M. Bradley, N. P. Harberd, G. C. Whitelam, *Plant Physiol.* **1994**, *105*, 141–149.

[6]   A. Möglich, R. A. Ayers, K. Moffat, *J. of Mol. Biol.* **2009**, *385*, 1433–1444.

[7]   P. H. Quail, *Plant Cell Environ.* **1997**, *20*, 657–665.

[8]   L.-O. Essen, J. Mailliet, J. Hughes, *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 14709–14714.

[9]   Y. S. Ho, L. M. Burden, J. H. Hurley, *EMBO J.* **2000**, *19*, 5288–5299.

[10]   N. C. Rockwell, Y.-S. Su, Lagarias, *Annu. Rev. Plant Biol.* **2006**, *57*, 837–858.

[11]   A. Levskaya, A. A. Chevalier, J. J. Tabor, Z. B. Simpson, L. A. Lavery, M. Levy, E. A. Davidson, A. Scouras, A. D. Ellington, E. M. Marcotte, et al., *Nature* **2005**, *438*, 441–442.

[12]   N. C. Rockwell, Lagarias, *Plant Cell* **2006**, *18*, 4–14.

[13]   F. Andel, Lagarias, R. A. Mathies, *Biochemistry* **1996**, *35*, 15997–16008.

[14]   H. Foerstendorf, T. Lamparter, J. Hughes, W. Gärtner, F. Siebert, *Photochem. Photobiol.* **2000**, *71*, 655–661.

[15]   C. Kneip, P. Hildebrandt, W. Schlamann, S. E. Braslavsky, F. Mark, K. Schaffner, *Biochemistry* **1999**, *38*, 15185–15192.

[16]   W. Rüdiger, F. Thümmler, E. Cmiel, S. Schneider, *Proc. Natl. Acad. Sci. U.S.A.* **1983**, *80*, 6244–6248.

[17]   C. Schumann, R. Gross, M. M. N. Wolf, R. Diller, N. Michael, T. Lamparter, *Biophys. J.* **2008**, *94*, 3189–3197.

[18]   P. Schmidt, T. Gertsch, A. Remberg, W. Gärtner, S. E. Braslavsky, K. Schaffner, *Photochem. Photobiol.* **1998**, *68*, 754–761.

[19]   E. S. Burgie, T. Wang, A. N. Bussell, J. M. Walker, H. Li, R. D. Vierstra, *J. Biol. Chem.* **2014**, *289*, 24573–24587.

[20]   H. Takala, A. Björling, O. Berntsson, H. Lehtivuori, S. Niebling, M. Hoernke, I. Kosheleva, R. Henning, A. Menzel, J. A. Ihalainen, et al., *Nature* **2014**, *509*,

245–248.

[21]   D. Vitkup, D. Ringe, G. A. Petsko, M. Karplus, *Nat. Struct. Biol.* **2000**, *7*, 34–38.

[22]   G. Durin, A. Delaunay, C. Darnault, D. J. Heyes, A. Royant, X. Vernede, C. N. Hunter, M. Weik, D. Bourgeois, *Biophys. J.* **2009**, *96*, 1902–1910.

[23]   A. F. E. Hauck, S. J. O. Hardman, R. J. Kutta, G. M. Greetham, D. J. Heyes, N. S. Scrutton, *J. Biol. Chem.* **2014**, *289*, 17747–17757.

[24]   A. H. West, A. M. Stock, *Trends Biochem. Sci.* **2001**, *26*, 369–376.

[25]   K. Yeh, *Science* **1997**, *277*, 1505–1508.

[26]   E. Giraud, J. Fardoux, N. Fourrier, L. Hannibal, B. Genty, P. Bouyer, B. Dreyfus, A. Verméglio, *Nature* **2002**, *417*, 202–205.

[27]   M. Jaubert, J. Lavergne, J. Fardoux, L. Hannibal, L. Vuillet, J.-M. Adriano, P. Bouyer, D. Pignol, E. Giraud, A. Verméglio, *J. Biol. Chem.* **2007**, *282*, 7320–7328.

[28]   M. E. Auldridge, K. T. Forest, *Crit. Rev. Biochem. Mol. Biol.* **2015**, *46*, 67–88.

[29]   N. C. Rockwell, D. Duanmu, S. S. Martin, C. Bachy, D. C. Price, D. Bhattacharya, A. Z. Worden, Lagarias, *Proc. Natl. Acad. Sci. U.S.A.* **2014**, *111*, 3871–3876.

[30]   N. C. Rockwell, S. S. Martin, K. Feoktistova, Lagarias, *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, 11854–11859.

[31]   A. de Marco, L. Vigh, S. Diamant, P. Goloubinoff, *Cell Stress Chaper* **2005**, *10*, 329.

[32]   N. Oganesyan, I. Ankoudinova, S.-H. Kim, R. Kim, *Protein Expr. Purif.* **2007**, *52*, 280–285.

[33]   J. R. Blackwell, R. Horgan, *FEBS Lett.* **1991**, *295*, 10–12.

[34]   A. Remberg, I. Lindner, T. Lamparter, J. Hughes, C. Kneip, P. Hildebrandt, S. E. Braslavsky, W. Gärtner, K. Schaffner, *Biochemistry* **1997**, *36*, 13389–13395.

[35]   M. Schmitt, B. Dietzek, G. Hermann, J. Popp, *Laser & Photonics Review* **2007**, *1*, 57–78.

[36]   D. J. Heyes, A. V. Ruban, H. M. Wilks, C. N. Hunter, *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 11145–11150.

[37]   D. J. Heyes, A. V. Ruban, C. N. Hunter, *Biochemistry* **2003**, *42*, 523–528.

[38]   M. G. Müller, I. Lindner, I. Martin, W. Gärtner, A. R. Holzwarth, *Biophys. J.* **2008**, *94*, 4370–4382.

[39]   M. Bischoff, G. Hermann, S. Rentsch, D. Strehlow, *Biochemistry* **2001**, *40*, 181–186.

[40]   K. Heyne, J. Herbst, D. Stehlik, B. Esteban, T. Lamparter, J. Hughes, R. Diller, *Biophys. J.* **2002**, *82*, 1004–1016.

[41]   K. Anders, D. von Stetten, J. Mailliet, S. Kiontke, V. A. Sineshchekov, P. Hildebrandt, J. Hughes, L.-O. Essen, *Photochem. Photobiol.* **2011**, *87*, 160–173.

[42]   K. Anders, A. Gutt, W. Gärtner, L.-O. Essen, *J. Biol. Chem.* **2014**, *289*, 25590–25600.

[43]   U. K. LAEMMLI, *Nature* **1970**, *227*, 680–685.

[44]   J. J. Snellenburg, S. P. Laptenok, R. Seger, K. M. Mullen, I. H. M. van Stokkum, *J. Stat. Softw.* **2012**, *49*, DOI 10.18637/jss.v049.i03.

Chapter 3 Photochemical mechanism of an atypical algal phytochrome

Uzma Choudry[a]†, Derren J. Heyes[a] †, Samantha J. O. Hardman[a] †, Michiyo Sakuma[a], Igor V. Sazanovich[b], Joyce Woodhouse[c], Eugenio De La Mora[c], Martin N. Pedersen[d], Michael Wulff[d], Martin Weik[c], Giorgio Schirò[c], Nigel S. Scrutton* [a]

[a]   U. Choudry, Dr.D.J.Heyes, Dr.S.J.O.Hardman, M. Sakuma, Prof. N. S. Scrutton

Manchester Institute of Biotechnology, University of Manchester, 131 Princess Street, Manchester M1 7DN (UK)

E-mail: nigel.scrutton@manchester.ac.uk

[b]   Dr. I. V. Sazanovich

Central Laser Facility,Research Complex at Harwell, Science and Technology Facilities Council, Harwell Oxford, Didcot, OX11 0QX (UK)

[c]   Dr. G. Schirò, J. Woodhouse, Dr. E. De La Mora, Dr. M. Weik

CNRS, Université Grenoble Alpes, CEA - Institut de Biologie Structurale, Grenoble 38044 (France)

[d]   Dr. M. N. Pedersen, Dr. M. Wulff

European Synchrotron Radiation Facility, Grenoble 38044 (France)

†   These authors contributed equally to this work.

Abstract

Phytochromes are bilin-containing photoreceptors that are typically sensitive to the red / far-red region of the visible spectrum. Recently phytochromes from certain eukaryotic algae have become attractive targets for optogenetic applications because of their unique ability to respond to multiple wavelengths of light. Here, we have used a combination of time-resolved spectroscopy and structural approaches across picosecond to second timescales to map photochemical mechanisms and structural change in this atypical group of phytochromes. The photochemistry of an orange / far-red light sensitive algal phytochrome from *Dolihomastix tenuilepis* has been investigated using a combination of visible, infra-red and X-ray scattering probes. The entire photocycle, correlated with accompanying structural changes in the cofactor / protein, are reported. Our study identifies a complex photocycle for this atypical phytochrome. It also highlights a need to combine outcomes from a range of biophysical approaches to unravel complex photochemical and macromolecular processes in multi-domain photoreceptor proteins that are the basis of biological light-mediated signalling.

3.1 Introduction

The red / far-red light-sensitive phytochromes (Phys) are essential photoreceptors that regulate a myriad of photo-morphological processes in plants, bacteria, fungi and algae (e.g. seed germination, leaf-opening).[1,2] Phys contain a covalently bound bilin cofactor and typically photoconvert between a red light-absorbing P_R state and a far-red light-absorbing P_{FR} state.[1] Photoconversion between the two states is triggered by an ultrafast *E/Z* isomerization of the bilin tetrapyrrole, leading to slower protein structural change (microsecond to millisecond timescales) to form an active signalling conformation. The initial photoisomerisation has been studied in numerous Phys from different organisms using time resolved UV-Vis, infrared (IR) and Raman spectroscopies and is known to involve isomerization around the C15 and C16 double bond between ring C and D of the bilin.[1,3-9] X-ray

crystallography studies of the photosensory region of a bacterial Phy in conjunction with time-resolved X-ray scattering data have shown that this causes a subsequent refolding of a conserved tongue region leading to large-scale conformational change in the remainder of the protein.[10,11]

Although Phys have proven to be attractive targets for new applications in the field of optogenetics,[12,13] their use has been limited due to their sensitivity to only a narrow range of wavelengths (red and far-red light). However, it was recently discovered that Phys from certain eukaryotic algae exhibit diverse spectral properties that are capable of spanning most of the visible region.[14] The origin of this spectral diversity is currently unknown and it is unclear whether these algal Phys undergo similar photocycles to those found in typical Phys. In order to exploit fully the potential of algal Phys as new optogenetic tools, it is important to develop an understanding of the photochemistry and photodynamics that underpin biological signalling.

We have investigated in detail the kinetics and structural changes of both the bilin cofactor and the protein throughout the entire photocycle of the photosensory region of an algal Phy by using a combination of time-resolved laser photoexcitation spectroscopy across multiple timescales (continuously from picoseconds to seconds). Spectral changes in the visible region were monitored to report on significant chromophore changes whilst the mid-infra-red spectral region was used to report on finer structural changes in the bilin cofactor. These measurements have been combined with time-resolved small and wide angle X-ray scattering (S/WAXS) measurements to understand how these spectral changes are coupled to larger conformational changes in the protein.

3.2 Results and Discussion

The eukaryotic algal Phy from *Dolihomastix tenuilepis* (Dten) has a far-red light-absorbing, P_{FR}, state, with an absorption maximum at 714 nm and lower intensity features at 375 and 614 nm, and an orange light-absorbing, P_O, state with an absorption maximum at 598 nm and a lower intensity feature at 354 nm (Figure 1).[14] Time-resolved spectroscopy has been used to measure optical changes in the visible and infra-red (IR) regions that occur upon

photoconversion of both states of the protein. Measurements in the visible region spanned the complete time-range of the photoreaction (ps-ms). These measurements were repeated in H_2O and D_2O buffers to identify any kinetic isotope effects associated with each of the individual steps of the photocycle, and to provide direct comparisons to the IR data, collected on samples in D_2O buffer. The IR measurements were recorded over the first 400 µs of the reaction to monitor the initial structural changes of the reaction.
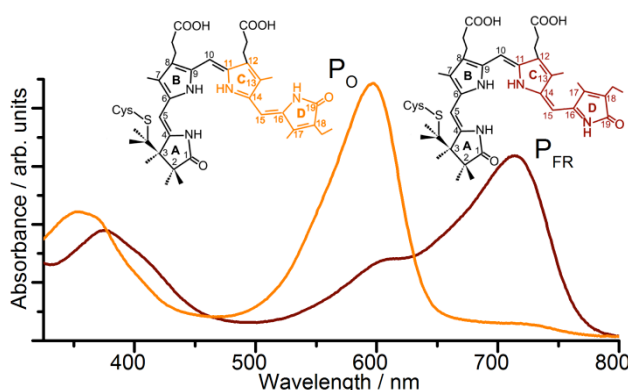

Figure 1. Absorption spectra, and structures of bilin chromophore, of Dten Phy in the orange-absorbing state and the far-red-absorbing states.

Global analysis has been used to model the time-resolved absorption difference spectra from the visible and IR measurements by fitting the data to a sequential, unbranched model to yield evolution associated difference spectra (EADS). Previous studies of similar photoreactions suggest that It is unlikely that any significant reaction branching will occur and that a simple sequential reaction scheme will fit the data adequately.[15,16] Major features in the IR spectral region monitored here include the stretches of the C=O bonds on rings A and D at ~1730 and ~1720 cm^{-1}, respectively and C=C stretches at ~1595, 1620, 1627, and 1648 cm^{-1}, which represent the bonds between rings B and C, A and B, C and D, and the vinyl group attached to ring D, respectively.[17] Features from the protein amide I group and amino acid side chains occur at ~1650 and ~1450 cm^{-1}.[18] Contour maps of the raw data along with the corresponding EADSs are shown in Figures 2 and 3 (below); more extensive data and analysis figures can be found in the Supporting Information (Figures S1-6).

The photoconversion of the P_O state gives rise to three resolvable steps in the visible region (represented by EADS 1- 4 in Figure 2). There is an initial bleach of the ground state absorption feature together with the appearance of a new

absorption peak at 670 nm and broad excited state absorption features between 400 and 500 nm (Figure 2B,D, EADS1). EADS1 then evolves into EADS2, which has an absorbance maximum at 660 nm and a proportionally similar sized excited state absorption band, with a lifetime of 116 ps (112 ps in D_2O).
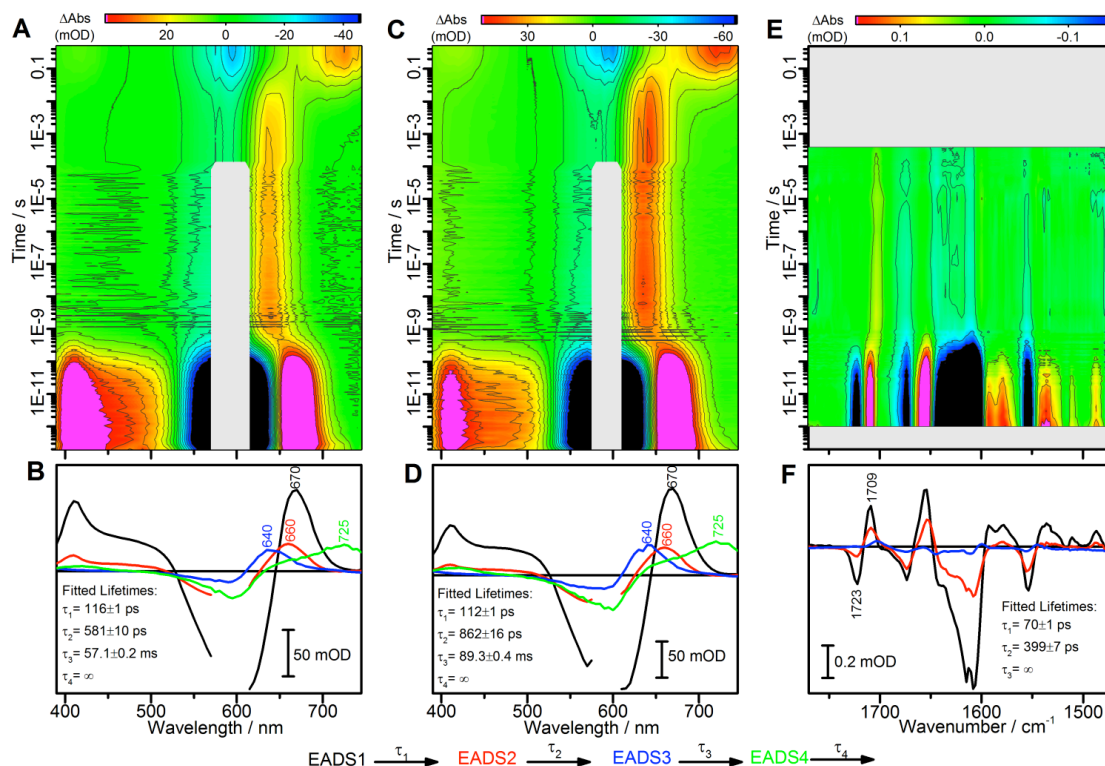

Figure 2. Contour plots of the raw data and evolution associated difference spectra (EADS) resulting from a global analysis of time-resolved visible and IR transient absorption data, collected after excitation of the P_O state. A and B, visible data collected on Dten Phy in H_2O buffer; C and D, visible data collected on Dten Phy in D_2O buffer;. E and F, IR data collected on Dten Phy in D_2O buffer. Visible data were recorded from 0.2 ps until no further spectral changes were observed (495 ms), IR data were recorded over a 1 ps – 390 μs window. The gap in the 580 – 605 nm range is due to scattered pump light obscuring the transient absorption signal.

In the IR region, although there are differences in the absolute intensities, the shape and positions of the spectral features are very similar between EADS1 and EADS2 (Figure 2F). Hence, it seems unlikely that this transition corresponds to the *E/Z* bond isomerization, which would be expected to produce noticeable changes in this spectral region. Instead, this kinetic component is likely to represent excited state relaxation, potentially of the excited state P_O which has not been isomerised assigned to excited state relaxation. The bleach feature at 1723 cm^{-1} and transient feature at 1709 cm^{-1} that is already present in EADS1 suggests a change in the C=O bond on ring D

and hence, it is possible that the *E/Z* isomerization of the bilin cofactor may have already occurred before the first measurement 1 ps after excitation. This assignment is supported by the residuals of the analysis of the visible data (Figures S1 and S3) for which it is possible to observe a very short lived component that cannot be resolved kinetically due to the lack of data points on this timescale, and proximity to the time resolution of the system (~ 0.2 ps). The next reaction intermediate, which has a visible absorption maximum at 640 nm (EADS3), is formed with a lifetime of 581 ps in H_2O (862 ps in D_2O). The IR spectra also display significant changes for this step, in both intensity and position of dominant spectral features There are changes across the full spectral range monitored, with apparently new features observed in regions assigned to C=C stretches of the bilin, as well as amide bands from the protein; these are likely to small-scale structural changes and rearrangements in and around the bilin cofactor. The final reaction step takes place beyond the time-range of the IR measurements, although a comparison of the final time-resolved difference spectrum with a steady-state difference spectrum of the P_O to P_{FR} conversion (Figure 3A) shows that further spectral evolution in the IR region occurs prior to the final P_{FR} formation. Hence, some structural changes and rearrangement of the bilin environment are beyond the 400 μs of these measurements.
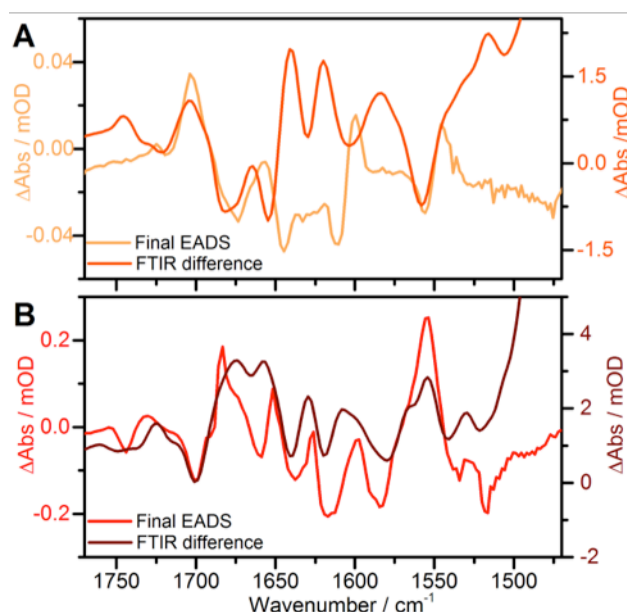

Figure 3. Comparison of the final EADS from global analysis of the time-resolved IR measurements (light coloured line, left axis) with the static infra-red difference spectra (dark coloured line, right axis) for **A**. the P_O to P_{FR} and **B** the P_{FR} to P_O photoconversions.

The absorbance feature with a maximum at 725 nm (EADS4), assigned to the final P_{FR} state, is then formed with a lifetime of 57.1 ms (89.3 ms in D_2O). Although all reaction lifetimes except the first are slower in D_2O compared with those in H_2O, the spectral shapes are virtually identical under both conditions. This implies that proton exchange does not affect the excited state properties or the isomerization but may influence some of the slower steps that are likely to involve changes to the conformation of the bilin cofactor and the protein itself.

In the reverse reaction direction photoexcitation of the P_{FR} state (Figure 4) results in quite different dynamics compared than those of the forward reaction, with five resolvable kinetic steps detected. In the visible region a fast, 0.3 ps lifetime, component corresponds to the relaxation of a highly electronically excited state (e.g. S_2) with intense ground state bleach and excited state absorption (EADS1) to a less electronically excited state (e.g. S_1), the spectra of which retains some ground state bleach, but displays virtually no excited state features (EADS2). In the time region covered by both the visible and in the IR measurements three transitions are observed. EADS2 converts to EADS3 with a lifetime of 5-6 ps. In both the visible and IR regions the changes result in a significant loss of spectral intensity.
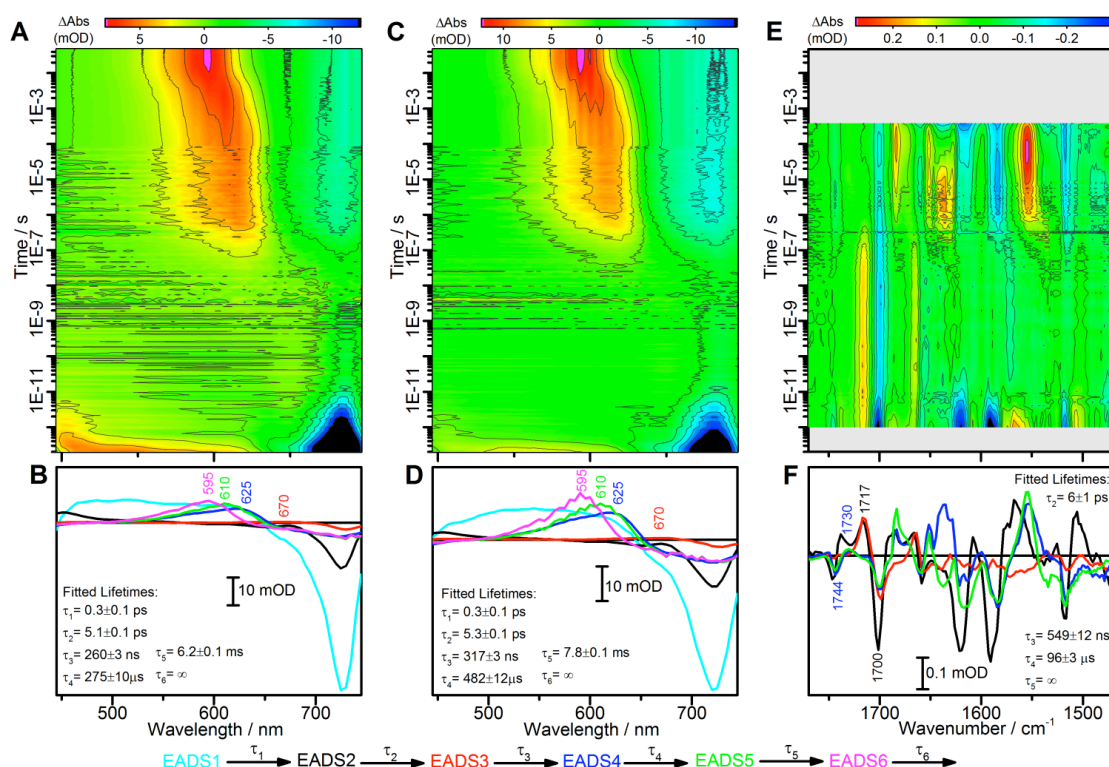
Figure 4. Contour maps of the raw data and evolution associated difference spectra (EADS) resulting from a global analysis of visible and infra-red transient absorption data, collected after excitation of the P_{fr} state. **A** and **B**, time-resolved visible data collected on Dten Phy in H_2O based solution; **C** and **D**, time-resolved visible data collected on Dten Phy in D_2O based solution; **E** and **F**, time-resolved IR data collected on Dten Phy in D_2O based solution. Visible data were recorded from 0.2 ps until no further spectral changes were observed (22.5 ms), IR data were recorded over a 1 ps – 390 μs window

In the visible region a very small positive feature is observed at around 670 nm, and the intensity of the ground state bleach is again reduced. This is likely due to an intermediate with a very slightly blue-shifted absorption feature. In the IR region the transition corresponds to a change from an initially excited state to a spectrum with few spectral features. The feature arising from the C=O of ring D again provides a useful marker of the reaction progress and the strong bleach feature at ~1700 cm^{-1} persists for the full 400 μs IR timecourse. In EADS3 this bleach feature appears to be coupled to a transient feature at ~1717 cm^{-1}, a shift in the opposite direction to that observed in the P_O to P_{FR} reaction. We therefore assign this 5 ps step to the photoisomerisation reaction. No spectral changes in either the visible or IR regions are then observed for several hundred nanoseconds. The transition from EADS3 to EADS4 occurs with a

lifetime of 280 ns in the visible data in H_2O (317 ns in D_2O in the visible and 549 ns in the IR). In the visible region the intermediate absorption band shifts to 625 nm and in the IR a new bleach/transient feature at 1744/1730 cm^{-1} appears. Based on previous findings [19] and the small kinetic isotope associated with this step it is likely that this feature is due to the deprotonation of one of the propionic side chains. The disappearance of the positive feature at 1717 cm^{-1} is assigned to the C=O of ring D, in keeping with previous studies of similar systems.[20]. The 625 nm absorbing intermediate described by EADS4 is then converted to an intermediate with an absorption maximum at 610 nm (EADS5) with a lifetime of 275 μs (482 μs in D_2O). Due to the kinetic isotope effect associated with this step it is likely to involve reprotonation of the bilin factor. The IR data were only collected up to 400 μs, which is close to the lifetime found in the visible datasets for this step. Consequently, it is unlikely that this step will be kinetically well resolved in the IR region, hence the 96 μs lifetime.

A comparison of the final component from the global analysis of the time-resolved IR data with the steady state difference spectra (Figure 3B) reveals that whilst there are many similarities between the spectra there are still a number of spectral features that evolve after 400 μs. The majority of these changes occur <1650 cm^{-1}, which is in the region where protein absorption is expected to make significant contributions. Hence, the slower steps are only likely to involve large protein conformational changes and not significant changes in the bilin environment. The final step resolved in the visible dataset show the formation of the final 595 nm absorbing P_O state (EADS6) with a lifetime of 6.2 ms (7.8 ms in D_2O).

Investigations of Phy systems with the well-studied red / far-red transitions suggest that in one direction the reaction consists of a picosecond isomerization of the bilin cofactor. This is then followed by single bond rotations leading to the breakage and subsequent reformation of some of the H-bond interactions with the protein environment. The other reaction direction also involves a picosecond isomerization step, but this is then followed by deprotonation of the bilin chromophore, other minor structural changes, and then reprotonation to form the final stable state.[21] It appears that despite the differences in wavelength sensitivity of Dten Phy a similar set of reactions occurs here. It is also known

that the red / far-red Phys are capable of undergoing a much slower dark reversion process, which can occur over minutes or days.[22-24] Therefore, measurements over the course of several hours were undertaken to identify any dark reactions from either of the two states of Dten Phy (Figures S7-8). The P_{FR} state fully converts to the P_O state within a few hours (lifetime of around 2 hours) in the dark. However, the P_O state does not undergo any significant dark reversion back to the P_{FR} state over a 24 hour time period.

In typical red / far-red Phys it is known that the slower step(s), which follow the initial photoisomerization, involve a series of large-scale conformational changes or domain motions in the protein.[10, 11, 25, 26] We have used time-resolved solution X-ray scattering measurements to characterize the nature and kinetics of the structural dynamics during the P_O to P_{FR} transition in Dten Phy. A large difference signal in the small/wide angle region ($q < 0.25$ $Å^{-1}$) was observed between the two states, indicating significant structural changes on the nanometre scale. The difference signal develops in both intensity and shape on the µs to second timescale (Figure 5A).
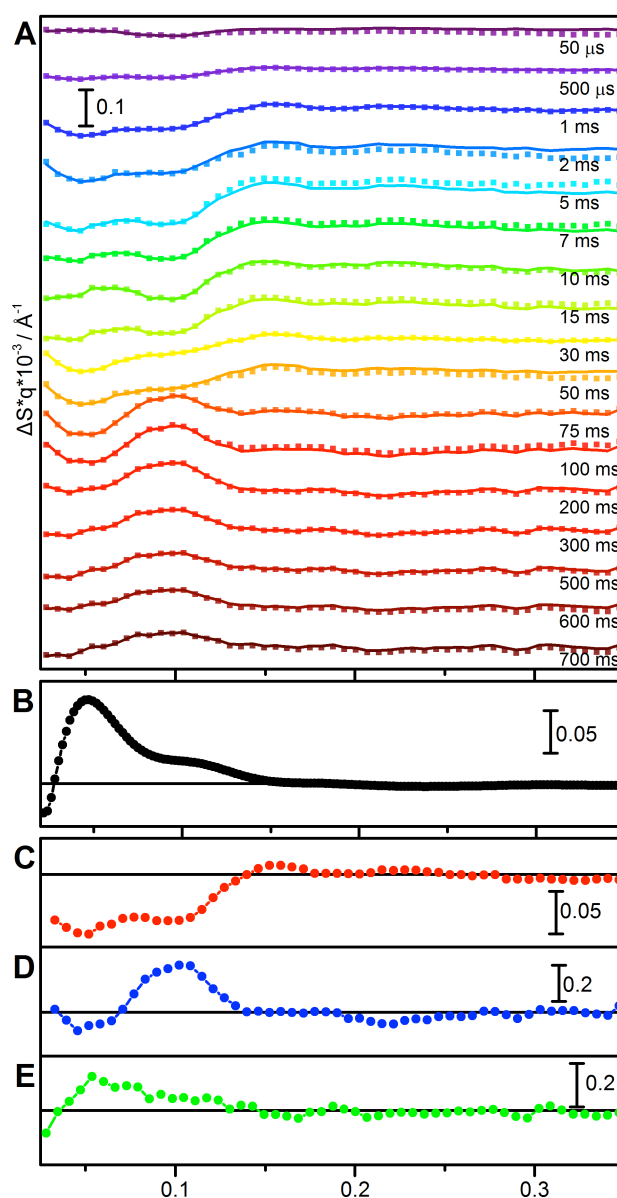
Figure 5. A. Laser-induced time-resolved X-ray scattering difference patterns for Dten Phy in solution, vertically offset for clarity, points are experimental data, lines are linear combinations of the first three basis patterns obtained by singular value decomposition. Note that scattering differences are multiplied by the scattering vector q to reduce differences in the amplitude scale in SAXS and WAXS regions. **B**. Static X-ray scattering difference between P_O and P_{FR} states and **C, D, E.** the three main basis patterns obtained by SVD analysis of the entire time resolved X-ray scattering dataset.

A singular value decomposition analysis revealed that all the structural data can be described by three main time-independent basis patterns (Figure 5C-E), which form a time-independent orthonormal basis on which the full set of time-dependent data can be decomposed. The first, which represents the shape of the signal developing on the μs to ms timescale, shows multiple structural changes across all regions. The second basis pattern describes the signal

developing in the tens of ms and contains an oscillation that is similar to the signature of the quaternary opening motion that has been previously described for the bacterial red / far-red Phy family.[11, 27] The third basis pattern is the signal of the final state, as revealed by the comparison with a static difference between P_{FR} and P_O stationary states (Figure 5B). A kinetic analysis of the whole data set in terms of the three main basis patterns reveals that the structural dynamics of Dten Phy evolve with three time constants of τ_1=8.7 ms, τ_2=60 ms and τ_3>500 ms. The second of these time constants correlates well with the 57.1 ms lifetime from the visible time-resolved data. Previous measurements have proposed that the X-ray scattering signal obtained for bacterial Phys represents an opening of the homodimer in the P_{FR} state, caused by an increase in the separation of the opposing PHY domains.[11, 27] The similar signal obtained for Dten Phy in the second basis pattern suggests that this motion is conserved in algal Phys on the ms timescale. However, additional structural dynamics are observed for Dten Phy on the μs timescale, which do not appear to be present in the bacterial Phys. Moreover, as further structural changes associated with P_O to P_{FR} transition occur on slower timescales than the optical changes it is likely that these final conformational changes in the protein do not significantly affect the environment of the bilin chromophore.

The presence of these large scale structural changes in the protein have also been confirmed by viscosity-dependence studies. Visible time-resolved experiments at a range of solvent viscosities (0 – 60% glycerol) revealed that the rates of the slower μs – ms steps are strongly dependent on the solvent viscosity (Figure 6). The lifetime of the final step observed in the P_O reaction increases 44-fold in 60 % glycerol, and the P_{FR} final lifetime 8-fold. This strongly indicates that these slower steps in the photocycle involve a series of solvent-coupled protein motions. Conversely, no such dependence on solvent viscosity could be observed on the sub-ns dynamics (Figure S9).
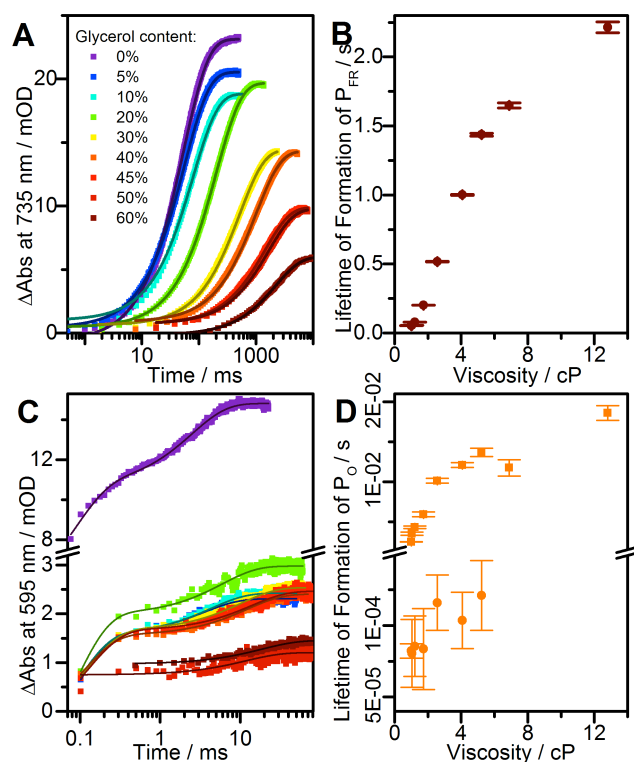
Figure 6. A. and **C**. The formation of P_{FR} and P_O states, following photoexcitation of P_O and P_{FR} states respectively at a range of glycerol concentrations. Data (points) were fitted to a single or double exponential function (lines). **B** and **D** The resulting lifetimes, with fitting errors, are shown as a function of viscosity.

3.3 Conclusions

Taken together, the time-resolved approaches described here have allowed us to investigate the entire photocycle of an atypical algal Phy that is sensitive to orange / far-red light (Fig. 7). It is obvious that the photochemical mechanism of Dten Phy is much more complex than initially observed from time-resolved visible spectroscopy measurements. It has only been possible to unravel these hidden complexities by using additional IR and X-ray scattering probes to study the initial photoisomerisation and slower, structural changes in the protein, respectively. Similar to the red / far-red photocycle in typical Phys it appears as though the orange / far-red photocycle of Dten Phy involves an initial photoisomerisation around the C15-C16 double bond between ring C and D of the bilin on the ps timescale. Following photoisomerisation, the formation of the final P_{FR} or P_O state involves multiple small and large-scale structural changes in the protein, not all of which are linked to spectral changes associated with the bilin cofactor. Time-resolved X-ray scattering measurements have identified 3

kinetically distinct structural transitions, one of which is similar to the opening of the dimer interface that has been observed previously in bacterial Phys.[11, 27] This study illustrates how crucial it is to use multiple biophysical tools in order to obtain a complete mechanistic understanding of these complex light-activated proteins.
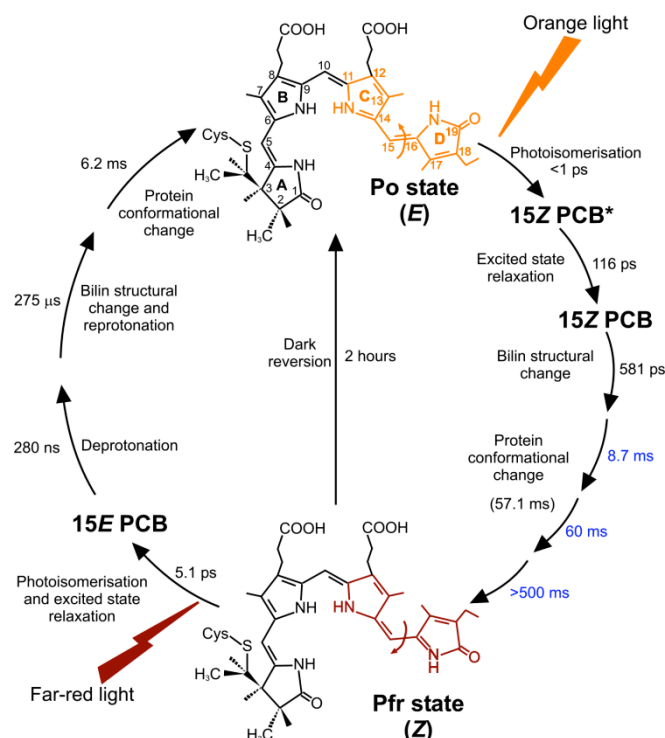
Figure 7. Schematic overview of Dten Phy photocycle, which involves changes to the PCB (phycocyanobilin) chromophore. Lifetimes derived from visible transient absorption (ps – ms) in H_2O are shown in black and the protein structural changes identified from S/WAXS (ms-s) are shown in blue.

3.4 Experimental Section

3.4.1 Sample Preparation and Basic Characterization

The *Dolihomastix tenuilepis* gene encoding the photosensory regions PAS-GAF and PHY was ordered from GeneArt, subcloned into pET21a. The vector was then co-transformed along with pCola Duet vector containing pcyA and HO1 genes into HMS 174 cells. Cultures were grown at 30 °C, 190 rpm until an OD of 0.1 at 600 nm was achieved. The temperature was then dropped to 25 °C. Cultures were induced using 100 μM IPTG at an OD of between 0.4 and 0.6 at 600 nm, and grown overnight. Nickel-IDA resin was used as the first step purification. The resin was equilibrated with 5 column volumes of 'buffer A' (50 mM Tris pH 7.5, 300 mM NaCl, 5 mM imidazole, 7.5 % glycerol, 1 % β-

mercaptoethanol) and incubated with the protein for 3 to 4 hours. The protein flow through was collected and the resin washed with 4 column volumes of 'buffer B', 50 mM Tris pH 7.5 300 mM NaCl, 30 mM imidazole, 7.5 % glycerol, 1 % β-mercaptoethanol. The protein was eluted with 'buffer C', 50 mM Tris pH 7.5, 300 mM NaCl, 250 mM imidazole, 7.5% glycerol, 1 % β-mercaptoethanol. Size exclusion was carried out on ÄKTA Pure (GE Healthcare) using a Hi Load 26/60 200 (GE Healthcare) column. The gel filtration column was equilibrated with two column volumes of gel filtration equilibration 'buffer D', 50 mM Tris pH 7.5, 300 mM NaCl, 7.5 % glycerol, 1% β-mercaptoethanol before the protein was injected. After injection, the protein was eluted in two-column volume gel filtration equilibration buffer. The relevant fractions were analyzed using SDS-PAGE to pool together the fractions containing the purest protein. The protein was then concentrated using a Sartorius Stedim Biotech vivaspin column with a 30 kDa cut-off, flash-frozen and stored at -80 °C. UV/Vis absorption spectroscopy was performed using a Cary 50 UV-Vis Spectrophotometer, Varian, Agilent Technologies. All spectra were recorded from 300 to 800 nm. Thorlabs mounted high power LEDs (of either 530 or 735 nm) were used in the dark to convert the phytochromes into their different isomeric states. For studies of the dark thermal conversion over time, the cuvette was left inside a fully covered spectrometer. Steady state difference spectra in the infra-red region were collected using a Brucker Vertex 80 FTIR, measurements were recorded on samples (~100 µM) in D_2O buffer (50 mM Tris pD 7.5, 300 mM NaCl, 7.5 % glycerol, 1% β-mercaptoethanol) contained between CaF_2 windows separated by a 100 µm Teflon spacer. Difference spectra were recorded for each state (P_O or P_{FR}) using the opposite state (P_{FR} or P_O respectively) as the background measurement. For the viscosity studies 200 µl of 10mg/ml protein was added 800 µl of 5, 10, 20, 30, 40, 45, 50 and 60 per cent glycerol. Data from the photoconversion measurements were used as the 0% glycerol sample. Data were fitted with a single exponential (P_{FR} formation and P_O formation at 50 and 60% glycerol) or double exponential (P_O formation <45% glycerol) function

3.4.2 Nanosecond laser flash photolysis

The Dten Phy photodynamics were studied using a laser system consisting of an optical parameter oscillator (OPO) pumped by a Q-switched Nd: YAG laser (Brilliant B, Quantel). Excitation wavelengths were generated using the OPO

pumped by the third harmonic of the laser. A quartz cuvette with a 2 mm path length for pump-excitation and 10 mm path length for probing of the was used. Samples had an O.D. of 0.8 at 735 nm in the 10 mm path. TA spectra were recorded using an LKS-60 Flash-Photolysis instrument (Applied Photophysics Ltd.). Kinetics were recorded in 5 nm increments across the visible spectral region. Samples illuminated with a 530 nm LED to generate the P_{FR} state were then photoexcited at 430 nm with 15 mJ pulse energy (this excitation wavelength was selected due to the limited pump wavelength range of the laser system) kinetics were recorded over a timescale of 0.05 – 47.5 ms. Samples were illuminated with a 735 nm LED to generate the P_O state were then photoexcited at 595 nm with 9 mJ pulse energy, kinetics were recorded over a timescale of 0.1 – 495 ms.

3.4.3 Ultrafast spectroscopy

Time-resolved visible spectroscopy was performed using a Ti:sapphire amplifier system (Newport Spectra Physics, Solstice Ace) producing 800 nm pulses at 1 kHz with 100 fs pulse duration. A TOPAS Prime optical parameter amplifier (OPA) with associated NirUVis unit was pumped with a portion of the output of the amplifier to generate the pump beams. A broadband ultrafast pump-probe transient absorbance spectrometer 'Helios' (Ultrafast Systems LLC) was used to collect data over a 3 ns time window with a time resolution of around 250 fs. The probe beam consisted of a white light continuum generated in a CaF_2 crystal and absorbance changes were monitored between 390 and 745 nm. A broadband sub-nanosecond pump-probe transient absorption spectrometer 'Eos' (Ultrafast Systems LLC) was used to collect data on longer timescales. A 2 kHz white-light continuum fiber laser was used to generate the probe pulses, the delay between pump and probe was controlled electronically. Samples reservoirs were illuminated with 530 or 735 nm LEDs and flowed through a 0.2 mm path length quartz cuvette. Samples had an O.D. of around 0.5 at 735 nm in the 0.2 mm path Excitation wavelengths of 430 nm and 595 nm, with 0.5 uJ pulse energy, were used over timescales up to 82 µs. The pre-excitation data were subtracted and spectral chirp corrected for before the analysis.

Time-resolve infra-red spectroscopy was performed using the TRMPS set-up of the ULTRA-LIFEtime system[28] at the Central Laser Facility, STFC, Rutherford

Appleton Laboratory, UK. This uses a 100 kHz ultrafast laser based on a custom dual Yb:KGW system (Pharos, Light Conversion). Samples in D_2O buffer (50 mM Tris pD 7.5, 300 mM NaCl, 7.5 % glycerol, 1% β-mercaptoethanol) were contained between two CaF_2 windows, separated by a teflon spacer to give a pathlength of approximately 50 μm. Samples reservoirs were illuminated with 530 or 735 nm LEDs and flowed through the cell, the sample holder was rastered to avoid sample damage. Excitation wavelengths of 575 nm at 1 kHz repetition rate and 0.6 μJ pump energy, and 735 nm at 500 Hz repetition rate and 2 μJ pump energy, were used. Data were collected for approximately 35 and 100 minutes respectively per dataset The excitation beam was set at the magic angle with respect to the IR probe beam. Difference spectra were generated relative to the ground state in the spectral window 1470-1770 cm^{-1} at time delays ranging between 1 ps and 390 μs. Pixel to wavenumber calibration was performed using a polystyrene standard. Infra-red pump-probe transient absorption data from the TRMPS set-up were averaged 0.1 μs on either side of each 10 μs "TRMP" step after the first 10 μs of data. Thus 105 (P_{FR} reaction) or 57 (P_O reaction) time points were spread over the first 10 μs timerange, and thereafter timepoints were every 10 μs up to 390 μs (39 in total)

3.4.4 Global Analysis

Global Analysis was performed using the open-source software Glotaran.[29] This procedure reduces the matrix of change in absorbance as a function of time and wavelength, to a model of one or more exponentially decaying time components, each with a corresponding difference spectrum (EADS). All datasets were fitted to a sequential, unbranched model. Errors quoted with the lifetime values are the standard errors calculated during the global analysis. Visible spectrum data sets from the nanosecond laser flash photolysis and ultrafast transient absorption techniques were combined into one dataset. The wavelength points from the ultrafast data were removed so that the data matched the those of the laser flash photolysis data (5 nm steps). The datasets were then scaled to match intensities of the major spectral features in overlapping time regions.

3.4.5 Time Resolved Small and Wide Angle X-ray Scattering

TR-S/WAXS experiments have been performed at the beamline ID09 of the European Synchrotron Radiation facility. A sample of Dten Phy (250 µM) was continuously illuminated by a 730 nm LED to stabilise the P_O form, flowed by a peristaltic pump (Gilson Minipulse 3) into a 2 mm quartz capillary and excited by a 5 ns laser pulse (EKSPLA) at 595 nm (1 mJ/mm^2). The X-ray scattering signal was monitored at different time delays after laser excitation by 20 µs X-ray pulse train. Scattering patterns were normalized the water peak (about 2.2 Å^{-1}) before making differences between laser-on and laser-off patterns. The time evolution of the solvent heating peak at 1.8 Å^{-1} was checked, and did not affect the scattering patterns in the q range where protein difference signal used for analysis was present (0-0.3 Å^{-1}). Singular value decomposition analysis of time resolved scattering dataset was performed using an in-house developed Python based code. Visual inspection and autocorrelation analysis of both basis patterns and amplitude vectors[30] indicated the first three basis patterns containing significant information, with the rest containing essentially only random noise. A linear combination of the first three basis patterns was then used to fit the time evolution of the scattering signal and to extract time constants.

3.5 Supplementary Information

3.5.1 Transient Absorption Data and Analysis

Figure S1. For the P_O to P_{FR} conversion, in the visible region, in H_2O based buffer: **A.** Contour plot of the raw data; **B. and C.** Difference spectra at selected time points; **D.** Kinetic traces (points) showing fits from global analysis (lines) at selected wavelengths with residuals from the fit shown below; **E.** Evolution Associated Difference Spectra (EADS) normalised to the intensity of the ground state bleach (ca. 595 nm); **F. and G.** Singular value decomposition of the residual matrix resulting from global analysis.

The gap in the <82 µs delay data in the 575 – 610 nm range is due to scattered pump light obscuring the transient absorption signal. Datasets covering 0.2 ps – 2.9 ns , 1.1

ns – 82 µs, and 0.138 – 495 ms were combined. The majority of the residual matrix lacks any significant structure, implying a good fit has been obtained. To obtain a good fit in the global analysis it was necessary to use 5 components, although the spectra labeled nEADSX and nEADS3 are in fact virtually identical, the lifetime of conversion between them corresponds to the step in the data between ultrafast measurements, contaminated by scattered pump light, and the laser flash photolysis measurements. As a result the spectra nEADSX, and the corresponding lifetime have been omitted from discussion of the photoconversion.
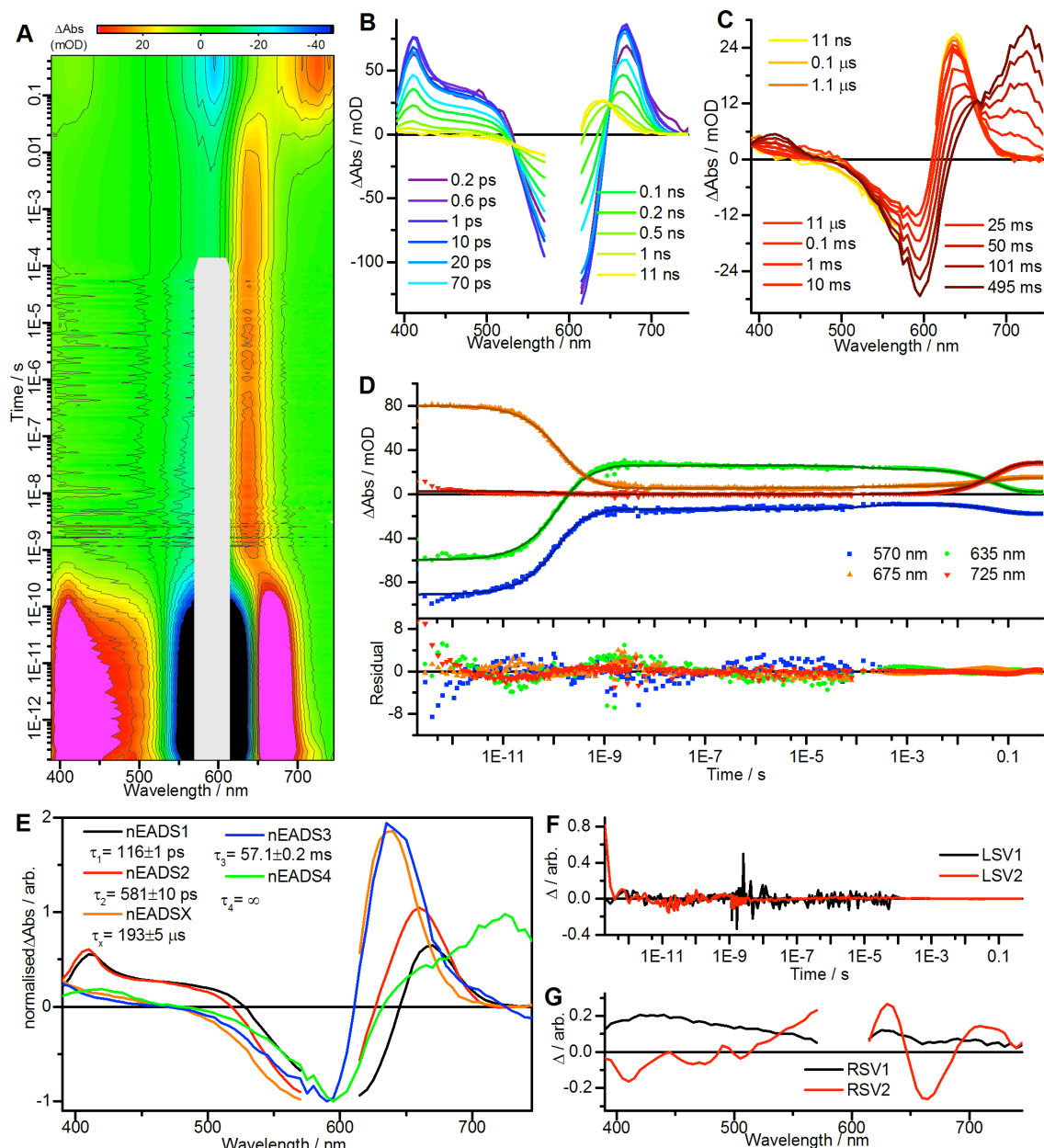
Figure S2. For the P_{FR} to P_O conversion, in the visible region, in H_2O based buffer: **A.** Contour plot of the raw data; **B. and C.** Difference spectra at selected time points; **D.** Kinetic traces (points) showing fits from global analysis (lines) at selected wavelengths with residuals from the fit shown below; **E.** Evolution Associated Difference Spectra (EADS) normalised to the intensity of the ground state bleach (ca. 725 nm); **F. and G.** Singular value decomposition of the residual matrix resulting from global analysis. Datasets covering 0.2 ps – 2.9 ns , 0.5ns – 82 µs, and 0.052 – 47.5 ms were combined. The residual matrix lacks any significant structure, implying a good fit has been obtained.

Figure S3. For the P_O to P_{FR} conversion, in the visible region, in D_2O based buffer: **A.** Contour plot of the raw data; **B. and C.** Difference spectra at selected time points; **D.** Kinetic traces (points) showing fits from global analysis (lines) at selected wavelengths with residuals from the fit shown below; **E.** Evolution Associated Difference Spectra (EADS) normalised to the intensity of the ground state bleach (ca. 595 nm); **F. and G.** Singular value decomposition of the residual matrix resulting from global analysis.

The gap in the <82 µs delay data in the 580 – 605 nm range is due to scattered pump light obscuring the transient absorption signal. Datasets covering 0.2 ps – 2.9 ns , 0.5 ns – 82 µs, and 0.138 – 495 ms were combined. The majority of the residual matrix lacks any significant structure, implying a good fit has been obtained. To obtain a good fit in the global analysis it was necessary to use 5 components, although the spectra labeled nEADSX and nEADS3 are in fact virtually identical, the lifetime of conversion

between them corresponds to the step in the data between ultrafast measurements, contaminated by scattered pump light, and the laser flash photolysis measurements. As a result the spectra nEADSX, and the corresponding lifetime have been omitted from discussion of the photoconversion.
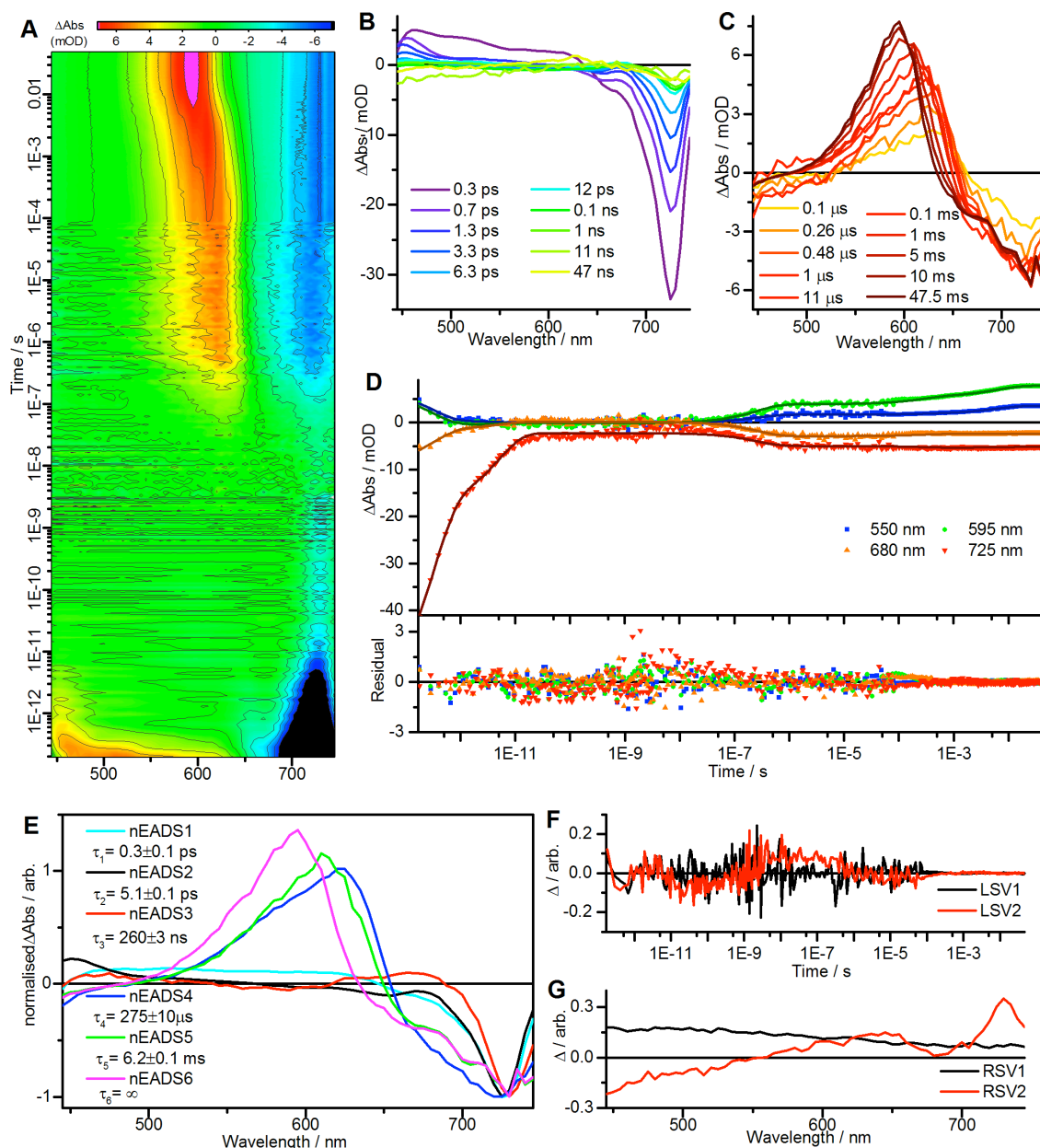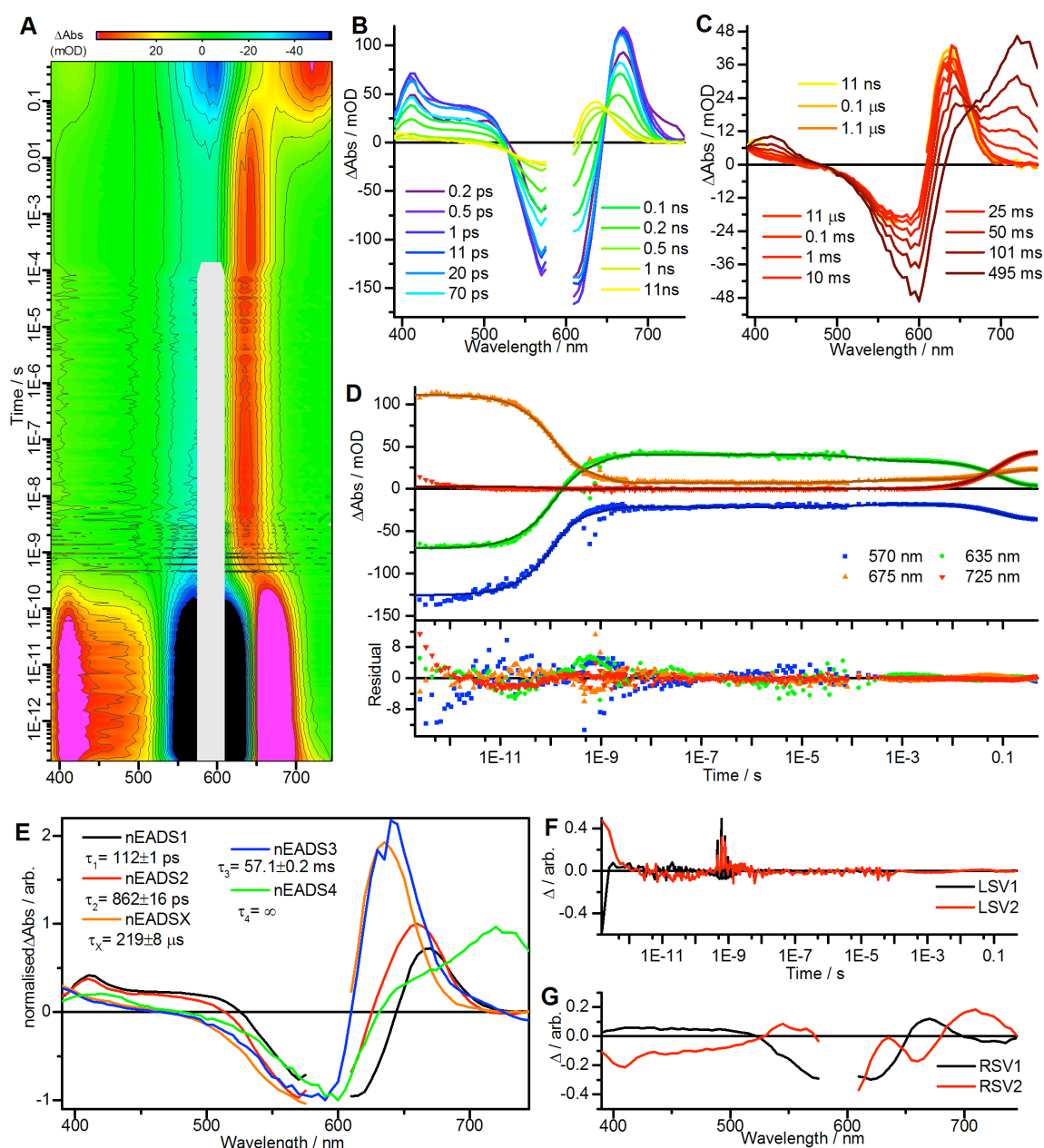
Figure S4. For the P_{FR} to P_O conversion, in the visible region, in D_2O based buffer: **A.** Contour plot of the raw data; **B. and C.** Difference spectra at selected time points; **D.** Kinetic traces (points) showing fits from global analysis (lines) at selected wavelengths with residuals from the fit shown below; **E.** Evolution Associated Difference Spectra (EADS) normalised to the intensity of the ground state bleach (ca. 725 nm); **F. and G.** Singular value decomposition of the residual matrix resulting from global analysis.

Datasets covering 0.2 ps – 2.9 ns , 0.5ns – 82 µs, and 0.052 – 47.5 ms were combined. The residual matrix lacks any significant structure, implying a good fit has been obtained.

Figure S5. For the P_O to P_{FR} conversion, in the IR region, in D_2O based buffer: **A.** Contour plot of the raw data; **B. and C.** Difference spectra at selected time points; **D.** Kinetic traces (points) showing fits from global analysis (lines) at selected wavenumbers with residuals from the fit shown below; **E.** Evolution Associated Difference Spectra (EADS) normalised to the intensity of the maximum intensity bleach; **F. and G.** Singular value decomposition of the residual matrix resulting from global analysis.
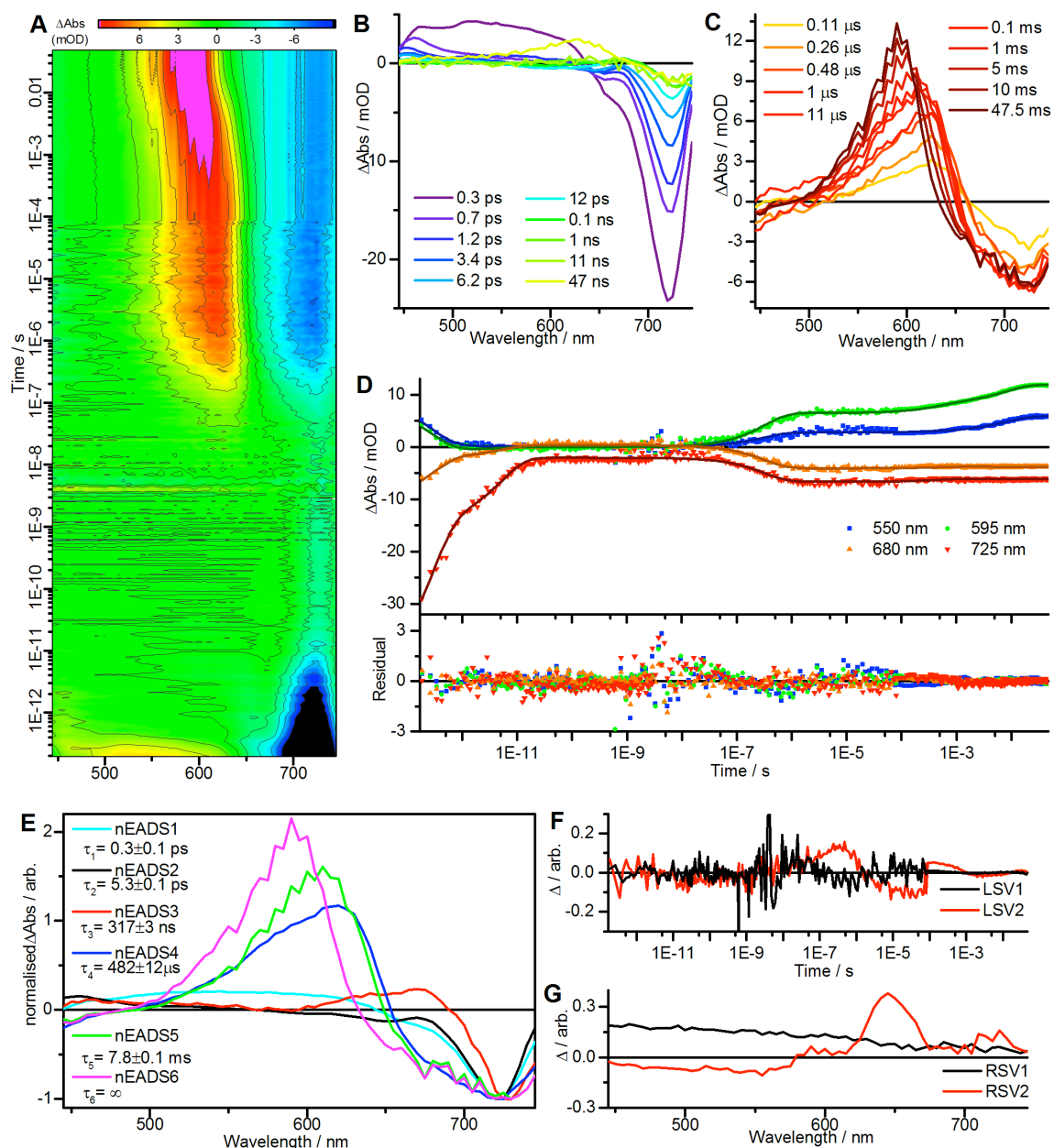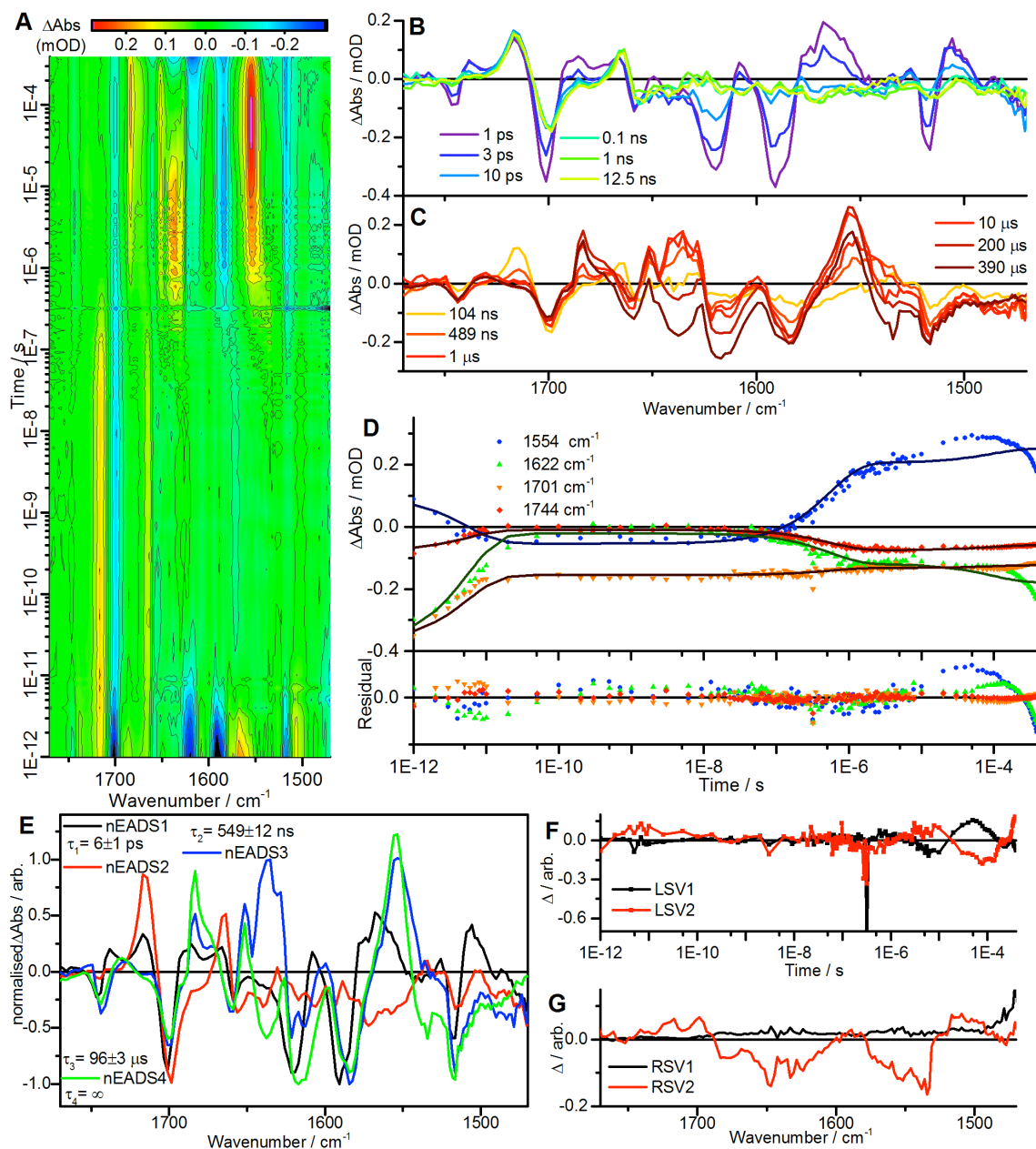
Figure S6. For the P_{FR} to P_O conversion, in the IR region, in D_2O based buffer: **A.** Contour plot of the raw data; **B. and C.** Difference spectra at selected time points; **D.** Kinetic traces (points) showing fits from global analysis (lines) at selected wavenumbers with residuals from the fit shown below; **E.** Evolution Associated Difference Spectra (EADS) normalised to the intensity of the maximum intensity bleach; **F. and G.** Singular value decomposition of the residual matrix resulting from global analysis

3.5.2 Dark Reactions

Dten Phy was fully converted into either the P_O or P_{FR} state, changes in absorbance across the visible spectrum were then recorded every 10 or 15 minutes over a 24 hour period, or until no further changes were observed. (Figure S7)
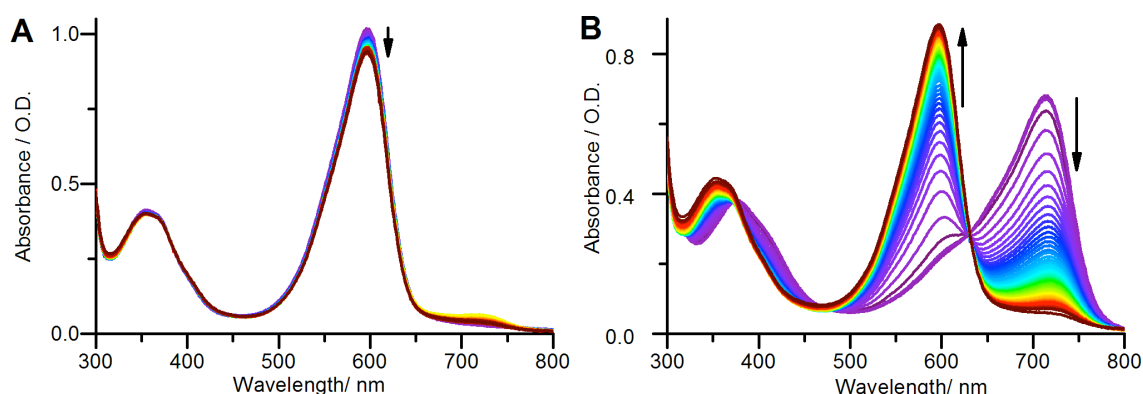

Figure S7. Spectral evolution of the P_O state (A) and P_{FR} state (B) in dark conditions. Changes at 598 and 714 nm (the absorption maxima of P_O and P_{FR} respectively) are shown in figure S8. There was no significant change in the P_O state over the timescale measured. In contrast the P_{FR} state reverted to the P_O state with a lifetime of around 2 hours.
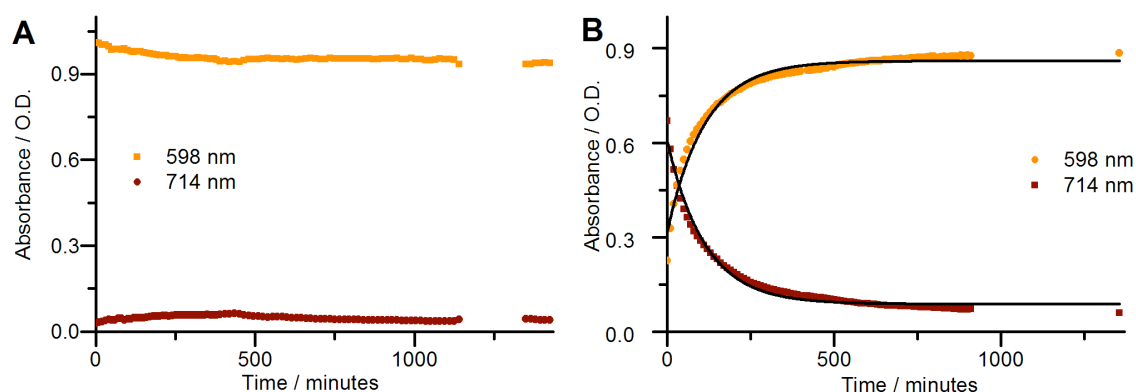
Figure S8. Kinetics of change in absorbance at 598 and 714 nm, starting from **A.** the P_O state and **B.** P_{FR} state in dark conditions. A single exponential was fitted to the data shown in panel B using a shared lifetime value for both kinetic traces, which yielded a lifetime of 113 ±2 minutes.

3.5.3 Solvent Viscosity Studies:

The effect of solvent viscosity on the faster timescale processes was also investigated, sample preparation and experimental procedures were as described in the main text. Data from the photoconversion measurements were used as the 0% glycerol sample. Glycerol solutions scatter light more efficiently than aqueous solutions, so there is a large region in the area of the pump wavelength where the change in absorbance is obscured. Kinetics were monitored by integration over a range of wavelengths covering the positive transient signal from the first reaction intermediate. Data were then fitted with a single exponential in both datasets.

There is no noticeable viscosity dependence in the loss of the first intermediate after photoexcitation of the P_O state. There is also no noticeable trend in the lifetime of formation of the first transient species after excitation of the P_{FR} state, however above 30% glycerol (2.6 cP) formation was almost entirely inhibited over the 2 µs timescale monitored.
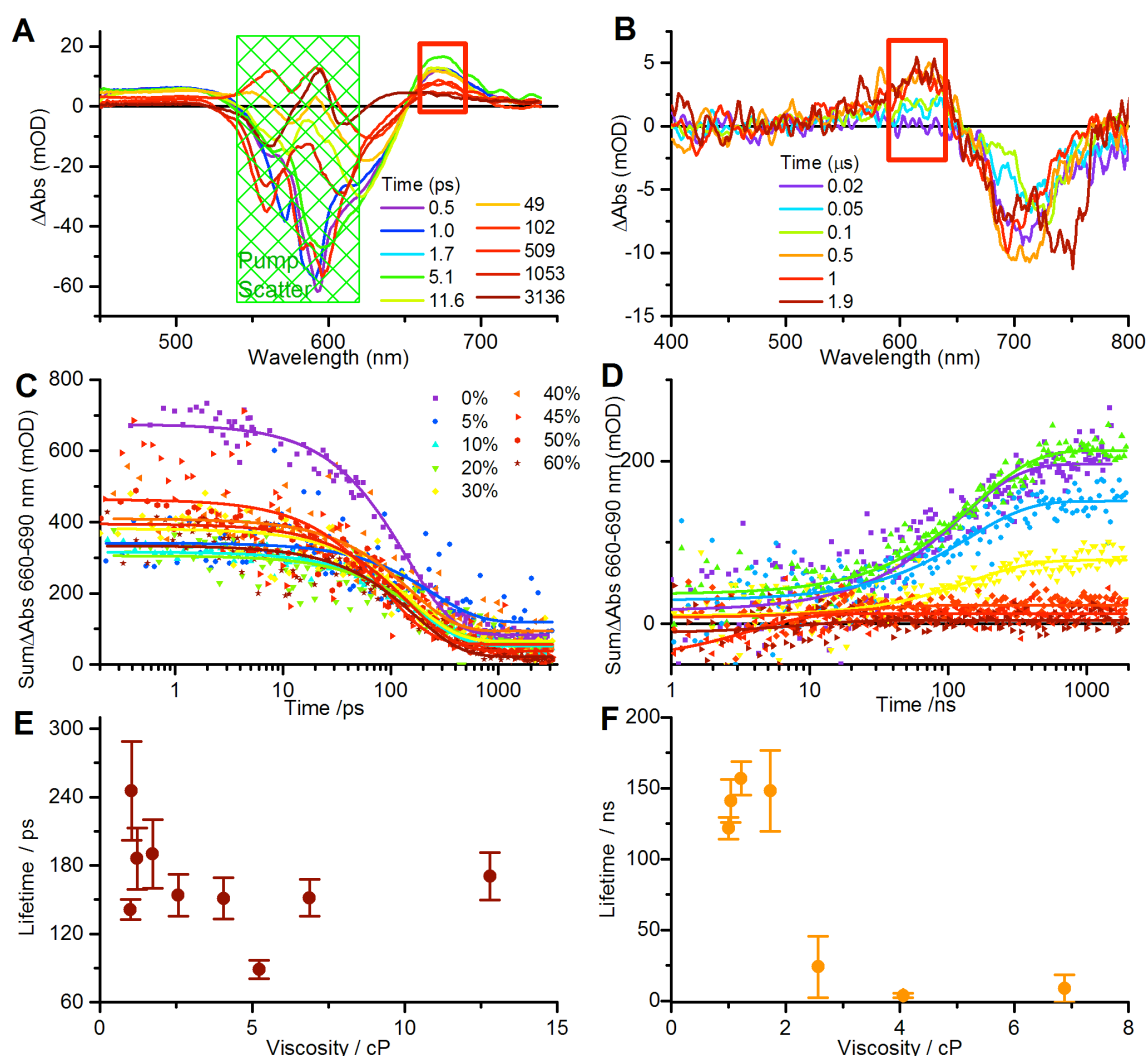
Figure S9. Difference spectra at selected timepoints after photoexcitation of the **A.** P_O and **B.** P_{FR} states at 5% glycerol concentration. Glycerol solutions scatter light more efficiently than aqueous solutions, so there is a large region in the area of the pump wavelength where the change in absorbance is obscured. Kinetic traces at a range of glycerol concentrations after photoexcitation of the **C.** P_O and **D.** P_{FR} states. Data were fitted to a single exponential function and plotted as a function of solvent viscosity (**E** and **F**).

3.6 Author Contributions

U.C., D.J.H., S.J.O.H., and G. S. wrote the paper with input from all authors; U.C. and M.S. prepared and samples; U.C. performed basic characterization; U.C. and D.J.H performed laser flash photolysis measurements; S.J.O.H performed visible pump probe spectroscopy; D.J.H, S.J.O.H., U.C. and I.S. performed infra-red time resolved measurements; G.S., D.J.H, S.J.O.H., J.W., E.D.L.M., M.P., M.W. and M.W., performed WAXS measurements; S.J.O.H. analyzed the visible and infra-red spectroscopy data; G.S analyzed the WAXS data.; N.S.S. and D.J.H conceived of and designed the work

3.7 Acknowledgements

3.8 References

[1]   N. C. Rockwell, Y.-S. Su, Lagarias, *Annu. Rev. Plant Biol.* **2006**, *57*, 837–858.

[2]   N. C. Rockwell, Lagarias, *Chem. Eur. J. of Chem. Phys.* **2010**, *11*, 1172–1180.

[3]   F. Andel, Lagarias, R. A. Mathies, *Biochemistry* **1996**, *35*, 15997–16008.

[4]   H. Foerstendorf, T. Lamparter, J. Hughes, W. Gärtner, F. Siebert, *Photochem. Photobiol.* **2000**, *71*, 655–661.

[5]   A. Remberg, I. Lindner, T. Lamparter, J. Hughes, C. Kneip, P. Hildebrandt, S. E. Braslavsky, W. Gärtner, K. Schaffner, *Biochemistry* **1997**, *36*, 13389–13395.

[6]   C. Kneip, P. Hildebrandt, W. Schlamann, S. E. Braslavsky, F. Mark, K. Schaffner, *Biochemistry* **1999**, *38*, 15185–15192.

 [7]   W. Rüdiger, F. Thümmler, E. Cmiel, S. Schneider, *Proc. Natl. Acad. Sci. U.S.A.* **1983**, *80*, 6244–6248.

[8]   C. Schumann, R. Gross, M. M. N. Wolf, R. Diller, N. Michael, T. Lamparter, *Biophys. J.* **2008**, *94*, 3189–3197.

[9]   P. Schmidt, T. Gertsch, A. Remberg, W. Gärtner, S. E. Braslavsky, K. Schaffner, *Photochem. Photobiol.* **1998**, *68*, 754–761.

[10]   E. S. Burgie, T. Wang, A. N. Bussell, J. M. Walker, H. Li, R. D. Vierstra, *J. Biol. Chem.* **2014**, *289*, 24573–24587.

[11]   H. Takala, A. Björling, O. Berntsson, H. Lehtivuori, S. Niebling, M. Hoernke, I. Kosheleva, R. Henning, A. Menzel, J. A. Ihalainen, et al., *Nature* **2014**, *509*, 245–248.

[12]   J. J. Tabor, A. Levskaya, C. A. Voigt, *J. of Mol. Biol.* **2011**, *405*, 315–324.

[13]   S. R. Schmidl, R. U. Sheth, A. Wu, J. J. Tabor, *ACS Synth Biol* **2014**, *3*, 820–831.

[14]   N. C. Rockwell, D. Duanmu, S. S. Martin, C. Bachy, D. C. Price, D. Bhattacharya, A. Z. Worden, Lagarias, *Proc. Natl. Acad. Sci. U.S.A.* **2014**, *111*, 3871–3876.

[15]   A. F. E. Hauck, S. J. O. Hardman, R. J. Kutta, G. M. Greetham, D. J. Heyes, N. S. Scrutton, J. Biol. Chem. **2014**, *289*, 17747-17757.

[16]   S. J. O. Hardman, A. F. E. Hauck, I. P. Clark, D. J. Heyes, N. S. Scrutton, Biophys. J. **2014**, *107*, 2195-2203.

[17]   A. Takiden, F. Velazquez-Escobar, J. Dragelj, A. L. Woelke, E.-W. Knapp, P. Piwowarski, F. Bart, P. Hildebrandt, M. A. Mroginski, *Photochem. Photobiol.* **2017**, *93*, 713-723.

[18]   F. Velazquez Escobar, T. Utesch, R. Narikawa, M. Ikeuchi, M. A. Mroginski, W. Gärtner, P. Hildebrandt, *Biochemistry* **2013**, *52*, 4871-4880.

 [19]   F. Velazquez Escobar, P. Piwowarski, J. Salewski, N. Michael, M. Fernandez Lopez, A. Rupp, B. M. Qureshi, P. Scheerer, F. Bartl, N. Frankenberg-Dinkel, F. Siebert, M. Andrea Mroginski, P. Hildebrandt, *Nat Chem* **2015**, *7*, 423-430.

[20]   T. Stensitzki, Y. Yang, A. L. Wölke, E.-W. Knapp, J. Hughes, M. A. Mroginski, K. Heyne, *Photochem. Photobiol.* **2017**, 93, 703-712.

[21]   C. Song, T. Rohmer, M. Tiersch, J. Zaanen, J. Hughes, J. Matysik, *Photochem. Photobiol.* **2013**, *89*, 259-273.

[22]   E. Giraud, J. Fardoux, N. Fourrier, L. Hannibal, B. Genty, P. Bouyer, B. Dreyfus, A. Verméglio, *Nature* **2002**, *417*, 202–205.

[23]   B. Karniol, R. D. Vierstra, *Proceedings of the National Academy of Sciences* **2003**, *100*, 2807–2812.

[24]   R. Tasler, T. Moises, N. Frankenberg-Dinkel, *FEBS J.* **2005**, *272*, 1927–1936.

[25]   A. T. Ulijasz, G. Cornilescu, C. C. Cornilescu, J. Zhang, M. Rivera, J. L. Markley, R. D. Vierstra, *Nature* **2010**, *463*, 250–254.

[26]   D. J. Heyes, B. Khara, M. Sakuma, S. J. O. Hardman, R. O'Cualain, S. E. J. Rigby, N. S. Scrutton, *PLoS ONE* **2012**, *7*, e52418.

[27]   A. Bjorling, O. Berntsson, H. Lehtivuori, H. Takala, A. J. Hughes, M. Panman, M. Hoernke, S. Niebling, L. Henry, R. Henning, et al., *Science Advances* **2016**, *2*, e1600920–e1600920.

[28]   G. M. Greetham, P. M. Donaldson, C. Nation, I. V. Sazanovich, I. P. Clark, D. J. Shaw, A. W. Parker, M. Towrie, Appl. Spectrosc. 2016, 70, 645-653.

[29]   J. J. Snellenburg, S. P. Laptenok, R. Seger, K. M. Mullen, I. H. M. van Stokkum, J. Stat. Softw. 2012, 49, 1-22.

[30]   E.R.   Henry,   J.   Hofrichter,   Meth.   Enzymol.   1992,   210,   129-19

Chapter 4. Design of new chimaeric phytochromes for optogenetic control of gene expression.

Abstract

Phytochromes have been the subject of numerous detailed photochemical and structural studies. Although phytochromes have proven to be an attractive target for new applications in the field of optogenetics, their use has been limited due to the availability of only a limited range of wavelength sensitivities. Several atypical phytochromes that exhibit broad spectral properties have been discovered recently. Here, we align a selection of these novel atypical phytochromes with two-component system (TCS) histidine kinases (HKs) to design functional chimaeric phytochromes as part of a multi-wavelength optogenetic toolkit. The photosensory core modules (PCM) of *Deinococcus tenupilis* (Dten) and *Nostoc punctiforme* (Npf) phytochromes were fused to the HK from the cyanobacterial phytochrome Cph1 and to the HK from the *Rhizobium meliloti* TCS, FixL. Time resolved spectroscopy was used to compare the photochemistry of the Npf-FixLHK chimaera with that of the wild type protein, and the *in-vitro* activity of the chimaeras were measured using a phosphorylation assay. FixL HK and the target promoter, nifA (nifAP), of the cognate response regulator FixJ were used in reporter assays to test the *in-vivo* activity of the FixL HK chimaeras. This work illustrates the potential application of producing chimaeric light-responsive phytochromes for optogenetic control of gene expression.

4.1 Introduction

Biological processes such as metabolism and cellular differentiation are controlled by precise spatial and temporal regulation of gene expression patterns.[1-3] Therefore, quantitative methods of controlling gene expression are necessary for studying and engineering biological systems. In many biological and biotechnological applications, which require external control of gene expression, chemical molecules are used. However, this approach has several limitations. Firstly, good spatial resolution cannot be achieved with chemical inducers because the rate of change in gene expression is limited by the

diffusion rate of the inducer. Secondly, once the chemical inducer is added, there is no temporal control over its action. Thirdly, chemical inducers can be modified or degraded by the cells. It is difficult to maintain a constant chemical signal in changing culture environment (*e.g.* change in cell density). Finally, commonly used inducers can be toxic and non-specific, altering host pathways.[4]

Optogenetics is the use of light to control gene expression, which in turn controls a physiological function.[5-9] Compared with chemical approaches, optogenetics allows unprecedented spatial and temporal control over gene expression and biological processes. This is achievable because light can be delivered with precision in space and time, and at specific wavelengths and intensity. This approach is unlikely to perturb other host pathways, as non-photosynthetic model organisms show minimal physiological responses to light.[4] Several light switchable proteins have been developed into optogenetic tools for synthetic regulation of gene expression in *Escherichia coli*, cyanobacteria, yeast, zebra fish, mammalian cells, and liver, kidneys and brain cells of live mice.[10-16] Opsins have been a well-exploited tool in the field of neuroscience, where it has been used to study neural networks and brain function.[9,17,18] Specific neurons can be activated or inhibited by using light with spatial and temporal precision to control ion flux and action potentials. These investigations provided detailed insights into the function of neural networks and allowed the study of brain function and neural diseases.[17,19,20] Optogenetics has also been applied to study cell signalling through the control of genes responsible for expressing enzymes or proteins vital in signalling pathways of specialised cells.[4,17,18,21]

Although several multichromatic gene regulatory systems have been recently developed, the wavelength range of these optogenetic tools is limited, only absorbing either in the red or blue region of the visible spectrum. Phytochromes and cyanobacteriochromes are tetrapyrrole photoreceptors that switch between two different signalling states in response to different wavelengths of light. The diverse spectral properties of cyanobacteriochromes, which cover the entire visible region, have led to their exploitation for use in optogenetics applications.[22,23]

Phytochromes are multi-domain proteins that primarily respond to red and far-red light. They can be split into the photosensory region, which is responsible for sensing the light signal and transducing it to a secondary effector region. The photosensory region is made up of a number of PAS, GAF and PHY domains, and the effector region is made up of the HK and ATPase domains.[24-28] The light signal is captured by the bilin chromophore within the PCM of phytochromes, which leads to an isomerisation and changes in the secondary structure, from α-helix to β-sheet, of a tongue region, which links the PCM to the HK domain. This shortens the tongue, which triggers ATPase activity and facilitates the autophosphorylation of the HK domain. This is followed by phosphotransfer onto a response regulator (RR), which is often involved in modulation of gene expression.[29-34] The modular architecture of phytochromes make them ideal targets for fusion with other domains in a 'mix and match' manner to couple a range of light sensitivities with a range of output responses.

Phytochromes have been developed for control over living cells through promoter regulation in yeast cells and protein dimerisation regulation in mammalian cells, via interactions with phytochrome interaction factors (PIFs).[35,36] The red/far-red phytochrome Cph1 from cyanobacterial *Synechocystis sp.* has been extensively used for light-activated gene expression, although as the target promoters of the Cph1 partner protein, Rcp1, are unknown,[10] function was coupled with the osmosensing HK EnvZ from *E. coli*. This system was further used for multichromatic regulation of gene expression with the green/red CcaS cyanobacteriochrome.[12] The UV sensitive UirS cyanobacteriochrome has also been developed for optogenetic purposes.[37]

Here, we select atypical phytochromes from the eukaryotic algae, *Deinococcus tenupilis* (Dten) and the cyanobacterium, *Nostoc punctiforme* (Npf) to develop optogenetic tools that can be coupled to two-component system (TCS) HKs. As these phytochromes have unusual photocycles, which are different to the typical red/far-red phytochromes, this will allow us to diversify the existing optogenetic toolkit.[23] As the response regulators and target promoters of these phytochromes are as yet unknown, we coupled the phytochromes with

homologous HKs of TCSs with known RRs and promoters.[38,39] Restriction digest, ligation and Infusion cloning methods were used to assemble these systems. *In-vitro* studies to measure the autophosphorylation and phosphotransfer steps using ^{32}P-radiolabelling assays confirmed that the Npf Cph HK and Npf FixL HK chimaeras were active. Npf FixL HK, its cognate response regulator FixJ and promoter nifAP were refactored into the pSR33.4 and pSR59.8 system designed by Schmidl *et al*. 2014 for *in-vivo* studies to follow the activity of this chimaera as an optogenetic tool through modulation of a GFP reporter.[40] We also measured the kinetics of formation of the two states of Npf FixL HK to compare with the wild type Npf phytochromes in order to confirm that the chimaera photochemistry was the same as the wild type PCM.

4.2 Results and Discussion

4.2.1 Design of Chimaeric Phytochromes

Dten and Npf phytochromes were selected for the design of 'plug and play' optogenetic gene expression regulators as these proteins have previously been expressed, purified, and characterised. As Cph1 HK and its cognate RR, Rcp1, have been extensively studied and is also from a phytochrome, this was deemed to be a good effector domain to be used as a proof-of-principle for chimaera construction (Figure 4.1A). However, as target promoters are not known for Rcp1, FixL HK and PhoR, which are most homologous to Cph1 HK in sequence, was also selected for the purpose of chimaera construction (Figure 4.1A). It is also a well-studied TCS with known RR and target promoter and has previously been used in producing other chimaeras.[41,42] Studies have previously revealed that the length and sequence of the helix linker joining the PCM to the HK is essential for signal translation from the PCM to the HK.[10] Upon alignment with the EnvZ HK sequence, a library of Cph1 constructs was produced by varying the number of Cph1 amino acids included by retaining or removing the last 9 amino acids and this revealed that the most active chimaera consisted of the first 517 amino acids of Cph1.[10] Based on this information, sequences of the PCMs were aligned with selected TCS HKs to identify crossover points, where the two regions could be fused together (Figure 4.1B).

A
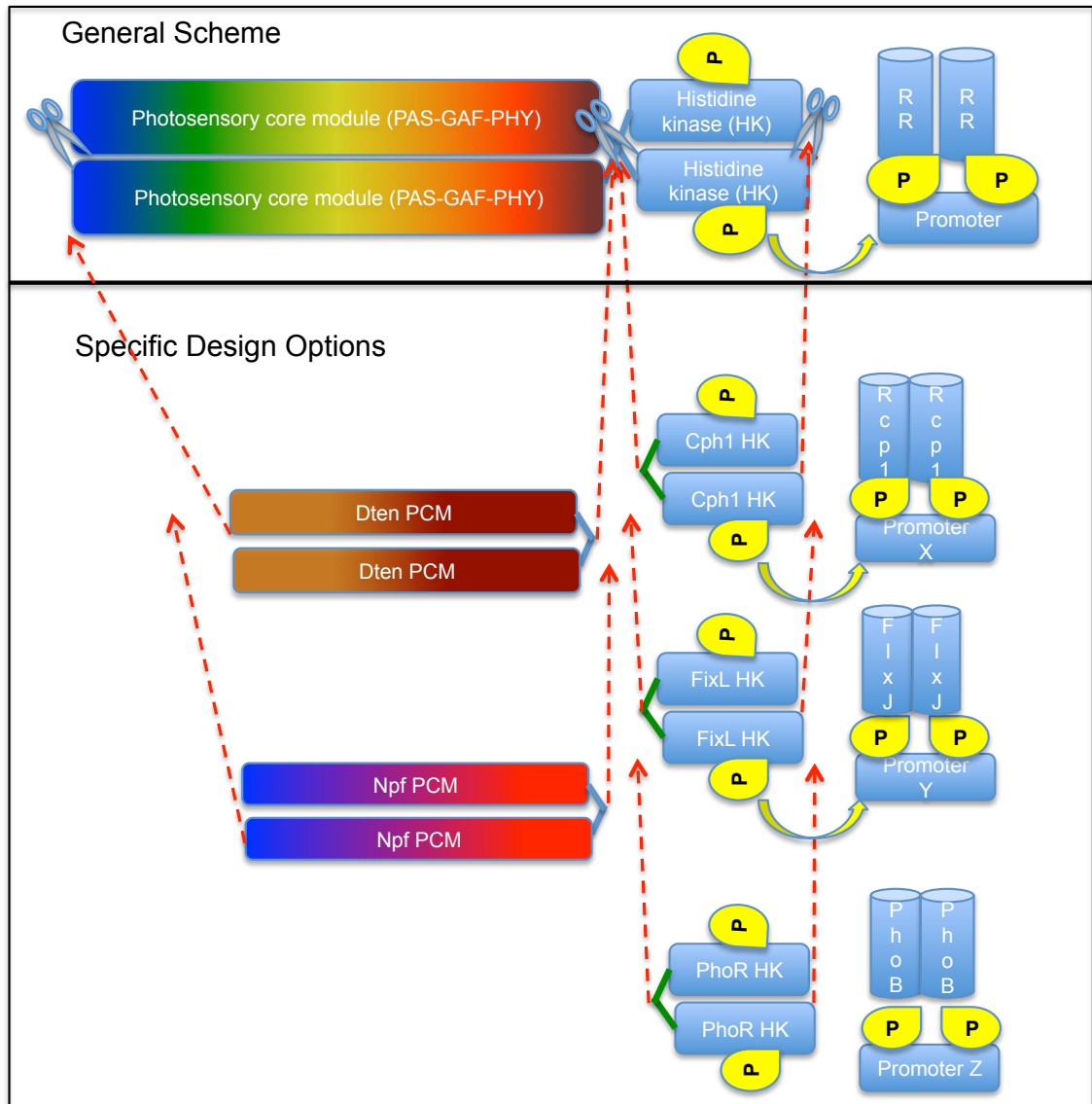
B

Cph1 linker and HK:	.......**ELAQLARN**LERSNADLKKFAYIASHDLQEPLNQVSNY........
FixL linker and HK:	.......**ELARLARLN**EMGEMASTLAHELNQPLSAIANYSHGC........
Dten PCM – Cph1 HK:	.... QQLGILLQETTESNEASS - **ELAQLARN**LERSNADLK........
NPF PCM – Cph1 HK:	.... ESAIALSNAIVGIVLSKAD**ELA - QLARN**LERSNADLK........
Dten PCM – FixL HK :	.......QQLGILLQETTESNEASS - **ELARLARLN**EMGEMAST........
Npf PCM – FixL HK:	.......ESAIALSNAIVGIVLSKAD**ELA - RLARLN**EMGEMAST........

Figure 4.1. **Design of chimaeric phytochromes. A.** Atypical phytochrome PCMs of Dten and Npf phytochromes were fused with TCS HK of Cph1, FixL and PhoR at a crossover point in the linker region, which is conserved in both length and sequence. Blue lines and green lines represent the linking regions of the phytochromes and HKs, respectively. Red arrows indicate the HK unit ligation. The blue cylinders represent response regulators, autophosphorylation and phosphotransfer is illustrated with yellow circles. **B.** Alignment of Cph1 HK kinase with FixL HK to identify a crossover point which was then used as a reference for the alignment of the chosen atypical phytochrome PCMs with the FixL HK and Cph1 HK to decide on fusion points.

Homologous residues (overlap) are shown in bold. The dash in the chimaeric sequences represents the point of fusion.

4.2.2 Expression and Initial Characterisation of Chimaeric Phytochromes

Expression trials were conducted to check the expression levels and solubility of the chimaeras. SDS-PAGE gels showing the results of these trials (Figure 4.2) confirm that His$_6$-tagged protein could be produced and purified. The ~100 kDa band in lanes 4 and 10 for Dten Cph1 HK, and 85 kDa band in lanes 7 and 12 (Figure 4.2A), were analysed using tandem mass spectrometry (MS/MS). Protein sequence matches with the Dten and Npf phytochromes and Cph1 HK for the bands confirmed expression of the respective chimaeric proteins. Similarly, Dten FixL HK and Npf FixL HK were successfully purified using Ni^{2+}-affinity chromatography as indicated by the presence of a ~100 kDa band in lanes 4 and 7 on the SDS-PAGE gel (Figure 4.2B). These bands were also excised and analysed using MS/MS to further confirm that they were resulting from the correct chimaeras.
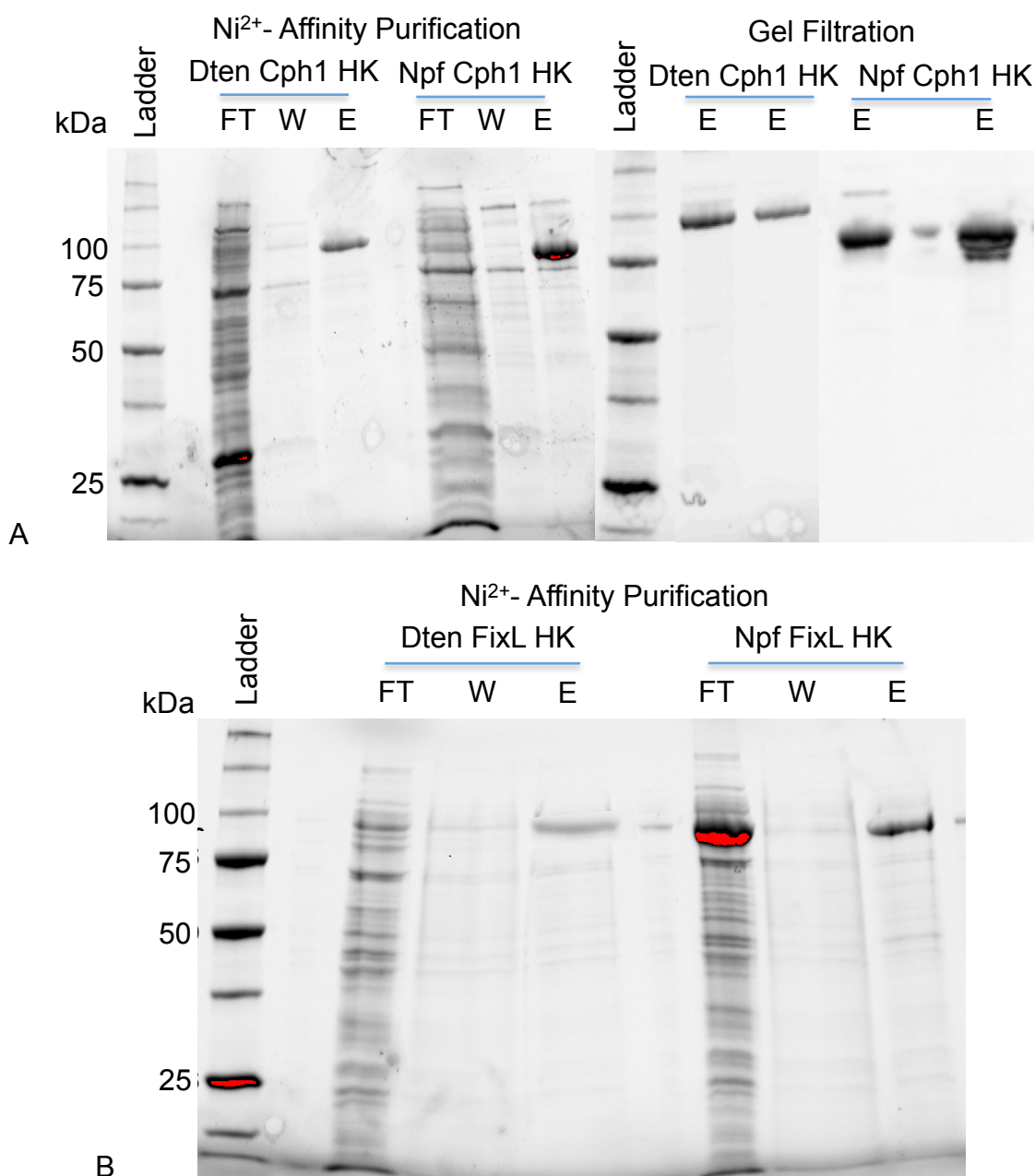
Figure 4.2. SDS-PAGE gels showing production of the chimaeric phytochromes.
A. Expression, Ni^{2+}- affinity purification and gel filtration of 94 kDa Dten Cph1 HK and 85 kDa Npf Cph1 HK chimaeras. Ladders are shown with the molecular weights. Flow through (FT), wash (W) and elute (E) fractions are analysed on the gel. The expected ~100 kDa Dten Cph1 HK band and ~80 kDa Npf Cph1 HK band can be seen in the expression and purification samples. **B.** Expression trials of 94 kDa Dten FixL HK and 84 kDa Npf FixL HK. FT, W and E fractions are analysed on the gel. ~100 kDa bands in the lanes E demonstrate successful purification of chimaeras.

To ensure that the chimaeras retained the original photochemistry, the absorption spectrum of each chimaera was recorded (Figure 4.3). The Dten Phy switches between orange and far-red light-absorbing states. The absorbance spectra of the Dten chimaeras with Cph1 HK and FixL HK were very similar to the Dten PCM in both states. All samples had absorbance peaks at 598 nm and 714 nm, with smaller peaks in the blue region at 354 nm and 375 nm, respectively (Figure 4.3 A, B and C). The far-red light-absorbing state also has a lower intensity absorbance in the orange region at 614 nm. The only difference is a slight difference in the intensity of the far-red light-absorbing state in the blue region of the spectrum.
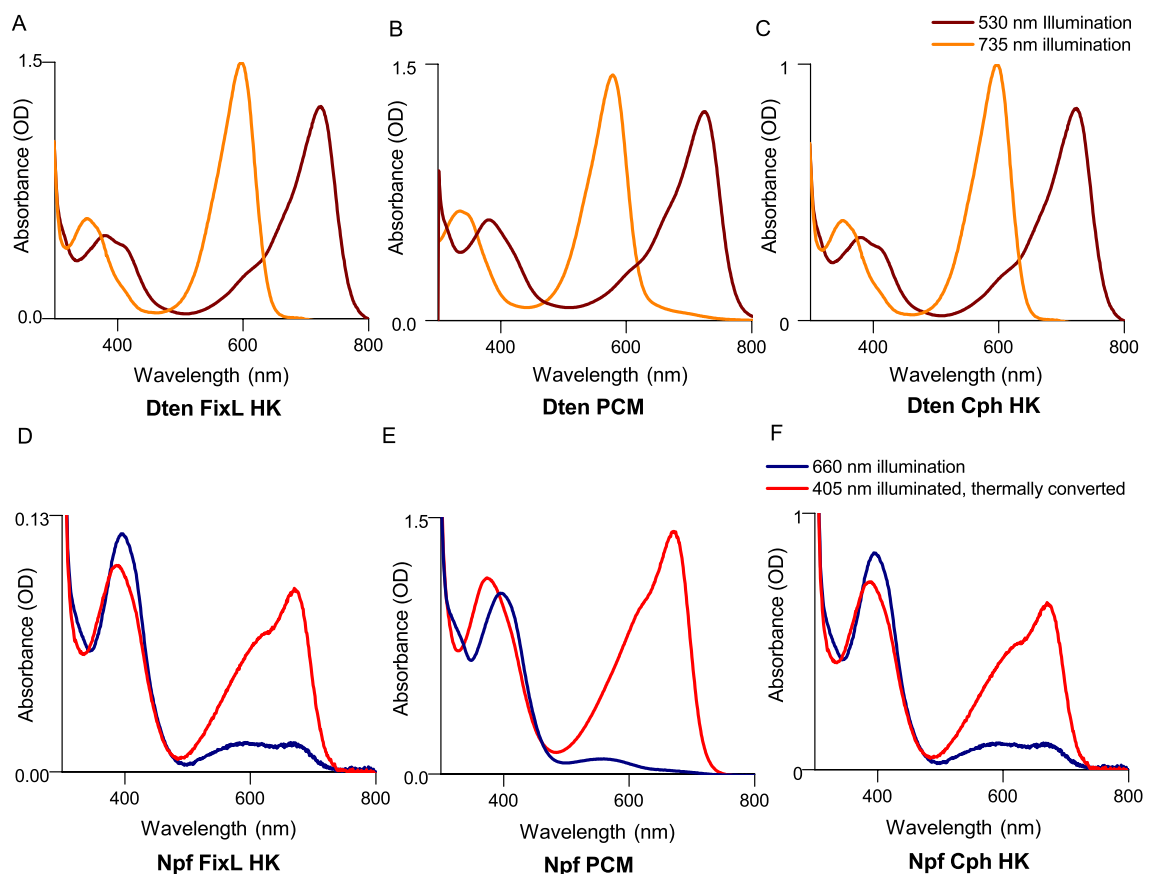
Figure 4.3. Steady state absorption spectra of chimaeric phytochromes. The steady state absorption spectra of the FixL HK (**A, D**) and Cph HK (**C, F**) chimeras are very similar to the wild type PCM (**B, E**) absorption spectra. Dten phytochrome absorbs maximally at 598 and 714 nm (**A, B, C**). Npf phytochrome absorbs maximally at 395 nm in the blue light-absorbing state, and 373 nm and 675 nm with low intensity absorbance at 625 nm in the red light-absorbing state (**D, E, F**).

It is known that the Npf PCM has a blue light-absorbing state, with an absorbance maximum at 395 nm and low intensity broad absorbance around 595 nm (Figure 4.3 D, E and F). Over a longer time course, this state builds into another state through thermal conversion (Figure 4.4), which absorbs maximally at 675 nm with additional absorbance at 625 nm and a smaller absorbance band in the blue region at 373 nm. The Npf FixL HK and Npf Cph HK have similar absorption spectra to the Npf PCM, although again there are slight differences in the blue region of the spectrum and the chimaeras have an additional low-intensity peak at around 690 nm in the blue light-absorbing state.
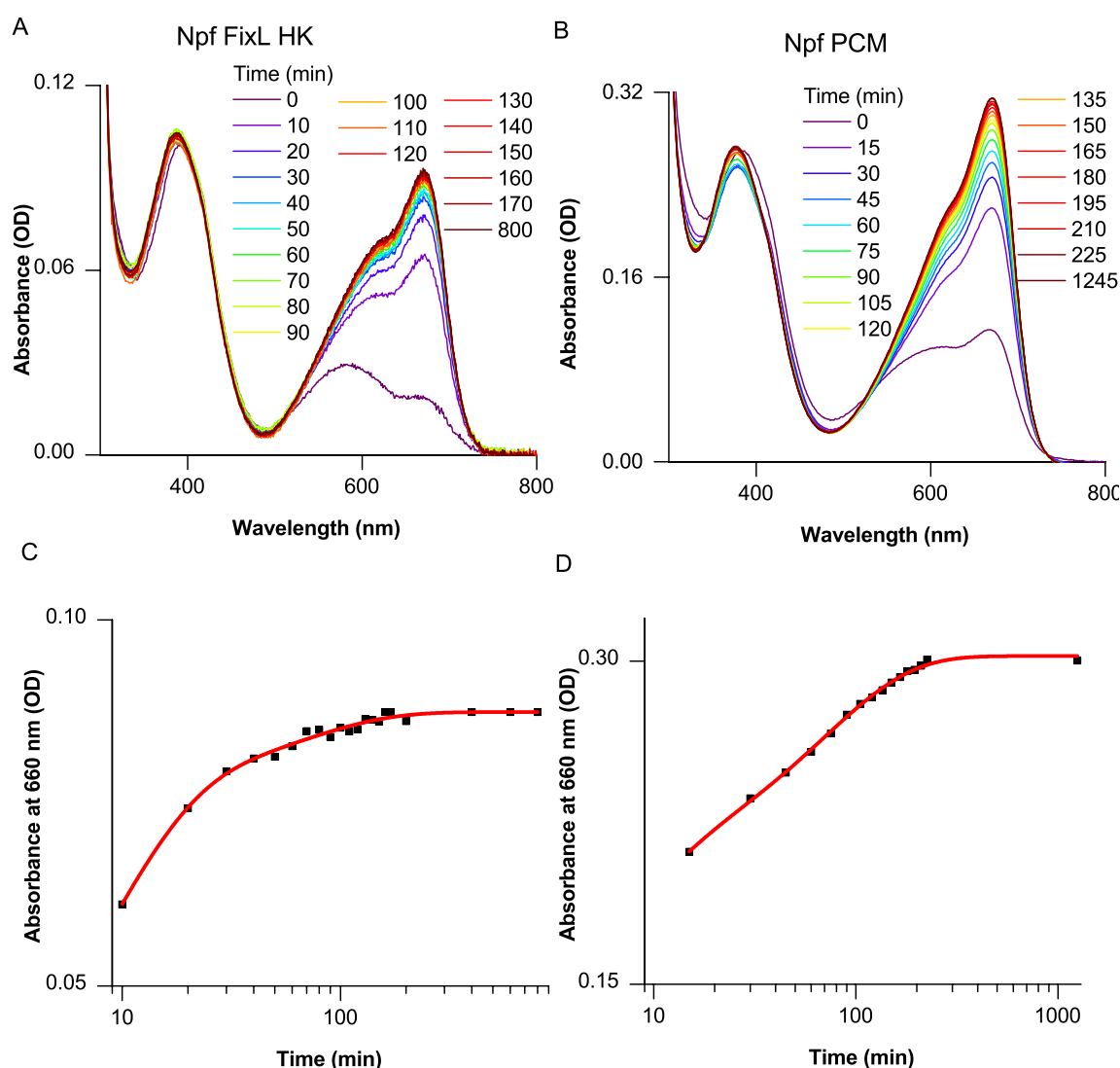
Figure 4.4. Kintetics of the thermal conversion step of Npf PCM and Npf FixL HK Pb > Pr photoreaction. A and **B**, absorbance spectra of the Npf FixL HK and Npf PCM respectively, captured at regular time intervals over several hours post excitation at 405 nm. **C** and **D**, kinetics of the slow, thermal step of the formation of Pr state of Npf FixL HK and Npf PCM respectively. Both datasets were fitted to a double exponential function.

The slow, thermal step of the Pb>Pr photoreaction was studied for both the chimaera and wild type phytochrome (Figure 4.4). The Pb state was excited at 405nm followed by incubation in the dark and absorbance spectra were recorded at regular intervals for a few hours (Figure 4.4 A and B). The kinetics of the formation of the Pr state were fitted to a double exponential function (Figure 4.4 C and D), which revealed lifetimes of 8 ± 0.4 min and 61 ±10 min for Npf FixL HK, and 6.3 ± 0.8min and 70 ± 3.3 min for Npf PCM. The kinetics of the thermal slower step of the chimaera is similar to the wild type Npf PCM.

4.2.3 *In-vitro* Phosphorylation Assays

As the PCMs and HKs were fused at the linker region, which has been reported to be essential for signal transduction,[10] it was necessary to check that the chimaeras were still able to function in terms of autophosphorylation and phosphotransfer activity. *In-vitro* tests were used to check whether the fused chimaeras were signalling active by following the autophosphorylation using a ^{32}P radiolabelling assay. The reactions were set-up using the kinase assay outlined in the methodology. The proteins used in the assay were all converted to one state at the beginning of the assay. The reactions were initiated with the addition of ATP and stopped at a range of times between 0 minutes and 30 minutes using SDS sample buffer. Proteins were resolved on 4-10% SDS polyacrylamide gels and trans-blotted to polyvinylidene difluoride (PVDF) membranes for autoradiography.

An increase in the radiation signal from ^{32}P over the time-course of the reaction in the top panel of Figure 4.5 shows that the Npf Cph HK is autophosphorylated and so it can be assumed that the linker is active and is transducing the signal from the PCM to the HK. Importantly, the thermally converted chimaera revealed significantly reduced autophosphorylation and phosphotransfer to Rcp1 activity (as shown in the bottom panel of Figure 4.5). . The lower level of phosphorylation in the thermally converted protein suggests that this is likely to be the signalling off state and the illuminated states are likely to be the signalling on state of the phytochrome. However, under 660 nm and 405 nm illumination there is no significant difference in autophosphorylation pattern.

This shows that these chimaeras do not respond to change in the illumination conditions.

Assay	405 nm Illumination	660 nm Illumination
Npf Cph1 HK	Reaction time (minutes) 0 1 2 5 10 20 30	Reaction time (minutes) 0 1 2 5 10 20 30
Npf Cph1 HK + Rcp1	Reaction time (minutes) 0 1 2 5 10 20 30 Thermally converted chimera Reaction time (minutes) 0 1 2 5 10 20 30	Reaction time (minutes) 0 1 2 5 10 20 30

Figure 4.5. Autoradiograph showing autophosphorylation and phosphotransfer activity of Npf Cph HK chimera. Autophosphorylation (top) and phosphotransfer activity (bottom) of Npf Cph HK chimaera under blue and red light illumination as reflected by the signal from ^{32}P. The reactions were started by the addition of ^{32}P-ATP and stopped at 0, 1, 2, 5, 10, 20 and 30 minutes. The top panel is the assay for Npf Cph HK chimaera only, following the autophosphorylation activity on the HK domain. The bottom panel is the assay for the Npf Cph HK chimaera in the presence of the cognate response regulator, Rcp1. The green arrow indicates the chimaeric phytochrome band and the amber arrow indicates the response regulator band.

Similarly, phosphorylation assays were set up for the Dten Cph HK chimaeras. As can be seen in the top panel of Figure 4.6, the Dten Cph HK chimaera has a constant level of ^{32}P incorporation over time. This indicates that although the

chimaera shows some phosphorylation activity, it is largely signalling inactive. There is no phosphotransfer activity, as can be seen from the lack of signal in the Rcp1 region of the assay shown in the bottom panel of Figure 4.6. In the wild type Dten protein there are two additional PAS domains, which make up the PCM. It is possible that the absence of these two domains in the present chimaeric proteins may have led to the chimaeras being signalling inactive. Consequently, chimaeric proteins with the two additional PAS domains attached between the PCM and the fused HK domain were synthesised and expressed. However, these were largely insoluble and did not bind to either nickel or Talon resin. It is possible that the expression of these chimaeric proteins could be optimised in the future using several different approaches (e.g. use of different cell strain for expression, change of plasmid, temperature of induction, addition of benzyl alcohol, change in culture conditions, etc.). Different purification methods (e.g. ammonium sulphate precipitation) could also be tested to find an optimal protein purification protocol.

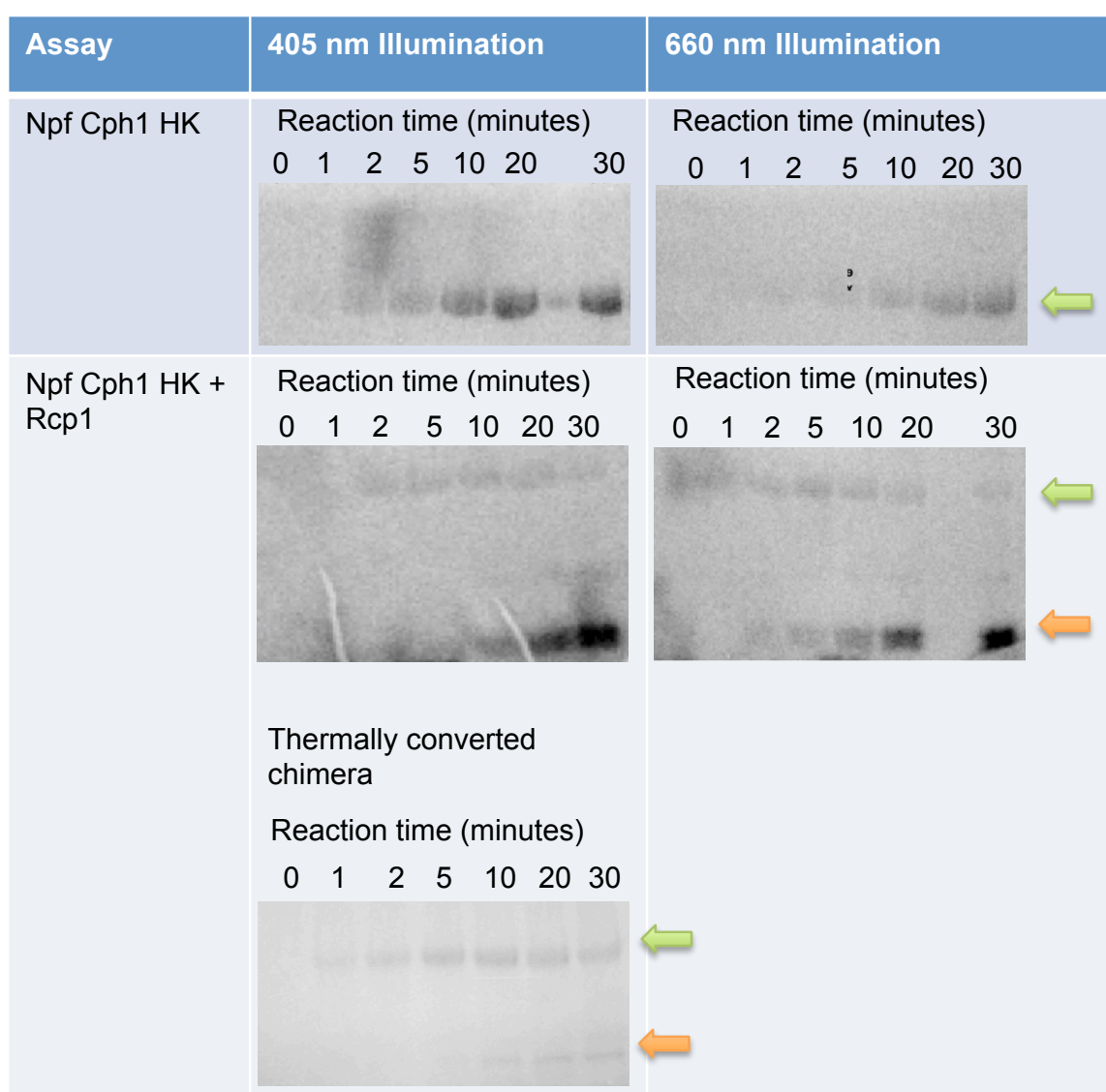| Assay	595 nm Illumination	735 nm Illumination
Dten Cph1 HK	Reaction time (minutes) 0 1 2 5 10 20 30	Reaction time (minutes) 0 1 2 5 10 20 30
Dten Cph1 HK + Rcp1	Reaction time (minutes) 0 1 2 5 10 20 30	Reaction time (minutes) 0 1 2 5 10 20 30

Figure 4.6. Autoradiograph showing autophosphorylation and phosphotransfer activity of Dten Cph HK chimera. Autophosphorylation (top) and phosphotransfer activity (bottom) of Dten Cph HK chimaera under orange and far-red light illumination as reflected by the signal from ^{32}P. The reactions were set-up using the kinase assay outlined in the methodology. The proteins used in the assay were all converted to one state at the beginning of the assay. The reactions were initiated with the addition of

ATP and stopped at a range of times between 0 minutes and 30 minutes. The top panel is the assay for the Dten Cph HK chimaera only, following the autophosphorylation activity on the HK domain. The bottom panel is the assay for the Dten Cph HK chimaera in the presence of the cognate response regulator, Rcp1. The green arrow indicates the chimaeric phytochrome band and the amber arrow indicates the response regulator band.

These assays were then repeated for the FixL chimaeras (Figures 4.7 and 4.8). The Npf FixL HK showed an increase in autophosphorylation over time after illumination at both wavelengths, as is evident by the increase in ^{32}P signal over time in the assay shown in the top panel of Figure 4.7. A similar pattern is seen for the phosphotransfer step to the FixJ partner protein in both states of the protein (Figure 4.7). Similar to the Npf Cph HK, the assay for the thermally converted Npf FixL HK revealed no phosphotransfer activity (bottom panel of Figure 4.7).
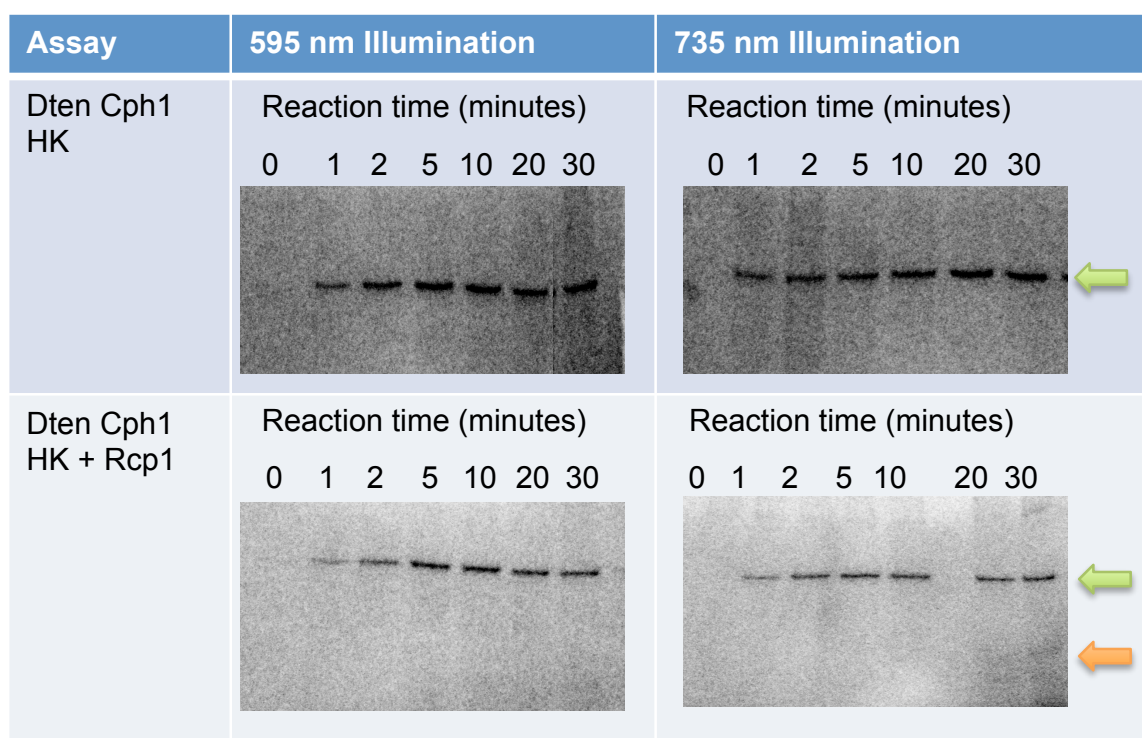
Figure 4.7. Autoradiograph showing autophosphorylation and phosphotransfer activity of Npf Cph HK chimera. Autophosphorylation (top) and phosphotransfer activity (bottom) of Npf FixL HK chimaera under blue and red light illumination as reflected by the signal from ^{32}P. The reactions were set-up using the kinase assay outlined in the methodology. The proteins used in the assay were all converted to one state at the beginning of the assay. The reactions were initiated with the addition of ATP and stopped at a range of times between 0 minutes and 30 minutes. The top panel is the assay for the Npf FixL HK chimaera only, following the autophosphorylation activity on the HK domain. The bottom panel is the assay for the Npf FixL HK chimaera in the presence of the cognate response regulator, FixJ. The green arrow indicates the chimaeric phytochrome band and the amber arrow indicates the response regulator band.

Figure 4.8 shows the phosphorylation assay for the Dten FixL HK. The autophosphorylation activity of the Dten FixL HK chimaera is confirmed by an

increase in ^{32}P signal over time in the top panel of Figure 4.8, but in the bottom panel the ^{32}P signal does not increase or decrease significantly over the course of the reaction so it can be concluded that the phosphotransfer activity of this chimaera can occur in both states of the protein. .
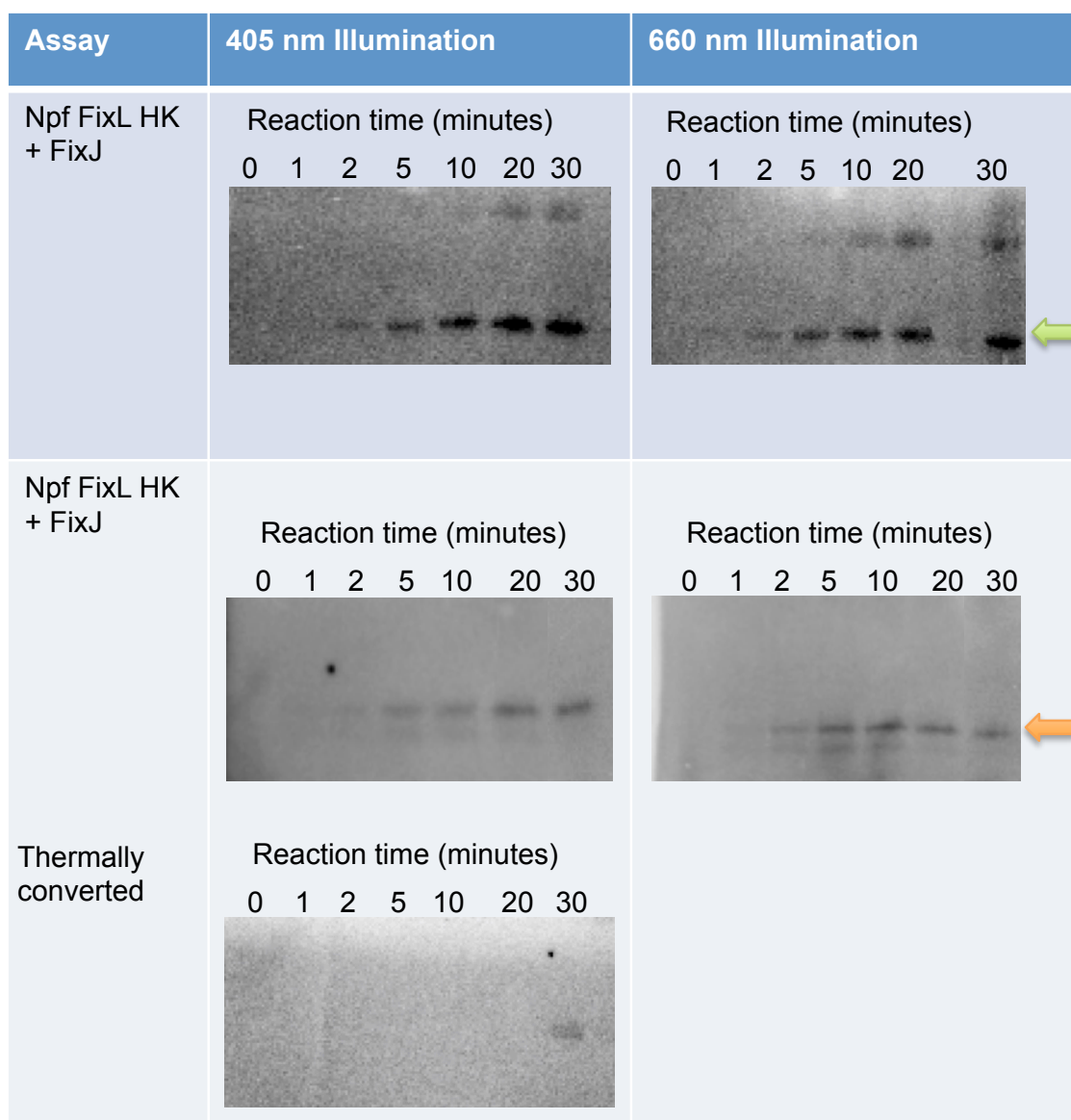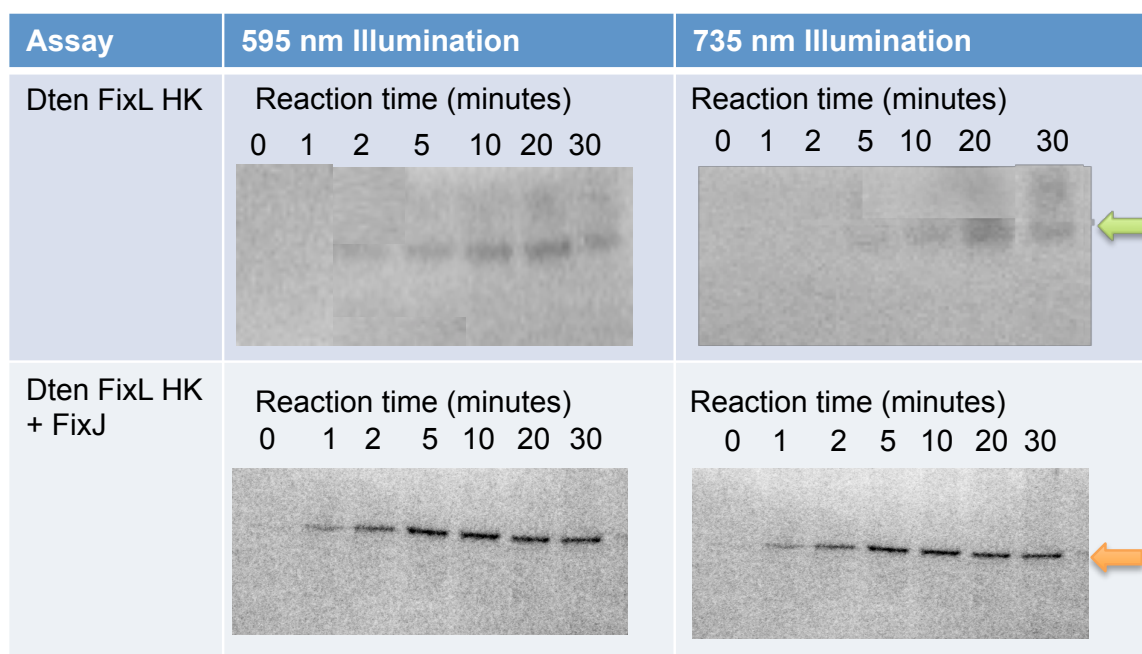

Figure 4.8. Autoradiograph showing autophosphorylation and phosphotransfer activity of Npf Cph HK chimera. Autophosphorylation (top) and phosphotransfer activity (bottom) of Dten FixL HK chimaera under orange and far-red light illumination as reflected by the signal from ^{32}P. The reactions were set-up using the kinase assay outlined in the methodology. The proteins used in the assay were all converted to one state at the beginning of the assay. The reactions were initiated with the addition of ATP and stopped at a range of times between 0 minutes and 30 minutes. The top panel is the assay for the Dten FixL HK chimaera only, following the autophosphorylation activity on the HK domain. The bottom panel is the assay for Dten FixL HK chimaera in the presence of the cognate response regulator, FixJ. The green arrow indicates the chimaeric phytochrome band and the amber arrow indicates the response regulator band.

4.2.4 Flash photolysis of Npf FixL HK

The *in-vitro* studies of the FixL chimaeras demonstrated that only the Npf FixL HK was active and could be taken forward to *in-vivo* studies. Therefore, the kinetics of this chimaera was studied to compare with the wild type Npf phytochrome. The kinetics of the formation of the two states of the protein were

measured for the Npf FixL HK chimaera to check whether the addition of the HK domains affected the kinetics of the photoreaction compared to that of the PCM. To do this the protein was all converted to a single form and then photoexcited with a laser pulse centred at either 660 nm and 420 nm, after which the absorbance changes were measured at 590 nm and 660 nm, respectively, to capture kinetics in the time window from 2.25 µs to 1 s. The kinetic transients were fitted using multiple different exponential functions (Figure 4.9) and the lifetimes are listed in Table 1.

The kinetic data for the formation of the 660 nm absorbing state after excitation at 420 nm was fitted to a double exponential function to yield lifetimes of 8.2 ms and 77 ms for Npf FixL HK and 8.0 ms and 59 ms for Npf PCM. Upon excitation at 600 nm the rate of decay of the 590 nm absorbing state was fitted to a triple exponential function to yield lifetimes of 64 µs, 629 µs and 97 ms for Npf FixL HK and 57 µs, 437 µs and 87 ms for Npf PCM. The lifetimes of the wild type protein and the chimaera are similar, but the lifetimes associated with slower processes appear to be slightly longer for the chimaera. These slower phases are likely to be associated with conformational changes in the protein[29,32] and therefore, the changes may be due to the addition of the bulky HK domain onto the chimaera, which is not present in the wild type PCM module.
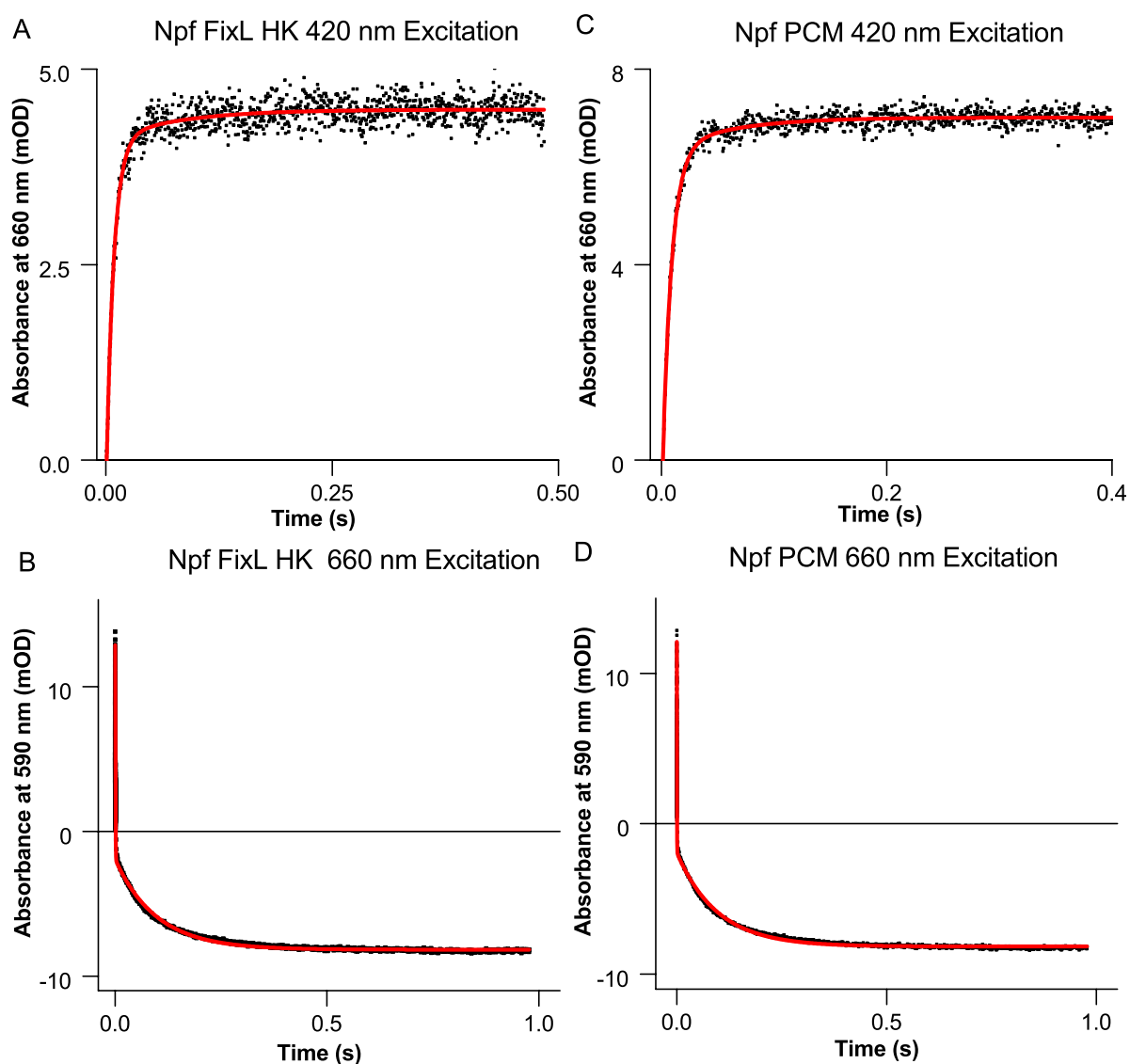
Figure 4.9. Npf PCM and Npf FixL HK kinetics of change in absorbance measured at 590 nm and 660 nm after photoexcitation at 660 nm and 420 nm respectively. A and **B**: kinetics of Npf FixL HK formation of 600 nm absorbing state and disappearance of 590 nm absorbing state, respectively. **C** and **D**: kinetics of Npf PCM formation of 600 nm absorbing state and disappearance of 590 nm absorbing state, respectively. The kinetic data for the formation of the 590 nm absorbing state was fitted using a double exponential function. The kinetic data for the formation of the 590 nm absorbing state was fitted using a quadruple exponential function.

Table1. Lifetimes of transition processes and intermediates in the formation of the 660 nm and 590 nm absorbing states. Standard errors from the fitting procedure are listed for the lifetimes.

	Npf FixL HK	Npf PCM
Kinetics of formation of 660 nm absorbing state	$t1 = 8.2 \pm 0.3$ ms	$t1 = 8.0 \pm 0.2$ ms
	$t2 = 77 \pm 14$ ms	$t2 = 59 \pm 7$ ms
Kinetics of formation of 590 nm absorbing state	$t1 = 64 \pm 2$ μs	$t1 = 57 \pm 2$ μs
	$t2 = 629 \pm 39$ μs	$t2 = 437 \pm 10$ μs
	$t3 = 97 \pm 7$ ms	$t3 = 87 \pm 1$ ms

To compare the relative efficiency of the Npf PCM and Npf FixL HK photoreactions, samples were converted to the red light-absorbing state and then sequentially excited using 420 nm laser pulses. The absorption spectrum of the sample was measured at regular intervals between the pulses, shown in Figure 4.10. A linear fit to the data revealed a change in absorption of 0.0012 per laser pulse for the Npf PCM and 0.0014 per laser pulse for Npf FixL HK. This shows that the photoreaction of the chimaera is as efficient as the wild type protein.
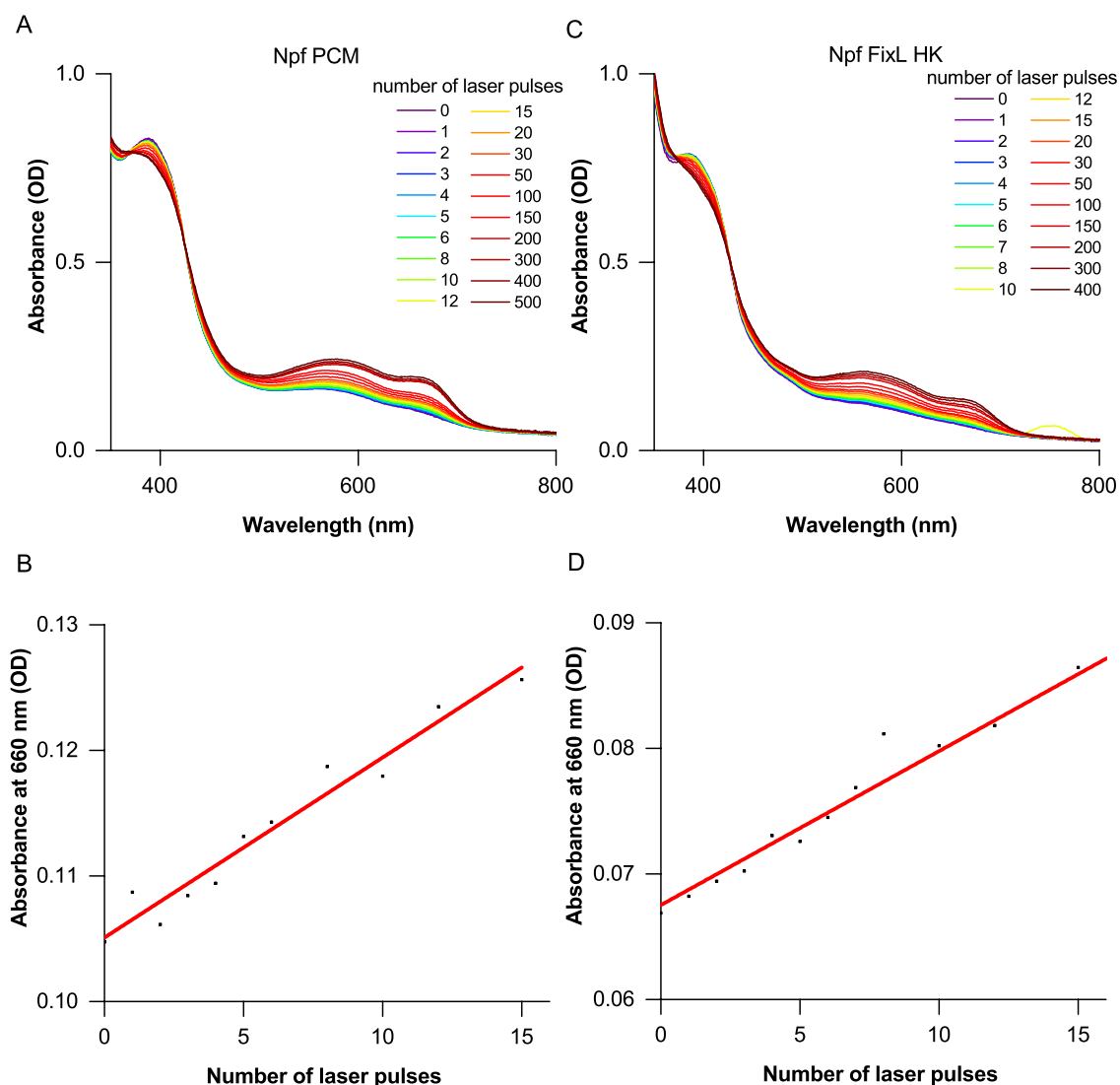
Figure 4.10. Comparison of quantum yield of 660 nm forming step of Npf PCM and Npf FixL HK. The top panels, **A** and **C**, shows the raw spectra for Npf PCM and Npf FixL HK respectively, obtained between sequential laser pulses used to pump the sample. The bottom panels, **B** and **D**, show the formation of the 660 nm absorbing state as the Npf PCM and Npf FixL HK samples, were pumped using increasing number of laser pulses, respectively. The data was fitted to a linear function, y=a + bx, fitted values for b of 0.0014 and 0.0012 were obtained for Npf PCM and Npf FixL HK samples, respectively.

4.2.5 *In-vivo* Assays

Once the *in-vitro* activity was confirmed, the *in-vivo* activity of the active chimaera could then be investigated by using a GFP reporter assay. The Npf FixL HK and Npf Cph HK chimaeras were the only proteins to exhibit a robust light-dependent autophosphorylation and phosphotransfer activity *in-vitro*.

However, the target promoters for the response regulator of the Cph1 HK, Rcp1, are currently unknown. Therefore, we only took the Npf FixL HK into subsequent *in-vivo* studies. To do this, constructs were designed based on plasmid systems that had previously been used successfully in other work.[12] This involved removing genes for the phytochrome, response regulator and target promoter from the published Cph1 EnvZ – OmpR and SyCcaS - SyCcaR red/green light absorbing system plasmids pSR33.4 and pSR59.8 and replacing them with the Npf FixL HK chimaera phytochrome, its response regulator FixJ and the target promoter for FixJ, nifAP. Hence, the entire TCS system was assembled onto two plasmids to form plasmids pSR1.0 and pSR2.0 (Figure 4.11), respectively. Npf FixL HK was ligated into pSR33.4 to assemble pSR1.0, whereas FixJ and the target promoter, nifAP, were transformed onto pSR59.8 to assemble pSR2.0. The nifAP was placed in front of the GFP reporter gene. The constructs were assembled using overlap extensions PCRs and InFusion cloning as outlined in the methodology.
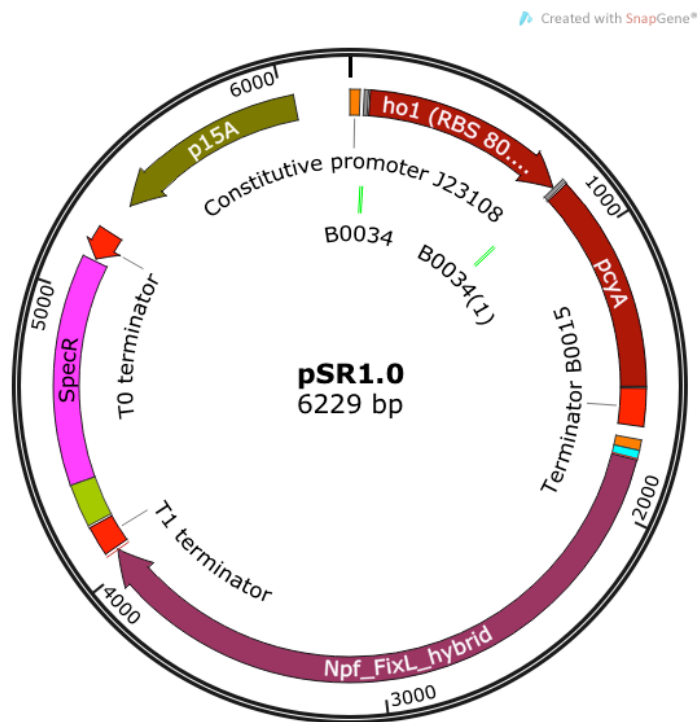
A
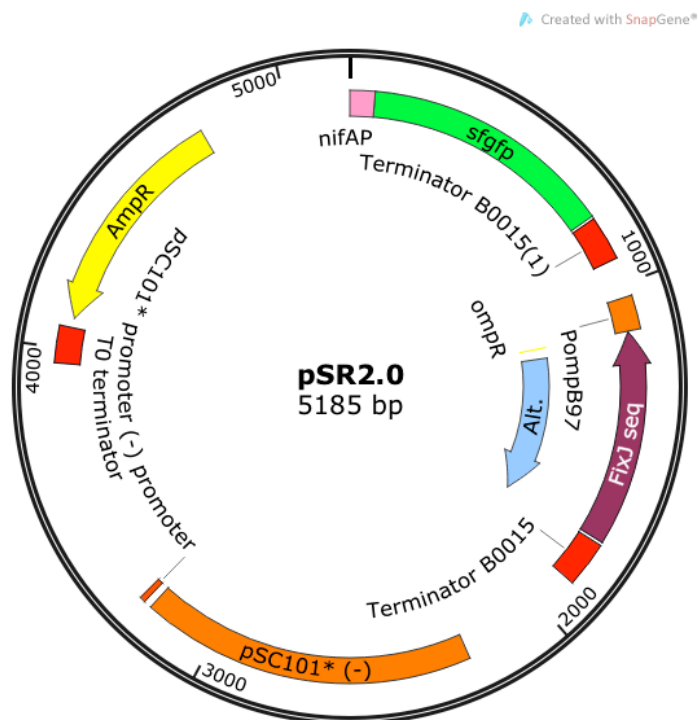
B

Figure 4.11. Plasmids constructed for *in-vivo* assays. Npf FixL HK, FixJ and NifAP genes refactored onto pSR33.4 (pSR1.0) and pSR59.8 (pSR2.0) plasmids, respectively. The *sfgfp* gene was placed under the control of nifAP, which is the target promoter of FixJ. The pSR1.0 plasmid also contains the genes for biliverdin production, *ho1* and *pcyA*. The *pcyA*, *ho1*, *npf fixL HK* and *fixJ* genes were placed under the

control of constitutive promoters. pSR1.0 is spectinomycin resistant and has a p15 origin or replication. pSR2.0 is ampicillin resistant and has a pSC101 origin of replication.

The pSR1.0 and pSR2.0 constructs were assembled and verified by DNA sequencing before being used in *in-vivo* assays. Cultures were set-up under blue and red light illumination to monitor the effect of the two different states on the promoter activation through the use of GFP reporter gene. The relative GFP fluorescence of the samples was measured and normalised (Figure 4.12). The chimaeric phytochrome was shown to be signalling active as a ~2 fold increase in relative fluorescence of GFP can be observed under 660 nm illumination compared to 405 nm illumination. The cells expressing GFP constitutively were used as a positive control and empty cells were used as a negative control. As expected, a higher level of GFP production was evident in the positive control and almost no fluorescence in the negative control. Surprisingly, GFP fluorescence in chimaera-containing cells reduced by ~20% over the 48 hours, which was not seen in any of the other cultures. In the future this could be further explored by setting up cultures illuminated with 405 nm LED for different lengths of time before being incubated in the dark to probe the *in-vivo* reporter gene expression difference in the thermally converted state of the chimaera. The 2-fold induction is similar to the levels of induction seen by Tabor *et al.* (2011) with the Cph1 EnvZ – OmpR and CcaS - CcaR Red-Green system (8.9-fold and 6.4-fold, respectively).[12] The 8.9 and 6.4-fold activation for red and green system were achieved through promoter optimisation.[43] Studies on the UirS - UirR system reported 4.4-fold activation.[37] Further optimisation studies could help improve the activity of the Npf FixL HK chimaera. Based on the *in-vitro* data, which showed significantly reduced ^{32}P signal in the thermally converted form, we would expect to see significantly reduced, or increased, reporter gene expression with this form of the protein.
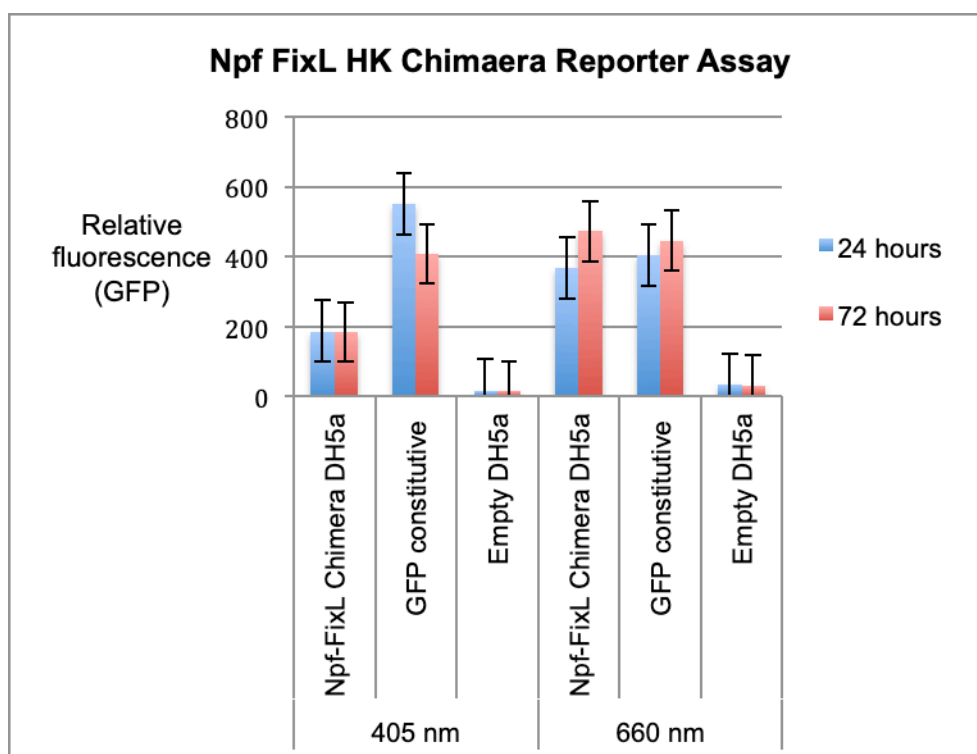
Figure 4.12. *In-vivo* **reporter assay set up to check Npf FixL HK chimaera's signalling activity.** GFP fluorescence was measured at 24 and 72 hours intervals to check the activity of the Npf FixL HK chimaera. Cells constitutively expressing GFP were used as a positive control and empty cells as a negative control. The results show ~2-fold activation of the chimaera under 660 nm illumination.

4.3 Conclusions

Here we have demonstrated that phytochromes can be paired with TCS HKs in a 'mix and match' manner to make functional chimaeras that can be used for optogenetic control of gene expression.[10,12] Further studies need to be conducted on the additional Dten FixL HK chimaeras to optimise its solubility and purification so that its use in optogenetics can be further explored. Additional chimaeras can be constructed using additional TCS HKs such as PhoR.[44] We have demonstrated that adding non-native / foreign TCS HK domains onto phytochrome PCMs does not affect their photochemistry and kinetics in unexpected ways and that certain chimaeras are active *in-vivo* using reporter gene assays. This may be a promising route for differential expression in biosynthetic pathways being engineered in synthetic biology to replace expensive chemical inducers and achieve finer control on gene expression within different parts of the pathway.

4.4 Methodology

4.4.1 Construction of chimaeras and refactoring of chimaeras into pSR system

Cph HK Chimaeras:

Npf Cph HK and Rcp1 genes were synthesised by GeneArt. To construct the Dten Cph HK chimera, Dten PCM was amplified out of pET21a using primers 1 and 2 (Table 2) and Cph HK was amplified out of Npf Cph HK pET21a using primers 3 and 4 (Table 2). CloneAmp polymerase protocol was followed for PCRs. The primers were designed with complimentary overhangs as per the InFusion^{TM} cloning protocol. Amplified products were DpnI digested for 3 hours; gel extracted and fused using InFusion cloning. This was transformed into Stellar cells. Colonies were selected and verified by sequencing. Verified colonies were used in further *in-vitro* and *in-vivo* studies.

Table 2. Primers used to construct Dten Cph HK chimera.

Primers used to construct Cph1 HK chimaeras:
Primer 1: AAACGACGGCCAGTGCATATGGCCGCTAAAATCTCGCG
Primer 2: GCTTGCTTCATTGGATTCCGTGGTTTCTTGC
Primer 3: TCCAATGAAGCAAGCTCTAACGTTCTAATGCGG
Primer 4: CACTGGCCGTCGTTTTAC

FixL HK Chimaeras:

FixJ, Dten FixL HK and Npf FixL HK chimaera genes were synthesised and cloned into pET21a by GeneArt.

4.4.2 *In-vivo* testing constructs

For *in-vivo* testing, published red green system (Cph1 EnvZ – OmpR and CcaS – CcaR) plasmids pSR33.4 and pSR59.8 were used as a backbone because these for optimised for the optogenetic system and could be used for comparison of our tools. pSR 33.4 and pSR59.8 were ordered from Addgene (deposited by Tabor lab, Schimdl *et al.* 2014).[43]

Npf FixL HK was amplified out of its vector using primers 5 and 6 (Table 3) and pSR33.4 was opened around the phytochrome gene using primers 7 and 8 (Table 3). The primers were designed to exclude the existing phytochrome system and consisted complimentary overhangs as required for InFusion cloning. The amplified vector was DpnI treated for 3 hours and both PCR products were gel extracted and fused using the InFusion Cloning kit to form plasmid pSR1.0. The Infusion product was transformed in Stellar cells and colonies were selected and verified with DNA sequencing. FixJ and nifAp were refactored into pSR59.8. FixJ was amplified out of its vector using primers 9 and 10 (Table 3). pSR59.8 was amplified using primers 11 and 12 for Npf FixL HK insertion and primers 13 and 14 for nifAp ligation (Table 3). The amplified vector was DpnI treated for 3 hours and both PCR products were gel extracted and fused using the InFusion Cloning kit to form plasmid pSR2.0. The Infusion product was transformed in Stellar cells and colonies were selected and verified with DNA sequencing.

Table 3. Primers for the construction of the pSR 1.0 and pSR 2.0 plasmid.

Primers used for refactoring the functional chimaera into pSR33.4 and PSR59.8 system:
Primer 5: CATAAGGCTCTAATACCATGAGCCAGCCGGAAAATACCAC
Primer 6: TTTGATGCCTCTAGTTTAATCATTTGCAACAATGCG
Primer 7: CGAGGAGCTCGGAGCACTAATCTAGAGGCATCAAATAAAA
Primer 8: TTCCGGCTGGCTCATGGTATTAGAGCCTTATGTGATAGCGGCGG
Primer 9: CCTTTGGGAGTACAAACAATGACCGATTATACCGTGCATATCGTGG
Primer 10: GCCTGGTCTAGATTATTATTATTAGCTCGGACCAAAACCGCC
Primer 11: GTTTTGGTCCGAGCTAATAATAATCTAGACCAGGC
Primer 12: GCACGGTATAATCGGTCATTGTTTGTACTCCCAAAGG
Primer 13: TAATATTAAGCGGGCGAGAAATGACTAAGGTGCTCCCAATGCGTAAAGGCGAAGAGC
Primer 14: TTAATATTAGGAGGTATGCGTATTGACAGTCGGAGAGCGATCAGAAATCATCCTTAGCGAAAGC

4.4.3 Expression and purification

The pET21a plasmids containing chimaera genes were co-expressed with the pCola Duet plasmid containing genes responsible for bilin synthesis, *PcyA* and *HO1,* encoding heme oxygenase-1 (HO1) and phycocyanobilin (PCB):ferredoxin (Fd) oxidoreductase (PcyA).

HMS174 cells were used for expression of chimeras. HMS174 cells containing pCola Duet plasmid were made competent using $CaCl_2$ and transformed with the chimaeric phytochrome DNA. Single colonies from agar plates were used for starter cultures, which were grown at 37 °C and used to inoculate 0.5 L cultures containing the appropriate antibiotics and 10 mM δ-ALA at OD_{600} 0.4 – 0.6. Upon inoculation the temperature was dropped to 30 °C. At OD_{600} 0.4 – 0.6 the cultures were induced with 100 μM IPTG and the cultures were then incubated at 26 °C overnight before being harvested.

The cell pellet was suspended in binding buffer (50 mM Tris pH 7.6, 300 mM NaCl, 5 mM Imidazole, 10% glycerol, 14 mM β-mercaptoethanol, pH 7.6) with protease inhibitor and passed through the cell disruptor (Constant Systems) once at 15 kpsi. The lysate was spun down and supernatant incubated with equilibrated (with binding buffer) Ni^{2+} resin for 3 hours minimum. The flow-through was then collected and the resin was washed with wash buffer (50 mM Tris pH 7.6, 300 mM NaCl, 50 mM imidazole, 10% glycerol, 14 mM β-mercaptoethanol, pH 7.6). Protein was eluted with elution buffer (50 mM Tris-HCl pH 7.6, 300 mM NaCl, 200 mM imidazole, 10 glycerol, 14 mM β-mercaptoethanol, pH 7.6) and concentrated. The protein was then purified using a GE Hi Load 26 60 200 pg gel filtration column and equilibration buffer (50mM Tris-HCl pH 7.6, 150 mM NaCl, 10% glycerol, 14mM β-mercaptoethanol, pH 7.6). Pure fractions were combined and concentrated.

4.4.4 Radiolabelling (*in-vitro)* assays

Radiolabelling assays were all conducted in a dark room at room temperature and under relevant illumination conditions using LEDs. Dark reactions were covered with foil as soon as the reactions were started. The reactions were set up using 10 μM phytochrome/ chimaeric phytochrome (all converted to one

state), 500 µM ATP and 0.13 µM 32-P ATP, 50 µM response regulator, 10% glycerol, 1X Protein Kinase Buffer (50 mM Tris-HCl pH 7.6, 300 mM NaCl, 1 mM β-mercaptoethanol, 2 mM Mg^{2+}, pH 8.1). ATP. Reactions were started with the addition of ^{32}P-ATP and stopped at 0, 1, 2, 5, 10, 20 and 30 minutes with the addition of 1x SDS sample buffer. Proteins were resolved on SDS–4-20% polyacrylamide gels and trans-blotted to PVDF membranes for autoradiography. The PVDF membrane was exposed to a GE storage phosphor screen for 60 - 180 minutes and visualized on a Typhoon biomolecular imager.

4.4.5 *In-vivo* Assays

In-vivo assays were carried out using DH5α cells. Plasmids pSR1.0 and pSR2.0 were co-transformed into the above cells. Colonies were selected to start experimental triplicates. Cells constitutively expressing GFP were used are positive control. Empty cells and cells harbouring plasmid pSR2.0 only were used as negative control, and to account for any leaky and non-light regulated GFP production, respectively. The cultures were grown under blue (405 nm), red (660 nm) and no illumination (dark) Apart from the GFP gene in pSR2.0, which is under the control of the light dependent nifAP promoter, all other genes are constitutive therefore once the cultures were inoculated with starter cultures the cultures were left to grow at 30 °C, 190 rpm overnight. The following morning cell density of cultures was measured and normalised to measure GFP fluorescence using a microtitre plate reader.

4.4.6 Flash photolysis studies of wild type and chimaera phytochromes

The phytochrome photodynamics were studied using a laser system consisting of an optical parameter oscillator (OPO) pumped by a Q-switched Nd: YAG laser (Brilliant B, Quantel). Excitation wavelengths were generated using the OPO pumped by the third harmonic of the laser. A quartz cuvette with a 2 mm path-length for pump-excitation and 10 mm path length for probing of the absorption spectra was used. Changes in absorption were recorded using an LKS-60 Flash-Photolysis instrument (Applied Photophysics Ltd.). Kinetics were recorded in 10 nm increments across the visible spectral region, with three

repeats at each probe wavelength. Protein activity was periodically checked against the initial protein activity. Samples illuminated with a 660 nm LED to generate the blue light absorbing state were then photoexcited at 420 nm, this excitation wavelength was selected due to the limited pump wavelength range of the laser system, kinetics were recorded over a timescale of 2.25 µs – 1 s. Samples illuminated with a 405 nm LED to generate the red light absorbing state were then photoexcited at 660 nm, kinetics were recorded over a timescale of 0.25 µs – 1s.

4.5 Reference List:

[1]   M. A. van der Horst, K. J. Hellingwerf, *Acc. Chem. Res.* **2004**, *37*, 13–20.

[2]   T. D. Lamb, *Prog Retin Eye Res* **2013**, *36*, 52–119.

[3]   M. Chen, J. Chory, C. Fankhauser, *Annu. Rev. Genet.* **2004**, *38*, 87–117.

[4]   *Nature Chemical Biology* **2014**, *10*, 483.

[5]   E. J. Olson, J. J. Tabor, **2014**, *10*, 502–511.

[6]   W. Bacchus, M. Fussenegger, *Curr. Opin. Biotechnol.* **2012**, *23*, 695–702.

[7]   G. Aston-Jones, K. Deisseroth, *Brain Research* **2013**, *1511*, 1–5.

[8]   A. Möglich, P. Hegemann, *Nature* **2013**, *500*, 406–408.

[9]   A. M. Packer, B. Roska, M. Häusser, *Nature Neuroscience* **2013**, *16*, 805–815.

[10]   A. Levskaya, A. A. Chevalier, J. J. Tabor, Z. B. Simpson, L. A. Lavery, M. Levy, E. A. Davidson, A. Scouras, A. D. Ellington, E. M. Marcotte, et al., *Nature* **2005**, *438*, 441–442.

[11]   A. Möglich, K. Moffat, *J. of Mol. Biol.* **2007**, *373*, 112–126.

[12]   J. J. Tabor, A. Levskaya, C. A. Voigt, *J. of Mol. Biol.* **2011**, *405*, 315–324.

[13]   M.-H. Ryu, M. Gomelsky, *ACS Synth. Biol.* **2014**, *3*, 802–810.

[14]   H. Liu, G. Gomez, S. Lin, S. Lin, C. Lin, *PLoS ONE* **2012**, *7*, e50738.

[15]   X. Wang, X. Chen, Y. Yang, *Nat. Methods* **2012**, *9*, 266–269.

[16]   S. Konermann, M. D. Brigham, A. Trevino, P. D. Hsu, M. Heidenreich, L. Cong, R. J. Platt, D. A. Scott, G. M. Church, F. Zhang, *Nature* **2013**, *500*, 472–476.

[17]   O. Yizhar, L. E. Fenno, T. J. Davidson, M. Mogri, K. Deisseroth, *Neuron* **2011**, *71*, 9–34.

[18]   F. Zhang, J. Vierock, O. Yizhar, L. E. Fenno, S. Tsunoda, A. Kianianmomeni, M. Prigge, A. Berndt, J. Cushman, J. Polle, et al., *Cell* **2011**, *147*, 1446–1457.

[19]   M. Avery, J. Nassi, J. Reynolds, *PLoS ONE* **2018**, *13*, e0205386.

[20]   P. Mahmoudi, H. Veladi, F. G. Pakdel, *J Med Signals Sens* **2017**, *7*, 71–79.

[21]   D. M. Shcherbakova, A. A. Shemetov, A. A. Kaberniuk, V. V. Verkhusha, *Annu Rev Biochem* **2015**, *84*, 519–550.

[22]   M. Ikeuchi, T. Ishizuka, *Photochem. Photobiol. Sci.* **2008**, *7*, 1159–1167.

[23]   N. C. Rockwell, D. Duanmu, S. S. Martin, C. Bachy, D. C. Price, D. Bhattacharya, A. Z. Worden, Lagarias, *Proc. Natl. Acad. Sci. U.S.A.* **2014**, *111*, 3871–3876.

[24]   P. H. Quail, *Plant Cell Environ.* **1997**, *20*, 657–665.

[25]   B. L. Montgomery, Lagarias, *Trends Plant Sci.* **2002**, *7*, 357–366.

[26]   B. Karniol, J. R. Wagner, J. M. Walker, R. D. Vierstra, *Biochem. J.* **2005**, *392*, 103–116.

[27]   A. Möglich, X. Yang, R. A. Ayers, K. Moffat, *Annu. Rev. Plant Biol.* **2010**, *61, 21–47.*

[28]   N. C. Rockwell, Y.-S. Su, Lagarias, *Annu. Rev. Plant Biol.* **2006**, *57, 837–858.*

[29]   H. Takala, A. Björling, O. Berntsson, H. Lehtivuori, S. Niebling, M. Hoernke, I. Kosheleva, R. Henning, A. Menzel, J. A. Ihalainen, et al., *Nature* **2014**, *509, 245–248.*

[30]   K. Anders, A. Gutt, W. Gärtner, L.-O. Essen, *J. Biol. Chem.* **2014**, *289, 25590–25600.*

[31]   S. Noack, N. Michael, R. Rosen, T. Lamparter, *Biochemistry* **2007**, *46, 4164–4176.*

[32]   A. Björling, O. Berntsson, H. Lehtivuori, H. Takala, A. J. Hughes, M. Panman, M. Hoernke, S. Niebling, L. Henry, R. Henning, et al., *Sci Adv* **2016**, *2, e1600920–e1600920.*

[33]   B. Esteban, M. Carrascal, J. Abian, T. Lamparter, *Biochemistry* **2005**, *44, 450–461.*

[34]   E. S. Burgie, T. Wang, A. N. Bussell, J. M. Walker, H. Li, R. D. Vierstra, *J. Biol. Chem.* **2014**, *289, 24573–24587.*

[35]   A. Levskaya, O. D. Weiner, W. A. Lim, C. A. Voigt, *Nature* **2009**, *461, 997–1001.*

[36]   S. Shimizu-Sato, E. Huq, J. M. Tepperman, P. H. Quail, *Nat. Biotechnol.* **2002**, *20, 1041–1044.*

[37]   P. Ramakrishnan, J. J. Tabor, *ACS Synth. Biol.* **2016**, *5, 733–740.*

[38]   E. H. S. Sousa, J. R. Tuckerman, A. C. S. Gondim, G. Gonzalez, M.-A. Gilles-Gonzalez, *Biochemistry* **2013**, *52, 456–465.*

[39]   M. A. Gilles-Gonzalez, G. S. Ditta, D. R. Helinski, *Nature* **1991**, *350, 170–172.*

[40]   S. R. Schmidl, R. U. Sheth, A. Wu, J. J. Tabor, *ACS Synth. Biol.* **2014**, *3, 820–831.*

[41]   S. Crosson, P. T. McGrath, C. Stephens, H. H. McAdams, L. Shapiro, *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102, 8018–8023.*

[42]   A. A. Iniesta, N. J. Hillson, L. Shapiro, *J. Bacteriol.* **2010**, *192, 3893–3902.*

[43]   S. R. Schmidl, R. U. Sheth, A. Wu, J. J. Tabor, *ACS Synth. Biol.* **2014**, *3, 820–831.*

[44]   S. Chakraborty, J. Sivaraman, K. Y. Leung, Y.-K. Mok, *J. Biol. Chem.* **2011**, *286, 39417–39430.*

Chapter 5 Discussion and Outlook

Photoreceptors have been of significant interest for exploitation in the field of optogenetics due to the unprecedented spatio-temporal control that can be achieved on physiological processes through the use of light. This has led to significant progress in the field of neuroscience as optogenetics has allowed the study of neural networks and brain function through controlled activation and inhibition of specific neurons. In biotechnology applications, the optogenetics approach can be used to replace expensive chemical inducers that do not allow for fine temporal control of gene expression.[1-4]

An attractive target for use in optogenetics applications are the phytochromes, which are found in plants, bacteria, fungi and algae, and are typically known to be red and far-red light absorbing photoreceptors.[1,5] However, a recent study reported a number of atypical phytochromes that have broad spectral properties and absorb across the entire visible spectrum.[6] This opened up the possibility of using these atypical phytochromes to build a toolkit of broad-spectrum optogenetic parts that can be coupled with multiple different output / physiological responses. The modular nature of phytochromes makes them ideal for being refactored in a 'mix and match' manner. This is especially interesting for applications in metabolic engineering and synthetic biology where the use of different wavelengths of lights could be used for differential control and fine-tuning of different parts of a biosynthetic pathway.

In the present work Dten and Npf phytochromes were two atypical phytochromes that could be successfully expressed, purified and shown to photoconvert between 2 states in initial spectroscopic characterisation experiments. A range of biophysical techniques, including time-resolved visible and IR spectroscopy, time resolved X-ray scattering and cryo-trapping studies were used to probe the mechanism of these photoreceptors to unravel how the photocycles compared to those of the typical red / far-red phytochromes that have been previously published in the literature. Spectrally distinct intermediate states were captured using time resolved spectroscopy over a range of timescales from femtoseconds to seconds. Cryo-trapping and X-ray scattering was used to identify intermediates that correlate with large-scale protein dynamics, whereas time-resolved IR spectroscopy was used to map the nature of chemical and bond changes occurring in the bilin chromophore on faster

timescales. These studies helped stitch together a detailed understanding of the photocycles of Dten and Npf phytochromes, which showed many similarities to those observed previously for the typical red-far-red phytochromes.[7-13]

This insight of the mechanism of these proteins showed that while Dten phytochrome switched between two states relatively fast, the Npf phytochrome had a more complicated photocycle involving a slower thermal component. It was envisaged that this could be leveraged in optogenetic applications to provide different levels of control (e.g. instant on - off vs. titrated switch). In recent years there have been several reports of phytochromes being re-purposed for optogenetic applications.[14-18] In this work we exploited the modular nature of phytochromes to build a 'mix and match' toolkit, which allows a number of wavelength sensitivities to be coupled to a range of output domains. We built a range of chimaeric phytochromes by coupling the PCMs of Dten and Npf phytochrome with FixL and Cph HK. *In-vitro* radiolabelling studies revealed that Npf Cph HK and Dten Cph HK chimaeras were active. However, as the target promoters of the Cph1 response regulator, Rcp1, are unknown, we could not proceed ahead into *in-vivo* studies with these. In contrast, the Npf Fixl HK chimaera revealed *in-vitro* activity and was taken into initial *in-vivo* tests with a GFP reporter, which showed promising results. The activity achieved here was of a similar level to that reported for the Cph1 – OmpR (optimised to 8.9 fold activation)/ CcaS – CcaR (optimised to 6.4 fold activation) and UirS – UirR system (4.4 fold activation). [17,19] The activity of the chimeras can be improved by optimising the downstream promoters.

Further studies will be required to optimise these novel chimaeric systems prior to their routine application for optogenetic purposes. For example, it may be necessary to experiment with different intensities of light, exploit light/dark cycles rather than continuous illumination or only start illuminating cultures at specific ODs. Ultimately, the aim would be to exploit these chimaeric phytochromes for the optogenetic control of entire biosynthetic pathways. With that in mind, in parallel to building chimaeras, preliminary work was started to design plasmid systems for the optogenetic control of the linalool biosynthesis pathway, which had previously been constructed in-house.[20] The plasmids were designed by refactoring the CcaS – CcaR system from the published red-

green Cph1 - EnvZ HK / CcaS – CcaR systems.[14] The logic here was that it was hoped that these more established systems could be used as a benchmark while the chimaeras were being constructed.

Overall, we have demonstrated the feasibility of this approach and shown that it is possible to produce functional chimaeras, which could then potentially be used for the optogenetic control of gene expression. Such optogenetic control of biosynthetic pathways has recently been demonstrated for regulation of biosynthesis of methionine biosynthesis in *Escherichia coli (E. coli)* by the use of CcaS CcaR system to control methionine synthase, MetE, production. MetE catalyses the final rate limiting step of methionine synthesis, which affects the culture growth.[18] Hence, there is precedent for using naturally occurring photoreceptor proteins for the optogenetic control of biosynthetic pathways, in a similar manner to the one described above for linalool production. Ultimately, the Npf FixL HK chimaeric phytochrome (and potentially others) would add to the use of these naturally occurring systems, and allow an optogenetic toolkit to be established for the light-dependent control of biosynthetic pathways for biotechnology applications.

Supplementary Information – future work

To prepare for demonstration of the optogenetic tools with biosynthetic pathways, the pSR43.6 and pSR58.6 plasmids (figure 5.1) used by the Tabor group were refactored onto the linalool synthesis system constructed in-house (plasmids shown in figure 5.2).[14] This would serve as a reference for comparison of the optogenetic tools that were engineered as part of this study. As shown in figure 5.1, the CcaR and its cognate promoter, PcpcG2-172, are encoded on the pSr58.6 plasmid, and the bilin biosynthesis genes, *ho1* and *pcyA,* along with CcaS are encoded on the pSR43.6 plasmid. The in-house linalool synthesis system was also encoded on two plasmids. In total, 7 genes encoding enzymes that are required for the production of a linalool precursor are placed on the pMVA plasmid. Geranyl-pyrophosphate synthase (GPPS) and linalool synthase (LinS) genes are encoded on the pBb_O plasmid, and both genes are under the control of a TetR/ TetA promoter.
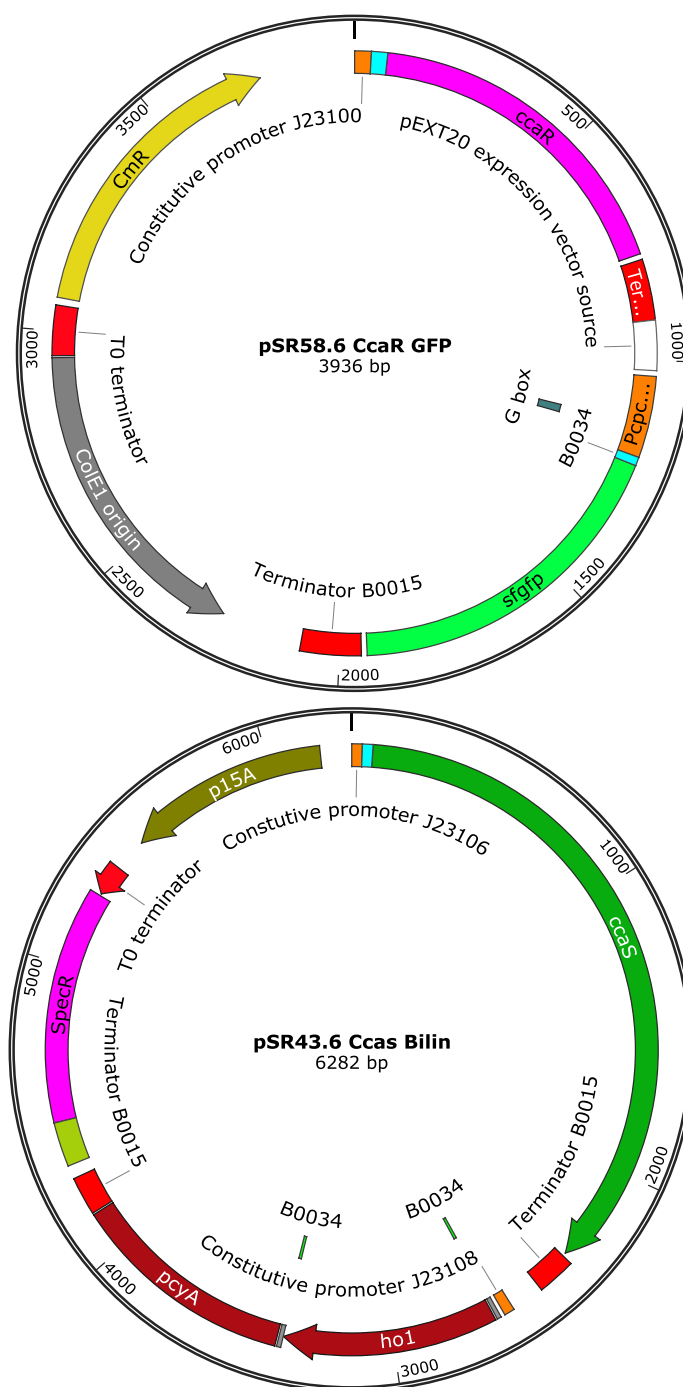
Figure 5.1. Published CcaS – CcaR system plasmids. pSR58.6 plasmid carries the gene encoding CcaR and its cognate promoter, along with chloramphenicol resistance and a ColE1 origin of replication. All genes other than the gene encoding GFP, which was under PcpcG2-172 control, were under the control of constitutive promoters. pSR43.6 plasmid carries genes encoding CcaS and those required for bilin biosynthesis, *pcyA* and *ho1,* along with spectinomycin resistance and a p15a origin of replication.

Figure 5.2. Linalool biosynthesis system. pMVA carries genes that encode enzymes for the synthesis of GPP, along with kanamycin resistance and a p15a origin or replication. pBb_O carries genes encoding GPPS and LinS, along with ampicillin resistance and a pBBR1 origin of replication

We aimed to control the production of linalool in an optogenetic manner by placing the *linS* gene under the control of the CcaR cognate promoter, PcpcG2-172. Primers were designed to move the entire CcaS – CcaR system, including the bilin biosynthesis genes, onto a refactored pBb_O plasmid to form

pBb_CcaS, CcaR as shown in figure 5.3. Although this work is still on-going, the majority of genes have now been cloned into this new plasmid and it is envisaged that this work can be continued by using this plasmid along with the pMVA plasmid to test the optogenetic control of linalool synthesis.
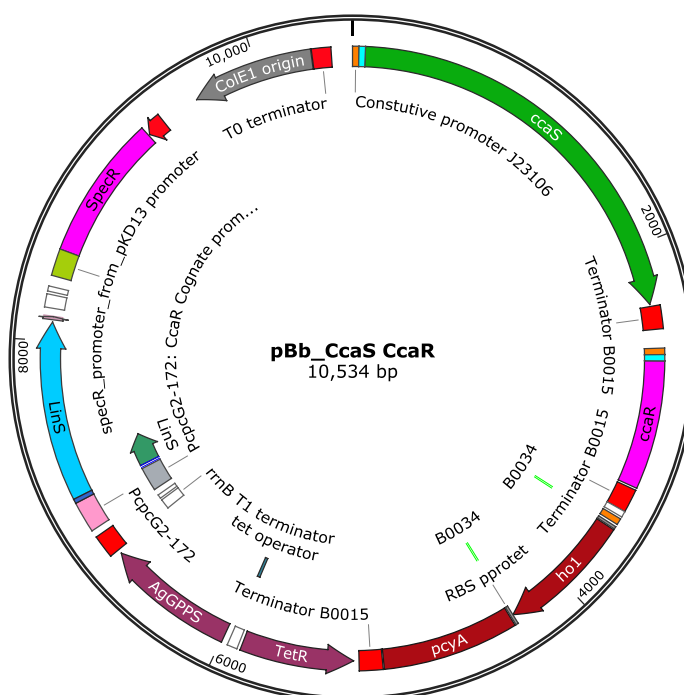

Figure 5.3. pBb_CcaS – CcaR plasmid. CcaS - CcaR system refactored onto the linalool systhesis pBb_O plasmid. CcaS - CcaR system refactored onto the pBb_O plasmid and change of origin of replication to ColE1. The linalool synthase gene, *linS*, was placed downstream of the PcpcG2-172 promoter.

Reference List:

[1]   N. C. Rockwell, Lagarias, *Chem. Eur. J. of Chem. Phys.* **2010**, *11*, 1172–1180.

[2]   F. Zhang, J. Vierock, O. Yizhar, L. E. Fenno, S. Tsunoda, A. Kianianmomeni, M. Prigge, A. Berndt, J. Cushman, J. Polle, et al., *Cell* **2011**, *147*, 1446–1457.

[3]   L. L. Looger, *Nat. Methods* **2011**, *9*, 43–44.

[4]   Y. Nihongaki, F. Kawano, T. Nakajima, M. Sato, *Nat. Biotechnol.* **2015**, *33*, 755–760.

[5]   N. C. Rockwell, Lagarias, *Plant Cell* **2006**, *18*, 4–14.

[6]   N. C. Rockwell, D. Duanmu, S. S. Martin, C. Bachy, D. C. Price, D. Bhattacharya, A. Z. Worden, Lagarias, *Proc. Natl. Acad. Sci. U.S.A.* **2014**, *111*, 3871–3876.

[7]   F. Andel, Lagarias, R. A. Mathies, *Biochemistry* **1996**, *35*, 15997–16008.

[8]   H. Foerstendorf, T. Lamparter, J. Hughes, W. Gärtner, F. Siebert, *Photochem. Photobiol.* **2000**, *71*, 655–661.

[9]   A. Remberg, I. Lindner, T. Lamparter, J. Hughes, C. Kneip, P. Hildebrandt, S. E. Braslavsky, W. Gärtner, K. Schaffner, *Biochemistry* **1997**, *36*, 13389–13395.

[10]   W. Rüdiger, F. Thümmler, E. Cmiel, S. Schneider, *Proc. Natl. Acad. Sci. U.S.A.* **1983**, *80*, 6244–6248.

[11]   C. Schumann, R. Gross, N. Michael, T. Lamparter, R. Diller, *Chem. Eur. J. of Chem. Phys.* **2007**, *8*, 1657–1663.

[12]   P. Schmidt, T. Gertsch, A. Remberg, W. Gärtner, S. E. Braslavsky, K. Schaffner, *Photochem. Photobiol.* **2008**, *68*, 754–761.

[13]   E. S. Burgie, R. D. Vierstra, *Plant Cell* **2015**, *26*, 4568–4583.

[14]   S. R. Schmidl, R. U. Sheth, A. Wu, J. J. Tabor, *ACS Synth. Biol.* **2014**, *3*, 820–831.

[15]   J. J. Tabor, A. Levskaya, C. A. Voigt, *J. of Mol. Biol.* **2011**, *405*, 315–324.

[16]   A. Levskaya, O. D. Weiner, W. A. Lim, C. A. Voigt, *Nature* **2009**, *461*, 997–1001.

[17]   P. Ramakrishnan, J. J. Tabor, *ACS Synth. Biol.* **2016**, *5*, 733–740.

[18]   A. Milias-Argeitis, M. Rullan, S. K. Aoki, P. Buchmann, M. Khammash, *Nat Commun.* **2016**, *7*, 12546.

[19]   S. R. Schmidl, R. U. Sheth, A. Wu, J. J. Tabor, *ACS Synth. Biol.* **2014**, *3*, 820–831.

[20]   N. G. H. Leferink, A. J. Jervis, Z. Zebec, H. S. Toogood, S. Hay, E. Takano, N. S. Scrutton, *ChemistrySelect* **2016**, *1*, 1893–1896.