Dipl.-Ing. Msc. Geoffrey Gourinchas

# Functional Cross-talk of Sensor and Effector Modules in Phytochrome Activated Diguanylyl Cyclases

## DISSERTATION

zur Erlangung des akademischen Grades

Doktor der Naturwissenschaften

eingereicht an der

## Technischen Universität Graz

Betreuer

Univ.-Prof. Mag. rer. nat. Dr. rer. nat. Peter Macheroux

Institut für Biochemie

Graz, August 2018

## EIDESSTATTLICHE ERKLÄRUNG

Ich erkläre an Eides statt, dass ich die vorliegende Arbeit selbstständig verfasst, andere als die angegebenen Quellen/Hilfsmittel nicht benutzt, und die den benutzten Quellen wörtlich und inhaltlich entnommenen Stellen als solche kenntlich gemacht habe. Das in TUGRAZonline hochgeladene Textdokument ist mit der vorliegenden Dissertation identisch.

| 01.08.2018 | |
| --- | --- |
| Datum | Unterschrift |

*"Nothing in biology makes sense, except in the light of evolution."*

Dobzhansky T. Biology teacher and geneticist.

# Aknowledgments

Three years of working on photoreceptors went extraordinary fast. I acquired a real sympathy for this field of research and I am going to miss working in dark conditions to study those astonishing molecular machines. Digging through all the scientific contributions in the field I had the opportunity to make during these three years, I can only be very thankful to all the persons whose have from direct or more indirect ways contributed to these achievements:

First of all, I am more than grateful to Andreas Winkler for trusting my capacities to work on this project. At an important corner of my career his trust and support allowed me to start this exciting PhD project which constituted three great years of my life. Not only for having taken care of my scientific education but also for all the advices that he shared with me for my professional development, I thank him very much. I want to thank him, but equally my colleague Stefan Etzl very much, for all those moments of trying to fix stuff in pitch dark conditions, in front of weird spectra and crappy HPLC baselines, at the synchrotron during the night, between frustrated enthusiasm to add new open questions and real excitement to solve some of them, these moments will stay in my mind forever. I would like to thanks the two bachelor students, Ursula Vide and Sara Trstenjak, who worked so hard on this project with me. Finally, what would be our research without a technical support?? I thank Elfriede Zenzmaier very much for all the technical help she provided during these three years but also for her constant energy and enthusiasm.

I equally want to thanks Peter Macheroux, for have taken the responsibility to officially supervising my PhD thesis and providing me with useful comments all along these three years, but also for constantly maintaining such a nice working atmosphere at the institute. In this regards, I would like to thank the entire Macheroux's group, all students who shared the lab facilities with us, and also Tinkara Kristovic of the secretary whom definitively have contributed to the nice working atmosphere at the institute. I thank also Karl Gruber and Stefan Schild for their nice implications in the yearly thesis committee meetings and their helpful comments on my research. I specially thank Karl Gruber to have allowed me to use the protein crystallization facility of the Karl-Franzens University of Graz. I must also thank all the persons that collaborated with us on the project, I think about David Buhrke, Peter Hildebrandt, Tobias Madl, Christoph Göbl, Christian Gruber, and others.

Also very important during these three years was the constant support from my family, my girlfriend Lilith Burger and her family. It is never an easy thing to be away from Home, but I can say now that I found in Austria a second home. Their support and encouragements truly helped me to face the hard things of doing a PhD.

Limoges, France, 29 march 2018

# Contents

# List of Figures

# List of Tables

# List of Abbreviations

3D: three dimensional

ADP: Anisotropic Displacement Parameter

ATP: Adenosine Triphosphate

BLAST: Basic Local Alignment Search Tool

BLUF: Blue-Light Using FAD

BRET: Bioluminescence Resonance Energy Transfer

BV: Biliverdin

CBCR: Cyanobacteriochromes

CCD: Charge Coupled Device

c-di-GMP: bis-3',5'-Cyclic diguanylic acid

Cryo-EM: Cryo-electron microscopy

DESY: Deutsches Elektronen-Synchrotron

DGC: diguanylyl cyclase

DNA: Desoxyribonucleic acid

DSSP: Define Secondary Structure of Proteins

DTE: Dithioerythritol

ESI: Electrospray ionization

ESRF: European Synchrotron Radiation Facility

FRET: Fluorescence Resonance Energy Transfer

FTIR: Fourier-Transform Infrared Spectroscopy

GAF: cGMP phosphodiesterase–Adenylyl cyclase–FhlA

GTP: Guanosine triphosphate

HDX: Hydrogen Deuterium Exchange

HOOP: Hydrogen atom Out-Of-Plane rocking

HPLC: High-Pressure Liquid Chromatography

LB: Luria-Bertani

LED: Light-Emitting Diode

LOV: Light-Oxygen-Voltage

MAD: Multi-wavelength Anomalous Diffraction

MCP: Methyl-accepting chemotaxis protein

MeOH: Methanol

MLHL: Maximum Likelihood target

MS: Mass Spectrometry

NCS: non-crystallographic Symmetry

ND: Neutral Diffusion

NMR: Nuclear Magnetic Resonance

NTE: N-terminal

PadC: Phytochrome activated diguanylyl Cyclase

PAS: Per-ARNT-Sim

PCB: Phycocyanobilin

PDE: Phosphodiesterase

PEG: Polyethylene Glycol

pH: Potential hydrogen

PHY: Phytochrome associated

PIF: Phytochrome Interacting Factor

PPi: Pyrophosphate

pppGpG: Pentacoordinate Phosphoryl transition state

PSM: Photosensory Module

PΦB: Phytochromobilin

RMSD: Root Mean Square Deviation

SAD: Single-wavelength Anomalous Diffraction

SAXS: Small Angle X-ray Scattering

SDS: Sodiun Dodecyl Sulfate

TCA: Trichloroacetic acid

TEV: Tobacco Etch Virus

TLS: Translation/Libration/Screw

TOF: Time-of-flight

TROSY: Transverse Relaxation-Optimized Spectroscopy

UV: Ultraviolet

Vis: Visible

YESCA: Yeast Extracted Casamino Acids

*ZZEssa*: 5Z*syn*, 10Z*syn*, 15E*anti*

*ZZZssa*: 5Z*syn*, 10Z*syn*, 15Z*anti*

# Abstract

The development of organisms is constantly modulated by intricate regulatory networks. Such biological networks include an array of interacting components that monitor changes in environmental condition to trigger physiological adaptations. Light is a ubiquitous actuator which allosterically affects many lifestyle decisions in all kingdoms of life. The perception of light is allowed by a collection of modular proteins, called photoreceptors, adapted to the broad visible spectrum of light. Phytochromes, which respond to red and far-red light, have attracted special attention due to the deep tissue penetration of red light. Recently, they have been extensively used as building blocks for the generation of near-infrared based optogenetics tools to control biological process *in vivo* with high spatiotemporal precision. However, the complexity behind the light signal transduction mechanism in phytochromes remains largely unexplained and precludes the rational design of innovative phytochrome-modulated biological tools. In order to characterize the mechanism of long-range signal transduction in phytochrome, the focus of this thesis was set on naturally occurring Phytochrome activated diguanylyl Cyclase (PadC) to better understand the structural requirements of enzymatic effector modulation by phytochrome sensors based on naturally evolved systems. We found that fine-tuned conformational dynamics between the various functional domains constituting full-length PadCs cause a broad diversification of photoresponse. Supported by new mechanistic insights, the sensor-effector linker region which bridges the two functional modules is presented as a critical actor involved in light signal transduction and modulation of the diguanylyl cyclase (DGC) effector. The 3D structure of a PadC homolog in the dark-adapted state as well as in a signaling active conformation revealed the complexity of the molecular cross-talk between the phytochrome and the DGC that in the case of PadC appears to require structural asymmetry in the dimer. Single protomer activation allows the signal transduction to the DGC effector, thereby highlighting asymmetric phytochrome homodimer as a potential evolutionary differentiation of bacteriophytochromes to adapt to the modulation of specific effector domains. Beyond the apparent similarity of global photoactivation mechanisms between bacteriophytochromes and other phytochromes subfamilies, we show that phytochromes have evolved a fine-tuning of specific interactions between functional domains allowing an astonishing diversification of photoresponses and regulation of various effectors.

# Kurzusammenfassung

Die Entwicklung von Organismen wird ständig von komplexen Regulationmechanismen moduliert. Diese beinhalten eine Reihe von interagierenden Komponenten, welche Änderungen der Umgebungsbedingungen verarbeiten, um physiologische Anpassungen auszulösen. Licht ist ein allgegenwärtiger Aktuator, der unterschiedliche Adaptionen der Lebensweise in allen Bereichen der Natur ermöglicht. Die Wahrnehmung von Licht wird von einer Reihe modularer Proteine ermöglicht. Diese sogenannten Fotorezeptoren sind an das breite sichtbare Spektrum des Lichts angepasst. Phytochrome, welche auf Licht im roten und nahe infraroten Bereich reagieren, haben aufgrund des Eindringens von Rotlicht in Tiefengewebe besondere Aufmerksamkeit erregt. In letzter Zeit wurden sie intensiv als Bausteine für die Herstellung optogenetischer Werkzeuge im nahen Infrarot Bereich genutzt, um biologische Prozesse mit hoher zeitlicher und räumlicher Präzision *in vivo* zu kontrollieren. Allerdings sind die komplexen Prozesse der Weiterleitung der Lichtsignale innerhalb von Phytochromen immer noch großteils ungeklärt, wodurch die rationale Planung von innovativen biologischen Werkzeugen, die mithilfe von Phytochromen reguliert werden, verhindert wird. Um den Mechanismus der Signalweiterleitung bei Phytochromen gekoppelten enzymatischen Effektoren besser zu verstehen, wurde folgender Schwerpunkt gelegt: Natürlich vorkommende, von Phytochromen aktivierte Diguanylylzyklasen (PadC) wurden untersucht, um die strukturellen Erfordernisse von enzymatischer Effektormodulation durch Phytochrom-Sensoren zu verstehen. Dabei wurde herausgefunden, dass eine fein abgestimmte Dynamik zwischen den verschiedenen funktionellen Elementen, die ein PadC in seiner vollen Länge ausmachen, eine breite Diversifizierung der Lichtregulation des enzymatischen Effektors ermöglicht. Mithilfe der neuen mechanistischen Erkenntnisse wurde klar, dass der Sensor – Effektor Verbindungsbereich, der die zwei funktionalen Hälften koppelt, als kritischer Teil in der Weiterleitung von Lichtsignalen und der Modulation des Effektors von Diguanylylcyclase (DGC) agiert. Die 3D-Struktur eines PadC-Homologs unter Dunkelbedingungen sowie in einer aktivierten Form zeigt die Komplexität des molekularen Wechselspiels zwischen dem Phytochrom Sensor und des DGC Effektors, welcher im Fall von PadC eine strukturelle Asymmetrie im Dimer zu verlangen scheint. Die Aktivierung eines einzelnen Protomers ermöglicht die Signalweiterleitung zum DGC-Effektor. So werden asymmetrische Phytochrom-Heterodimer als potentielle evolutionäre Zweiteilung im Bereich der

Bacteriophytochrome hervorgehoben, die eine Modulation von spezifischen Effektor-Funktionen ermöglichen. Nicht nur die Ähnlichkeit der globalen Fotoaktivierungsmechanismen zwischen Bakteriophytochromen und anderen Untergruppen von Phytochromen wird gezeigt, sondern auch die Entwicklung der Phytochrome zu einer genauen  Abstimmung spezifische Interaktionen zwischen Funktionsbereichen. Dies ermöglicht eine erstaunliche Diversifizierung von Lichtreaktionen und die Regulierung eines selben Effektors.

# Publications

This thesis is based on the following research papers:

1. <u>Gourinchas G</u>, Etzl S, Göbl C, Vide U, Madl T, Winkler A, Long-range allosteric signaling in red light-regulated diguanylyl cyclases, *Science Advances*, **2017**, 3;3(3):e1602498

   DOI: 10.1126/sciadv.1602498

2. <u>Gourinchas G</u>, Heintz U, Winkler A, Asymmetric activation of a homodimeric red light regulated photoreceptor, *eLife* 7, e34815 (**2018**).

   DOI: 10.7554/eLife.34815

3. <u>Gourinchas G</u>, Vide U, Winkler A, Functional inter- and intra-protomeric molecular crosstalks involved in phytochrome photoresponse, **2018** *(in preparation)*

# Contribution report

Paper 1.    I contributed to the elaboration of the main project together with my supervisor. I notably modified the plasmid pT7-ho1 to establish our system of phytochrome and biliverdin chromophore co-expression in *E. coli*. I cloned, expressed, and purified various PadC homologs and characterized their spectroscopic properties as well as their capacities of DGC photoregulation by establishing an *in vivo* screening method based on previously published methods. I participated to the HPLC-based quantification of the DGC activity for part of the analyzed constructs. I crystallized the full-length *Is*PadC construct and solved the structure together with my supervisor. I also solved the structure of the PSMcc variant as well as the *Is*PadC structure with GTP bound. I recorded the HDX-MS analysis of *Is*PadC and PSMcc variant and analyzed the data together with my supervisor. I participated to the redaction of the manuscript, I wrote the materials and methods part and designing part of the pictures.

Paper 2.    I cloned, expressed, and purified all the *Is*PadC coiled-coil linker variants. I characterized their spectroscopic properties and analyzed the *in vivo* and *in vitro* DGC photoregulation capacities. I crystallized and solved the structure of *Is*PadC$^{Reg2}$. I recorded the HDX-MS analysis of *Is*PadC$^{Reg2}$ and analyzed the data. I made a major contribution to write the manuscript together with my supervisor. To complement the published data I included in this thesis the HDX-MS analysis of the *Is*PadC$^{\Delta442-477::SG}$ variant to support the structural asymmetry occurring upon light illumination of this variant. I also included preliminary refined structural information about the *Is*PadC$^{Reg2}$ GTP soaked structure to highlight the allosteric modulation of the DGC dimer conformation upon GTP binding. Finally I included UV/Vis spectra of *Is*PadC$^{Reg2}$ crystals measured at room temperature to observe the effect of light illumination on the *Is*PadC$^{Reg2}$ crystals.

Paper 3.    I co-designed the project with my supervisor and had a major role in the design of the synthetic constructs and the planning of the cloning, expression, and purification mainly performed by a bachelor student whose I supervised. I

characterized the spectroscopic properties and DGC regulation capacities of the generated constructs. I wrote the paper.

# 1. Introduction

## 1.1.    Photoreceptors in all kingdoms of life

At the end of the nineteenth century, W. Kühne proved for the first time that proteins, which he named *visual purple*, are located within the rod cells of the retina and have a central role in vision (1). A few years later, G. Wald determined the chemical nature of the chromophoric light-absorbing molecule of these proteins to be the 11-*cis* isomer of the derivative of vitamin A, called retinal (2). The holoprotein is able to absorb photons to trigger visual phototransduction by promoting the hyperpolarization of cell membranes and the release of transmitters. This biological system, called rhodopsin (3), was the first characterized protein *photoreceptor*. Nowadays, a plethora of various photoreceptor families have been described and characterized (4). Although for a long time they have been considered rather as peculiarities of phototrophic organisms, the discovery of red light (5), as well as blue light (6) photoreceptors in bacteria challenged the fact that only phototrophic organisms can sense light. From unicellular algae (7) and prokaryotic organisms (8) where photoreceptors modulate phototaxis and photophobic responses to light, to complex eukaryotic organisms where they modulate networks of photomorphogenesis in plants (9) and entrainment of circadian rhythms in animals (10), photoreceptors act as critical relay of ambient light changes to trigger metabolic adaptations in all kingdoms of life.

Photoreceptors frequently feature a structural assembly between a photosensory module (PSM) including the structural domains allowing the photochemistry of the chromophore to happen, and an output module which triggers a biological response following its stimulation by the photoactivated PSM. During evolution, photoreceptors have largely diversified in order to respond to the broad range of light wavelengths (11). As a consequence, photoreceptors are able to bind specific chromophore molecules in order to sense different light qualities from UV (12) to far-red light (13) regime (Fig. 1). This photochromic diversity makes them very attractive for biological applications by employing their PSMs as building blocks to generate light controlled biological functions (14, 15). However, the molecular signal transduction triggered by light, sometimes over very long distances, remains largely unexplained. This significantly limits the rational design of innovative light modulated biological tools.
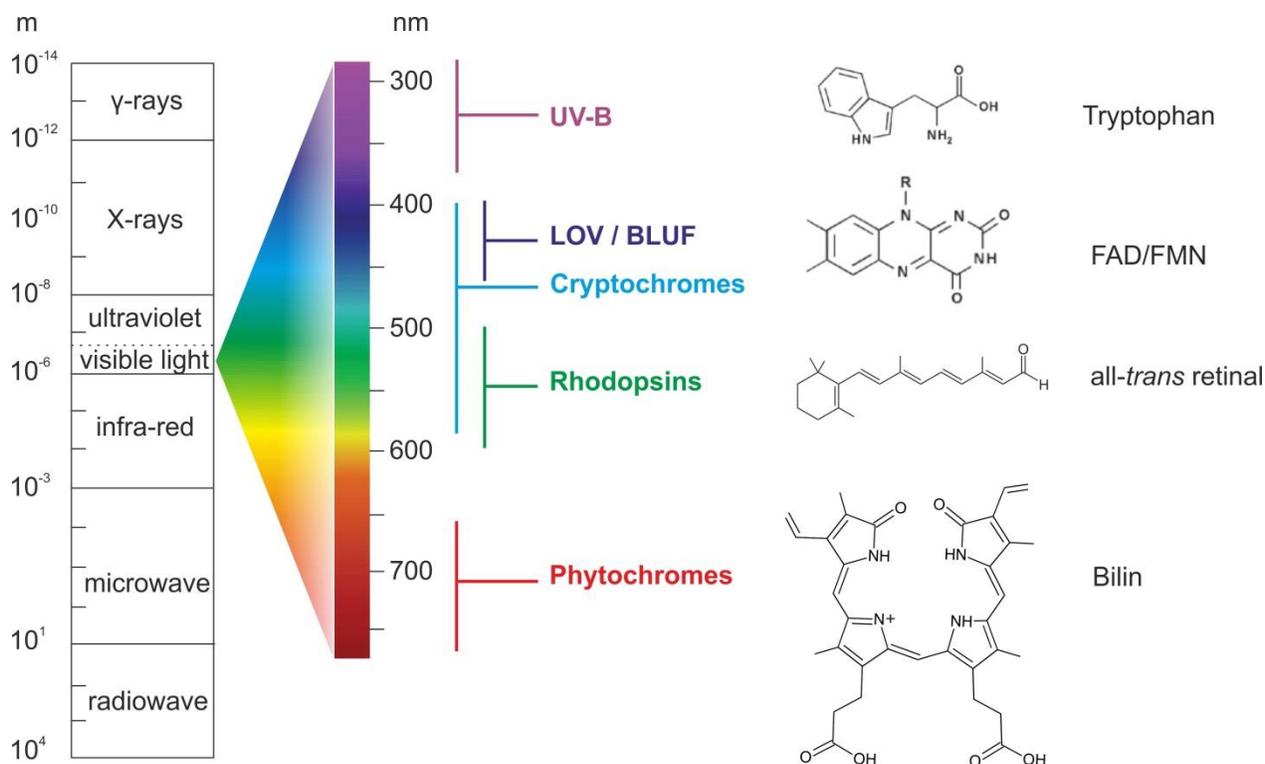
**Figure 1. Overview of photoreceptor families and their co-factors.** Photoreceptors are highly diversified within the visible spectrum of ambient light. Their photochromic diversity is essentially due to the binding of various small molecules responsible for light absorption properties. While UV sensitive photoreceptors can sense UV-B wavelength by their intrinsic tryptophan residues, other photoreceptors (e.g. phytochromes) require the binding of chromophoric molecules with larger conjugated systems.

Due to the deep tissue penetration property of red light, and the ubiquitous presence of biliverdin in mammalian cells, near-infrared based photoreceptors (e.g. phytochromes) recently raised interest for the high spatiotemporal control of biological events *in vivo*. Naturally occurring phytochrome and phytochrome-like sensors have evolved to couple red light detection to the regulation of a plethora of functions ranging from the allosteric regulation of enzymatic effectors (8, 16) to the modulation of gene expression by specific interactions with transcription factors (17) (Fig. 2).
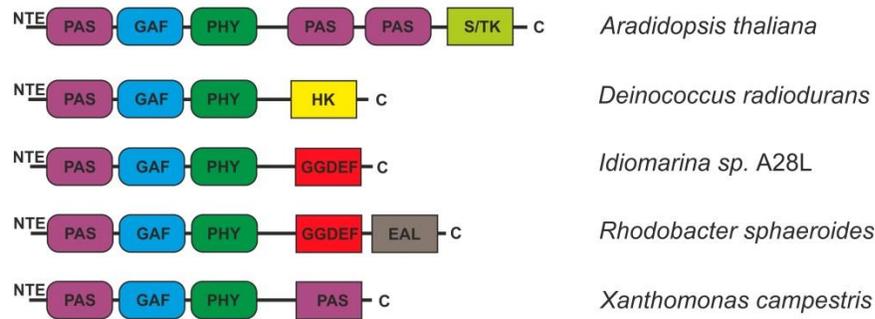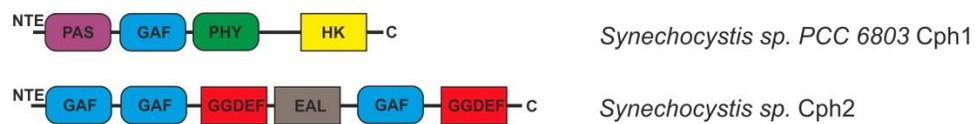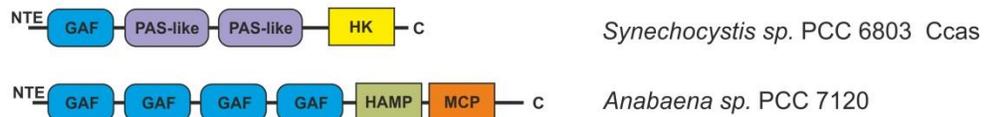
**Figure 2. Overview of various phytochrome subfamilies.** (**A**) Domains organization of phytochrome members of various subfamilies. Structural domains that belong to the phytochrome sensors are represented as rounded boxes while structural domains that belong to the effector domains are represented as straight boxes. For all of these phytochrome species, the bilin chromophore is bound within the GAF domain via a thioether linkage with a cysteine residue either located in the N-terminal extension of the PAS domain (NTE) or in the GAF domain. PAS, Per-ARNT-Sim; GAF, cGMP phosphodiesterase-Adenylyl cyclase-FhlA; PHY, Phytochrome associated domain; S/TK, serine/threonine kinase; HK, histidine kinase; GGDEF, diguanylyl cyclase domain; EAL, phosphodiesterase domain; HAMP, four-helical coiled-coil domain present in Histidine kinases, Adenylate cyclases, Methyl accepting proteins, and Phosphatases; MCP, Methyl-accepting chemotaxis protein.

Among phytochrome-regulated biological functions, the modulation of enzymatic effectors often involved in allosteric regulation of biological pathways, are promising targets to the development of optogenetics tools enabling the precise spatiotemporal modulation and study of these pathways in vivo (16, 18, 19). Interestingly, a closer look on diguanylyl cyclase effectors (GGDEF domains) reveals the astonishing natural modularity of coupling between a same effector and various photosensory domains and also light independent sensory domains (Fig. 3).
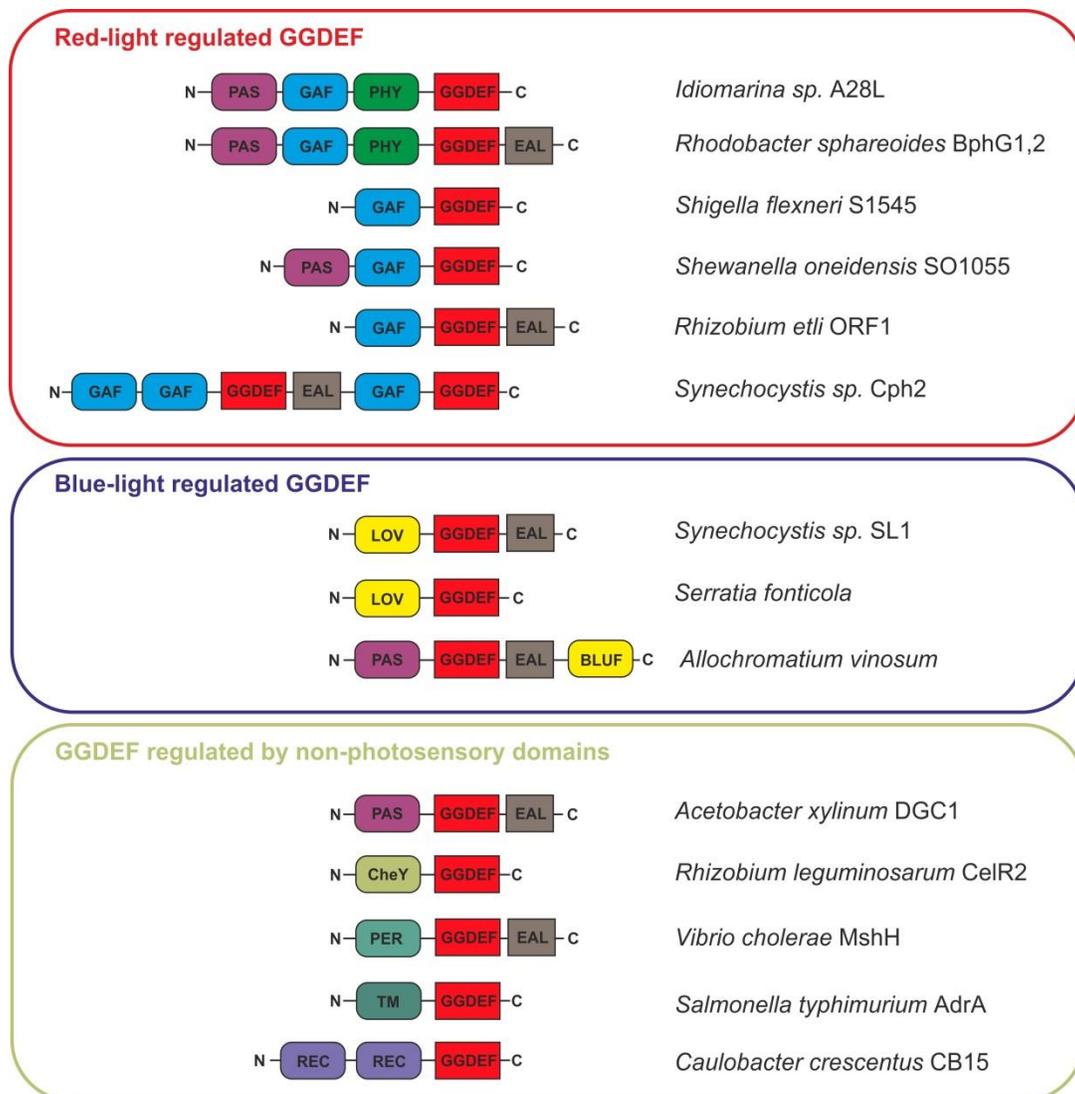
**Figure 3. Overview of sensory domains naturally coupled to GGDEF effector.** GGDEF domains responsible of diguanylyl cyclase activity are found naturally regulated by photosensory and non-photosensory domains. Red and blue light photosensory domains are the main classes of photoreceptors coupled to GGDEF regulation. LOV, Light-Oxygen-Voltage; BLUF, Blue-Light-Using-FAD; CheY, chemotaxis response regulator receiver domain; PER, periplasmic sensor module; TM, transmembrane domain; REC, phosphorylation receiver domain; EAL, phosphodiesterase domain.

In order to better understand how nature achieved to couple the modulation of a same effector domain to various sensor domains, we focused on naturally occurring Phytochrome – GGDEF couples, called in this thesis Phytochrome activated diguanylyl Cyclases (PadC), whose *Is*PadC shows in Fig. 3 is one member, which feature red light modulation of guanosine triphosphate (GTP) conversion to bis-(3'-5')-cyclic dimeric guanosine monophosphate (c-di-

GMP) (Fig. 4). From the biochemical and structural characterization of such naturally occurring systems we can better understand the structural prerequisites to achieve different photomodulation of DGC effector by phytochrome sensors.
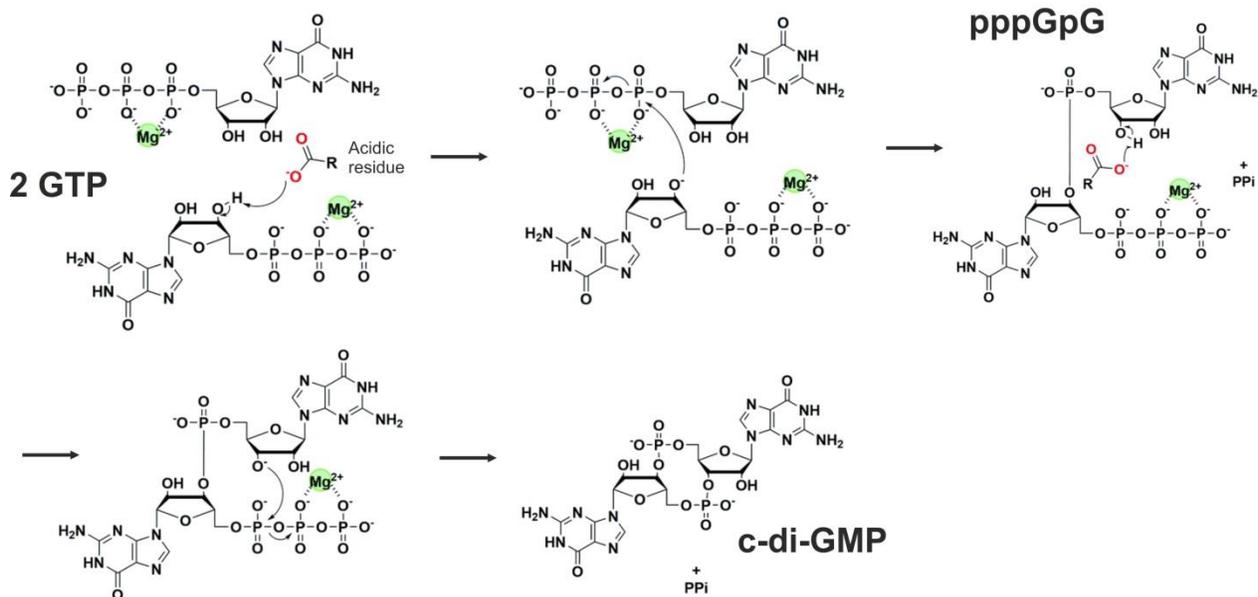


**Figure 4. General scheme of the mechanism of c-di-GMP synthesis.** Upon GTP binding into the active site of DGCs, an acidic residue of the DGC play the role of catalytic base and deprotonate the hydroxyl group of the ribose C3. A phosphodiester bond is then formed between the deprotonated hydroxyl group of the ribose C3 of one GTP and the deprotonated hydroxyl group of the α-phosphate group of the second GTP that is coordinated by a magnesium ion. This reaction forms the pentocoordinate phosphoryl intermediate state (pppGpG). The same reaction of phosphodiester bond formation occurs on the pppGpG intermediate to form the finale c-di-GMP molecule. The full reaction releases two molecules of pyrophosphate (PPi). Adapted with permission from Opoku-Temeng C. *et al.,*, 2016, (20), doi: 10.1039/c6cc03439j.

## 1.2.    The bilin chromophore senses the red light

Phytochromes and phytochrome-related proteins are able to sense red and far-red light by binding a specific linear tetrapyrrole co-factor derived from heme and belonging to the family of bilin molecules (21). The binding region is situated in a cGMP phosphodiesterase/adenylyl cyclase/FhlA (GAF) domain (22) belonging to the subfamily of bilin-lyase-specific-GAF in the case of phytochrome proteins (23). Mutagenesis studies revealed that a cysteine in this region is engaged in a covalent interaction with the bilin chromophore via a thioether linkage (24, 25).

Bacteriophytochromes (26) as well as some in cyanobacteriochromes conserve a different cysteine residue engaged in the adduct linkage with the bilin (27) in a typical N-terminal region while plants and cyanobacterial orthologs feature the adduct-forming cysteine within the GAF domain. Interestingly, the evolutionary separation of plant, cyanobacterial, and bacterial phytochromes have adapted their chromophore binding affinities to specific classes of bilin resulting in shifted UV/Vis absorption characteristics in the red light region of the visible light spectrum (5). Phytochromobilin (PΦB) is bound in plant phytochromes while phycocyanobilin (PCB) is preferred by some cyanobacterial phytochromes and unicellular algae, whereas biliverdin (BV), which is a precursor in the PCB and PΦB biosynthesis pathway (28), is preferentially bound by bacteriophytochromes (Fig. 5).
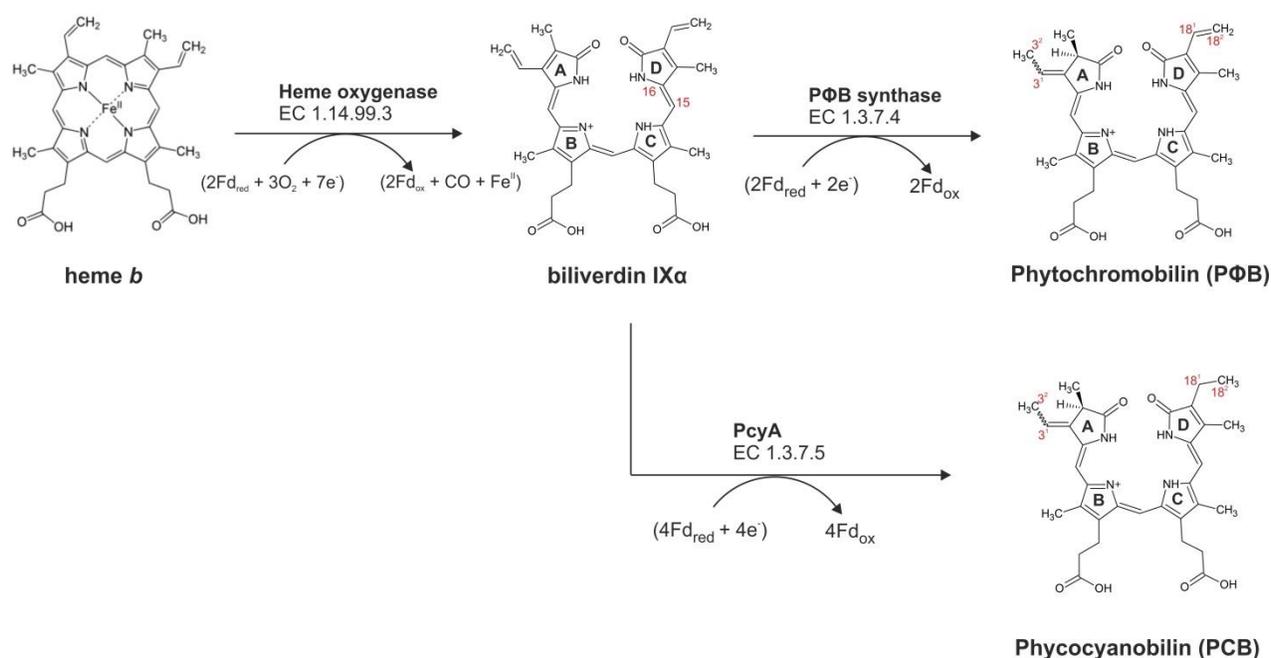


**Figure 5. Simplified scheme of the bilin pigment biosynthesis.** The initial reaction of the biosynthesis of bilins is the cleavage of the heme macrocycle usually regiospecifically at the α-*meso*carbon by the heme oxygenase to yield biliverdin IXα. This initial reaction requires seven electrons and three molecules of molecular oxygen and produces carbon monoxide (CO), free ferrous iron ($Fe^{II}$), oxidized ferredoxin ($Fd_{ox}$), and biliverdin. To produce Phytochromobilin (PΦB), the PΦB synthase catalyzes the reduction of the biliverdin A-ring $2,3,3^1,3^2$ diene system. The reaction requires two electrons and produces two oxidized ferredoxin and the phytochromobilin. To produce phycocyanobilin, PcyA catalyzes the successive reduction of the biliverdin $18^1,18^2$ double bond, that leads to the semireduced intermediate $18^1,18^2$-dihydrobiliverdin (not shown on this scheme for simplification) which is then further reduced at the A-ring $2,3,3^1,3^2$ diene system to form the final phycocyanobilin. The reaction requires the transfer of four electrons.

In bacteriophytochromes the BV adduct is formed between the Cys thiol and the $C3^2$ carbon of BV (29), while in PCB and PΦB the bond occurs between the Cys thiol and the $C3^1$ carbon (30) resulting in a difference in the chemical structure of the two chromophores upon covalent bonding reflected by a difference on the overall conjugated system. In addition to the fact that PΦB has an additional reduction of the $3^1,3^2$ bond, and PCB has two additional reductions of the $3^1,3^2$ and $18^1,18^2$ bonds respectively, PCB and PΦB adducts show a shift of their absorbance maximum around 630 and 665 nm respectively (Fig. 6). The molecular structure of phytochromes in the dark-adapted state (Pr-state) revealed a 5*Zsyn*, 10*Zsyn*, 15*Zanti* (*ZZZssa* or 15*Z*) configuration of the tetrapyrrole (31) with a protonated pyrroleninic nitrogen and the D-ring substantially tilted out of the plane formed by the A-B-C pyrrole rings. For all phytochrome molecular structures available nowadays, the chromophore is found buried in its binding pocket within the GAF fold and features electrostatic contacts with surrounding amino acids. The propionate groups of the ring B and C anchor the chromophore in its binding pocket with hydrogen bonding involving conserved polar residues (32–34).
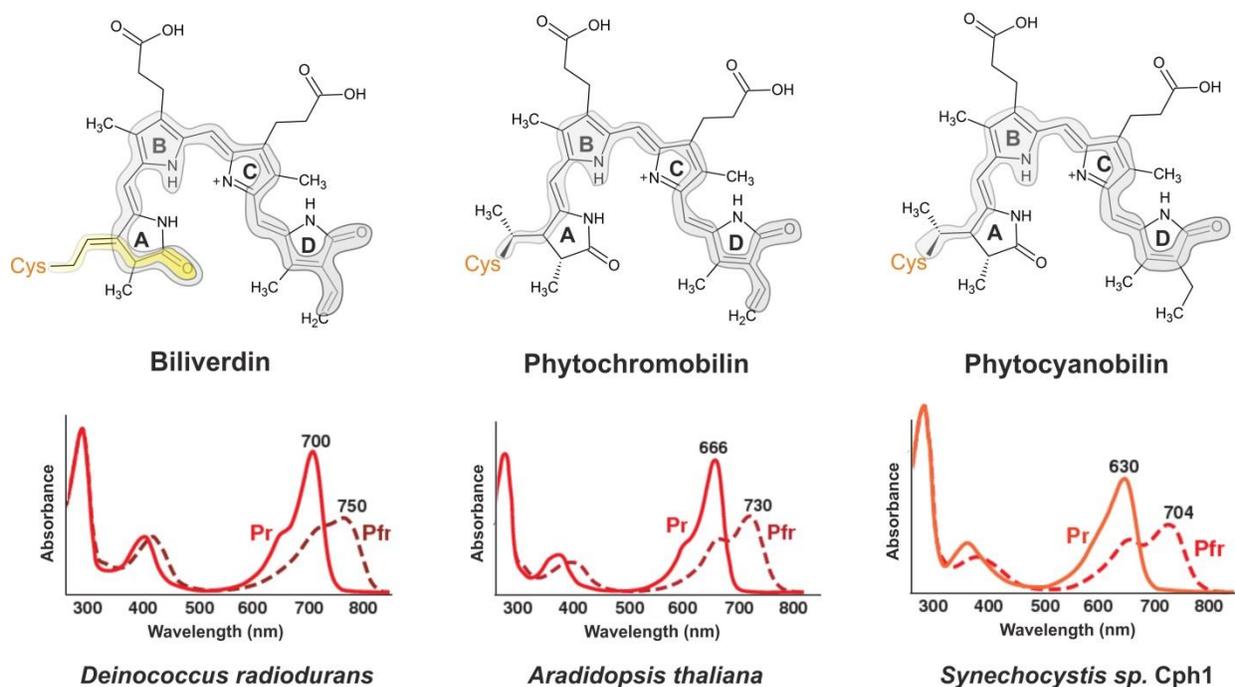


**Figure 6. Covalent linkage of bilin chromophores.** Cys represent the cysteine residue involved in the covalent linkage with the bilin chromophore. The double bond adjacent to the bilin carbon atom involved in the covalent linkage site became reduced during the coupling reaction. In biliverdin, this double bond is part of a short π-electron system (highlighted in yellow), but not part of the long π-electron system (highlighted in gray). Since the spectral properties are mostly determined by the long π-electron system, covalent adduct does not alter the absorbance maximum of biliverdin. However in PCB and PΦB, this double bond is part of the long π-electron system that leads

to a shift of the chromophore absorbance maximum to lower wavelengths. Adapted with permission from Lamparter T. 2004 (35), doi: 0.1016/j.febslet.2004.07.050.

Moreover, the bilin configuration is stabilized by hydrogen interactions involving the protonated pyrrole nitrogen and polar groups of adjacent residues within a highly conserved Pro-Ala-Ser-Asp-Ile-Pro (PASDIP) motif, and with a uniquely fixed water called pyrrole water assumed to assist the chromophore photoconversion (36). Upon light illumination and absorption of the energy from a photon of red light, the tetrapyrrole isomerizes along the C15=C16 double bond resulting in a 5*Zsyn*, 10*Zsyn*, 15*Eanti* (*ZZEssa* or 15*E*) configuration that ultimately leads to the rotation of the D-ring about 180° (37, 38). Following this local change, the chromophore environment, then the overall protein structure, rearrange to form the final Pfr-photoproduct (39). The Pfr-state conformation can then recover back to the Pr-resting state thermally or by means of far-red light illumination. Using recently solved structures of the dark and photoactivated states of bacteriophytochromes, the proposed mechanism of phytochrome spectral shift and thermal reversion is based on a sliding motion of the bilin chromophore in the GAF pocket and refolding of the typical PHY-tongue element. The net effect is a reorientation of the B and C rings propionates and different rotamerization of surroundings residues to accommodate the motion (34, 36, 40, 41). The route taken by this sliding motion depends of course on the various residues in the vicinity of the chromophore and leads to nuances in the photoconversion routes occurring in different phytochrome species. In bacteriophytochromes, studies based on time-resolved and low-temperature trapping vibrational spectroscopy techniques allowed to describe a simple set of intermediate states populated during Pr to Pfr photoconversion (Fig. 7). Upon illumination, the bilin chromophore in 15*Z* configuration in its protein environment forming the typical Pr-state isomerizes to the 15*E* configuration on a picosecond time scale. This leads to the photoreversible primary photoproduct of Pr called lumi-R (42). The lumi-R state is then converted to a second intermediate state called meta-R which can be subdivided in different meta-R species (meta-Ra, meta-Rc) featuring different protonation states of the pyrrole rings nitrogen and structural rearrangements in the surrounding protein environment (43). Finally, the meta-Rc state is converted to the finale photoproduct where the protein environment features a full Pfr-state (44).
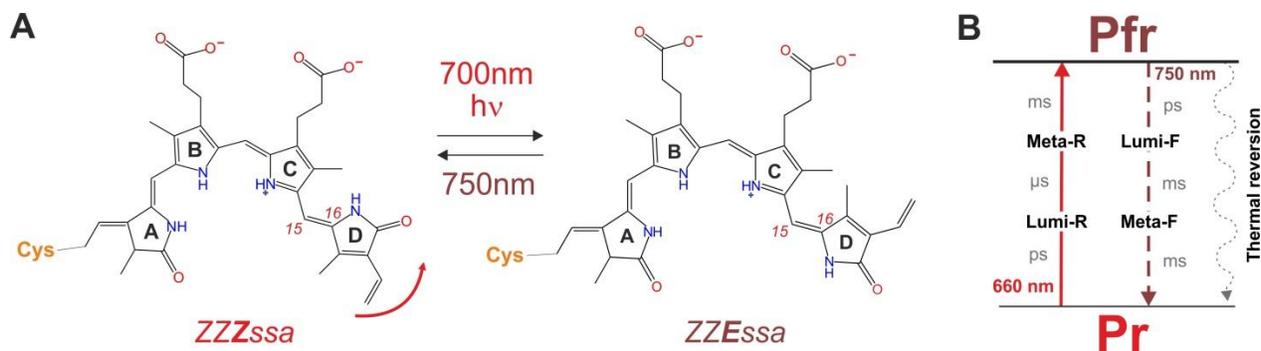
**Figure 7. Photoconversion in bacteriophytochromes.** (**A**) Photoisomerization of the C15=C16 double bond of the methine bridge situated between pyrrole rings C and D leads to a rotation of ~180° of the D-ring to form the 15*E* isomer of the covalently linked biliverdin. (**B**) Simplified schematic representation of Pr to Pfr photoconversion in bacteriophytochromes. Upon red or far red light illumination, different intermediate states of the chromophore environment are populated. 15*Z* to 15*E* isomerization of the BV chromophore happens at the picosecond time scale while structural rearrangements of the chromophore environment to form the final Pfr-photoproduct happen at the microsecond to millisecond timescales.

## 1.3. Red light influences important lifestyle decisions

Due to their high abundance in plants, phytochromes have been detected and extracted from plant tissues already at the end of the 19th century (22). In plants, phytochromes have been characterized in the regulation of the broad network of photomorphogenesis including the control of seed germination, stem/petiole/hypocotyl elongation, chloroplast biogenesis, leaf expansion, flowering time as well as shade avoidance response (9, 46, 47). In microorganisms, phytochromes have first been discovered in phototrophic bacteria where light sensing was assumed to help bacteria move towards ideal illuminated conditions to harvest solar energy (48). However, in the early 20th century, the successive discovery of phytochrome sensors in *Deinococcus radiodurans* and *Pseudomonas aeruginosa*, two non-phototrophic eubacteria, extended the use of environmental light to the surviving of chemotrophs organisms (5). Afterwards, several studies demonstrated the common presence of various photoreceptors families in non-phototrophic bacteria (48–51). One of the reason for chemotroph bacteria to sense light is to protect themselves from UV radiation damage (52) by triggering the production of protective pigments or their displacement out of the light source. For instance in *D. radiodurans,* the absorption of red and far-red light by a bacteriophytochrome regulates the synthesis of the

carotenoid deinoxanthin which protects the cell growth from detrimental effects of light (5). The property of phytochromes to photointerconvert between red and far-red light absorbing forms allows the phytochromes population to monitor constantly the light intensity and quality that is present in the microorganism environment.

Several characterized bacteriophytochromes feature a red light modulation of a histidine kinase output domain involved in a two-component phosphorelay system triggering the transfer of a phosphate from ATP to the histidine kinase domain (49, 53). The phosphate is then transferred to a response regulator which upon binding to DNA or a protein target affects a functional response by phosphorylation (54). However, many bacteriophytochromes have been reported to have no kinase outputs. Bacteriophytochrome from phototrophic bacteria *Bradyrhizobium sp*, *Rhodopseudomonas palustris,* or *Xanthomonas campestris* have non-enzymatic output domains but feature C-terminal PAS domains assumed to mediate protein-protein interactions that regulate expression of photosynthesis genes (55). The bacteriophytochrome found in *Xanthomonas campestris* for which the full-length molecular structure has been solved (56) shows a red light dependency of its virulence. Indeed, red light illumination of *X. campestris* attenuates its virulence and downregulates xanthan exopolysaccharide production and biofilm formation known as virulence factors (57). Some phytochromes are found coupled to signaling output involved in chemotaxis (e.g. modulation of HAMP (histidine kinases, adenylate cyclases, methyl binding proteins, phosphatases), and MCP (methyl-accepting chemotaxis protein) domains (58–64). HAMP and MCP domains feature helical motifs involved in the regulation of phosphorylation or methylation state of dimeric receptors. Their modulation by red light photoreceptors have been for instance shown for the cyanobacteriochrome *Sy*PixJ where motility driven by the type IV pilus is induced by red light (65). For photosynthetic organisms like cyanobacteria, the ability to position themselves in ideal conditions of light and nutriment is essential. Notably because too much light produces free radicals damaging for the survival of the cells (58).

Other bacteriophytochromes have been found coupled to GGDEF domain alone, or GGDEF in combination with EAL domain as red light regulated output modules (23, 32, 66–69). GGDEF domains are responsible for diguanylate cyclase (DGC) activity, and EAL domains are responsible for phosphodiesterase (PDE) activity that respectively synthetize and degrade c-di-GMP (70) (Fig. 8). This bacterial second messenger has been attributed as central modulator of

various cellular responses, ranging from motility, sessility, virulence, to cell cycle progression regulation. The exact implications of c-di-GMP for each biological output are still an active area of investigation (71). The presence of regulated GGDEF and EAL domains in bacteria has often been attributed to the control of transitions between surface attachment and detachment that reflect the transition from single cellular usually motile, to multicellular usually sessile bacterial lifestyle. These multicellular bacterial communities produce an extracellular matrix, called biofilm, composed of excreted polysaccharides, proteins and DNA (72). Bacterial biofilms are at the origin of a major health issue as they can form on medical devices, such as catheters or implants that might cause chronic infections difficult to treat (73). This is one of the reasons why precise regulation of nucleotidyl cyclases, like diguanylate cyclases, represent interesting pharmacological targets nowadays (74).



**Figure 8. C-di-GMP is an essential bacterial lifestyle modulator.** Diguanylyl cyclases are naturally coupled to various sensor domains that constantly monitor environmental cues to regulate important physiological responses via the production of c-di-GMP. Adapted with permission from Krasteva, P.V., *et al.*, 2012, (75) doi: 10.1002/pro.2093.


## 1.4.     Phytochrome structure and function

Bacteriophytochromes can generally be expressed in *Escherichia coli* in large quantities and can be purified with high yield. The *in vivo* co-expression of the appropriate bilin generated

by biosynthetic enzymes, or its addition in pure form after phytochrome purification, allow the reconstitution of the functional holoprotein. This advantage allowed many biophysical and structure-based analyses of phytochromes. In 2005, Wagner *et al.,* (31) were the first to solve the 3D structure of the biliverdin binding GAF domain of a phytochrome using the PAS-GAF core of *D. radiodurans* bacteriophytochrome. Two years after, Wagner *et al.,* (33) were able to solve a high-resolution structure of the same PAS-GAF core as a dimer. These structures and many others confirmed afterwards the predicted binding of biliverdin by a thioether linkage between a cysteine located upstream of the PAS domain and the $C3^2$ carbon of the A-ring vinyl group (24, 31, 33) (*cf.* Fig. 6 and Fig. 9). The rest of the open-chain tetrapyrrole is buried within the GAF binding pocket framed by a six-stranded β-sheet and two helices (Fig. 9). Opposite to the β-sheet a three-helix bundle is often involved in the formation of the PSM dimer interface. Interestingly, in the phytochrome super-family a strong sequence similarity is observed in the vicinity of the chromophore, and the main core of the PAS-GAF region is remarkably well structurally conserved. Consequently, the bilin chromophore is restrained in a mostly identical configuration (15*Z*) even among phytochromes with diverging chromophore linkage site. The interface between PAS and GAF domains is strengthened by a typical structural feature, called the figure-of-eight knot, which lassoes the N-terminal helix upstream of the PAS domain (31) (Fig. 9).
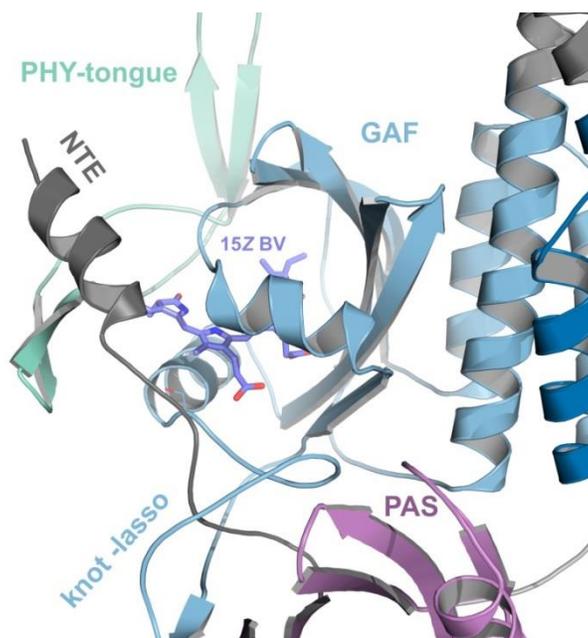


**Figure 9. Bacteriophytochrome chromophore binding site.** Close-up view on the chromophore binding site of the canonical bacteriophytochrome *Is*PadC of *Idiomarina sp.* A28L (PDB 5LLW (32)). NTE element, PHY-tongue region, GAF, and PAS domain are colored in gray, green, blue, and purple, respectively.

Usually in phytochromes, the GAF domain is directly followed by a single PHY domain. This domain features a central five stranded β-sheet with α-helices flanking both sides of the β-sheet. Aligned with the helical spine exiting the GAF domain, usually a long helical linker connects the PHY domain to the light regulated output module (Fig. 10). Due to the lack of full-length structures, structural details about the connection between the PSM and the sensor-effector linker, and between the sensor-effector linker and the output module are rather scarce (32, 56, 76, 77). Interestingly, the geometry of the succession of the PAS, GAF, and PHY domains within the phytochrome dimer and particularly the helical spine geometry shows a quite flexible dimer interface probably adapted to specific signaling responses as observed by comparison between the dark-adapted structures of various phytochrome dimers (32, 78, 79) (Fig. 10).
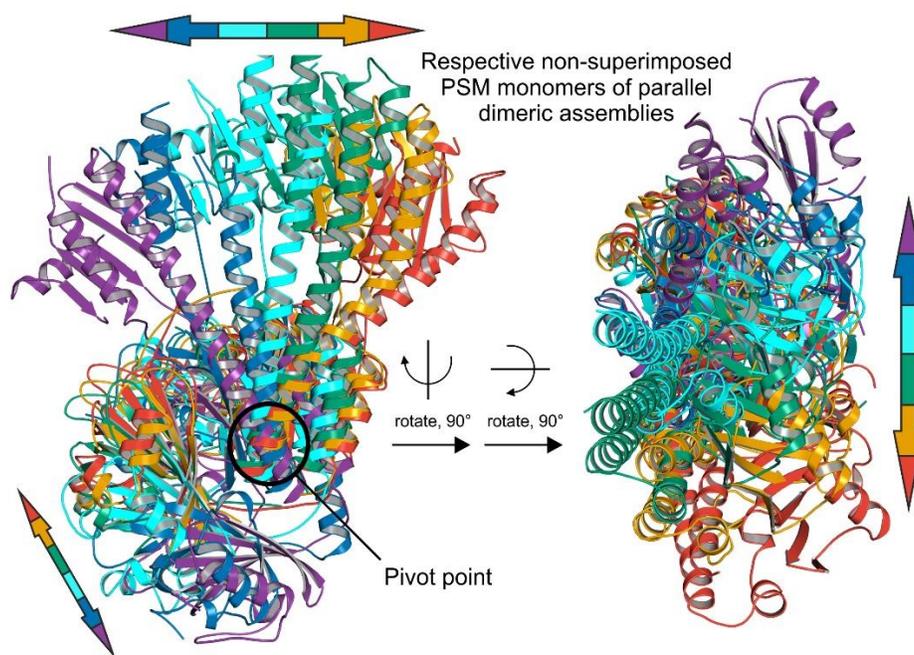


**Figure 10. Structural plasticity of phytochrome dimerization.** The structural plasticity of the central helical spine and consequently the overall parallel dimeric phytochrome assembly is even more pronounced when comparing the non-superposed monomers. In this case, monomers of PSM modules have been aligned to the PAS-GAF-PHY monomer of *Is*PadC chain A (cyan) (32) and the respective other monomers are displayed (rmsd of 3.2, 2.3, 2.2, 2.2 and 1.2 Å for PDB 3G6O_A (80) (violet), 4Q0J_A (34) (orange), 5AKP_B (56) (green), 5C5K_B and 4OUR_B (81) (blue), respectively). No clustering of the PHY domain orientation with respect to the $P_r$- or $P_{fr}$-state character of the tongue can be observed. However, the structural differences among various phytochrome dimers are clearly non-random and occur along a specific trajectory that corresponds to a hinge-like motion at the dimer interface. Interestingly, the two extremes of this hinge-like motion correspond to structures obtained for $P_{fr}$-state crystals (violet and red). In contrast, the only parallel phytochrome structures with adjacent C-terminal domains (green and

cyan) cluster in the middle of the overall trajectory. While this suggests that the more pronounced dimer rotation of other PSM assemblies might be due to missing interactions of their output modules and linker regions, the characteristic structural transition reflecting the plasticity of the PHY domain dimerization is very likely functionally relevant for phytochromes in general.


The PHY domain also connects the GAF domain near the chromophore binding site through a structural hairpin, called the PHY-tongue (30, 82). In canonical phytochromes featuring a Pr-state as ground state, the PHY-tongue shows two anti-parallel β-strands (Fig. 11). They are involved in highly conserved interactions near the bilin, including a salt bridge between an arginine of a conserved Pro-Arg-X-Ser-Phe (PRXSF) motif and the aspartate of the PASDIP motif, and hydrophobic interactions between the GAF domain and a conserved tryptophan of the PHY-tongue element (83). From the dark-adapted Pr-state structures of phytochromes, it appears that the β-strand region of the PHY-tongue are structurally similar as the main interactions stabilizing the conformation of the PHY-tongue come from the highly conserved PRXSF and WXG motifs interactions with the PASDIP motif of the GAF domain (29, 32, 39, 81, 83). However, the looped tip of the PHY-tongue element is much more variable in length and composition and one can observe very different sets of interactions between this part of the PHY-tongue and the NTE region in various phytochrome species (30, 32, 34, 81) (Fig. 11). Importantly, the previously observed flexibility at phytochromes PSM dimer interfaces (Fig. 10 and 11) is dependent of the degree of freedom of the PHY domains which might be related to the tuned interactions of the PHY-tongue element with the NTE region. The evolutionary adaptation of the PHY-tongue element and NTE region is most probably involved in the broad diversity of phytochrome photoresponses.
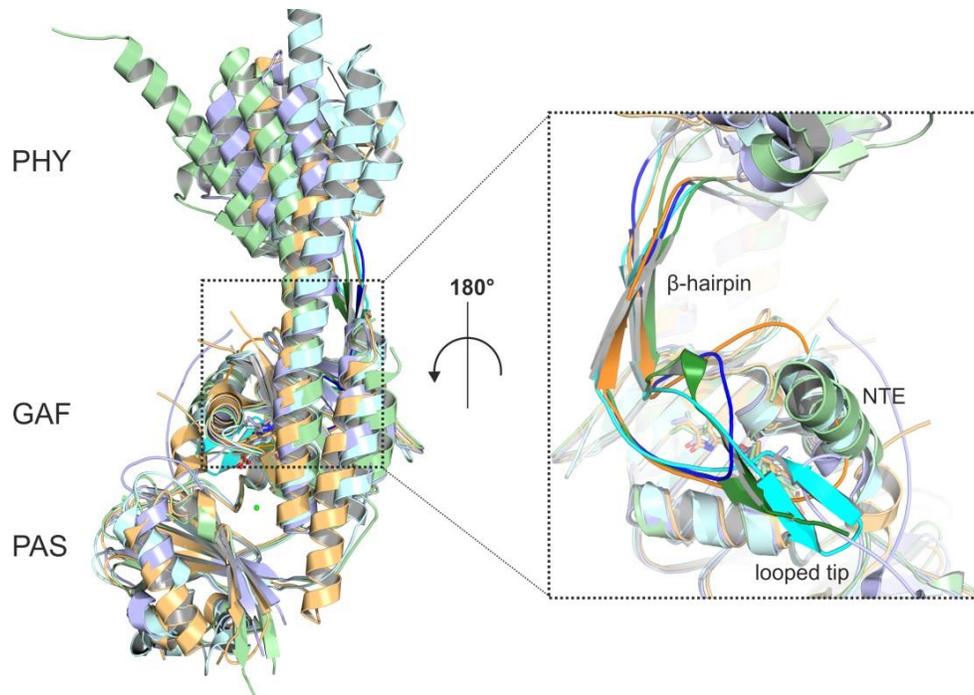
**Figure 11. Modularity of the PHY-tongue looped tip.** In the Pr-state, the β-hairpin structure of the PHY-tongue element is highly conserved in phytochromes. However, interactions between the PHY-tongue looped tip and the NTE region appear quite modular and influence the flexibility of the PHY dimer interface as observed for various phytochrome PSMs aligned based on the PAS-GAF core of one protomer (left side). *Is*PadC (PDB 5LLW (32)) is colored in cyan, *At*PhyB (PDB 4OUR (81)) is colored in orange, *Dr*BphP (PDB 4Q0J (34)) is colored in blue, and *Sy*Cph1 (PDB 2VEA (30)) is colored in green. Dark version of the colors is used to highlight the PHY-tongue region while light version of the colors is used for the rest of the PSM.

Strikingly, successively in 2008, 2012, and 2014, the determination of 3D structures of phytochromes PSMs revealed the critical role of the PHY-tongue element in the signal transduction mechanism (39, 82, 84). In these structures, instead of a two-stranded β-sheet the PHY-tongue element forms an α-helix. The net effect of this structural refolding is the modification of the PHY-tongue contact points with the GAF domain (Fig. 12). Especially the salt bridge between the PRXSF and PASDIP motif, described above in the text, is broken and replaced by the aspartate of the PASDIP motif interacting with the conserved serine of the PRXSF motif (Fig. 12). Hydrophobic contacts of the PHY-tongue β-strand and the GAF domain are also changed and a study based on the *Syn*Cph2 phytochrome proposed a tryptophan switch scenario whereby the hydrophobic interaction of the conserved tryptophan of the PHY-tongue with the GAF domain switch to another tryptophan or aromatic residue located on the opposite β-

strand of the PHY-tongue element (83) (e.g. W451 and Y472 for *Dr*BphP as described in Fig. 12). Various mutational studies on the PHY-tongue region highlighted the importance of this structural element for the proper photoconversion from Pr- to Pfr-states even with respect to its limited contacts in the vicinity of the chromophore (30, 34, 81, 83).
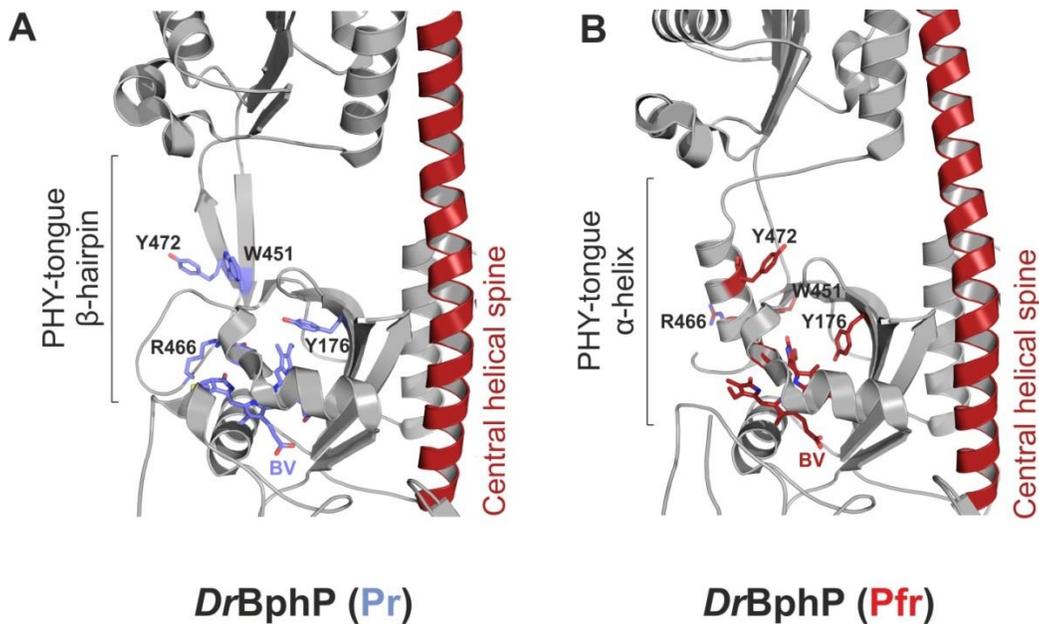


**Figure 12. The light signal transduction involves the refolding of the PHY-tongue element.** Close-up views on the PHY-tongue element from *Dr*BphP in Pr-state (PDB 4Q0J (34)) (**A**) and Pfr-state (PDB 5C5K (85)) (**B**), respectively. Important residues around the biliverdin chromophore are highlighted in sticks representation. Upon red light illumination, the β-hairpin conformation of the PHY-tongue is destabilized and one part of the PHY-tongue get unstructured while the other part refolds to an α-helix.

Although the dark-adapted state structure of various phytochromes are very precious to understand the light signal transduction, what is even more convincing are the paired structures of the dark-adapted and photoactivated states of phytochromes, like the recent study of Takala *et al.*(39) and Song *et al,*.(86), that allows a direct observation of the structural rearrangements occurring upon light activation in a specific phytochrome species (Fig. 13). Recent crystallographic (32, 56, 76), cryo-EM models (34, 87), and SAXS envelope of full-length phytochromes (88) revealed the extensive dimer interface in head-to-head phytochrome dimers that are especially important for effector domains that need to work in concert (32, 76, 88). Chromophore isomerization and subsequent refolding of the PHY-tongue β-hairpin to an α-helix appear to be a common driving force in phytochromes featuring an arrangement of PAS, GAF,

and PHY domains to modulate the dimer interface and trigger conformation rearrangements at the effector domain. However, head-to-tail dimers are also observed for phytochromes featuring output domains which are not necessarily acting as dimers (83, 84), or for phytochrome PSMs that are preferentially monomeric in solution (81, 84).
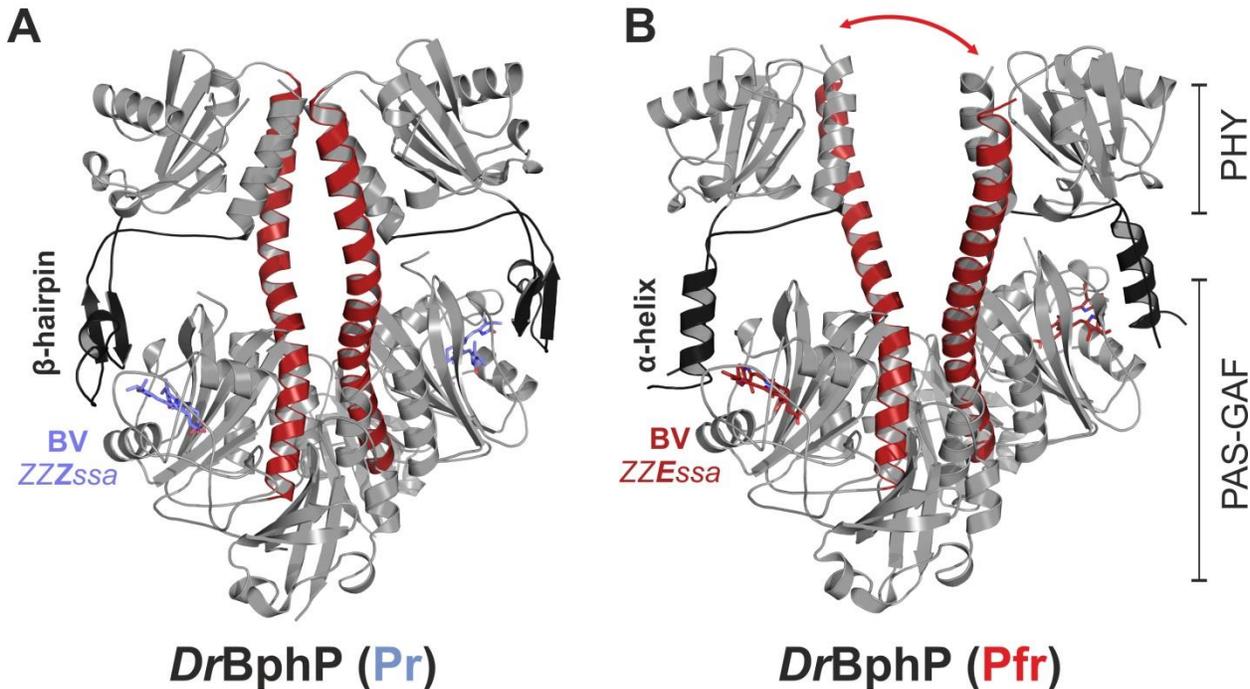


**Figure 13. Signal transduction in dimeric phytochrome photosensory module.** Structures of the dimeric phytochrome photosensory module of *Dr*BphP in the Pr-state (PDB 4QOJ (34)) (**A**), and in the Pfr-state (PDB 5C5K (85)) (**B**), respectively. Upon red light illumination and *trans*-isomerization of the BV chromophore, local structural rearrangements around the rotating D-ring destabilize important interactions maintaining the β-hairpin conformation of the PHY-tongue that results in an increase of conformational flexibility at the PHY dimer interface. How this structural rearrangement affects the output effector in the context of the full-length molecule remains currently poorly understood.

## 1.5. The current understanding about the light signal transduction in phytochromes

From studies using phytochromes loaded with bilins co-factor locked in the 15*Z* or 15*E* configuration (89, 90) and by the use of NMR to study the changes in and around the bilins co-factors (91, 92) the initial event of the molecular signal transduction in phytochromes has been

attributed to the 15*Z* to 15*E* isomerization of the bilin chromophore and subsequent rotation of the D-ring of about ~180°. This local changes is then transduced to the GAF domain where surrounding residues adopt different rotamer arrangements as observed from dark and light states structure of phytochromes (82, 88, 93). These structural studies reveal how the bilin co-factor and the residues within the bilin binding pocket work in concert to elicit the primary photochemistry and subsequent thermal relaxation movement of phytochromes (39, 44). These local residues changes around the chromophore are assumed to weaken the connection between the PHY-tongue element and the GAF domain by notably affecting the salt bridge between a conserved Asp residue of the PASDIP motif in the GAF domain and the conserved Arg residues of the PRXSF motif in the PHY-tongue element presented earlier in the introduction. It results that the PHY-tongue strand previously distal to the GAF domain refolds to an α-helical structure and became proximal to the GAF domain establishing new hydrophobic contacts and hydrogen bonds with residues of the GAF domain (39, 80, 83, 86). The refolding of the PHY-tongue element modulates the dimer interface of the overall assembly transducing the molecular light signal to downstream effector domains. Initial Pr- and Pfr-states structures for the well characterized *Dr*BphP showed a spreading apart of the two PHY domains which eventually would affect the dimer interface of the output module (Fig. 13) (34, 39, 88). However, the recent characterization of an activated conformation of the phytochrome studied in this work (32, 77) in line with additional structural information on full-length *Dr*BphP (88) suggests more subtle changes at the dimer interface to transduce the molecular signal that promote differences in overall conformational dynamics in phytochromes to tune the PSM photoactivation to the required photoresponse (Fig. 14).
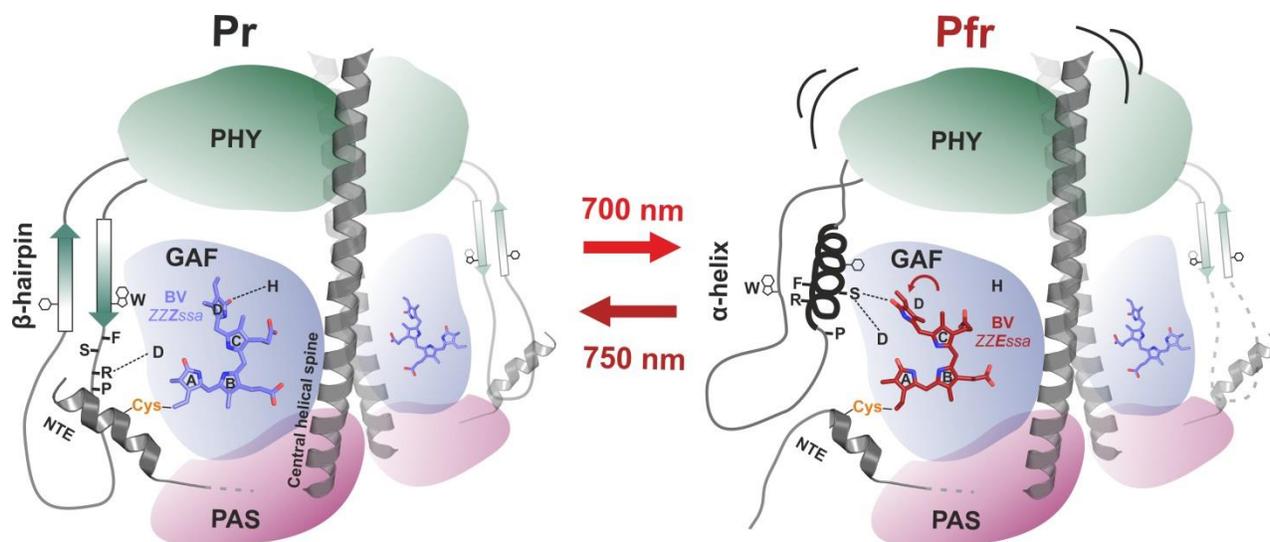
**Figure 14. Simplified representation of the current concept of molecular signal transduction in phytochrome sensors.** For many phytochrome sensors characterized so far, a common concept of the early steps of the signal transduction mechanism can be represented. Photoisomerization of the BV chromophore destabilizes molecular interactions between the NTE element, the GAF domain, and the PHY-tongue element that trigger the refolding of part of the PHY-tongue to an α-helix and increase the conformational flexibility at the PSM dimer interface. Both symmetric Pfr/Pfr and asymmetric Pfr/Pr activated dimers have been observed in phytochromes and therefore the PHY-tongue element of the symmetry related protomer of the activated conformation (right side) is represented as dashed line to indicate the modularity of the PHY-tongue element in the context of the activated phytochrome dimer. Upon activation of the PSM, the molecular signal is transduced to the sensor-effector linker element and eventually to the effector output.

# 1.6. The benefit of phytochromes; near-infrared based biological tools for fluorescent reporters, biosensors, and optogenetic applications

The fundamental benefit of understanding molecular details of the signal transduction mechanism in phytochromes and phytochrome-like sensors is the possibility to rationally couple phytochrome sensors to output effectors and modulate biological functions *in vivo* (14). The biliverdin chromophore, used by bacteriophytochromes, is ubiquitous in mammalian tissues allowing the generation of the holoproteins *in vivo*. Moreover, the conjugated system of bilin chromophores covalently linked to the phytochrome sensor environments generates a near-infrared shifted absorption spectrum within the near-infrared tissue transparency window that is

of great advantage for *in vivo* light stimulation of whole cells or animals. In fact, near-infrared light wavelengths feature a low absorbance by hemoglobin, melanin, and water, and therefore phytochromes are ideal candidates as fluorescent proteins, biosensors and reporters for optical imaging, detection, and *in vivo* light control of biological functions (18, 94–96).

The natural coupling of phytochromes to the regulation of gene transcription, enzymatic activity modulation, protein-protein interactions, pigmentation modulation, redox-sensing, phototaxis, circadian rhythms, virulence, and other functions, allows using phytochromes as building blocks for modulating various biological functions. In 2008, Wagner *et al.,* (97) discovered that mutating residues interacting with the bilin chromophore makes the protein highly fluorescent. In phytochrome PSMs engineered by truncation of the PHY domain, the inability of the chromophores to undergo full photoconversion from Pr to Pfr states leads to the increase of fluorescent quantum yields (Fig. 15A). These findings attracted special attention on phytochromes to engineer them as fluorescent biological reporters (98). The fact that the emission and excitation spectrum of such fluorescent probes are in the red and far-red light range of the visible light spectrum allow to extend their use to opaque tissue or even to whole living organisms. The first successful fluorescent reporter obtained from bacteriophytochrome, named IFP1.4, was engineered from the PAS-GAF core of *Dr*BphP and has proven to be useful for whole-body liver visualization (99). Some years later, an improved version of this fluorescent reporter which was able to bind the endogenous biliverdin of mammalian cells efficiently was engineered from the PAS-GAF core of *Rp*BphP2 and called iRFP713 (100, 101).

Fluorescent phytochromes also hold promises for their engineering as biosensors for the *in vivo* study of protein-protein interactions. In this case, phytochromes are engineered to be fused to a sensing part which receive the Fluorescence Energy Resonance Transfer (FRET) or Bioluminescence Resonance Energy Transfer (BRET) signal providing the readout of protein-protein interactions (102). Fluorescent phytochromes have notably been engineered to monitor the level of intracellular proteins based on their post-translational modifications. For instance, fluorescence ubiquitination-based cell cycle progression indicators have been generated by fusion between a fluorescent phytochrome and cell cycle regulating proteins that enable the non-invasive detection of cell cycle progression by microscopy and flow cytometry methods (103) (Fig. 15B). The preferred method to generate phytochrome biosensors is based on protein splitting into fragments which alone are not functional but regain their functionality when they

are in close proximity (102). Based on previous designs, several splitting sites have been empirically described for bacteriophytochromes (104–106). However, the method still requires a substantial effort of random mutagenesis studies and large screenings efforts.

For optogenetic purposes, plant phytochromes have been by far the most engineered phytochrome species. In general, effort is put on the design of plant phytochromes with modified photoactivated state stability and signaling properties to potentially benefit the development of plants with agricultural advantages (107). The property of plant phytochromes to bind PIFs transcription factors with high affinity in a light dependent manner (108) has notably been exploited to design light switchable gene expression systems (Fig. 15C).

Recently, the confirmation of the existence of a large repertoire of photoreceptors in chemotrophic organisms and their association with enzymatic output activity regulation has attracted special attention. Substitution of output enzymatic effector by biologically related ones have been reported successful in some cases (16, 76, 109–111) (Fig. 15D). However, the fusion between phytochromes PSM and specific enzymatic effectors remains unpredictable in most of the cases and some important structural elements like the sensor-effector linker region requires a particular attention as it has been reported as crucial for the output effector activity regulation (32, 109, 111, 112). The overall dynamics of phytochromes and their complex dimer interface need to be investigated and understood in order to couple phytochromes sensor to specific enzymatic effectors in a rational way. Moreover, the photochromic diversity in phytochromes and the evolution of PSMs to tune chromophore photoconversion properties remain largely unexplained but represent an interesting perspective for engineering phytochromes with reduced spectral overlap between Pr and Pfr absorption.
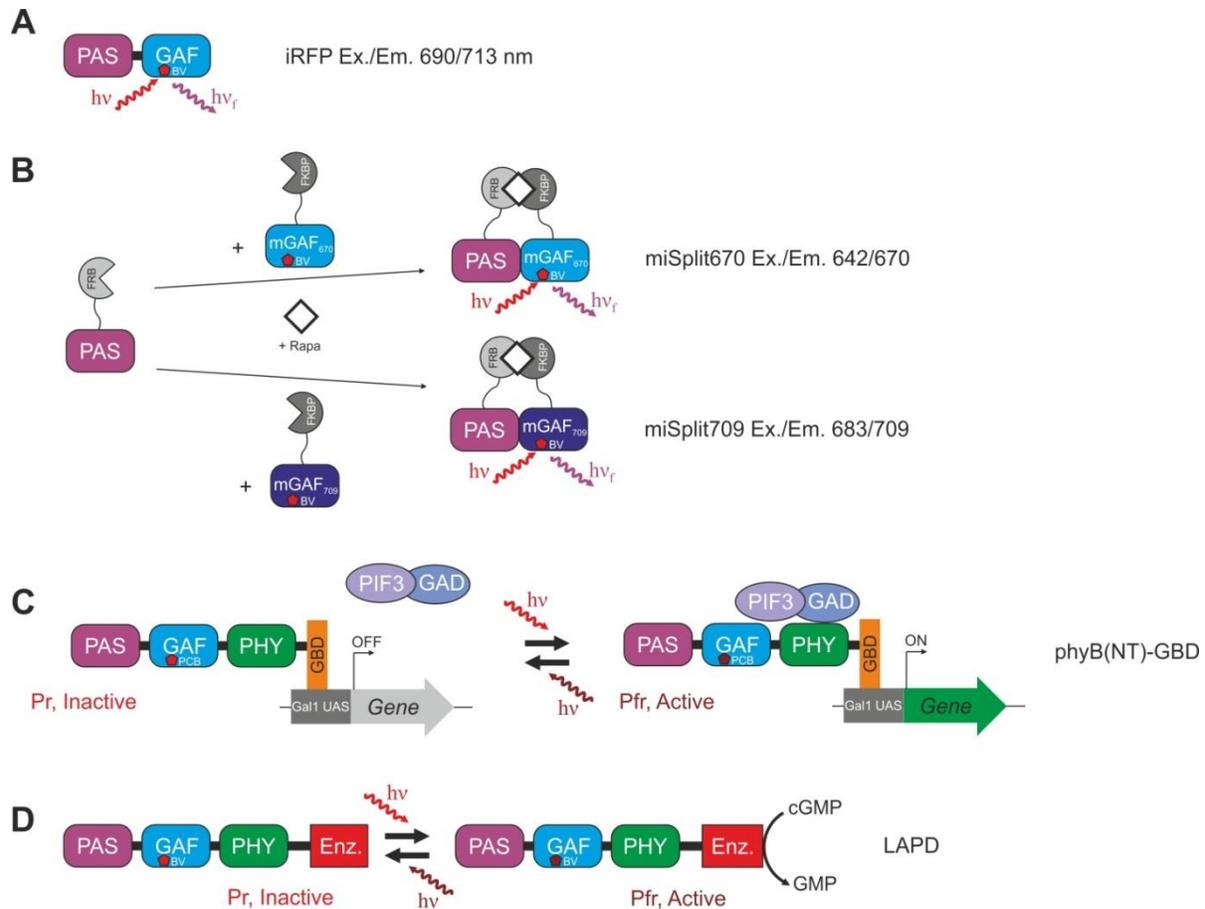
**Figure 15. Overview of various strategies based on red light absorption by phytochromes.** (**A**) iRFP (infra-Red Fluorescent Protein) based on phytochromes are generated by truncating the phytochrome sensor to its PAS-GAF core and screening mutations leading to high fluorescence property by random mutagenesis. The example shows correspond to iRFP generated from *Rhodopseudomonas palustris* (*Rp*BphP2) (101). (**B**) Two biomolecular fluorescence complementation (BiFC) monomeric miSplit reporters. The case of miSplit670 and miSplit709 are shown as example (103). The generation of monomeric iRFP exploit the monomeric property of the naturally occurring *Rp*BphP1 (84). The two miSplits share the same PAS fragment that can interact with either mGAF$_{670}$ or mGAF$_{709}$ fragment producing the fluorescence signal corresponding to complemented miSplit670 or miSplit709, respectively. *In vivo*, transiently transfected HeLa cells express cytoplasmic FRB-PAS together with either cytoplasmic FKBP-mGAF$_{670}$ or nuclear FKBP-mGAF$_{709}$. The addition of Rapamycin (Rapa) to the media promotes the interaction between FRB and FKBP that leads to the functional miSplits and the corresponding fluorescence signal. (**C**) The natural coupling of phytochromes to enzymatic effectors allows the use of these naturally occurring constructs, as well as new fusions synthetically generated, to modulate physiological pathways *in vivo*. The case of LAPD corresponding to the synthetic fusion between *Pa*BPhy phytochrome sensor and *Hs*PDE2A phosphodiesterase effector is shown as example (111). (**D**) Light-responsive gene promoter system (108). The N-terminal part of phytochrome sensor (phy-B) from *Aradidopsis thaliana* is fused to a GAL4 DNA-binding domain (phyB(NT)-GBD) that recognizes the GAL1 upstream activation sequence. When the PhyB sensor is in Pr state in darkness condition, the interaction between PIF3 and phyB does not occur and the expression of expression of the target gene is off.

Upon red light illumination, phyB is converted to its Pfr active conformation that can bind the PIF3 moiety of the PIF3-GAD fusion protein. The photoactivated phyB(NT)-GBD thereby recruits the PIF3-GAD protein to the target promoter where the transcriptional activation domain of PIF3-GAD induces transcription of the target gene.

# 1.7.    An integrative structural biology approach to better understand the dynamics of sensor-effector communication

In order to understand such a complex machinery that photoreceptors, and notably phytochromes, are, we need to be able to look at high resolution pictures of what happens at the molecular level during the light signal transduction mechanism. High resolution static pictures of the full-length systems conformation can be obtained by X-ray crystallography. However, the dynamic perspective of the system cannot be fully appreciated from X-ray structures. Therefore, we used a combination of X-ray crystallography with Hydrogen Deuterium eXchange couples to Mass Spectrometry (HDX-MS) analysis in order to appreciate the conformational dynamics of the full-length phytochromes upon red light illumination.

Short introductions to both of these methods are presented in the following two chapters.

## 1.7.1.    X-ray crystallography

The basic principle of protein X-ray crystallography is to record diffraction images of a focused X-ray beam upon interaction with a highly ordered protein crystal. From the diffraction data, it is possible to reconstitute the electron density representing the structure of the molecule by Fourier's methods. Within the electron density around each atom of the molecule, one can then build the corresponding 3D arrangement of the atoms of the molecule of interest (Fig. 16).
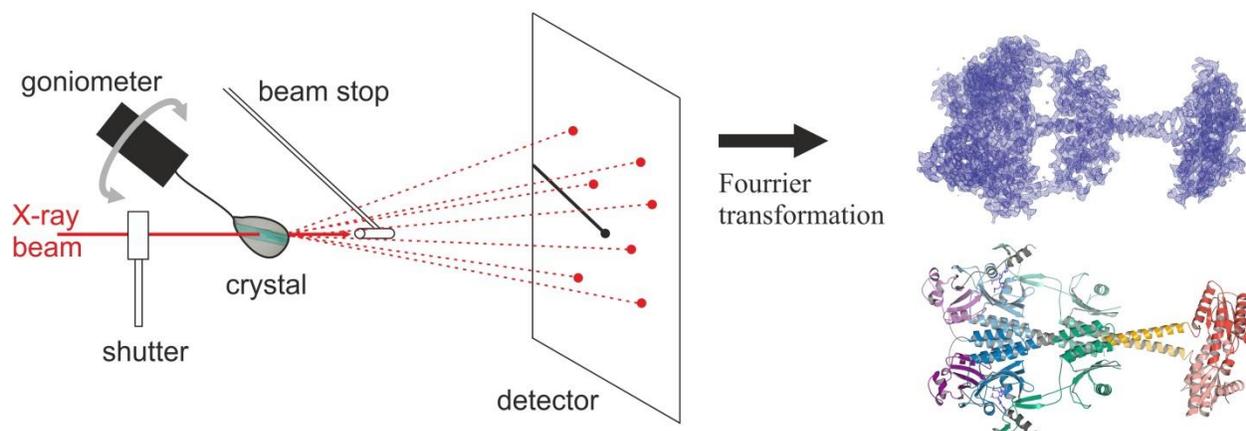
**Figure 16. Basic principle of X-ray protein structure determination.** A protein crystal is mounted on a rotatable goniometer and is exposed to a collimated, intense X-ray beam in the 5-20 keV energy range. Individual diffraction images are recorded on a detector during small rotation increments of the crystal. Non diffracted X-rays are stopped on a beam stop to prevent degradation of the detector. The diffraction data set is then transformed into an electron density by means of Fourrier transformation when the amplitudes and the phases for each diffraction are known. Depending on the quality of the electron density obtained, the protein structure can be automatically or manually built into the electron density (blue mesh).

X-ray diffraction is the result of high energy X-ray electromagnetic radiations of very short wavelength that interacts with the electrons of the protein atoms primarily through its oscillating electric field vector (E) (Fig. 17). The electric field (E) is perpendicular to the propagation direction of the wave, and oscillates with a frequency corresponding to the energy and the inverse of the wavelength of the X-rays (Fig. 16A). The interaction of the oscillating electric field vector with electrons induces oscillations of the electrons which emit themselves electromagnetic waves of the same frequency. The phenomenon of diffraction is the net result of the phase difference between these emitted scattered waves.

In an ordered crystal, we can slice the crystal in parallel reflecting planes which all pass by the same lattice points and which are defined by their miller indices (h,k,l) (Fig. 17B). If scattered waves between two parallel reflecting planes are in phase they will interfere constructively to form a scattered wave of higher amplitude that will eventually form a diffraction spot on the detector if the Bragg's law is respected (Fig. 17B).
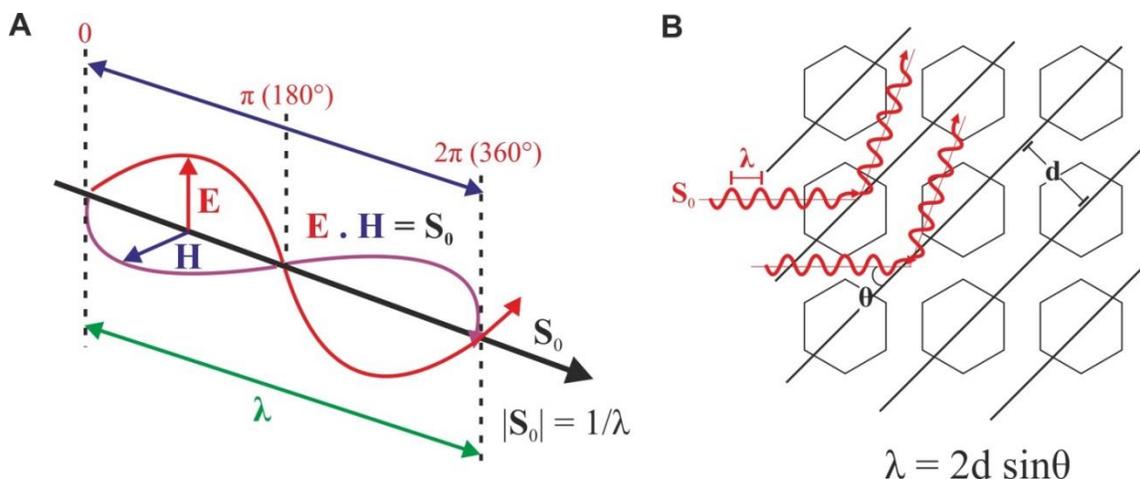
**Figure 17. Anatomy of an electromagnetic wave and its interactions with a protein crystal.** (**A**) The electric field vector E is perpendicular to the magnetic field vector H, and both oscillate perpendicular to the propagation vector $S_0$ of the wave of wavelength $\lambda$. The vector $S_0$ in the direction of wave propagation has a magnitude of $1/\lambda$. The interaction of the electric field vector E with dielectric matter is about six orders of magnitude larger than the interaction of the magnetic field vector H, thus the effects of H in X-ray scattering can be neglected. Panel A is reproduced with permission from *Biomolecular Crystallography* by Bernhard Rupp © 2009-2014 Garland Science/Taylor & Francis LLC. (**B**) When considering an ordered protein crystal, parallel reflecting planes of miller indices (h,k,l) can be draw between the same lattices points. If two scattering waves interacting with atoms of two parallel planes are in phase they interfere constructively following the Bragg's law ($\lambda = 2d \sin \theta$) to form a scattered wave of higher amplitude. $\theta$, angle formed between the incident wave and the reflecting plane. d, distance between two reflecting planes. $\lambda$, wavelength of the X-ray beam.

To relate the planes in the crystal lattice to the points in the diffraction pattern, we can represent the scattered wave in the reciprocal space using the Ewald's sphere representation (Fig. 18). The scattering angle formed between the incident wave and the scattered wave is defined to be $2\theta$ and a circle of radius $1/\lambda$ is drawn. The point A formed at the intersection between the Ewald's sphere and the scattered wave fulfills the Bragg's law and will form a diffraction spot on the detector. Upon rotation of the crystal under the X-ray beam, other scattered waves coming from different atoms will cross the Ewald's sphere and the resulting spot are recorded on the detector (Fig. 18).

$$\lambda = 2 \sin \theta / OA$$
$$<=> \lambda = 2d \sin \theta$$

Diffraction spots
of a protein crystal
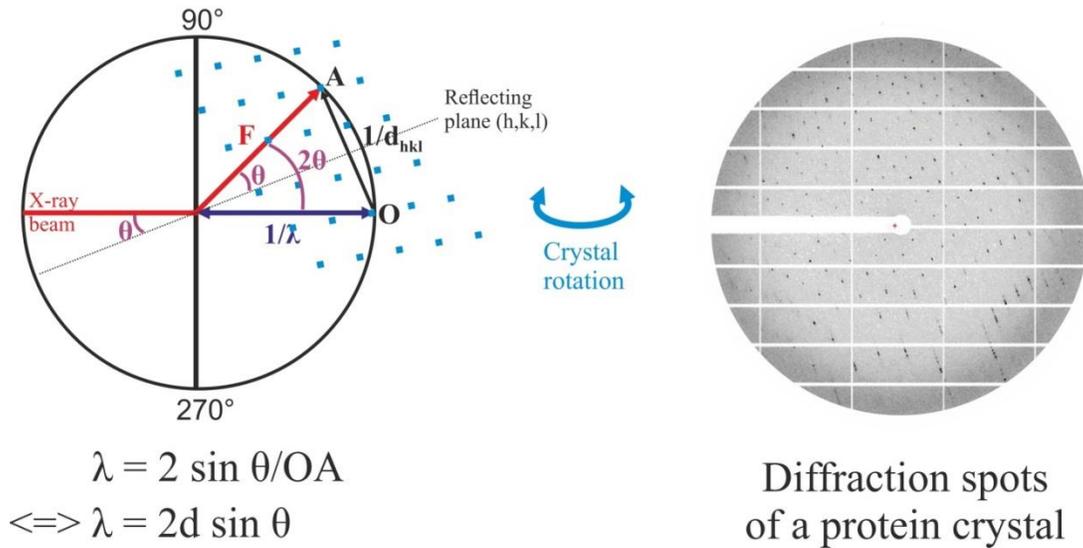
**Figure 18. Representation of X-ray diffraction in the reciprocal space.** (**A**) It is possible to represent any plane wave and its phase relation to a fixed origin as a vector F in the complex plane. The magnitude of the vector F is given by the amplitude of the wave $|F|$, and its direction by the phase angle $\theta$ relative to a fixed origin. Using this vector representation of the waves, and knowing that constructive interference is only true if the Bragg's law is respected, diffraction spots recorded on the detector results come from the crossing point between the scattered wave and the Ewald's sphere of radius $1/\lambda$.

During the X-ray diffraction experiment, the intensities and the position of the scattered reflections are measured. Therefore, for each diffraction we can assign the index (h,k,l), and the intensity of the reflection that is proportional to the square of the structure factor amplitude $|F_{(h,k,l)}|$.

The structure factor ($F_{(h,k,l)}$) for a reflection $h,k,l$ is a complex number resulting of a summation over all atoms $j$ with $x,y,z$ their fractional coordinates. $f(j)$ represents the scattering factor of atom $j$ and depends on the kind of atom and the diffraction angle of the corresponding reflection $(h,k,l)$:

$$F(h, k, l) = \sum_{j=1}^{atoms} f_{(j)} \exp[2\pi \cdot i(hx_{(j)} + ky_{(j)} + lZ_{(j)})]$$

The Fourier transformation of the structure factor $F_{(h,k,l)}$ allows to transform the diffraction pattern in the reciprocal space into the real space of electron density ($\rho_{(x,y,z)}$) which is

the sum of the contributions to the point (*xyz*) of a wave scattered from a plane (*hkl*) whose amplitude depends on the number of electrons in the plane. That gives the following equation where *V* is the volume of the unit cell:

$$\rho(x,y,z) = \frac{1}{V} \sum_h \sum_k \sum_l \left| Fhkl \right| \exp(2\pi i(hx + ky + lz) + i\Phi(hkl))$$

To calculate ρ*(x,y,z)* we need the amplitude of each diffraction that is easily accessible from the square root of the intensities of each reflection constituting the collected diffraction pattern as mentioned earlier in the text, however, the term Φ*(k,k,l)* representing the phases of each reflections (*h,k,l*) is lost during the previous summation of the scattered waves and need to be determined experimentally or by means of model phases.

Solving the phase problem is one of the fundamental steps in solving a protein structure. If related proteins with high sequence identities, or high resolution structures of parts of the target protein, are known, the classical method to overcome the phase problem is to use the phases information of that structural information as restraint for the starting phases used for the map reconstructions of the target protein (113). The whole process is called molecular replacement and although being time-efficient, if not enough similarity exists between the search probe and the protein of interest, strong bias can be introduced in the electron density because the phase information dominates the electron density reconstruction.

To overcome this issue, the method of single-wavelength anomalous diffraction (SAD) as used in this thesis to solve the first full-length structure of *Is*PadC, can be employed. To do so, heavy atoms need to be introduced into the protein, like selenomethionine in this case which features the heavy atom of selenium. The imaginary constituents of the atomic scattering factor that are dependent of the wavelength but not of the scattering angle can be perturbed by tuning the X-ray beam wavelength to generate an abnormal scattering factor for the heavy atoms. Friedel's law that holds $\left| F(hkl) \right| = \left| F(-h-k-l) \right|$ is no longer true that give rise to anomalous difference that can be used to localize the anomalous scatterers (Fig. 19). The placement of these atoms and the anomalous difference associated are then used together to estimate crystallographic

phases for the entire structure. An electron-density map can then be calculated using these initial phases information and the average of structure factors, which is usually improved by density-modification techniques afterwards (114).
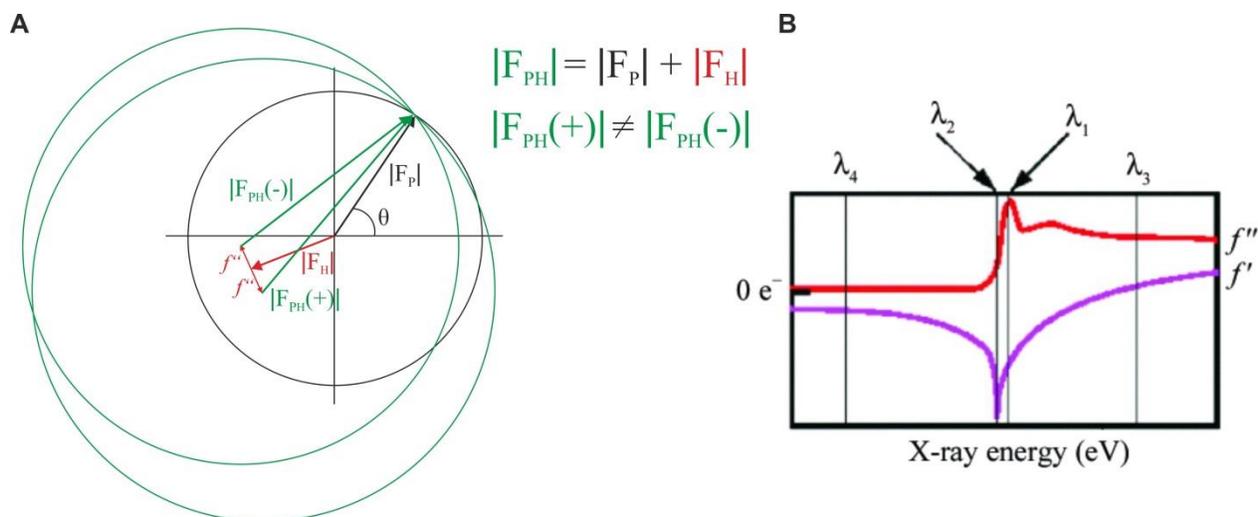


**Figure 19. Experimental phasing.** (**A**) For a protein loaded with heavy atoms, the reflection $|F_{PH}|$ is equal to the reflections coming from the atoms of the protein ($|F_P|$) plus the reflections coming from the heavy atoms ($|F_H|$). By tuning the wavelength of the X-ray source used, one can specifically tune the imaginary absorption term (*f''*) of the heavy atoms that has for consequence to makes the Friedel's law not true that renders possible the localization of the anomalous scatterers. (**B**) Typical absorption curve for an anomalous scatterer. Adapted with permission from (115), Taylor G.L. 2010. doi: 10.1107/S0907444910006694.

## 1.7.2.    Hydrogen-Deuterium eXchange coupled to Mass Spectrometry (HDX-MS)

The better understanding of light signal transduction mechanisms occurring in photoreceptors requires the integration of structural information with that of the conformational dynamics that is involved in the molecular communication between sensors and effectors. Since X-ray crystallography only allows us to observe static snapshots of the protein in different conformational states with limitations to obtain structural information of the protein under constant illumination, we combined the structural details obtained by X-ray crystallography with information derived from the global conformational dynamics observed by HDX-MS.

The basic principle of this technique relies on the fact that exposure of the protein to $D_2O$ induces rapid exchange of the hydrogen at the amide bond of disordered regions lacking stable secondary structure, while more folded elements feature more protected amide bond hydrogen atoms (Fig. 20).
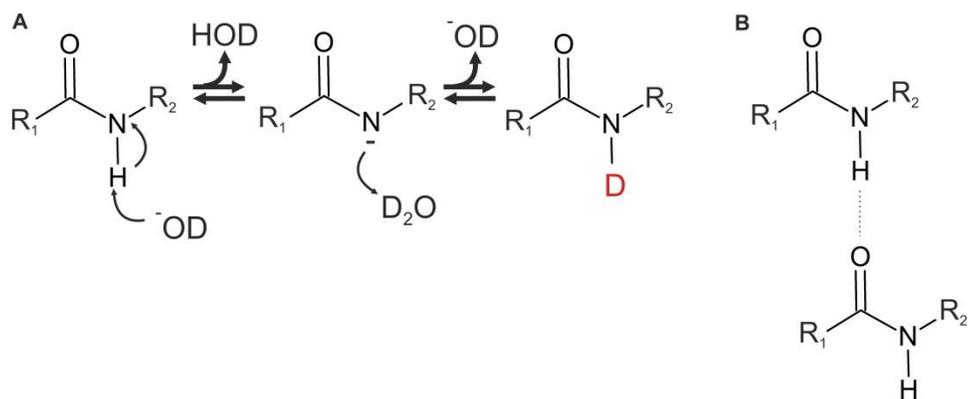


**Figure 20. H/D exchange in solution.** (**A**) Mechanism of base catalyzed backbone amide H/D exchange in solution. (**B**) Amide bond involved in stable secondary structure are protected and exchange at slower exchange kinetics compared to amide bond of unfolded or flexible regions.

Therefore, the exchange with $D_2O$ in stable secondary structure elements results in slower exchange kinetics mediated by the decrease of conformational dynamics in the protein. The exchange reaction is then quenched by means of temperature and pH decrease to avoid back exchange of $H_2O$ with $D_2O$, and the protein is digested into small peptides by the non-selective proteolytic activity of pepsin. The peptides are afterwards separated by reverse-phase liquid chromatography before being analyzed by Electrospray Ionization Mass Spectrometry (ESI-MS) analysis. Cleaved peptides obtained from pepsin activity are previously identified by tandem MS/MS analysis. Incorporation of deuterium into the peptides shifts the mass of the peptide allowing us to precisely estimate the number of deuterons incorporated during the incubation time (Fig. 21).
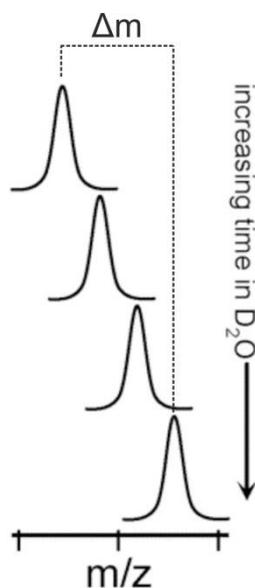
**Figure 21. Time dependent deuterium incorporation.** Upon increasing incubation time in $D_2O$, more hydrogen of the amide bonds are exchanged with deuterium ions leading to a mass shift of the analyzed peptides.

It is then possible to determine the relative deuterium uptake for each analyzed peptide as a function of the labeling time *t* according to the following equation where *m(t)* is the centroid mass of the peptide, and $m_0$ is the corresponding reference data for unlabeled samples, respectively:

$$Relative\ deuteration\ level\ (D_{rel}(t)) = m(t) - m(0)$$

The difference between the relative deuterium incorporation at the amide bond hydrogens in two different states of the protein [$\Delta D_{rel} = D_{rel}(state\ 2) - D_{rel}(state\ 1)$] enables to localize the regions occurring an increase of conformational dynamics, a loss of secondary structure, or rather a stabilization of structural elements (Fig. 22). Hydrogens of some amino-acids side chains are also exchanging with deuterium, however, the incorporated deuterium quickly back exchanges with HOD at this position that renders the extraction of structural information from the amino-acids side chains exchange impossible by HDX-MS.
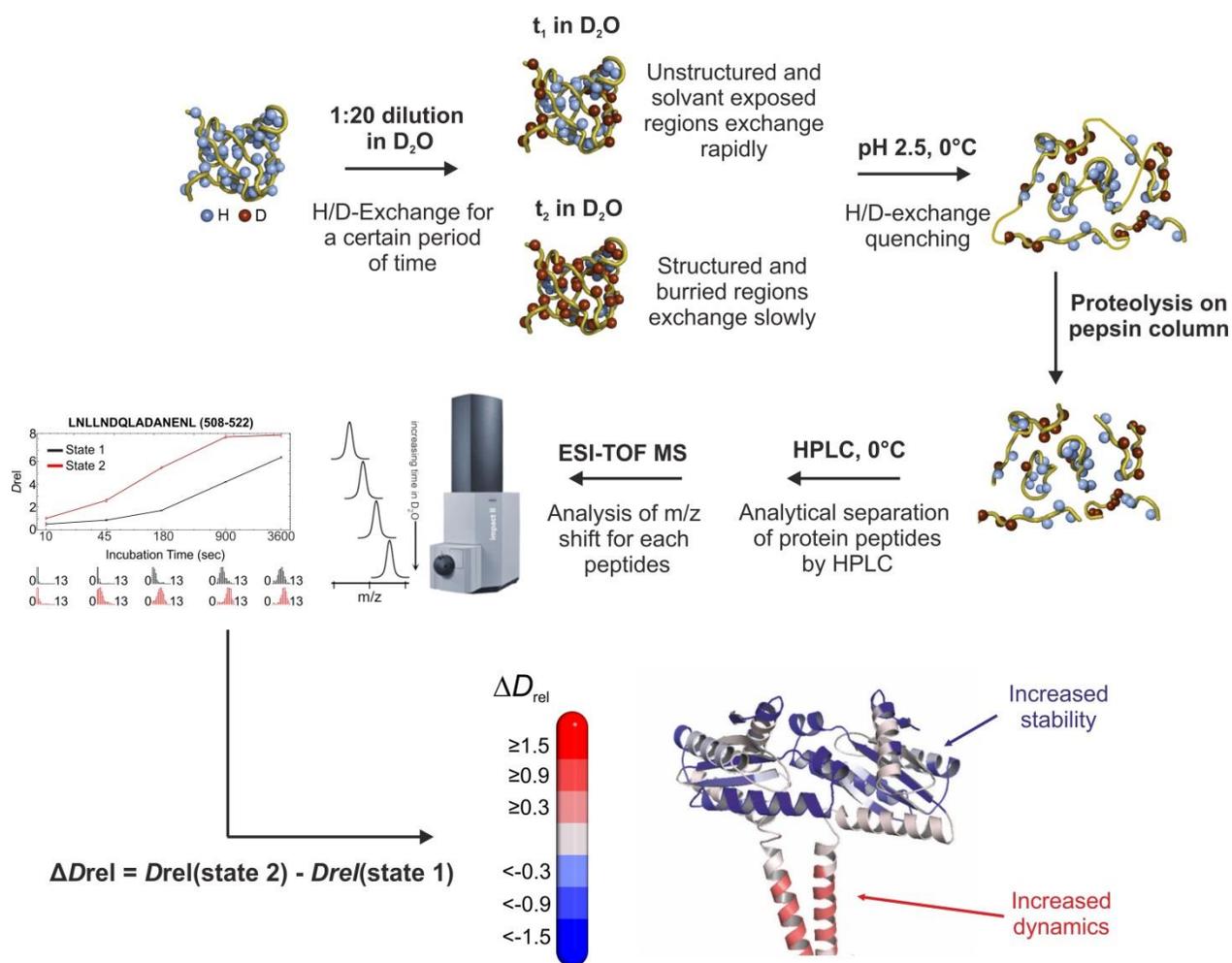
**Figure 22. Workflow of Hydrogen/Deuterium exchange analysis coupled to mass spectrometry.** The protein in buffer solution is diluted at least 1:20 in $D_2O$ buffer and incubated at 20°C for various time lengths of incubation from sec to hours. Unstructured and solvent exposed regions are quickly exchanging at the amide bond hydrogen, whereas the structured and buried regions are exchanging slowly. After the desired time of incubation, the H/D exchange is quenched by lowering the pH and the temperature. The protein is then digested into smaller peptides by pepsin before being separated by HPLC and analyzed by mass spectrometry. The incorporation of deuterium into the peptides shift its m/z ratio analyzed by ESI-TOF MS that allows determining the number of deuterium incorporated into the peptides. The same analysis of the protein in different states (e.g. in dark and under light illumination) allows to obtain two different deuteration profiles for each peptide and the difference of deuterium incorporation profile lead to either a positive difference meaning that the peptide occurred an increase of conformational dynamics, or a negative difference meaning that the peptide occurred a stabilization.

# 2. Results

## 2.1.   Long-range   allosteric   signaling   in   red   light-regulated diguanylyl cyclases

Results presented in this chapter are mainly based on the following publication:

# Longe-range allosteric signaling in red light-regulated diguanylyl cyclases

Gourinchas G[1], Etzl S[1], Göbl C[2,3], Vide U[1], Madl T[2,3,4], Winkler A[1,*]

[1]*Institute of Biochemistry, Graz University of Technology, Petersgasse 12/II, 8010 Graz, Austria.*

[2]*Center for Integrated Protein Science Munich, Technische Universität München, Department of Chemistry, Lichtenbergstraße 4, 85748 Garching, Germany.*

[3]*Institute of Structural Biology, Helmholtz Zentrum München, Ingolstädter Landstraße 1, 85764 Neuherberg, Germany.*

[4]*Institute of Molecular Biology and Biochemistry, Medical University of Graz, Harrachgasse 21/III, 8010 Graz, Austria.*

*\*Corresponding authorEmail: andreas.winkler@tugraz.at*

## 2.1.1.    PadCs homologs present different regulation capacities of the GGDEF effector

Diguanylyl cyclases are important enzymes involved in the production of c-di-GMP acting as important bacterial second messengers (70, 71, 116). In the recent decades, the investigation of c-di-GMP regulatory network has revealed a growing number of c-di-GMP implications in cellular functions, including regulation of the cell cycle, differentiation, biofilm formation, motility, virulence, and other biological responses (71, 117, 118). Diguanylyl cyclases (DGC), characterized by their highly conserved Gly-Gly-(Asp/Glu)-Glu-Phe (GGDEF) sequence at the active site, and phosphodiesterases (PDE) characterized by their highly conserved Glu-Ala-Leu active site (EAL) are broadly represented in all major bacterial phyla where they catalyze the synthesis and degradation of c-di-GMP, respectively (70, 71, 116). The adjustment of the cellular levels of c-di-GMP is achieved by a variety of regulatory mechanisms which couple signals to DGC and PDE activity (116, 119). Regulation usually exploits the dimeric structure of the catalysts, whose active centers is formed at the dimer interfaces. Typical phytochromes acting as a dimer and featuring an arrangement of PAS - GAF - PHY domains have evolved to modulate the catalytic activity of GGDEF alone or GGDEF – EAL tandem by red light (66) (Fig. 23).

These naturally occurring systems, called PadCs in this thesis, are the system of choice studied here in order to improve the understanding of the light signal integration and transduction from the phytochrome sensor to the catalytic modulation of the DGC effector. Following a BLAST search from one PadC sequence, several PadC homologs were found. The various homologs evolutionary differentiate in two main branches; the phytochrome activated GGDEFs, and the phytochrome linked GGDEF – EAL tandems (Fig. 23). The EAL domain features the degradation of c-di-GMP to 5'-phosphoguanylyl-(3'-5')-guanosine (pGpG) (111). In this thesis, we only focus on the group of phytochromes modulating a GGDEF domain only.
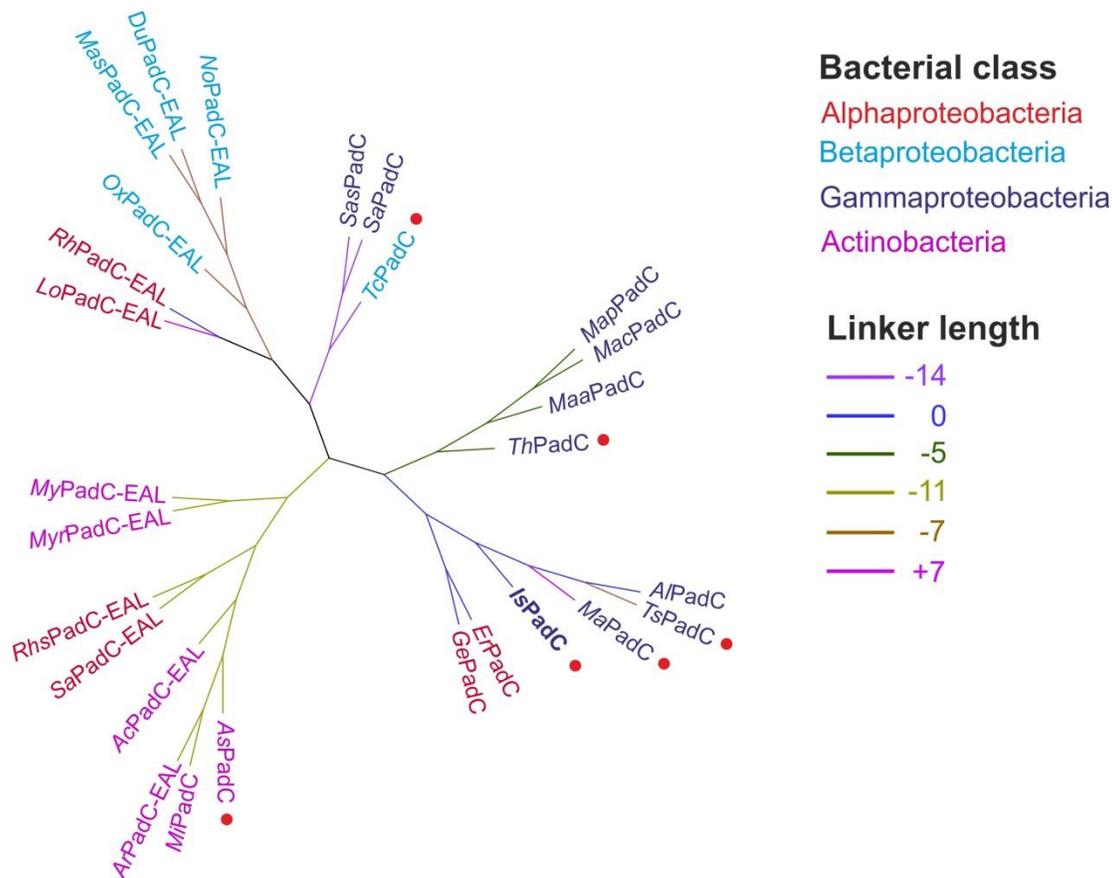
**Figure 23. Phylogenetic tree of naturally occurring PadCs.** (**A**) Unrooted phylogenetic tree representing several PadC homologs. Naturally occurring coupling of phytochrome with diguanylyl cyclases (GGDEF) and phytochrome with tandem diguanylyl cyclases – phosphodiesterases (EAL) are denoted as PadC and PadC-EAL, respectively. Branches are colored according to the coiled-coil sensor-effector linker length differences when taking the *Is*PadC linker as reference (*cf.* Appendix Fig. 1). Names of the proteins are colored according to the bacterial class they belong to. *Is – Idiomarina species A28L* and WP_007419415; *Ma – Marinimicrobium agarilyticum* and WP_027329460; *Ts – Thioalkalivibrio species ALMg3* and WP_026331574; *Al – Alkalilimnicola ehrlichii* and WP_083761846.1 ; *Th – Thiohalospira halophila* and WP_093427120.1; *Maa – Marinospirillum alkaliphilum* and WP_072324728.1 ; *Mac – Marinospirillum celere* and WP_091958617.1 ; *Map – Marinobacter persicus* and WP_091706258.1 ; *Tc – Thauera chlorobenzoica* and WP_075148475.1 ; *Sa – Salinisphaera hydrothermalis* and WP_ 084188582.1; *Sas – Salinisphaera shabanensis* and WP_084623642.1 ; *Er – Erythrobacter sp. QSSC1-22B* and WP_082992438.1 ; *Ge – Gemmobacter changlensis* and WP_107665948.1 ; *As – Agromyces sp. Leaf222* and WP_082462169.1 ; *Mi – Microbacterium pygmaeum* and WP_091492622.1 ; *Ar – Arthrobacter sp. CAL618* and WP_ 024476458.1; *Ac – Actinotalea ferrariae* and WP_052022258.1 ; *Saf – Salinihabitans flavidus* and WP_093117713.1 ; *Rhs – Rhodobacter sphaeroides* and WP_061350905.1 ; *Myr – Mycobacterium sp. Root135* and WP_056547734.1 ; *My – Mycobacterium chlorophenolicum* and WP_060937607.1 ; *Lo – Loktanella maricola* and WP_100368535.1; *Rh – Rhodobacter aestuarii* and WP_076485451.1; *Ox – Oxalobacteraceae bacterium IMCC9480* and WP_040724740.1 ; *Mas – Massilia consociate BSC265* and WP_051933242.1; *Du – Duganella sp. Leaf61* and WP_056147085.1 ; *No – Noviherbaspirillum suwonense* and WP_099914387.1.

Several conserved sequence motifs among bacteriophytochromes are also observable for PadC homologs. Notably within the GAF domain, residues in the vicinity of the chromophore identifiable by sequence alignments with the sequence of high resolution phytochrome structures appear to be highly conserved most probably to conserve the initial 15Z configuration of the chromophore (Appendix Fig. 1). The PASDIP sequence motif containing the Asp residue engaged in a salt bridge with the Arg residue of the highly conserved PRXSF motif in the PHY-tongue element is strictly conserved in all PadC homologs suggesting rather similar interactions between the PHY-tongue and the GAF domain (Appendix Fig. 1). Nevertheless, some homologs present untypical residue changes. *As*PadC presents a Thr instead of the highly conserved His residue interacting with the D-ring of 15Z BV (His282 in *Is*PadC sequence). *Maa*PadC features a glycine instead of the highly conserved cysteine residue involved in the covalent linkage with the BV (Cys17 in *Is*PadC sequence), however this latter variation results potentially of a sequencing artefact as a single base exchange can differentiate a codon coding for a Cys or a Gly residue. In the tongue region of the PHY domain, the WXG motif and the Trp447 and Trp473 (numeration based on the *Is*PadC sequence) are strictly conserved suggesting that the hydrophobic interactions of Trp473 with the GAF domain potentially conserve a similar tryptophan switch scenario occurring in the PHY-tongue region upon PHY-tongue refolding as observed for other phytochromes (39, 83, 85, 86). Compared to the PSM domains, the DGC presents overall strong sequence conservation. However, the linker element connecting the PSM to the DGC presents an interesting variability. Indeed, from residue 501 to 528, corresponding to the *Is*PadC sequence (Fig. 24), a conserved coiled-coil character can be predicted among the various homologs (Appendix Fig. 1). Coiled-coil sequences are characterized by a succession of heptad units defined by two helical turns represented by 7 residues labelled from *a* to *g* with *a* and *d* positions occupied by hydrophobic residues to maintain the coiled-coil oligomer structure whereas the rest of the positions are occupied mostly by polar and charged residues (120–122) (Fig. 25).
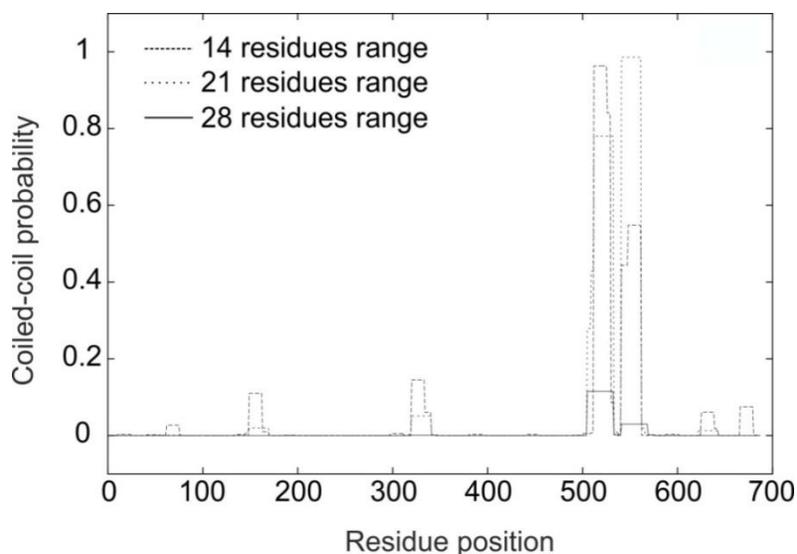
**Figure 24. Prediction of coiled-coil character in *Is*PadC sequence.** Coiled-coil prediction was performed over the entire *Is*PadC sequence using the program COILS (123) and is represented as coiled-coil character probability over residue ranges of 14, 21, and 28 residues. The region between residues 500-530 corresponding to the expected coiled-coil linker in *Is*PadC features a confident prediction of coiled-coil character.
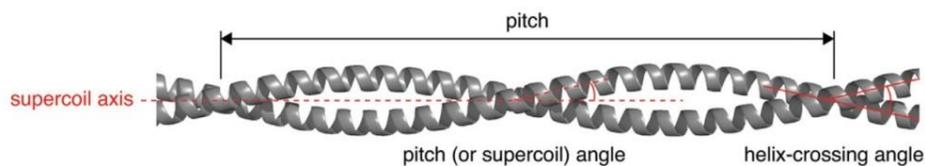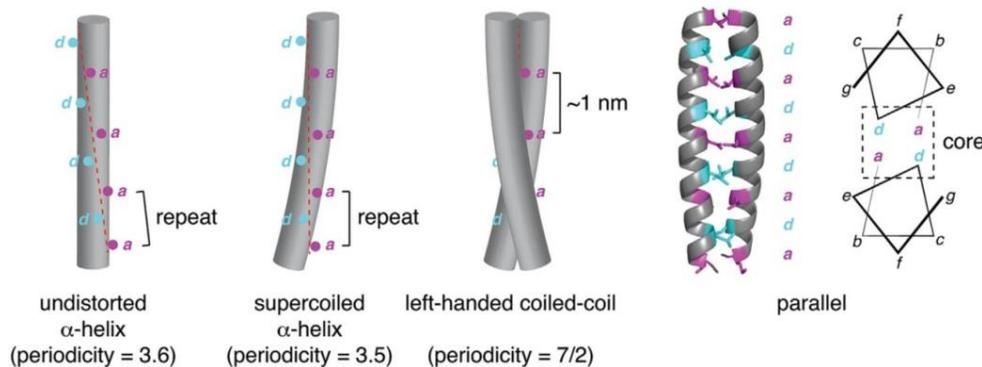


**Figure 25. Properties of coiled-coil architecture.** The basic parameters of the coiled-coil are its pitch (periodicity of the supercoil), the associated pitch (or supercoil) angle (angle of the helix with respect to the supercoil axis), and the helix-crossing angle (angle at which the two helices cross each other). The canonical coiled-coil is characterized

by a heptad repeat in which hydrophobic residues are conserved at positions *a* and *d,* and charged or polar residues are present at the other positions. Undistorted α-helices cannot pack together in a side-by-side arrangement due to the non-integral periodicity of the helix (3.6). By supercoiling the helices, the periodicity is reduced to 3.5, leading to the 7/2 periodicity of a canonical, left-handed coiled-coil, with each heptad repeat measuring ~1 nm along the coiled-coil. Reproduced with permission from Truebestein L. *et al.*, 2016 (122), doi: 10.1002/bies.201600062.

Strikingly, the coiled-coil linker region in PadC homologs varies in terms of length and composition. This evolutionary differentiation led to the clustering of PadCs in groups with different linker lengths. Interestingly, these clusters also tend to conserve similar sequence peculiarities in the linker region like conserved glycines in the helical pattern (e.g *Ts*PadC, *Tsb*PadC, *Tsc*PadC), or charged residue-rich heptad unit compositions (e.g. *Ma*PadC, *Msp*PadC, *Tth*PadC) (Appendix Fig. 1). Taking *Is*PadC coiled-coil linker region as reference, we observed that PadC homologs usually feature variations of multiple of 7 residues in their predicted coiled-coil linker lengths. However, some PadC homologs present atypical linker length variations (e.g. *Th*PadC, *As*PadC, respectively -5, and -11 compared to *Is*PadC coiled-coil sequence). Due to the supercoiled character of the coiled-coil dimer helices (Fig. 25), a difference of 7 residues (two helical turns) might not necessarily maintain a same structural alignment between phytochrome sensor and GGDEF effector.

Based on the instability index determined by the software ProtParam (124), we selected full-length PadCs with the smallest instability probability to clone and express the construct in *E. coli*. Since selected PadC homologs were usually expressing well in *E. coli* BL21 (DE3) cells co-expressing the heme oxygenase-1 the capacities of photoresponse and GTP catalysis were initially analyzed for all the selected homologs (Fig. 26-27). All the purified PadC homologs featured a canonical phytochrome character observable by the absorption of the phytochrome Q-band that showed a maximum of absorption around 700 nm representative of the Pr-state in dark condition (Fig. 26A). All spectra shift towards the far-red light wavelength with a maximum of absorption around 750 nm upon red light illumination (660 nm, 45 mW cm$^{-2}$) and recover thermally to their initial dark state spectra with different recovery kinetics (Fig. 26B). The broad diversity of Pfr-state stability and time constants of thermal reversion to the Pr-state reflects the difference of global assembly and conformational rearrangements of the various homologs during the process of photoconversion (Fig 26B and Appendix Table 1). Difference spectra between illuminated and dark conditions show that the maximum of red light absorption is slightly shifted

among the various homologs and even strongly blue-shifted for two of them (*Tc*PadC and *Th*PadC). As both of these blue-shifted homologs successfully loaded biliverdin during protein co-expression with the heme oxygenase 1, and no obvious variation of important residues around the BV D-ring is observed, it remains currently uncertain what structural features shift the Pr-absorption spectra.
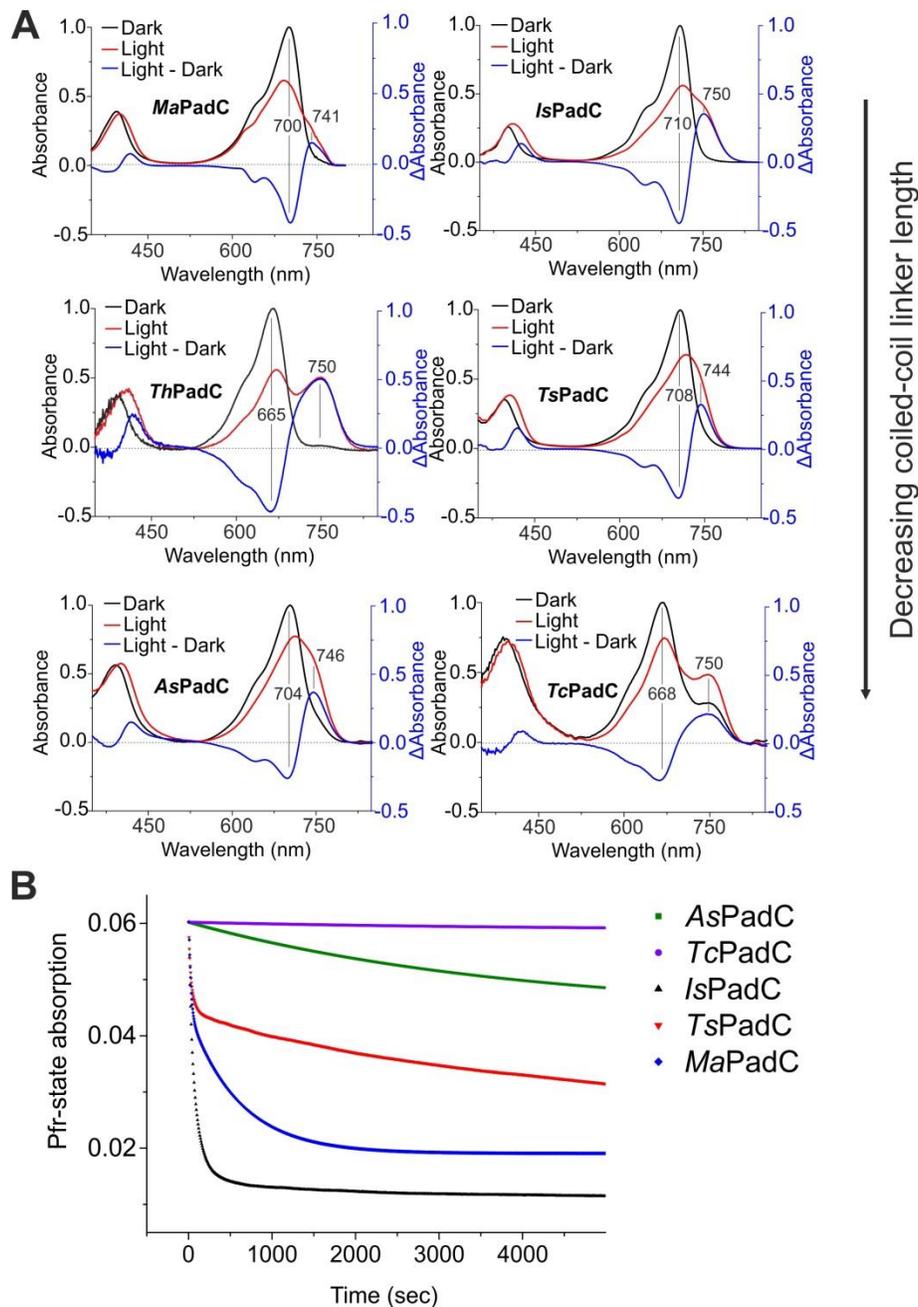


**Figure 26. Overview of the naturally occurring PadCs characterized.** (**A**) UV/Vis spectra of 6 PadC homologs characterized in this thesis which are indicated by a red spot in panel A. The dark-adapted spectrum is colored in black, the steady-state illuminated spectrum is colored in red, and the difference between illuminated and dark-state

spectra is colored in blue. Spectra have been scaled based on the Pr-state Q-band absorption maximum. (**B**) Thermal reversion of the illuminated steady-state spectrum to the Pr ground state.

A congo-red *in vivo* screening approach was adapted to rapidly characterize c-di-GMP production in *E. coli* BL21 (DE3) overexpressing the PadC of interest and co-expressing the heme oxygenase-1 (*cf.* Materials and Methods part). *E. coli* cells containing active DGCs show a red coloration of the colony due to the c-di-GMP induced production of cellulose binding the congo-red dye (125). Strikingly, we observed that the various PadC homologs featuring the same arrangement of functional domains have very different photomodulation capacities highlighting the importance of the coupling between the PSM sensors and the DGC effectors and indicates an important effect of the coiled-coil sensor-effector linker length and composition on the signal transduction mechanism (Fig. 27).
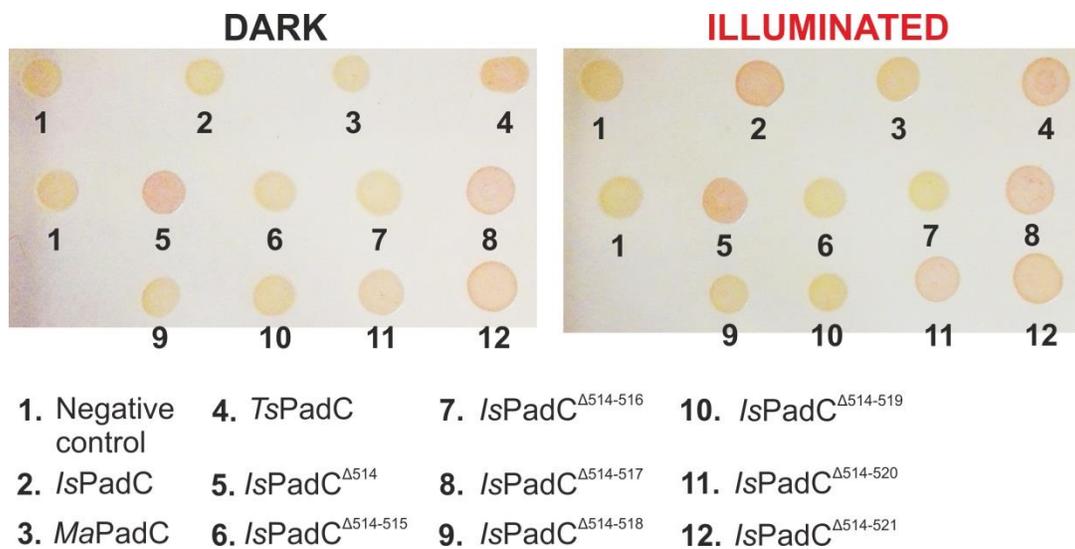


**Figure 27. Diguanylyl cyclase activity measurements.** Congo-red based screening of *in vivo* diguanylyl cyclase activity. Red phenotypes of bacterial colonies are characteristic of cellulose production triggered by c-di-GMP production and complexing with the congo-red dye present in the media. From position 5 to 12 are *Is*PadC coiled-coil linker deletion variants of respectively 1 to 8 amino acids which highlight the importance of the coiled-coil linker length and composition on the photoregulation of the DGC effector.

The intensities of red coloration observed on the congo-red based assay upon DGC activity are also dependent on the over-expression of the PadCs tested. Therefore differences of PadC expression *in vivo* might generate different DGC regulation comparison between different

PadCs in comparison to *in vitro* measurements using purified protein. However, we usually observed a very good correlation between the *in vivo* screening and the *in vitro* quantification of DGC activity (Fig. 28 and Appendix Table 2). As observed for *Is*PadC (Fig. 28) characterized PadCs usually show an apparent substrate inhibition of DGC activity at high concentration of substrate GTP indicating a complex cross-talk between the DGC and the phytochrome sensor.
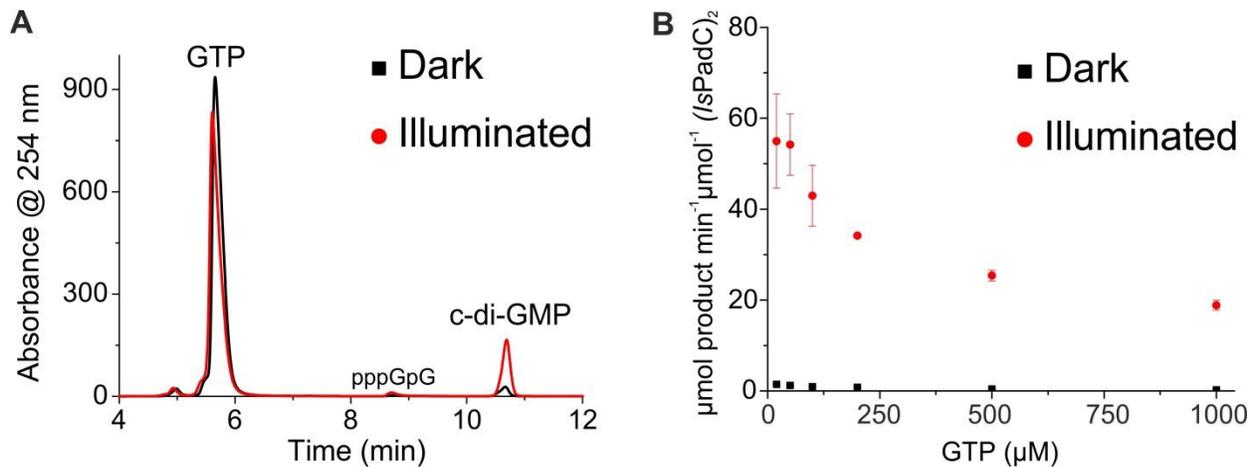


**Figure 28. HPLC-based assay of DGC activity.** (**A**) HPLC-based assay for the *in vitro* characterization of light-regulated enzymatic activity in *Is*PadC. The retention times of GTP, pppGpG, and c-di-GMP are indicated. (**B**) Initial rates of c-di-GMP production by *Is*PadC at different GTP concentration in dark and under illumination. The characteristic GTP concentration dependence of initial rates of c-di-GMP formation is indicative of substrate inhibition, which is observed for both light- and dark-state activities. Product formation was quantified in triplicate for several reaction times and the standard deviation of individual points contributed to the error estimation of the linear fit that is used to calculate the initial rate of product formation. The standard error of the estimate from the linear regression is shown as error bar for each GTP concentration.

If all those PadCs have a same arrangement of functional domains in common, they share their major discrepancy in the sensor-effector linker region (Appendix Fig. 1). To that extent, we decided to look closer at the functional prerequisites of the sensor-effector linker region. The respective deletion of 4 residues (1 helical turn; *Is*PadC$^{\Delta514-517}$) and 7 residues (2 helical turns; *Is*PadC$^{\Delta514-520}$) were assumed to conserve a relatively similar structural alignment between PSM and DGC, however, their reduced coiled-coil character obviously affected the stability of the coiled-coil dimer in its dark state conformation and prevented the inhibition of DGC activity in the dark (Fig. 29 and Appendix Table 2). The deletion of only 3 residues (*Is*PadC$^{\Delta514-516}$) apparently strongly disturbs the relative positioning of the DGC protomers leading to the preferential production of the linear pppGpG intermediate. Interestingly, the coiled-coil – DGC

fragment of *Is*PadC (ccDGC) also shows only a preferentially low conversion of GTP to pppGpG indicating that the stand-alone coiled-coil dimer arrangement observed in the dark state structure of *Is*PadC is not sufficient to undergo the full catalytic motion of GTP turnover to c-di-GMP (Appendix Table 2).
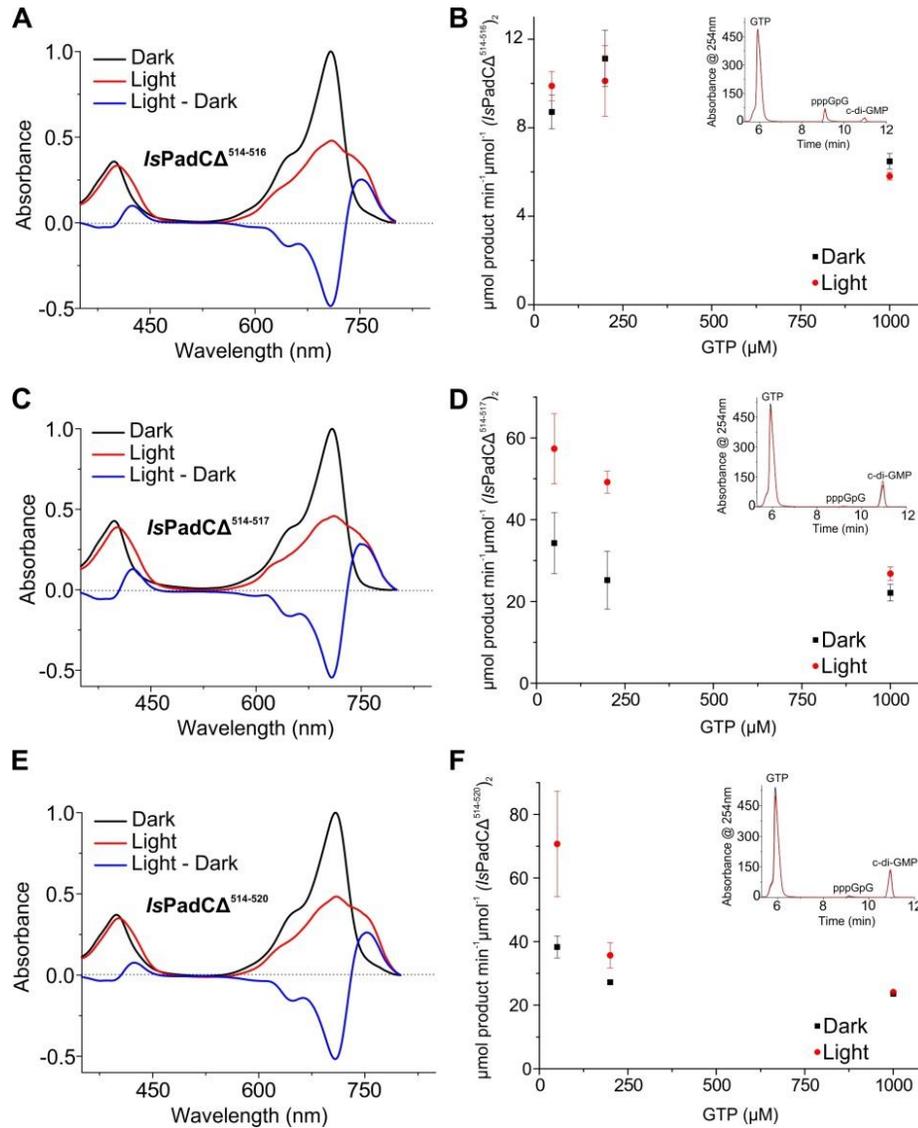


**Figure 29. Spectroscopic and kinetic characterization of *Is*PadC deletion variants.** UV-Vis spectra of different variants in their dark-adapted states (Pr, black lines) and after red light illumination (Pfr, red lines). The constructs shown correspond to deletions of 3, 4 and 7 residues in the coiled-coil region; *Is*PadC$^{\Delta 514-516}$ (**A**), *Is*PadC$^{\Delta 514-517}$ (**C**) and *Is*PadC$^{\Delta 514-520}$ (**E**), respectively. Panels **B**, **D**, and **F** show the results of the kinetic characterization of *Is*PadC$^{\Delta 514-516}$ (**B**), *Is*PadC$^{\Delta 514-517}$ (**D**), and *Is*PadC$^{\Delta 514-520}$ (**F**) obtained from the HPLC analysis of GTP conversion at different GTP concentrations. Product formation was quantified in triplicate for several reaction times and the standard deviation of individual points contributed to the error estimation of the linear fit that is used to calculate the initial rate of product formation. The standard error of the estimate from the linear regression is shown as error bar

for each GTP concentration. The insets in all panels show a representative chromatogram of the HPLC analysis revealing the preferential formation of the linear pppGpG intermediate for $Is$PadC$^{\Delta514-516}$, while the other constructs show $Is$PadC-like c-di-GMP formation.

## 2.1.2. The *Is*PadC full-length structure reveals the importance of the sensor-effector linker element

From several expressed and purified PadC homologs, $Is$PadC which also features the highest dynamic range of c-di-GMP production was so far the only construct to actually crystallize as full-length protein in the dark-adapted state. After the optimization of the crystallization conditions in hanging drop vapor diffusion setup at 289 K equilibrated against a reservoir solution containing 0.1 M bis-Tris (pH 5.5), 0.1 M ammonium acetate, and 17 % polyethylene glycol 10,000 it was possible to obtain crystals diffracting to 3.0 Å revealing a parallel dimer conformation of the full-length $Is$PadC confirmed in solution by Small Angle X-ray Scattering (SAXS) analysis (Fig. 30). UV/Vis absorption spectroscopy on $Is$PadC crystal confirmed the Pr-state of the chromophore in its protein environment (Fig. 31).

We addressed the relative energetic contribution of individual domains to the full-length parallel $Is$PadC dimer interface by performing a PISA analysis (126). Interestingly, the calculated $\Delta G^{diss}$ values, which indicate the free energy of assembly dissociation, revealed that a major contribution to the overall stability of the assembly is provided by the coiled-coil linker (residues 501-528, $\Delta G^{diss} = 8.0$ kcal mol$^{-1}$). The sum of additional contacts provided by the PSM altogether (residues 8-500, $\Delta G^{diss} = 8.8$ kcal mol$^{-1}$) contribute only to a similar extent to the overall stability of the phytochrome dimer (PAS-GAF-PHY-coiled-coil, $\Delta G^{diss} = 22.6$ kcal mol$^{-1}$).
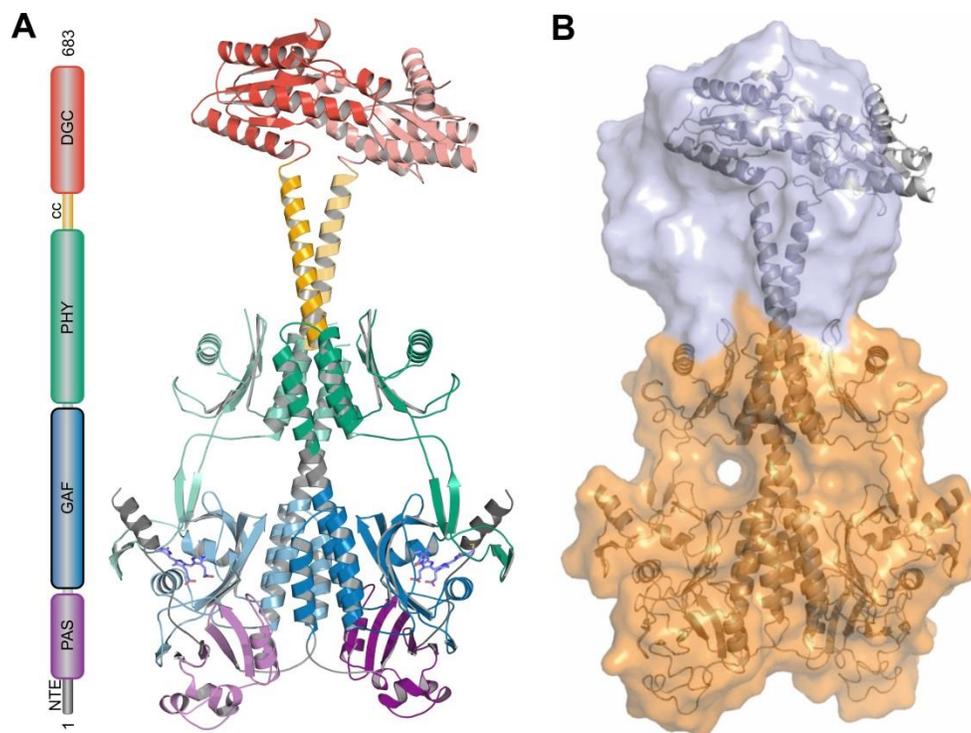
**Figure 30. The dark-adapted full-length structure of *Is*PadC.** (**A**) Dark-adapted full-length structure of *Is*PadC. Individual domains are colored in dark grey, violet, blue, green, orange and red for the N-terminal extension (NTE), Per–ARNT–Sim (PAS), cGMP phosphodiesterase–Adenylyl cyclase–FhlA (GAF), phytochrome-associated (PHY), coiled-coil (cc) and diguanylyl cyclase (DGC) domains, respectively. (**B**) SAXS-based structural model for the *Is*PadC dimer in its dark-adapted $P_r$-state. The surface represents the conformational space sampled by the GGDEF domains in the 7 best structures according to the fit between the experimental and back-calculated SAXS data (*cf.* Appendix Fig. 2). Structures are aligned to the PAS-GAF-PHY domains.

The molecular arrangement of the domains in the crystal lattice was obtained by molecular replacement using search models for PAS-GAF, PHY, and GGDEF domains from *R. Palustris Rp*A3015 (PDB 4S21 (79)), *R. Palustris Rp*BphP3 (PDB 4R70 (79)), and *Pseudomonas aeruginosa* WpsR (PDB 3I5B (127)), respectively. The search models were adapted to the *Is*PadC sequence using PHENIX Sculptor (128). A much better defined electronic density of the assembly was obtained by solving the phase problem using a combination of the molecular replacement information with single-wavelength anomalous dispersion data obtained from crystals of selenomethionine-labeled *Is*PadC. The final refined structure shows that two molecules of *Is*PadC are present in the asymmetric unit that are arranged in a parallel dimer architecture with a central two-fold symmetry axis slightly perturbed by a kink of the GGDEF dimer (Fig. 30A).
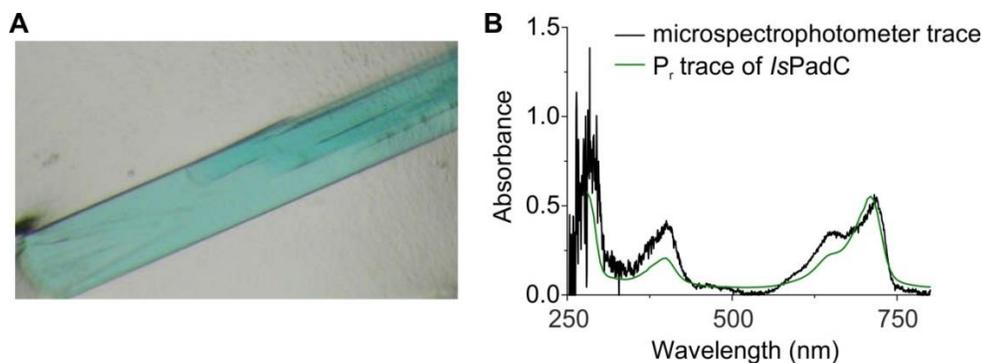
**Figure 31. *Is*PadC full-length crystal and spectrum.** (**A**) Typical plate elongated crystal of full-length *Is*PadC in dark-adapted conditions. The green coloration comes from the covalent binding of the BV chromophore. (**B**) UV/Vis spectrum of *Is*PadC crystal recorded with a CCD-based microspectrophotometer. The crystal spectrum in dark overlaps the dark-state solution spectrum of *Is*PadC.

The same arrangement of a parallel dimer was also obtained for the molecular structure of *Is*PadC with its GGDEF domain truncated (PSMcc) which diffracted to higher resolution (2.4 Å). PSMcc crystals resulted from the proteolytic degradation of full-length *Is*PadC in the same crystallization condition (Fig. 32).
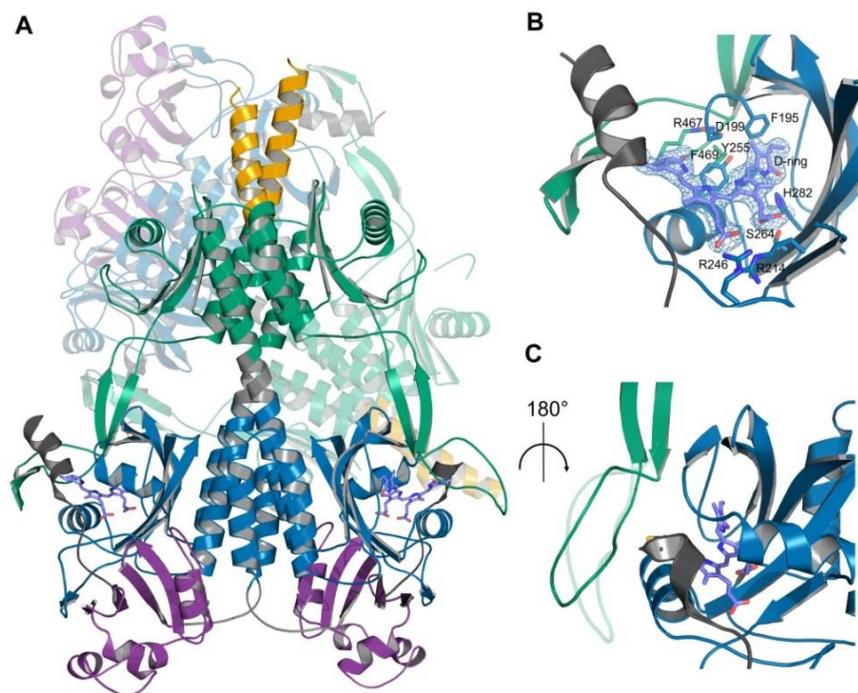


**Figure 32. Characterization of the *Is*PadC PSMcc variant.** (**A**) Cartoon representation of the asymmetric unit content of the crystallized PSMcc variant (residues 1-528). The dimer of chains C and D is shown in forefront with

individual domains colored in dark grey, violet, blue, green, and orange for the NTE, PAS, GAF, PHY, and coiled-coil domains, respectively. The biliverdin cofactor is shown as light blue stick model. Chains A and B in the background are shown in transparency. (**B**) Close up view of the biliverdin binding pocket of the higher resolution PSMcc structure. The $2F_o-F_c$ electron density map contoured at $1\sigma$ around the cofactor is shown as light blue mesh. Functionally relevant residues are shown as stick models. (**C**) Close up view of the PHY tongue region of the protomer D. The view is rotated relative to panel A to better visualize the loss of the β–hairpin extension in the tongue region of chain D. Superposition with the conformation of the tongue region observed in all other chains (chain B shown in transparent mode for comparison) highlights the characteristic differences. These structural rearrangements are induced by crystal contacts with a neighboring molecule and support the conformational flexibility of the tongue region.

Interestingly, the PSMcc structure was solved as a head-to-tail tetramer and one protomer adopted a different β-hairpin PHY-tongue structure with fewer contacts between the PHY-tongue looped tip and the NTE region (Fig. 32).

Several structural characteristics observed for deposited Pr-state phytochrome structures (29, 34, 39, 56, 81, 82) are also observed for the *Is*PadC structure. Notably, the extended N-terminal helix (NTE) featuring the covalent attachment site of the BV and continuing to the PAS domain through the typical figure-of-eight knot (129) from the GAF domain interacts with conserved hydrophobic contacts with the PHY-tongue region adopting the β-hairpin structure typical of Pr-state phytochromes. Moreover, the conformation of important conserved residues in the vicinity of the BV co-factor in its 15*Z* configuration support the Pr-state conformation of the molecule (39, 85) (Fig. 33A). Overall the global assembly of the PSM closely resembles the Pr-state structures of other phytochrome PSMs with, however, a rather extended dimer interface at the central helical spine connecting the GAF to the PHY domains, and at the PHY regions that are directly linked to the coiled-coil linker region (Fig. 33B). The C-terminal GGDEF domain is placed in an orientation where the half active sites are situated at the interface of the GGDEF dimer (Fig. 33C). This DGC dimer arrangement resembles the one in the complex with an unproductive GTP analogue (PDB 4H54) (130).
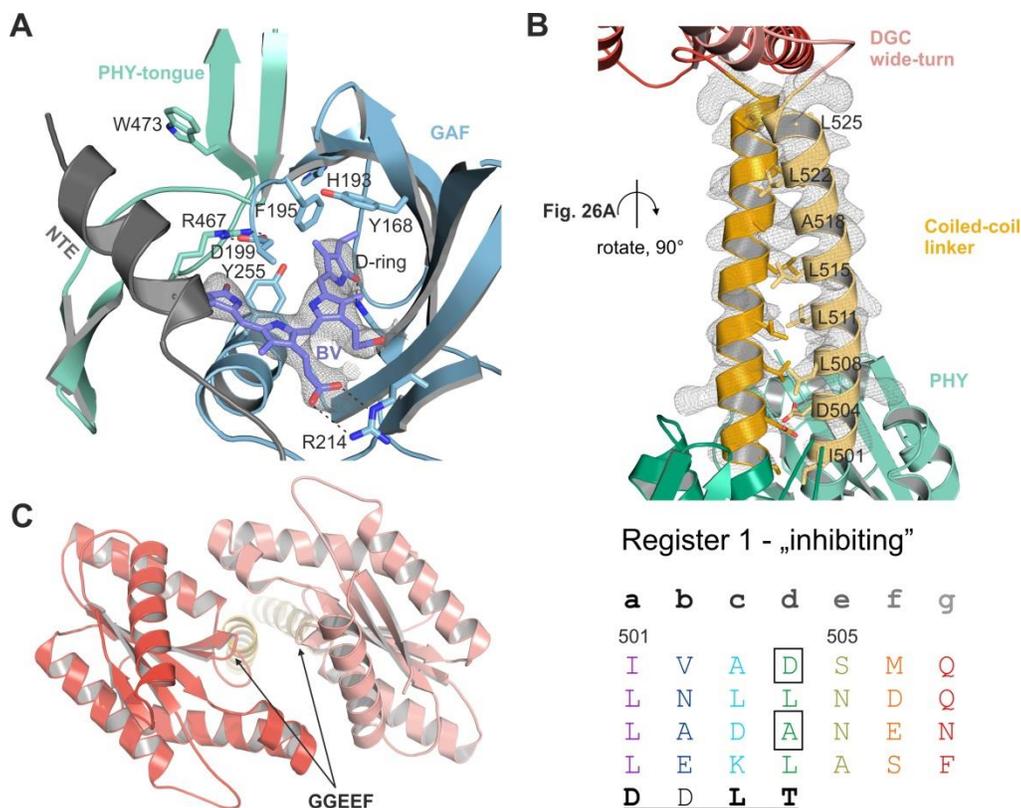
**Figure 33. Structural details of the *Is*PadC full-length structure.** (**A**) Close-up view of the BV chromophore Pr-state environment. The BV and important surrounding residues are represented in sticks. The *2F_o-F_c* electron density map contoured at 1σ around the chromophore is shown as light gray mesh. (**B**) Close-up view of the sensor-effector coiled-coil linker. The *2F_o-F_c* electron density map contoured at 1σ along the residues 501-528 is shown as light gray mesh. Schematic representation of the *Is*PadC coiled-coil linker sequence and succession of heptad repeats are shown below. Letters from *a* to *g* represent one full heptad repeat. Destabilizing residues in the coiled-coil sequence are boxed. The conserved DXLT motif of the DGC wide turn is underlined. (**C**) Top view of the DGC dimer. The arrows indicate the loop of the active site containing the GGEEF sequence.

For the first time, this full-length structure allows to appreciate the coupling between a phytochrome and the DGC effector. From residue Leu501 onwards, a typical coiled-coil signature extends over 3.5 heptad repeats and ends at the conserved wide-turn structure preceding the helix α0 of the GGDEF domain (Fig. 33B).

To look at a productive dimer of DGC bound to GTP in the context of our full-length PadC, we soaked *Is*PadC crystals with GTP and solved the structure by molecular replacement using individual domains of *Is*PadC as search models (Appendix Table 3). Interestingly, the

binding of GTP results in an increase of symmetry within the DGC dimer. This also displaces the coiled-coil dimer arrangement slightly (Fig. 34).
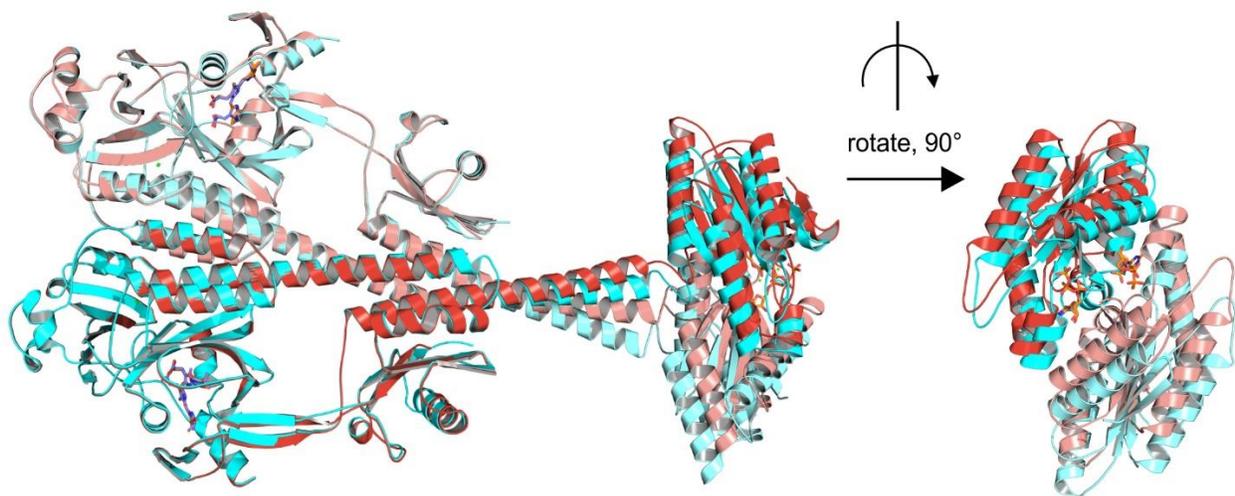


**Figure 34. Effect of substrate binding on the overall architecture of *Is*PadC.** Superposition of the full-length *Is*PadC structure obtained after GTP soaking (colored in red) with native *Is*PadC (colored in cyan). The two substrate molecules bound at the interface of the DGC dimer are shown as orange stick models. The superposition reveals a virtually identical PSM dimer (rmsd = 0.4 Å for 840 aligned Ca atoms) and only the linker region and the DGC dimerization show characteristic rearrangements upon GTP binding (rmsd = 2.0 Å for 306 Ca atoms of the dimeric DGC assembly) that are essentially restricted to a more pronounced bending of the coiled-coil linker coupled to a dimer rearrangement of the GGDEF domains that does not affect individual DGC domains (rmsd of 0.4 Å for 140 aligned Ca atoms in chain B).

Considering the relatively long distance between the C3 ribose hydroxyl group of one GTP and the alpha phosphate group of the other GTP that are involved in the formation of the first phosphodiester bond leading to the pentacoordinate phosphoryl transition state (pppGpG), it is most probable that other DGC dimer configurations are needed for the formation of the c-di-GMP and the release of the second PPi (*cf.* Fig 4). However, considering the high GTP concentration needed for soaking and the observed GTP inhibition of PadC activity, the GTP bound DGC dimer could also reflect a substrate inhibited DGC dimer.

Due to its position as a molecular bridge between the PSM and the DGC dimers, the coiled-coil linker has an important contribution to the relative placement of the two DGC moieties in respect to each other. Regarding the truncation of residues within the coiled-coil sequence from one residue up to two helical turns, it is obvious that not only the length but also

the composition of the coiled-coil linker dramatically influences its stability and promotes variations in the dynamic range of DGC photomodulation (Fig. 30 and Appendix Table 2).

Since all of these results indicate a fine-tuned dynamics at the coiled-coil linker region, we employed the method of Hydrogen-Deuterium eXchange coupled to Mass Spectrometry (HDX-MS) to analyze and compare the global conformational dynamics in dark condition and upon red light illumination of the full-length *Is*PadC.

## 2.1.3.    Conformational changes upon light signal transduction in full-length *Is*PadC

As described in the corresponding introductory paragraph, HDX-MS provides information on deuterium accessibility to the amide bond hydrogen thus on the stability of secondary structure elements of macromolecules. This enables the assessment of structural fluctuations and changes in conformational dynamics occurring between two different functional states (131), and hence has proven specifically powerful to study photoactivated proteins (132) but also many other systems. A preliminary trypsin digestion of *Is*PadC full-length in dark and under illuminated conditions indicated an increase of conformational dynamics upon light illumination notably within the coiled-coil linker region (Fig. 35).
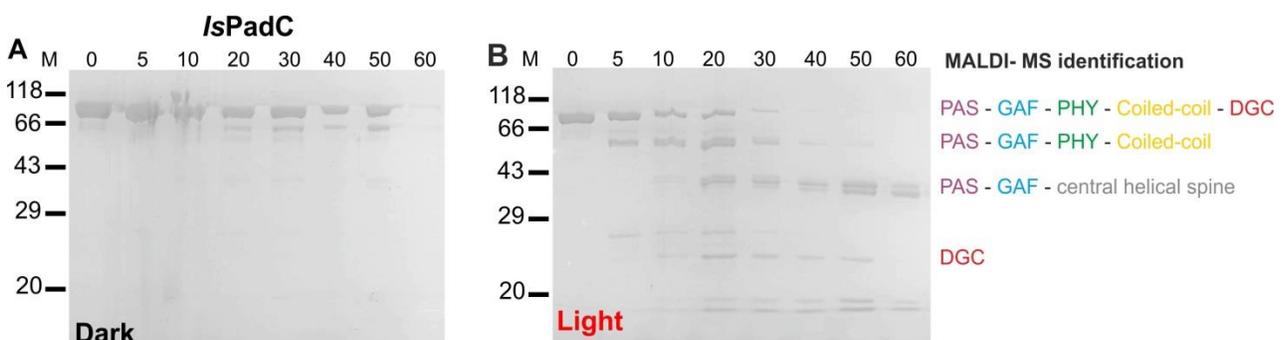


**Figure 35. Time course of tryptic digests of *Is*PadC, *Ma*PadC and *Ts*PadC under dark- and light-conditions.** Panels A-B show Coomassie-blue stained 12 % SDS-PAGE gels with 5 µg of *Is*PadC loaded per lane after trypsin incubation in the dark over a total of 60 min (**A**) and with constant red light illumination over the same time frame (**B**). A ratio of trypsin to *Is*PadC corresponding to 1/100 was used. An initial peptide identification has been performed by MALDI-MS and the band assignments are represented.

Indeed, the HDX-MS analysis of *Is*PadC as well as of the PSMcc truncation in both dark conditions and under constant red light illumination (630 nm, 0.7 mW cm$^{-2}$) showed an increase of deuterium uptake in several regions (Fig. 36 and Appendix Fig. 3-4). The rapid increase of deuterium incorporation at the PHY-tongue region notably for peptides comprised between positions 440-470 observed upon red light illumination supports the unfolding of the PHY-tongue region β-hairpin upon 15*E* isomerization of the BV (Fig. 36A-B). Interestingly, PHY-tongue peptides comprised within position 470-480 show less deuterium exchange that is indicating of subtle stabilization of this part of the PHY-tongue. Moreover, the rather broad deuteration profile of peptides in this part of the PHY-tongue indicates different deuterium incorporation kinetics within the same region in *Is*PadC homodimer originating potentially from different structural ensembles (Appendix Fig. 3-4). The increase of deuterium incorporation in GAF regions surrounding the BV co-factor and containing important residues shown to be affected upon BV photoisomerization in other bacteriophytochromes, as well as in the region containing the PASDIP motif involved in the salt bridge with the Arg of the PRXSF PHY-tongue motif, supports an unfolding of the β-hairpin PHY-tongue region upon 15*Z* to 15*E* isomerization of the BV in *Is*PadC. Apparently, the unfolding of the PHY-tongue from its β-hairpin configuration releases structural restraints on the PSM dimer assembly with the net effect of increasing the conformational dynamics at the central helical spine region and at the associated coiled-coil linker reflected by the increase in deuterium incorporation for longer deuteration times in these regions (Fig. 36).
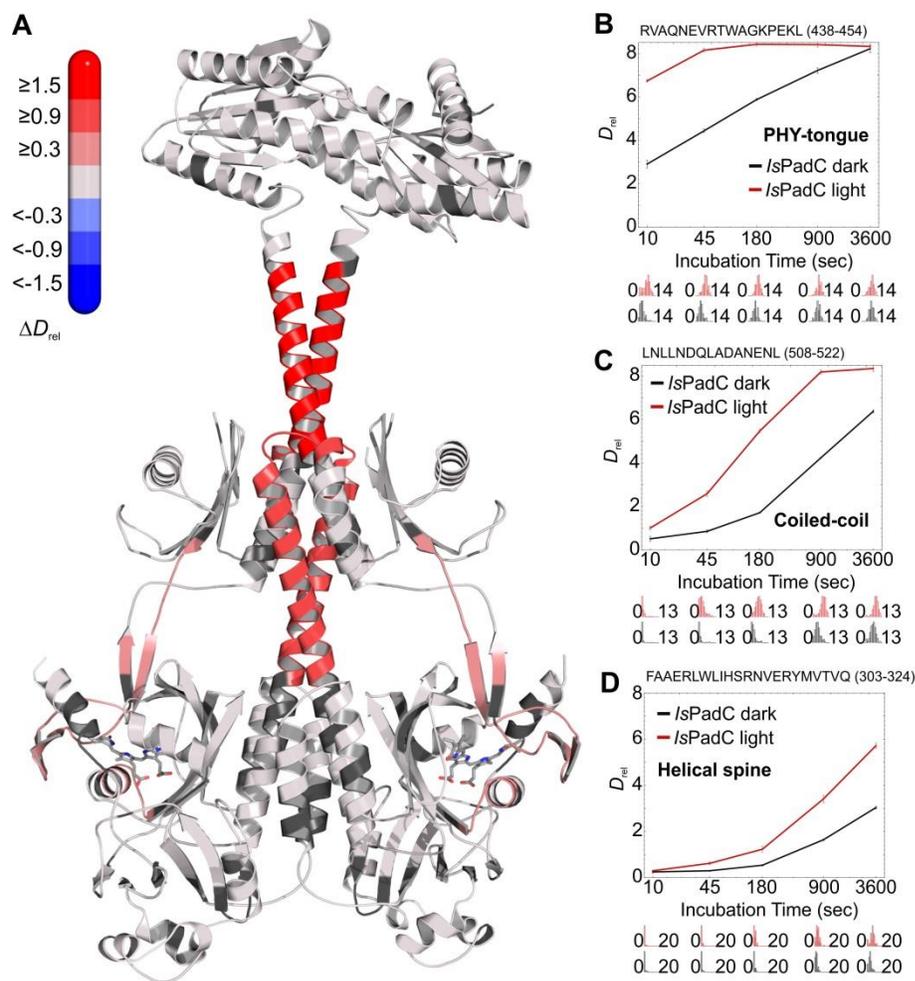
**Figure 36. Light-induced changes in conformational dynamics of full-length *Is*PadC observed by HDX-MS.**
(**A**) *Is*PadC structure colored according to changes in relative deuterium incorporation ($\Delta D_{rel}$) between light-state and dark-adapted *Is*PadC after 15 min deuteration ($D_{rel}$ of *Is*PadC$_{light}$ − $D_{rel}$ of *Is*PadC$_{dark}$). The scale bar in the top left corner indicates the changes in $\Delta D_{rel}$ with blue colors corresponding to reduced deuterium incorporation and red colors reflecting increased exchange of amide protons upon red light illumination. The biliverdin cofactor is shown as gray stick model. (**B-D**) Deuterium uptake curves of selected *Is*PadC peptides with $D_{rel}$ plotted against the deuteration time for light- and dark-state HDX-MS experiments. The lower parts show software estimated abundance distributions of individual deuterated species on a scale from undeuterated to all exchangeable amides deuterated. Deuteration plots for a peptide corresponding to the PHY tongue element (B), the PSM-DGC linker (C) and the internal helical spine linking the GAF and PHY domains (D). $D_{rel}$ values are shown as the mean of three independent measurements, and error bars correspond to the standard deviation. An overview of all analyzed peptides is provided in Appendix Fig. 4.

Interestingly, the same deuteration profile was observed for the corresponding regions as cited above for the PSMcc truncated variant (Appendix Fig. 3-4). This indicated that the secondary structure in the helical spine and coiled-coil regions are conserved, and suggested that the signal transduction is based on subtle conformational changes within those structural regions. Interestingly, a variant of *Is*PadC with the PHY-tongue region (residues 442-477) replaced by Ser-Gly loop (*Is*PadC$^{\Delta442-477::SG}$) showed an increase of DGC activity already in dark compared to the wild-type protein supporting the inhibiting effect of the Pr-state β-hairpin structure of the PHY-tongue element on the global molecular communication between phytochrome and DGC (Fig. 37). Moreover, the fact that the variant still feature some regulation of DGC activity between dark and illuminated conditions highlights the importance of the PSM central helical spine in the signal transduction mechanism and correlates the increase of conformational dynamics observed in this region upon illumination (Fig. 37B and Appendix Table 2).
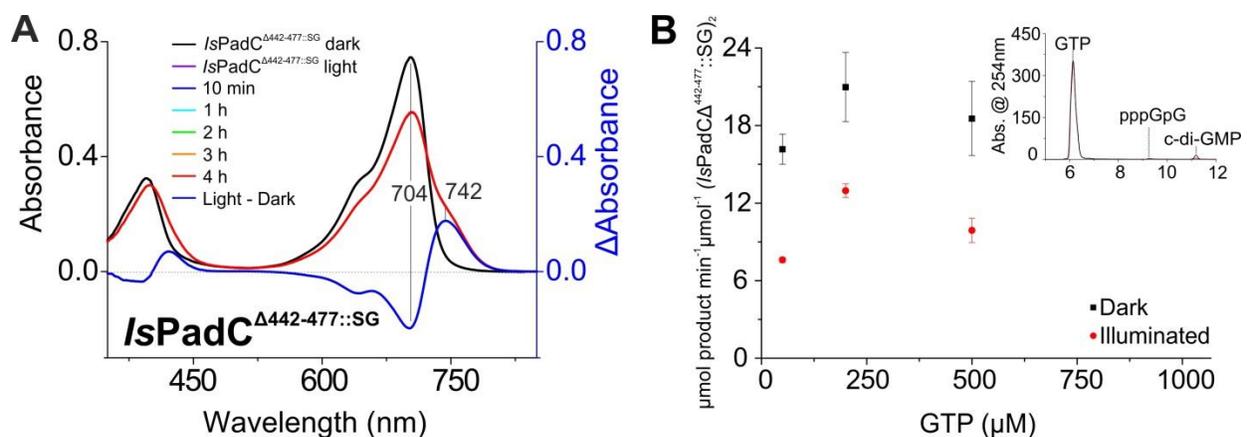


**Figure 37. *Is*PadC$^{\Delta442-477::SG}$ characterization.** (**A**) UV/Vis absorption spectrum of *Is*PadC$^{\Delta442-477::SG}$. Spectrum after various time points are recorded and show that 15*E* BV does not revert back to the Pr ground state. (**B**) Kinetic characterization of *Is*PadC$^{\Delta442-477::SG}$ obtained from the HPLC analysis of GTP conversion at different GTP concentrations. Product formation was quantified in triplicate for several reaction times and the standard deviation of individual points contributed to the error estimation of the linear fit that is used to calculate the initial rate of product formation. The standard error of the estimate from the linear regression is shown as error bar for each GTP concentration. The inset in the panel shows a representative chromatogram of the HPLC analysis.

By comparing our dark state conformational register in the coiled-coil linker dimer to a synthetic fusion described as constitutively active between a DGC and a typical GCN4 leucine zipper coiled-coil (127) (PDB 3I5B), we observed that another conformational register can be populated within the *Is*PadC coiled-coil linker and that could be associated to a positive

modulation of the DGC activity (Fig. 38A-D). The structural alignment of *Is*PadC and GCN4-DGC based on the conserved DXLT motif of the DGC domain reveals that the Leu501 in an *a* position of the first heptad repeat in *Is*PadC sequence is shifted to a *d* position in the GCN4-DGC sequence (Fig. 38 A-E). In order to confirm the existence of the register 1 (inhibiting register observed in the dark state structure) and the register 2 (stimulating register deduced from the alignment with the constitutively active GCN4-DGC fusion) within the *Is*PadC coiled-coil linker region, we replaced the destabilizing residues in *a* and *d* positions in the register 1 (D504L, A518L, resulting in the construct *Is*PadC$^{Reg1}$) and 2 (S505V, A526V generating the constructs *Is*PadC$^{Reg2}$) by hydrophobic residues (*c.f.* boxed residues in Fig. 38). Asparagine residues in *a* position of the stimulating register is assumed to stabilize coiled-coil dimer allowing however a certain conformational flexibility to the dimer (133). To observe the effect of their replacement by stabilizing hydrophobic interactions we also generated the *Is*PadC$^{Reg2.a}$ variant derived from the *Is*PadC$^{Reg2}$ variant with the Asn512 and Asn519 replaced by valines (Fig. 38).
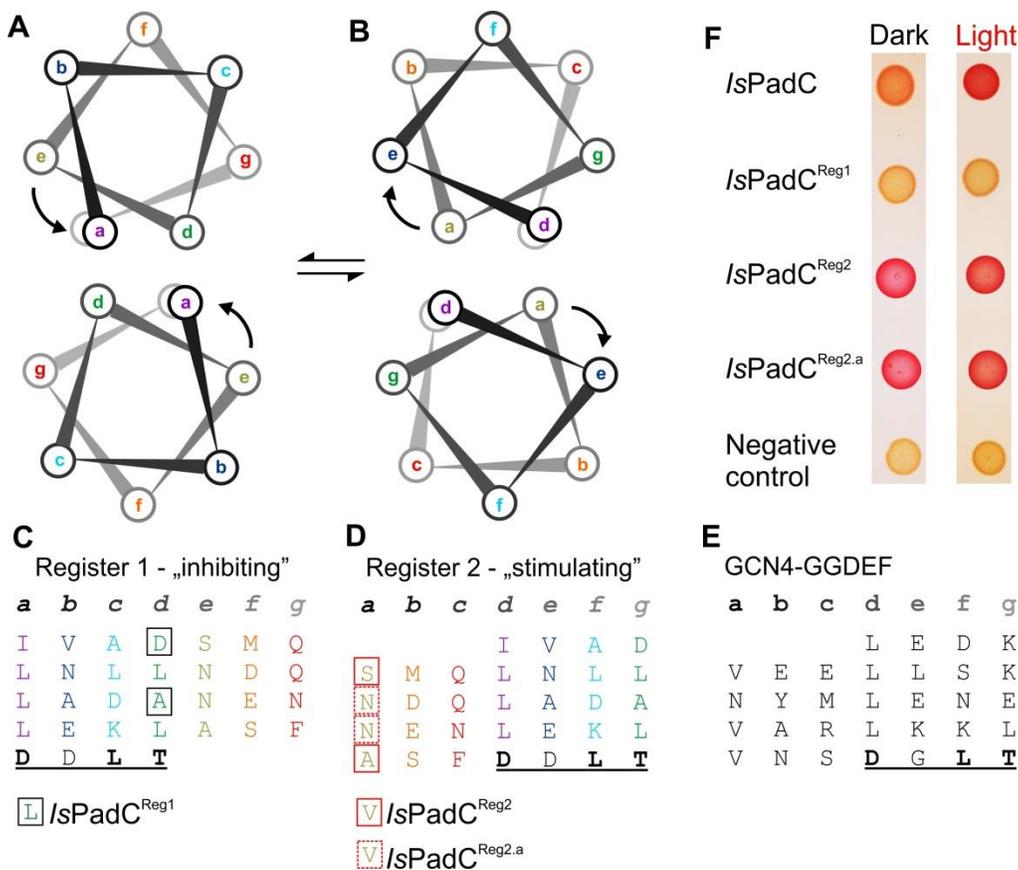


**Figure 38. Coiled-coil architecture of the sensor-effector linker.** (**A**) Heptad register observed in the crystal structure of *Is*PadC (Register 1 – inhibiting register). A conformational rearrangement of the heptad positions *e* to *a* within the coiled-coil populates the "Register 2 – stimulating register" architecture (panel **B**). Panels (**C**) and (**D**)

show the heptad units of the sensor-effector linker in register 1 and 2, respectively, rainbow-colored according to the heptad repeats of register 1. Coiled-coil destabilizing residues are boxed. The highly conserved DXLT motif of GGDEF domains is underlined. (**E**) For comparison, also the heptad units of a super-active, artificial GCN4-GGDEF fusion (127) are shown. (**F**) The assignment of active and inhibited states to register 2 and 1, respectively, is confirmed by the observed DGC activity in a cell-based screening system. Wild-type *Is*PadC shows the expected increase in DGC activity upon red light illumination, as seen by the red coloration of the cells. In contrast, an *Is*PadC variant stabilizing register 1 can no longer be activated upon illumination, whereas the register 2 stabilizing variants are constitutively active. *Is*PadC D504L, A518L (*Is*PadC$^{Reg1}$); *Is*PadC S505V, A526V (*Is*PadC$^{Reg2}$); and *Is*PadC S505V, N512V, N519V, A526V (*Is*PadC$^{Reg2.a}$).

An initial *in vivo* screening of DGC activity, which was subsequently confirmed by the *in vitro* quantification of c-di-GMP production for both purified variants, showed that the coiled-coil conformational register observed in the dark-adapted *Is*PadC structure (register 1) is the inhibiting register of DGC activity whereas the register 2 is the stimulating register of DGC activity (Fig. 38F and Appendix Table 2). Indeed, mutating the destabilizing residues Asp504 and Ala518 by leucine to generate *Is*PadC$^{Reg1}$ led to constitutively inactive variant while the mutation of Ser505 and Ala526 by valines to generate *Is*PadC$^{Reg2}$ led to a constitutively active variant featuring only a rather low 2-fold activation between dark and light conditions. The mutation of the remaining Asn512 and Asn519 by valines in *Is*PadC$^{Reg2.a}$ led to a constitutively active variant not photoregulated highlighting the importance of the Asn residues in *a* position of the stimulating register. The conformational dynamics of the PSM is therefore finely tuned to modulate the dynamics of coiled-coil register switching, enabling the modulation of DGC activity by influencing the positioning of the GGDEF catalytic residues and the dynamics of the GGDEF dimer. Obviously, within the PSM Pr/Pr homodimer, the stable β-hairpin conformation of the PHY-tongue restrains the conformation of the PSM dimer maintaining a stable inhibiting register at the coiled-coil linker that can only barely switch to the stimulating register resulting in some basal dark state DGC activity. However, upon illumination and isomerization of the biliverdin chromophore, important interactions between the PHY-tongue element and the GAF domain are destabilized with the net effect of relaxing the β-hairpin structure of the PHY-tongue. This unfolding event promotes an increase in conformational dynamics at the central helical spine that propagates to the PHY dimer interface enabling the switching of the coiled-coil conformational register towards the stimulating register resulting in DGC activation. Further support for this general mechanism description comes from a variant of *Is*PadC with the PHY-tongue region

deleted ($Is$PadC$^{\Delta442\text{-}477::SG}$) that features a substantial increase of dark state activity highlighting the inhibiting function of the β-hairpin conformation of the PHY-tongue in Pr-state (*cf.* Fig 37 and Appendix Table 2).

It results from our X-ray and HDX-MS approach complemented with our functional analysis of naturally occurring and synthetic variants of PadCs that the full-length systems can be assimilated to the violin model of enzyme regulation (134) where the conformational dynamics of all the functional domains are involved in a complex molecular cross-talk that drives the allosteric modulation of the output effector (Fig. 39).
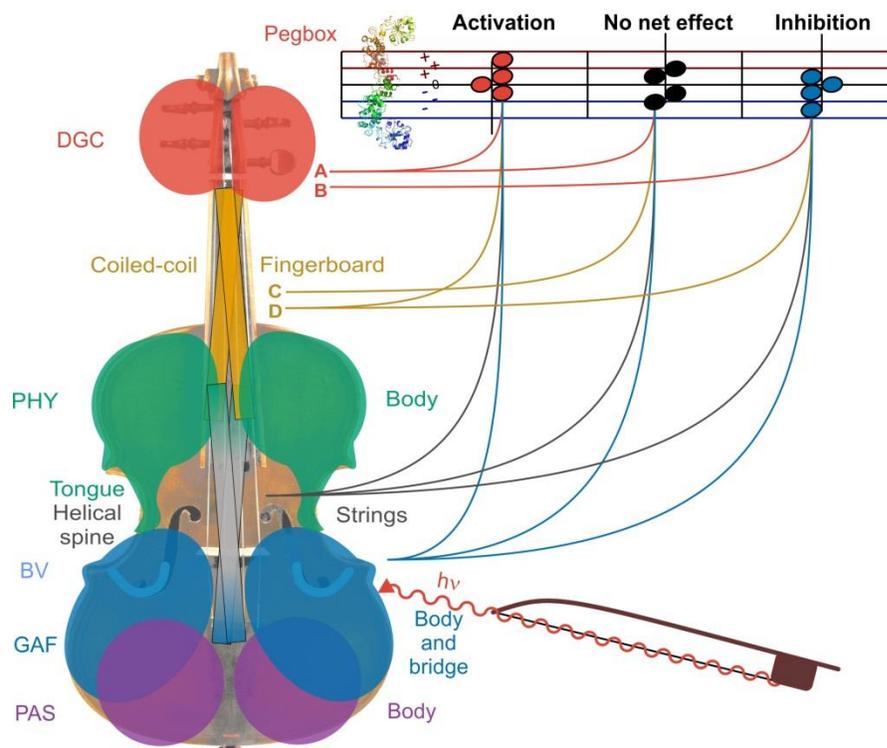


**Figure 39. Schematic model of signal integration pathways in phytochrome linked enzymatic effectors.** The characteristic structure of $Is$PadC and regulatory properties of closely related homologs suggest a model of signal transduction corresponding to the violin model (134). Instead of a linear cascade of structural changes resulting in enzyme activation, the conformational dynamics of the whole system define the population of functionally relevant states leading to either activation or inactivation with similar overall architectures (135). In the case of the phytochrome-violin the pegbox corresponds to the effector domain whose activity is tuned by the sensory module. Hitting the right chords on the enzymatic activity clef for stimulating GTP turnover is more complex than a specific actuation of the fingerboard, which would correspond to, for example, variation of its length (yellow lines and linkers C and D), and additionally depends on properties of the strings (gray), the shape of the violin body (blue) as well as the effector-pegbox (red; e.g. two different DGC constructs A and B). The characteristic structural changes

observed for various phytochrome structures, which had also been interpreted as specific light-induced rearrangements (39, 85), reflect the structural plasticity of phytochrome dimerization. The latter, in turn, enables the modulation of the conformational dynamics of the overall system and thereby allows complex, evolutionary fine-tuning of the body of phytochrome-violin to optimize the output functionality as required by each system. The absence of characteristic structural changes, such as a defined rotation or a separation of the coiled-coil linker, enables the realization of systems featuring activation or inactivation within the same molecular architecture.

## 2.1.4.    Discussion

The large repertoire of naturally occurring phytochromes and phytochrome-like sensors couple to various output functionalities is promising for the engineering of *in vivo* light switchable biological tools. From the first low resolution envelope of a full-length phytochrome coupled to an histidine kinase effector (136) to the recent structures of the dark and illuminated state of phytochromes PSM (39, 86), common elements of the structural changes involved in the light signal transduction have been revealed and appear to conserve interesting similarities with the *Is*PadC phytochrome structures presented here. The initial photoisomerization characterized by the 15*Z* to a 15*E* configuration change of the covalently linked bilin chromophore is the common first step in the signal transduction process. Local changes in residue rotamers close to the D-pyrrole ring flipping ~180° are also globally conserved among phytochromes (39, 80, 86). Notably the resulting destabilization of important interactions between the PHY-tongue element and the GAF domain seems to be conserved in canonical phytochromes (79, 80, 82, 83, 137) as well as in *Is*PadC as assumed from our HDX-MS analysis that showed an increase of conformational dynamics at the PASDIP GAF motif in parallel to the strong increase in deuterium incorporation at the PHY-tongue region. However, what is happening upon PHY-tongue unfolding and/or refolding on the global conformation and output module activation remains difficult to understand for previously characterized phytochromes due to the lack of information on full-length systems (39, 88). With our structural, functional and conformational dynamics analysis of full-length *Is*PadC, we can now conclude that the unfolding of the PHY-tongue region increases the conformational flexibility along the PSM dimer interface and notably at the PHY dimer interface that enables the coiled-coil linker to switch from a conformational register inhibiting the DGC activity to a conformational register stimulating the DGC activity. However, the reconstitution of *Is*PadC dark-adapted structural envelope from SAXS

measurements reveals a high flexibility of the DGC dimer indicating that the catalytic and inhibited conformations of the DGC dimer might be a complex sampling of various coiled-coil linker – DGC dimer conformations. Due to its unique position as a connector between sensor and effector module and its apparent major influence on modulating the conformational dynamics of the DGC dimer, the sensor-effector linker region that features a coiled-coil character in all PadCs homologs has a primordial role in the signal transduction mechanism. Interestingly, the *Is*PadC variant with 7 residues truncated in its coiled-coil region features a different photoresponse than *Ts*PadC, a PadC homolog featuring a linker 7 residues shorter than *Is*PadC. This revealed that not only the linker length but also the composition are fine-tuned in respect of the PSM dynamics and photoactivation properties, suggesting a coevolution of phytochromes and linker elements to achieve a specific regulation of DGC effectors. The full-length system responds to the violin model where PSM and DGC form two poles influencing each other (134). As some of our truncations in the linker element result in low dynamic range of activation and quite different basal enzymatic activity levels, it is obvious that a precise understanding of the molecular dynamics in this structural element is required to be able to rationally design phytochrome – enzymatic effector couples with high dynamic ranges of photoactivation suitable for optogenetics applications. A recently described mechanism of coupling between a blue light photoreceptor and a histidine kinase effector shows an interesting parallel with the linker mechanism described here (112). Also other protein functions coupled to coiled-coil linker like (138) mediating cargo binding to dynein by BicD, modulation of interactions of the cell cycle regulator Nek2 (139), and transmembrane signaling and integration by HAMP domain (140) show similar changes in conformational registers within their coiled-coil linker regions opening the reflection on potentially similar mechanism of dimeric effector modulation by coiled-coil linkers. However, the divergence of activation mechanism proposed already among the family of photoreceptors rather suggests an evolutionary adaptation of coiled-coil linker to specific sensor dynamics. Our new functional insights on the sensor-effector linker modulating DGC activity highlights the crucial coordination between phytochrome sensors and linker elements potentially explaining the challenge resulting from the fusion of a phytochromes sensor with different effectors.

In order to better understand the molecular coupling between phytochrome PSMs and coiled-coil linkers we further characterized the variant *Is*PadC$^{Reg2}$ assumed to promote the stimulating conformation of the coiled-coil linker.

## 2.2. Asymmetric activation of a homodimeric red light regulated photoreceptor

The results presented in this chapter are mainly based on the following publication:

# Asymmetric activation mechanism of a homodimeric red light-regulated photoreceptor

Gourinchas G[1], Heintz U[2], Winkler A[1*]

[1]*Institute of Biochemistry, Graz University of Technology, Graz, Austria.*
[2]*Max Planck Institute for Medical Research, Heidelberg, Germany.*
*Corresponding author Email: andreas.winkler@tugraz.at*

## 2.2.1. Heterogeneous chromophore environments upon photoactivation of the *Is*PadC homodimer

A striking hallmark of all PadC homologs characterized so far is represented by their illuminated steady-state spectra that show only a partial photoconversion of the initial Pr-state population to Pfr-state compared to other bacteriophytochromes featuring homogeneous Pfr activation (Fig. 40A) (32). This property is however not a peculiarity of PadCs homologs as several other phytochrome species also present similar illuminated spectra (30, 66, 141, 142). In order to confirm the steady-state equilibrium of *Is*PadC illuminated spectrum, we applied increasing light intensities on various *Is*PadC concentrations starting from a homogeneous Pr/Pr homodimer population and recorded the steady-state spectrum for each light intensity (Fig. 40B). From this experiment, it appears that Pfr contribution observed at 750 nm increases accordingly to the amplification of the photon flux density (Fig. 40B). However, the increase of Pfr contribution flattened at a certain intensity of light and led to a similar plateau for every protein concentration measured. This observation was really contrasting the previous characterization of other bacteriophytochrome photoactivations (29, 34, 80) that feature more Pfr contribution in their steady-state illuminated spectra. Notably, the recent crystallization of *Dr*BphP PSM in a symmetrically activated conformation featuring a Pfr state at each protomer of the PSM homodimer (39) would suggest that *Is*PadC behaves heterogeneously in term of photoconversion in its chromophore environments under constant red light illumination. The fact that the profile of Pfr-contribution increase correlates with the measured DGC activity for every light intensities confirmed that the heterogeneity in chromophore environments is relevant for the signal transduction and activation of the DGC effector (Fig. 40B).
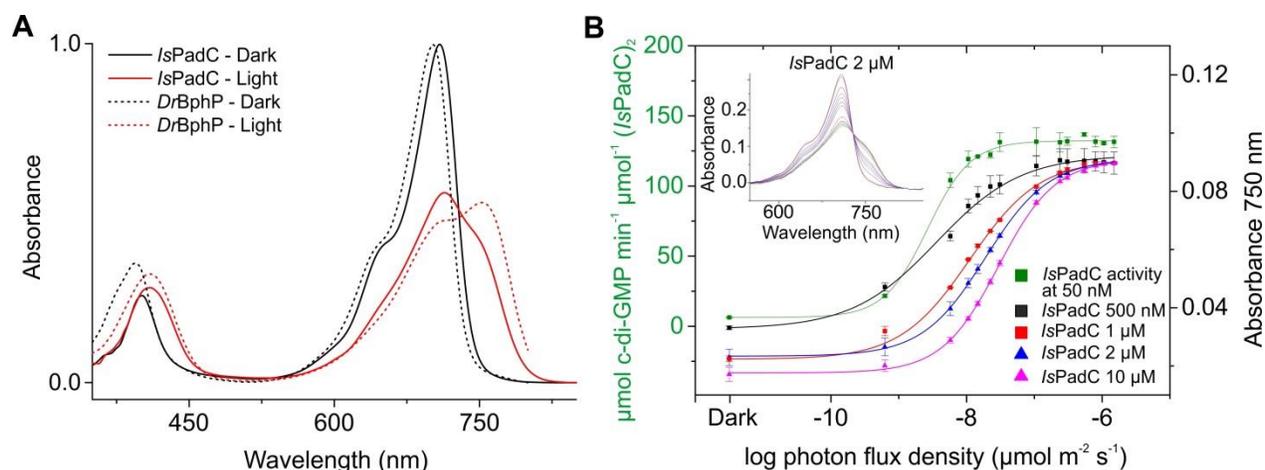
**Figure 40. Spectroscopic analysis of *Is*PadC Pfr conversion.** (**A**) UV/Vis characterization of *Is*PadC showing the dark state spectrum (dark) in comparison to one obtained directly after red light illumination (red). For comparison, spectra acquired under identical light and measurement setting using *D. radiodurans* phytochrome (*Dr*BphP) are shown. (**B**) Effect of red light photon flux density on enyzmatic activity and on the spectral contribution at 750 nm. To account for the different protein concentrations, all measurements were scaled based on the respective dark state absorption of the Pr peak. To account for reduced Pr contributions due to partial light activation by the measuring light at low protein concentrations, the 750 nm contributions were subsequently normalized to saturating red light conditions of the 2 µM sample. Full spectra at various flux densities are shown for the 2 µM *Is*PadC concentration in the inset.

In an attempt to better understand the heterogeneity in chromophore environments in *Is*PadC, the full-length homodimer was denatured under acidic conditions (143) (methanol/TCA 0.1 %) starting from the steady-state illuminated spectra to the Pr/Pr homodimer and the respective population of 15*Z* and 15*E* chromophores were followed over the time. Strikingly, we observed that more than 90 % of the chromophore population is present in 15*E* isomer under constant red light illumination (660 nm, 45 mW cm$^{-2}$) similarly to *Dr*BphP in the same conditions (Fig. 41A). Moreover, for the same protein concentration, decay rate constants and relative amplitudes of each phase of the thermal recovery in solution resemble those extracted from the 15*Z* isomer thermal recovery quenched by denaturation suggesting that every 15*E* chromophore under constant illumination contributes to the observed steady-state illuminated spectrum (Fig. 41B). Obviously, a similar population of 15*E* isomer of the population in the two bacteriophytochrome species under constant illumination did not led to a same illuminated spectrum. This might be due to different biliverdin environments of the two protomers

constituting of *Is*PadC homodimer potentially promoted by asymmetric photoconversion within the homodimer assembly (144, 145, 43).
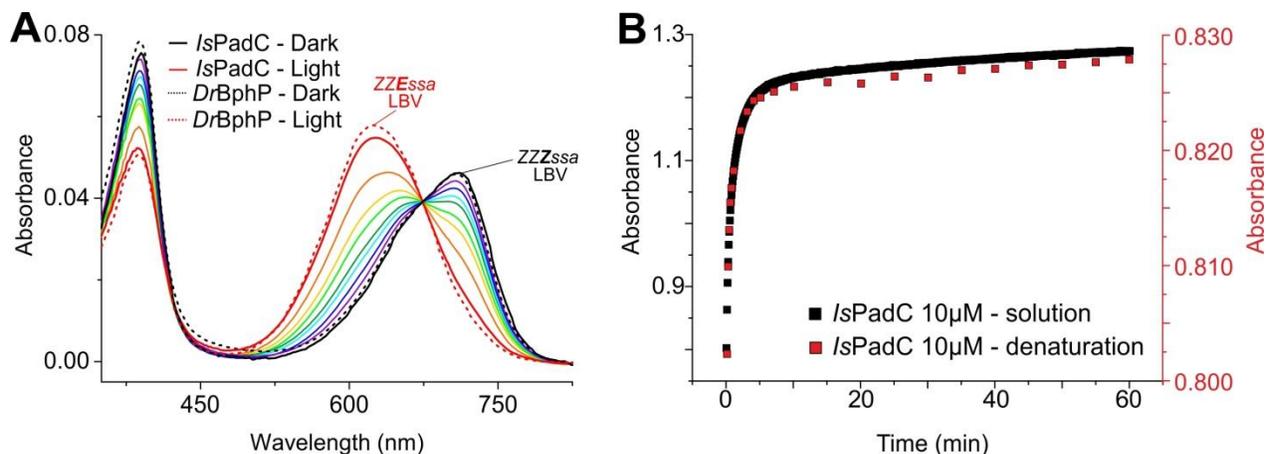


**Figure 41. BV isomer photoconversion.** (**A**) Thermal recovery of 15*E* biliverdin followed by denturation of 10 μM native samples at different time points. For comparison *Dr*BphP was denaturated in the dark and under illuminated conditions. (**B**) Changes in absorbance of the 15*Z* BV population (708 nm) during thermal recovery. Each time points are quenched by diluting *Is*PadC 10 µM 1:10 in methanol/TCA 0.1 %.

## 2.2.2.    The Pfr-state stability depends of the coiled-coil linker stability

The UV/Vis spectra of the three *Is*PadC coiled-coil register variants (*Is*PadC[Reg1], *Is*PadC[Reg2], and *Is*PadC[Reg2.a]) are very similar and present similar composite illuminated spectra with partial conversion to the Pfr-state visible by the absorption at 750 nm. However, the absorption maximum of the Q-band in dark condition is slightly shifted from 710 nm for *Is*PadC wildtype to 708 nm for *Is*PadC[Reg1] and *Is*PadC[Reg2] and to 707 nm for *Is*PadC[Reg2.a] that highlights the influence of the coiled-coil linker conformation on the chromophore environment (Fig. 42A-C). However, the generated mutations have a pronounced impact on the global dynamics of the full-length protein as observed by the striking difference of thermal recovery from the illuminated steady-state spectra to a Pr/Pr homodimer (Fig. 42D and Appendix Table 1). We observed that for the same protein concentration of 2 µM, the recovery kinetics of *Is*PadC and *Is*PadC[Reg2] can be accurately fitted with a sum of two exponential decay functions, while the recovery kinetics of *Is*PadC[Reg1] and *Is*PadC[Reg2.a] require the addition of a third exponential decay function reflecting a

phase with extremely slow recovery (Fig. 42D and Appendix Table 1) that highlights the critical influence of the coiled-coil linker conformation on the stability of the phytochrome photoactivated state. The analysis of the recovery rates and relative amplitude contributions to the overall amplitude extracted from the fitted thermal recovery data shows that recovery phases at 710 nm (recovery of Pr-state) and at 750 nm (depletion of Pfr-state) correlate for all constructs supporting the presence of spectrally closely related species in the illuminated steady-state spectra.
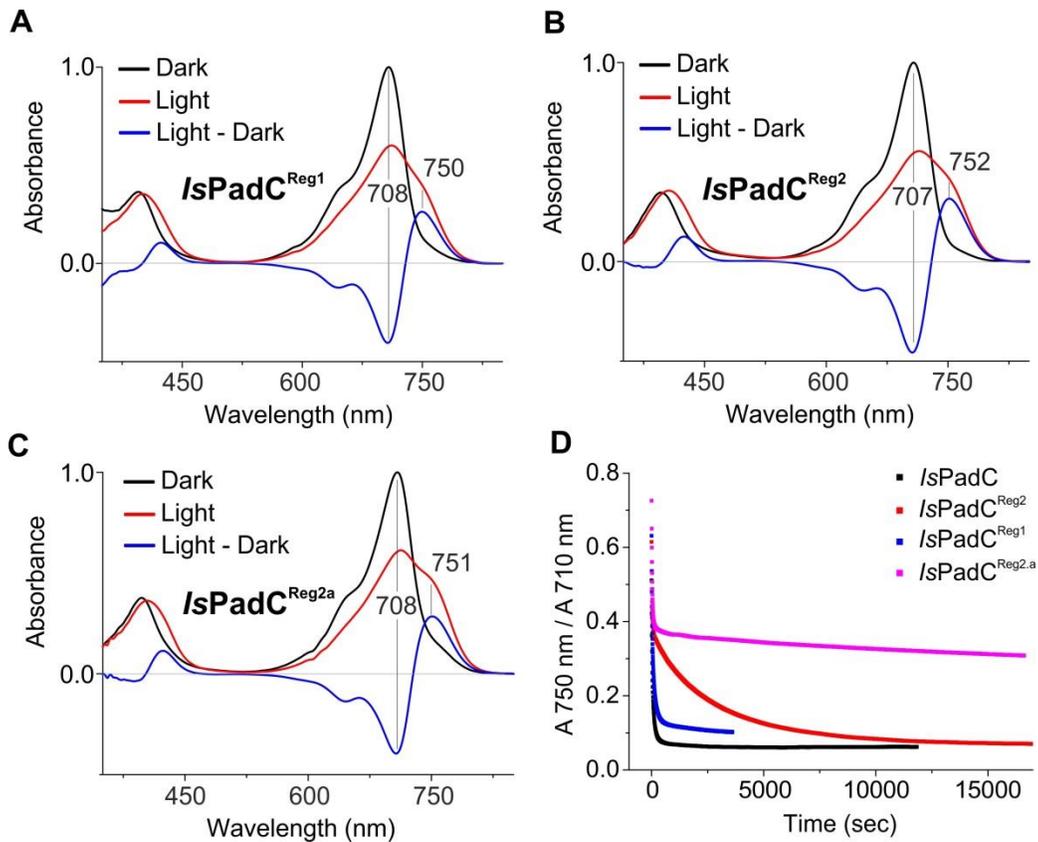


**Figure 42. Characterization of *Is*PadC coiled-coil register variants.** (**A-C**) UV/Vis spectrum of *Is*PadC[Reg1] (**A**), *Is*PadC[Reg2] (**B**), and *Is*PadC[Reg2.a] (**C**), respectively. A difference spectrum between illuminated steady-state and dark state is represented in blue. (**D**) Thermal recoveries of the various variants represented as the ratio of the absorbance at 750 nm over the absorbance at 710 nm.

What is particularly remarkable is the fact that the relative amplitude contribution of the fastest phase of the thermal recoveries remains constant for all the three variants (~36 %) while the two other phases feature strong variations in time constant and relative amplitude potentially meaning that the fastest phase is only barely affected by the difference of global dynamics in the

full-length molecule (Appendix Table 1). The contribution of one third of this phase in the overall amplitude of the recovery would suggest that the fastest phase of the recovery reflects the reversion to the Pr-state of a colored species with a lower molecular extinction coefficient than a classical Pfr-state while the other slower phases would reflect the reversion to Pr-state of the classical Pfr-state. From the literature, it appears that meta-R species are good candidates for this colored species with a lower extinction coefficient than Pfr and which are blue-shifted in the spectrum (43, 146, 147) suggesting that *Is*PadC features potentially an asymmetric Pfr/meta-R-like activated dimer in its steady-state illuminated spectrum.

Interestingly, a variant of *Is*PadC presenting a deletion of the PHY-tongue element (*Is*PadC$^{\Delta442-477::SG}$) features a composite UV/Vis spectrum of Pr and meta-R contributions in respect of the missing PHY-tongue element and previous observations of meta-R absorption (43, 97, 146) (*cf*. Fig. 37A). This spectroscopic asymmetry can also be seen by the fact that only parts of the chromophores are isomerized upon illumination and retained activated for days while the other chromophores radiatively decays as shown by the increase of fluorescence for this variant compared to the wildtype (Fig. 43A-B). Moreover, an initial characterization of *Is*PadC$^{\Delta442-477::SG}$ by HDX-MS revealed a bimodal deuterium uptake profile of peptides constituting of the GAF region containing the PASDIP motif that strongly supports two different biliverdin environments upon illumination of this variant (Fig. 43C). Interestingly, this asymmetry in the variant maintains the conformational dynamics increase in the coiled-coil linker already present in dark conditions in comparison to *Is*PadC wildtype (Fig. 43D). This increase of conformational dynamics in the linker region in dark conditions for the *Is*PadC$^{\Delta442-477::SG}$ variant correlates with a higher DGC activity and supports the inhibitory effect of the β-hairpin conformation of the PHY-tongue in Pr-state in the context of the full-length *Is*PadC (Fig. 43D and Appendix Table 2).
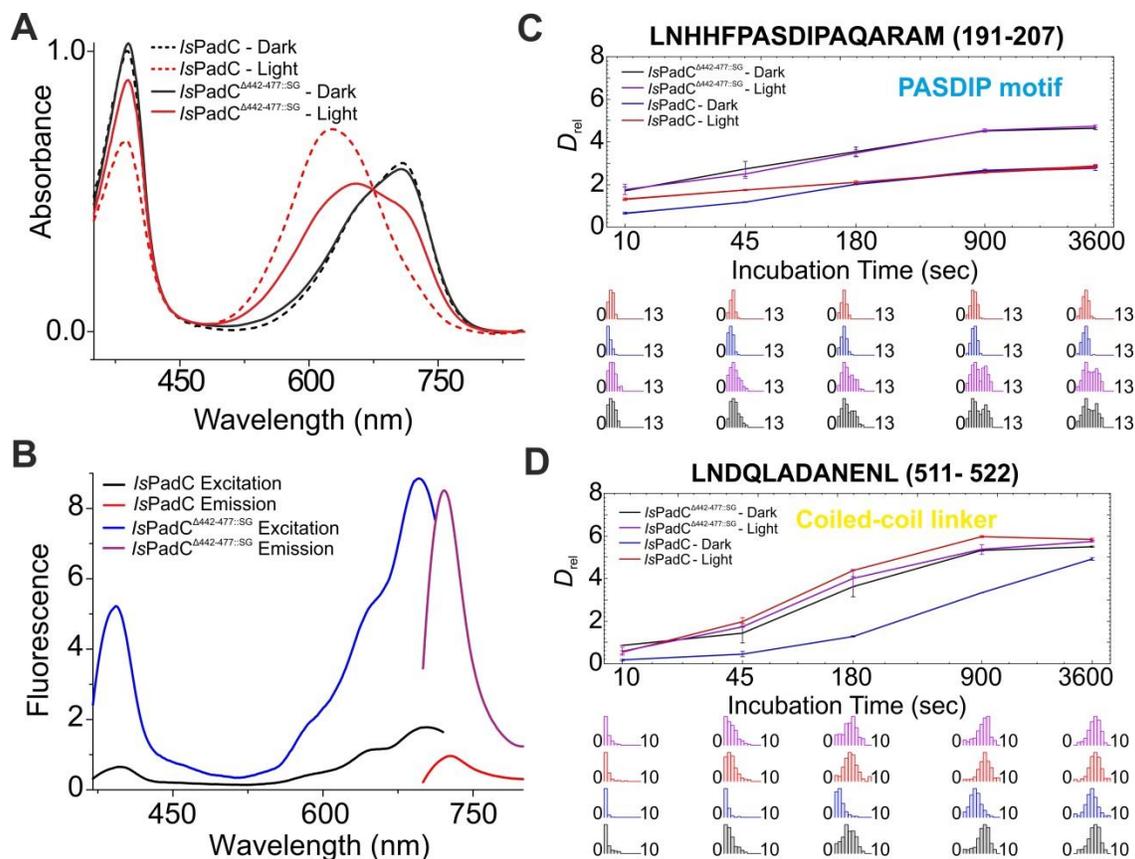
**Figure 43. Spectroscopic characterization of _Is_PadC$^{\Delta442\text{-}477::SG}$ variants and initial HDX-MS analysis.** (**A**) UV/Vis absorption spectrum of _Is_PadC$^{\Delta442\text{-}477::SG}$ variant denatured in acidic conditions (methanol/TCA 0.1 %) under both dark and light conditions (solid lines). The dashed lines represent the denaturation of _Is_PadC wildtype in the same conditions. (**B**) The 15$E$ isomerization of only half of the BV in _Is_PadC$^{\Delta442\text{-}477::SG}$ correlates with an increase in fluorescence at identical concentrations of proteins. (**C-D**) Comparison of deuterium uptake profiles of selected common peptides between _Is_PadC and _Is_PadC$^{\Delta442\text{-}477::SG}$. A peptide in the GAF region containing the PASDIP motif (**C**), and a peptide in the coiled-coil linker (**D**), respectively. The peptide containing the PASDIP motif (**C**) shows a bimodal deuterium uptake in both dark and light state in line with two different chromophore environments within the _Is_PadC$^{\Delta442\text{-}477::SG}$ homodimer.

Obviously, in the context of the full-length _Is_PadC, the PSM dimer interface promotes inter-protomeric interactions triggering allosteric effects on the chromophores environments that enable the photoconversion from Pr to Pfr-state in one protomer whereas restraining the photoconversion of the other protomer to an intermediate state between Pr and Pfr that in regard of our analysis would rather resemble a meta-R-like species. Meta-R species can however be differentiated into different meta-R intermediates species featuring slightly different protein environments and protonation states of the biliverdin (43, 146). Therefore, further experiments

described in the outlook of this thesis are necessary to elucidate the conformational state of *Is*PadC meta-R species present in the illuminated spectra.

## 2.2.3. The full-length structure of *Is*PadC$^{Reg2}$ reveals an asymmetrically activated conformation

The previously characterized *Is*PadC$^{Reg2}$ variant featuring high DGC activity in both dark and light conditions has been crystallized under dark conditions using the sitting-drop vapor diffusion method equilibrated at 293 K against a reservoir solution containing 0.1 M magnesium formate and 12 % (w/v) of PEG 3,350. After several rounds of crystal optimization (Fig. 44) plate-like crystals very similar to *Is*PadC dark state crystals, were obtained and diffraction data to 2.85 Å were collected.



ClassicsII H8
0.2 M Magnesium Formate
15% (w/v) PEG 3,350

ClassicsII H8
0.2 M Magnesium Formate
12% (w/v) PEG 3,350
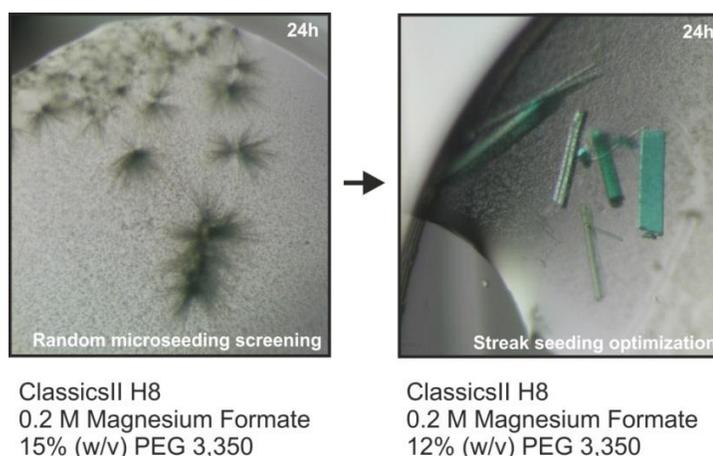
**Figure 44. *Is*PadC$^{Reg2}$ crystallization conditions optimization.** Promising crystal needles obtained after random microsseding screening (148) using thin crystal needles of *Is*PadC$^{Reg2}$ from another condition as seeds (left side) were in turn used for streak seeding 2 µl of the same condition with lowered PEG concentration (right side).

The structure was solved by molecular replacement using individual domains of *Is*PadC structure (32) and the coiled-coil linker and PHY-tongue regions were manually built within the refined electronic density. Even though the space group and the overall crystal lattice constants are similar for *Is*PadC and *Is*PadC$^{Reg2}$ structures including the conservation of similar protein contacts within the crystal (Appendix Fig. 5 and Appendix Table 3), we observed a pronounced

structural rearrangement of the overall *Is*PadC<sup>Reg2</sup> structure compared to the wildtype *Is*PadC structure. Notably, an increase of asymmetry is observed within the DGC dimer that features an opening of the dimer at one side that is obviously promoted by the transition of the coiled-coil linker to the expected stimulating register 2 conformation. Interestingly, the superposition of *Is*PadC<sup>Reg2</sup> structure on *Is*PadC dark-adapted and GTP bound structures shows that GTP binding promotes already the destabilization of the inhibiting register towards the stimulating register of the coiled-coil linker (Fig. 45).
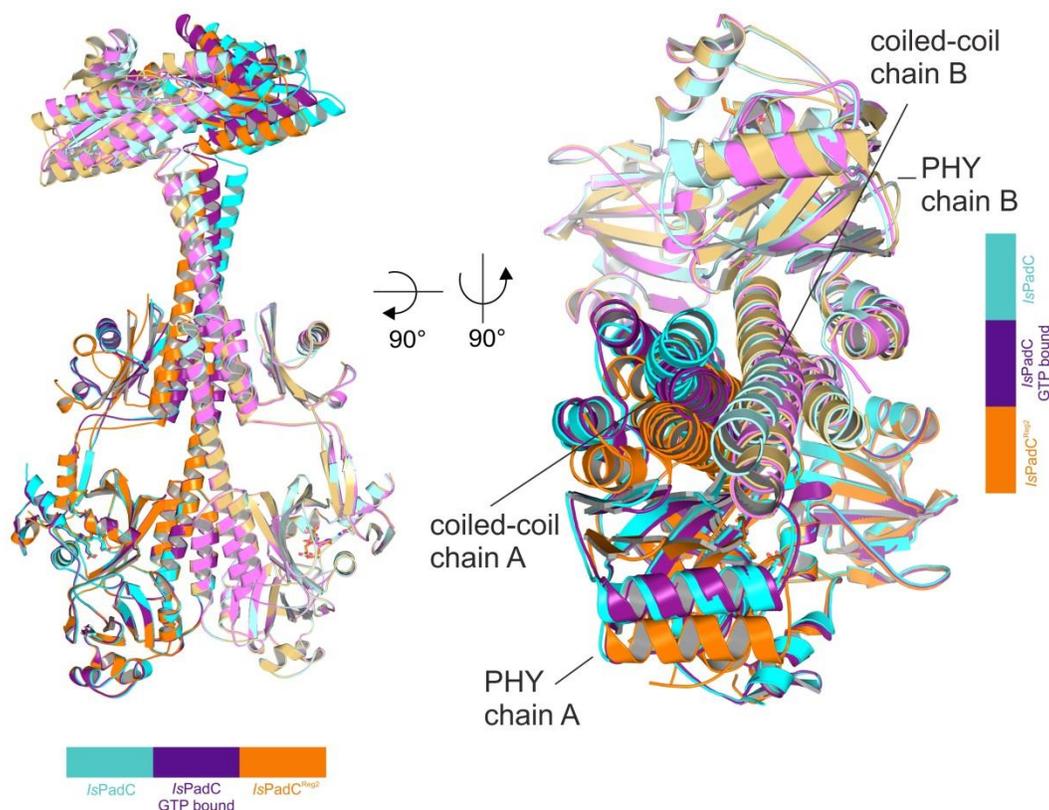


**Figure 45. Flexibility of the coiled-coil linker element in response to the molecular cross-talk between sensor and effector.** Superposition of *Is*PadC<sup>Reg2</sup> (orange) and *Is*PadC soaked with GTP (purple) structures onto the *Is*PadC dark-state (cyan) structure based on the PAS-GAF cores (residues 1–312) of the respective chains B. To focus on the coiled-coil linker conformation, the DGC dimer was removed for clarity from the rotated view. The superposition shows that GTP binding and the rearrangement of the DGC dimer feature structural changes that are similar to those of the stimulating coiled-coil linker conformation.

The asymmetry in the output module propagates to the PSM where a substantially different conformation of the dimer is observed presenting one PSM protomer in Pr-state superposing very well to the corresponding protomer of *Is*PadC (RMSD of 0.219 over 278 α-

carbon based on the superposition of PAS-GAF core of chains B (residues 1-312)), while the symmetry related protomer features a typical Pfr-state as previously observed for other phytochromes structures in activated state (39, 80, 85, 88) strongly differing from the Pr-state structure of the corresponding protomer in *Is*PadC structure (RMSD of 0.334 over 266 α-carbon based on the superposition of PAS-GAF core of chains A) (Fig. 46).
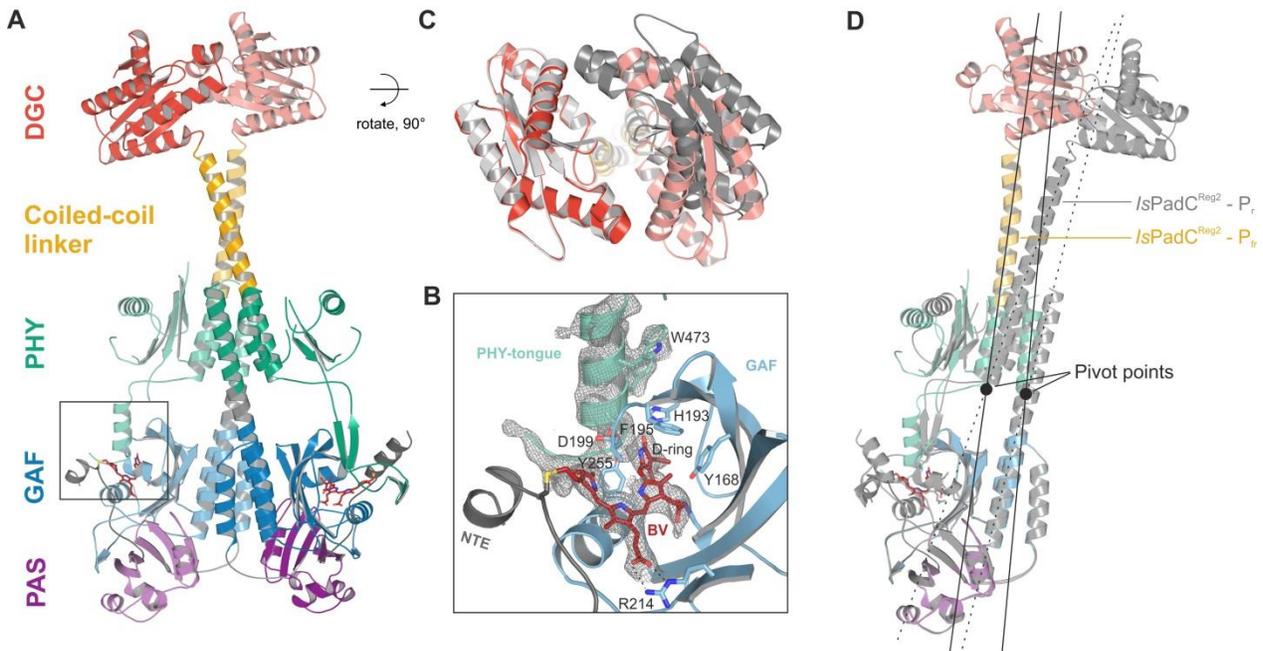


**Figure 46. Asymmetric activation of the *Is*PadC<sup>Reg2</sup> homodimer.** (**A**) Crystal structure of *Is*PadC[Reg2] in cartoon representation colored by individual domains. The two biliverdin chromophores are colored in dark-red and represented in sticks. (**B**) Close-up view of the biliverdin-binding pocket corresponding to the Pfr conformation, highlighting the secondary structure change of the PHY-tongue element. The $2F_o$-$F_c$ electron density map contoured at 1σ around the chromophore is shown as light gray mesh. Important residues are shown in sticks. For clarity residues 200-213 have been hidden from the view. (**C**) Top view of the asymmetric GGDEF dimer induced by the coiled-coil linker helix translation. The GGDEF dimers are superimposed by aligning protomer A of *Is*PadC[Reg2] to the corresponding protomer B (residues 529-683) and the moved copy of the GGDEF dimer is represented as gray cartoon. (**D**) Superposition of the Pfr protomer of *Is*PadC[Reg2] (colored as in A) and the Pr protomer of *Is*PadC[Reg2] (colored in light gray) based on the PAS-GAF core (residues 1-312) of the respective chain A and B. Arrows indicate the structural rearrangements associated with $P_{fr}$ formation. The pivot points and the associated lines show the hinge-like rearrangements at the central helical spine and the terminal PHY-helix extending into the coiled-coil linker.

A closer look to the chromophore environments shows that the chain A of *Is*PadC[Reg2] structure features an electron density fitting an α-helical conformation of the stretch of residues

465 to 477 while no electron density is observed for the residues 441 to 461 suggesting an unfolded state of this part of the PHY-tongue element (Fig. 47).
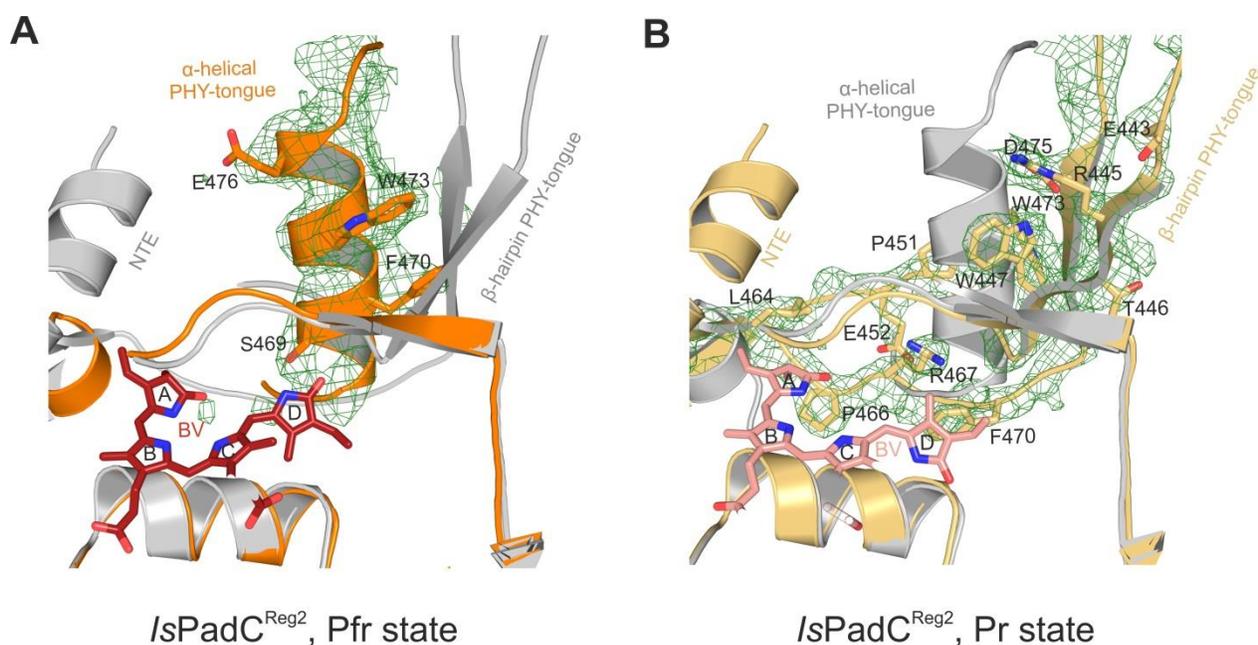


**Figure 47. Pr and Pfr models of the PHY-tongue regions of *Is*PadC^Reg2^.** (**A**), Polder-map (149) of the omitted PHY-tongue region (residues 441-479) of protomer A featuring a Pfr state. The PHY-tongue region adopting a Pfr environment is represented in cartoon and colored in orange. The PHY-tongue region of protomer B (colored in light gray) is superimposed based on the PAS-GAF core to highlight the difference between electron density in this region and the Pr conformation of the PHY-tongue. (**B**), Polder-map of the omitted PHY-tongue region (residues 441-479) of protomer B (in light orange) featuring a Pr state. The PHY-tongue region of the protomer A is superimposed based on the PAS-GAF core and colored in light gray. The polder-map is a difference electron density map contoured at 3σ and shown as green mesh.

The chromophore binding site of the protomer featuring the refolded PHY-tongue element reveals that the biliverdin in its 15*Z* configuration would not fit the electron density but would require the chromophore in its 15*E* configuration (Fig. 48).
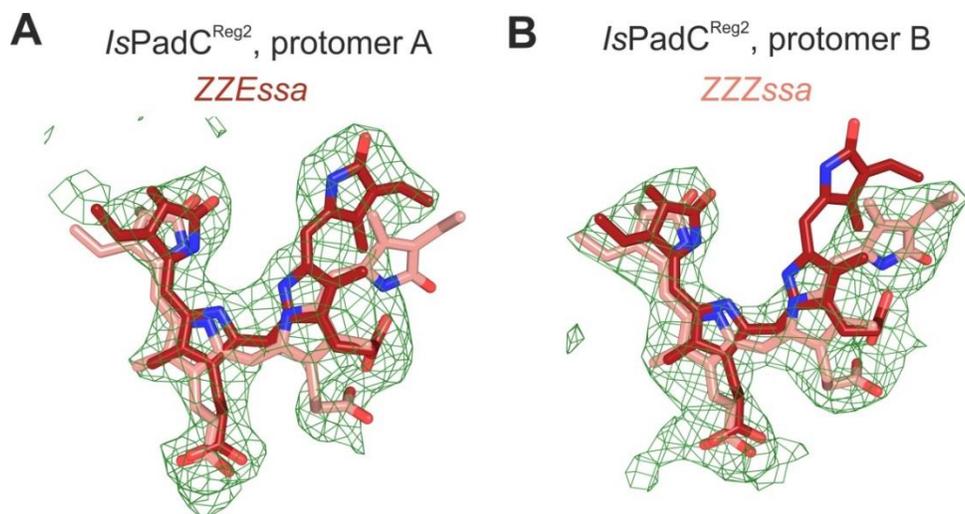
**Figure 48. Polder-maps generated for the cofactor binding site with chromophores omitted from both chains of *Is*PadC$^{Reg2}$ and *Is*PadC.** (**A-B**), Polder-map of chromophore binding sites in *Is*PadC$^{Reg2}$ chains A and B from left to the right. In each case the chromophore of the corresponding other protomer is superimposed based on the structural alignment of the PAS-GAF core (residues 1-312). The 15*E* chromophore is colored in dark red while the 15*Z* conformation is colored in light red.

The structural isomerization of the biliverdin is however only observed for this protomer (chain A) while the symmetry-related protomer (chain B) features the biliverdin in its classical 15*Z* conformer supporting an asymmetric activation of the homodimer (Fig. 46-48). In addition to the biliverdin isomerization and D-ring flip in one protomer, local residues around the D-ring (Tyr168, His193, and Phe195) rearrange similarly to corresponding residue positions in previously characterized Pfr-state structures of phytochromes further supporting the Pfr-state conformation of this protomer (Fig. 49). Correlating with the observed Pfr/Pr heterodimer the UV/Vis absorption spectra of *Is*PadC$^{Reg2}$ crystals in comparison to *Is*PadC scaled based on their Soret band absorption reveals an increase of Pfr contribution around 750 nm that clearly demonstrates that *Is*PadC$^{Reg2}$ crystals contain a mixture of Pr and Pfr population reflecting the refined Pfr/Pr heterodimer (Fig. 49C). Interestingly, a similar spectrum is observed for measurements performed at room temperature and showed a slight decrease of the absorption at 700 nm upon red light illumination of the crystals. This observation could potentially suggest that the 15*Z* chromophore can isomerize in the crystal upon illumination and potentially local residues around the D-ring rearrange to form a meta-R-like species (Fig. 49D). Attempts of collecting

diffraction data from *Is*PadC$^{Reg2}$ illuminated crystal were so far unsuccessful as the quality of the diffraction decreased quite strongly after illumination of the crystals.
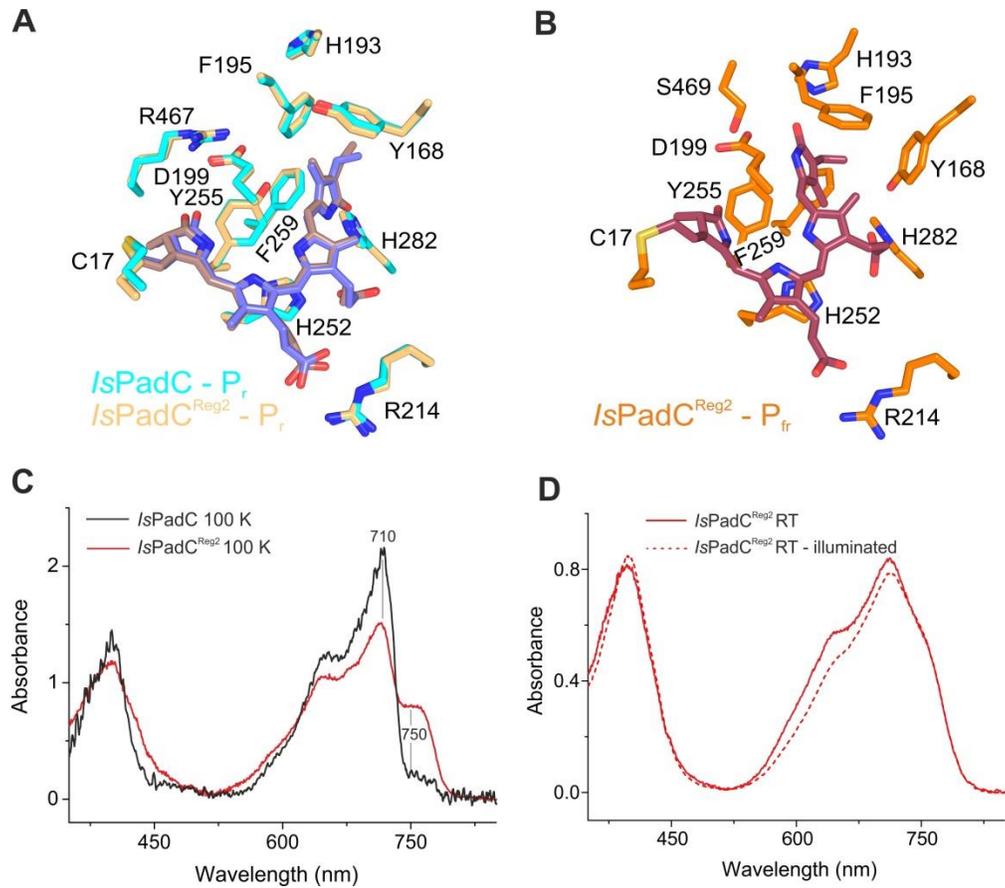


**Figure 49. Asymmetry in chromophores environments.** (**A-B**) Stick representation of residues around the biliverdin D-ring for the *Is*PadC P$_r$-state (cyan and BV in slate) superposed to the *Is*PadC$^{Reg2}$ Pr-state (light-orange and BV in light-red) (**A**) and the *Is*PadC$^{Reg2}$ P$_{fr}$-state (orange and BV in red) (**B**), respectively. The 15*E* configuration of biliverdin results in rotamer repositioning of Tyr168, Phe195, His193, Tyr255, and results in loss of the Asp199-Arg467 salt bridge between the GAF domain and the PHY-tongue. (**C**) A spectrum corresponding to *Is*PadC$^{Reg2}$ crystals is compared to a spectrum of *Is*PadC dark state crytals recorded at cryogenic temperatures (100 K). To account for preferred orientation of proteins in the crystal lattice various spectra at different positions of the protein crystal with respect to the light path were averaged (20 and 4 individual spectra have been averaged for *Is*PadC$^{Reg2}$ and *Is*PadC, respectively). The two averaged spectra are scaled such that they reflect the isosbestic point of P$_r$ and P$_{fr}$ states. (**D**) Spectrum of *Is*PadC$^{Reg2}$ crystal measured at room temperatrue (RT) in dark conditions (solid line) and after red light illumination (dashed line).

The net effect of the heterogeneous chromophores environments is the structural flexibility of the PHY domain linked to the refolded PHY-tongue element that results to a

pronounced rearrangement of the PHY dimer interface (*cf.* Fig. 46D). Since the C-terminal helix of the PHY domain of one protomer directly extends to the corresponding coiled-coil linker helix, the hinge-like motion of the central helical spine and PHY domain of the Pfr-state protomer drove the transition of the coiled-coil linker register to its stimulating register 2 in a motion that can be approximated by a quasi-translation along most of the coiled-coil structure (Fig. 50).
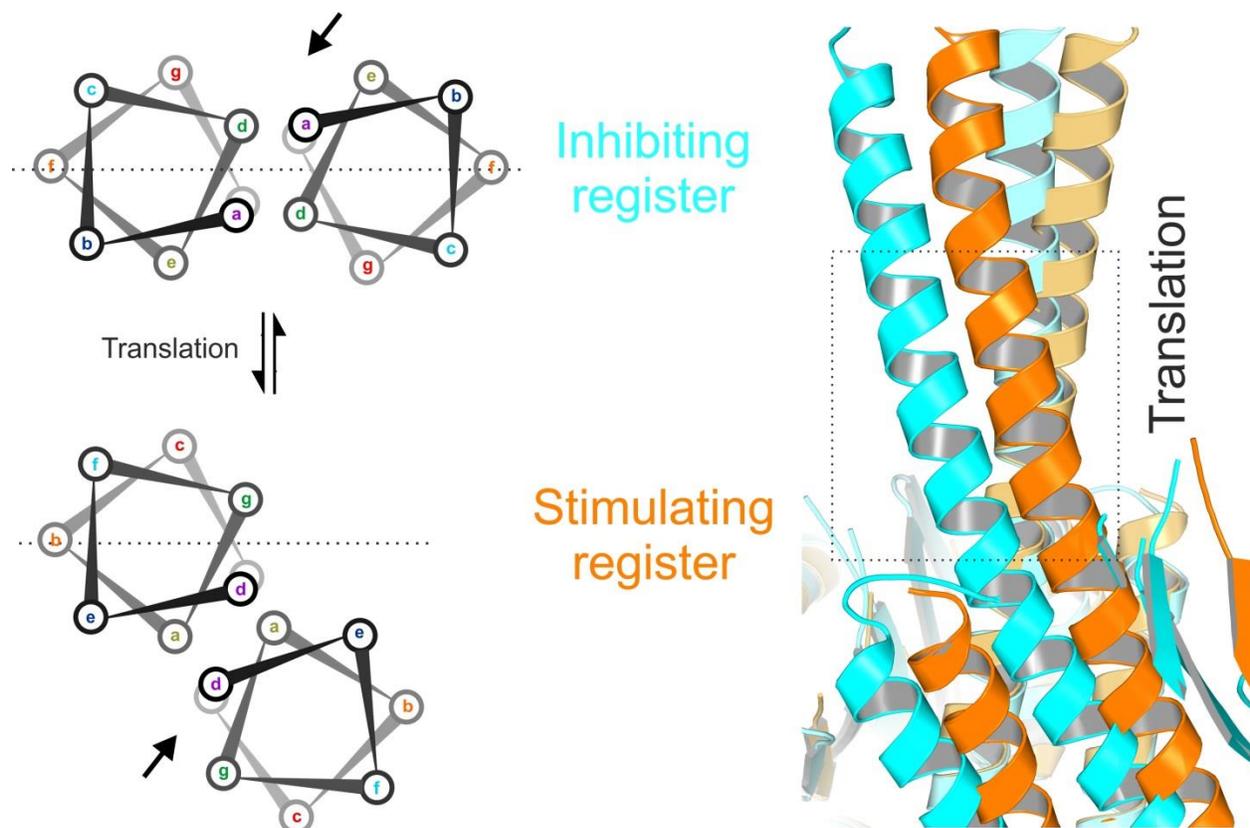


**Figure 50. Structural rearrangements in the coiled-coil dimer conformation upon asymmetric activation.** The stimulating register 2 conformation of the coiled-coil linker is populated by translation of one linker helix upon PHY-tongue refolding. Close-up view of the coiled-coil linker structure upon switching between the two conformational registers shows that the register 2 population is a complex hinge-like motion but can be approximated by a translation for most of the structural region.

This quasi-translational motion of the coiled-coil linker extends to the very end of the linker directly linked to the conserved DXLT motif of DGC (150) that has for consequence to modify the relative positioning of the typical wide-turn loop structure of the DGC containing the DXLT motif that directly influences the placing of the GGDEF catalytic sites. In this respect, the fact that *Is*PadC$^{Reg2}$ and *Is*PadC structures have been solved in a similar crystal lattice implies

that the conformational rearrangement observed for the DGC dimer are associated with the coiled-coil register transition due to the PSM asymmetric activation rather than being affected by lattice specific crystal contacts. As already mentioned, the quasi-translation motion of the coiled-coil linker led to an overall rotational motion of the associated DGC protomer generating a more opened DGC dimer at one side and a closing of the dimer at the other side (RMSD 0.370 over 141 α-carbon residues 529-683, of $Is$PadC$^{Reg2}$ chain B superimposed over $Is$PadC chain B) (Fig. 51).
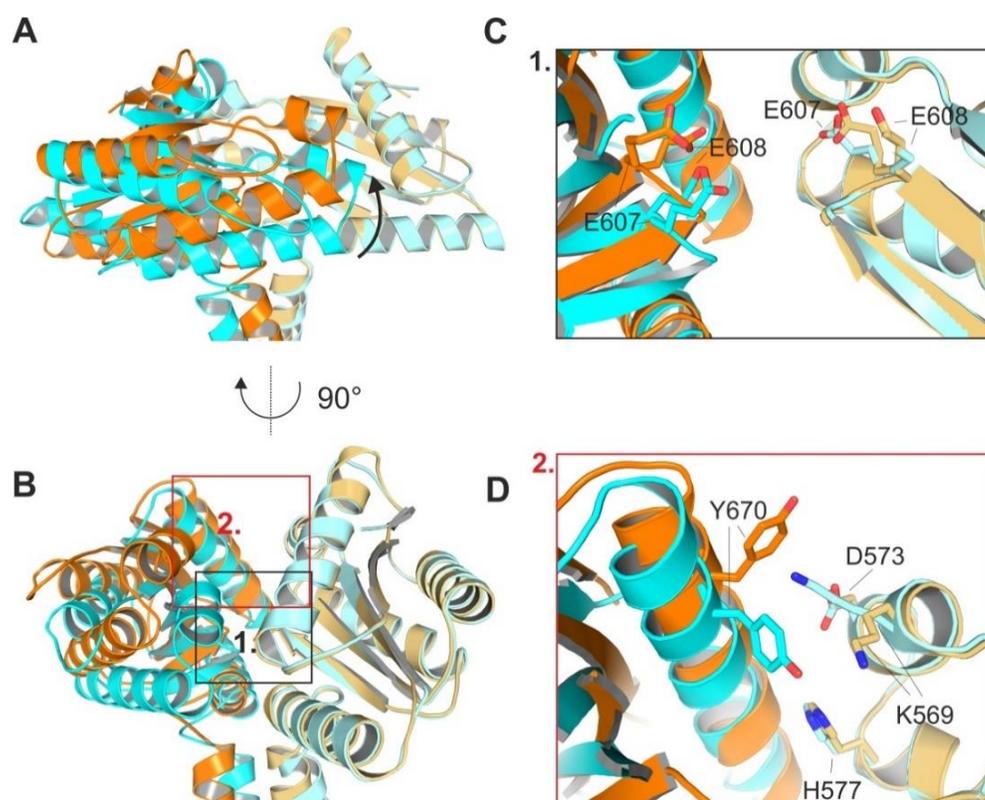


**Figure 51. Allosteric effects of the coiled-coil rearrangement on the conformation of the DGC dimer.** (**A**), Structural alignment of the $Is$PadC$^{Reg2}$ (orange) and the $Is$PadC (cyan) DGC dimers based on superposition of their P$_r$ state protomers (chains B) from residues 529 to 683 (RMSD of 0.370 Å over 141 α-carbons). (**B**), The rotated view shows the rotational displacement of the DGC dimer due to the associated coiled-coil linker helix translation. (**C**), The close up view of the GGEEF site shows that, upon coiled-coil linker register switching and rearrangement of the DGC dimer, the relative positioning of the two substrate coordinating sites is changed to potentially facilitate initiation of the catalytic cycle. (**D**), Close up view of the interface between helices α1/α2 of one DGC protomer and helix α4 of the other protomer. Important residues at the interface are highlighted in stick representation.

In the absence of a substrate or product bound in the active site of $Is$PadC$^{Reg2}$ structure it is difficult to attribute a signaling state to this new DGC dimer conformation. However, the fact that $Is$PadC$^{Reg2}$ is already much more active in dark might suggest that the open dimer conformation of DGCs either increases the affinity of GTP for the binding site and/or places the two GTP moieties in an energetically favorable position to catalyze the first phosphodiester bond formation. Moreover, none of the deposited inhibited DGC dimers (127, 150–153) feature a similar conformation as $Is$PadC$^{Reg2}$ DGC dimer indicating that the observed DGC configuration is catalytically relevant. Notably, the new DGC dimer position the two GGDEF active sites in a way that the distance between them is slightly shortened in comparison to the wildtype structure (32) potentially facilitating the interaction between the C3 hydroxyl group of one GTP and the alpha phosphate group of the respective other GTP molecule that would generate a favorable environment for the production of the pppGpG intermediate (150, 154). Nevertheless, considering the conformational dynamics that require the condensation of two GTP molecules into one c-di-GMP molecule through the intermediate production of pppGpG, the DGC dimer observed in $Is$PadC$^{Reg2}$ represents only one dimer conformation over the various regulatory DGC dimers potentially sampled by the dynamics of the full-length system. The soaking of $Is$PadC$^{Reg2}$ crystal with GTP 10 mM led to rather poor diffracting crystals, nevertheless, crystals soaked overnight generated a diffraction ~3 Å (not deposited in the PDB database) that shows a similar DGC dimer as $Is$PadC dark-adapted state GTP soaked structure (PDB 5LLX)(32) (Fig. 52). The electron density for the coiled-coil linker and PHY-tongue regions are quite weakly defined suggesting some structural rearrangements in these regions upon GTP binding in the DGC dimer of $Is$PadC$^{Reg2}$. The poor quality of the electron density map in these regions is probably due to heterogeneous conformations in the crystal lattice upon GTP binding suggesting that GTP conversion might potentially happens during soaking of the crystal.
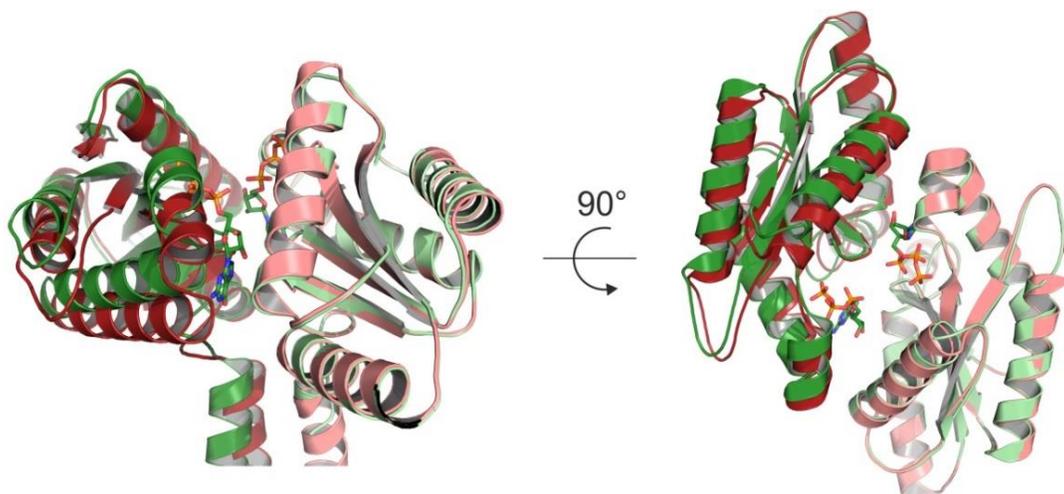
**Figure 52. Binding of GTP in *Is*PadC[Reg2] induces a re-closing of the DGC dimer.** Superposition of the DGC dimers from *Is*PadC (PDB 5LLX, (32)) (red) and *Is*PadC[Reg2] (green) GTP soaked crystals. GTP molecules of *Is*PadC soak structure are shown in sticks representation. The superposition is done on residues 528-683 of chains A. The binding of GTP in *Is*PadC[Reg2] induces a re-closing of the DGC dimer. Due to the rather weak resolution of the electron density map the two molecule of GTP have not be placed so far.

## 2.2.4. A complex sensor-effector cross-talk during signal transduction

To observe the effect of the differently stabilized coiled-coil variants of *Is*PadC, we analyzed the *in vitro* GTP conversion kinetics of *Is*PadC[Reg1], *Is*PadC[Reg2], and *Is*PadC[Reg2.a] by HPLC-based method. As observed for *Is*PadC wildtype, the coiled-coil register variants showed a slight inhibition of DGC activity upon increasing concentration of substrate, however, specific accumulation of pppGpG intermediate is not observed for those variants in comparison to previously characterized *Is*PadC variants with specific residues deletion in the coiled-coil linker (32) (Fig. 53).
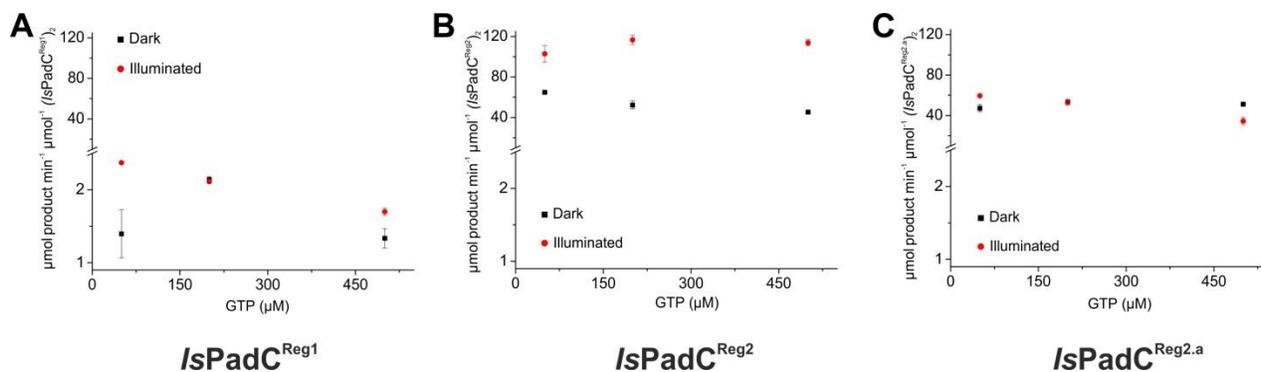
**Figure 53. Kinetics characterization of the coiled-coil register variants.** Kinetic characterization of $Is$PadC$^{Reg1}$ (**A**), $Is$PadC$^{Reg2}$ (**B**), $Is$PadC$^{Reg2.a}$ (**C**), obtained from the HPLC analysis of GTP conversion at different GTP concentrations. Product formation was quantified in triplicate for several reaction times and the standard deviation of individual points contributed to the error estimation of the linear fit that is used to calculate the initial rate of product formation. The standard error of the estimate from the linear regression is shown as error bar for each GTP concentration. The inset in the panel shows a representative chromatogram of the HPLC analysis.

Interestingly, $Is$PadC$^{Reg1}$ and $Is$PadC$^{Reg2.a}$ variants with a fully stabilized coiled-coil linker in register 1 and 2 respectively, do not present any photoregulation of DGC activity at a substrate concentration of 200 µM whereas $Is$PadC$^{Reg2}$ presents a residual 2-fold of photoregulation of DGC activity. Considering that $Is$PadC$^{Reg1}$ and $Is$PadC$^{Reg2.a}$ represent two extremes in coiled-coil stabilization in the inhibiting and stimulating conformational register respectively, $Is$PadC$^{Reg2}$ highlights together with $Is$PadC wildtype the crucial importance of the fine tuning of the coiled-coil sequence composition to enable enough conformational dynamics in the linker region for the modulation of DGC activity. In respect of these observations, the photoactivated state of the PSM might or might not correlate with DGC activation depending on the stabilization dynamics of the coiled-coil linker region.

To have a better representation of the impact of the coiled-coil linker stabilization on the PSM dynamics we performed a HDX-MS analysis of the full-length $Is$PadC$^{Reg2}$ variant and compared it to the previous analysis of $Is$PadC (32). Interestingly, both $Is$PadC and $Is$PadC$^{Reg2}$ present a similar dynamics of hydrogen to deuterium exchange when comparing deuterium exchange in dark and upon light illumination (Fig. 54).
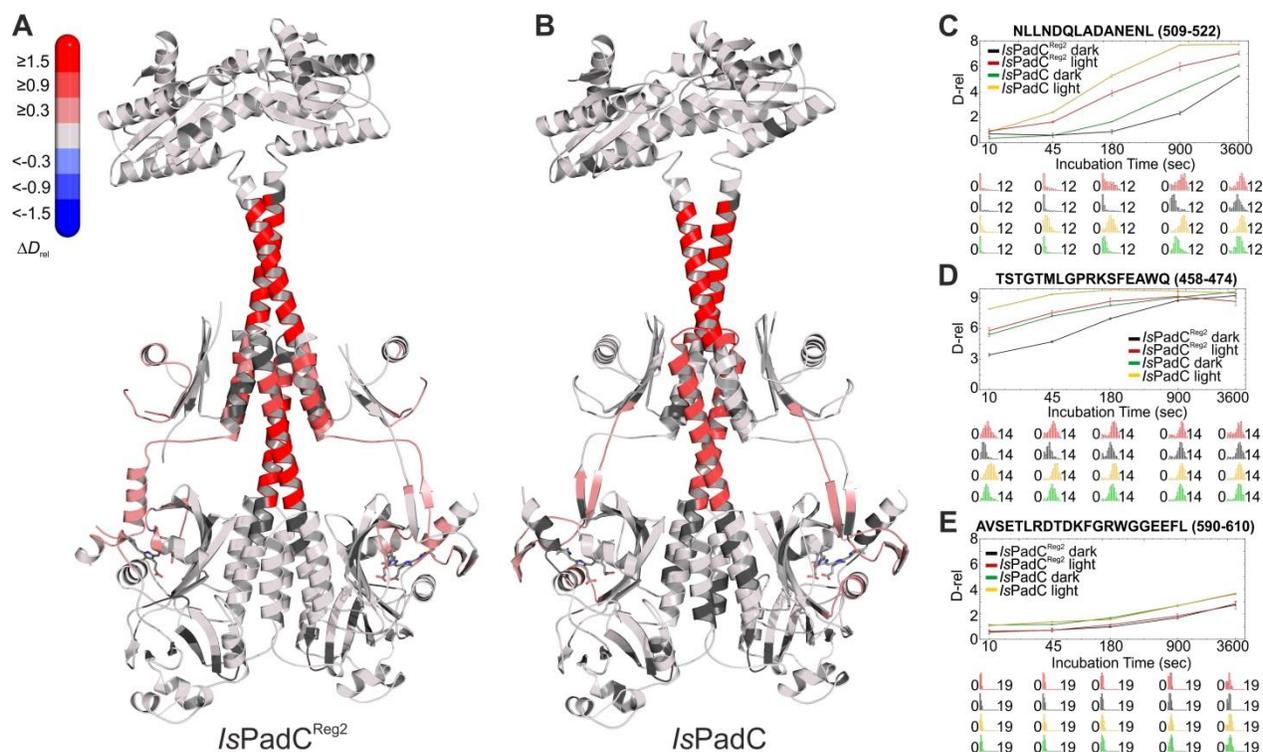
**Figure 54. Comparison of the changes in conformational dynamics upon illumination of *Is*PadC^Reg2 and *Is*PadC.** (**A-B**) Changes in conformational dynamics upon illumination of *Is*PadC^Reg2 (**A**) and *Is*PadC (**B**), respectively, evaluated by HDX-MS. Since time dependent deuterium uptake at the amide positions for different peptides correlates with the stabilization of both β-sheets and α-helices, the deuterium uptake of individual peptides can be related to the stability of the secondary structure elements in a defined region. Closer inspection of individual peptides (panels **C-E**), revealed that the absolute conformational dynamics of individual peptides differ substantially. The presented structures are colored according to the observed changes in relative deuterium incorporation ($\Delta D_{rel}$) between light-state and dark-adapted state ($D_{rel}$ light – $D_{rel}$ dark) after 15 min of deuteration. The changes in $\Delta D_{rel}$ are indicated by the scale in the top left corner with blue corresponding to reduced deuterium incorporation and red reflecting increased exchange of amide protons upon red light illumination. The biliverdin chromophore is represented as sticks and colored in gray. (**C-E**) Comparison of deuterium uptake profiles of *Is*PadC^Reg2 and *Is*PadC peptides in the coiled-coil linker region (C), in the PHY-tongue region (D), and in the GGEEF element of the DGC (E), respectively. $D_{rel}$ is plotted against the deuteration time for light- and dark-state HDX-MS experiments. The error indicators correspond to the sample standard deviation of triplicate measurements. The lower parts show software-estimated abundance distributions of individual deuterated species on a scale from undeuterated to all exchangeable amides deuterated.

Notably, the important regions highlighted previously for *Is*PadC conformational dynamics analysis (32) are also highlighted for *Is*PadC^Reg2. The increase of deuterium incorporation within the coiled-coil linker region observed previously for *Is*PadC has been

correlated with an increase of conformational dynamics enabling the stimulation of GTP turnover. However, when comparing the absolute level of deuterium incorporation of $IsPadC^{Reg2}$ to $IsPadC$ we observed in both dark state and red light illumination conditions an overall reduction of deuterium incorporation rather indicative of a decrease of conformational dynamics. This is especially observable at the PHY-tongue, coiled-coil linker, and GGEEF regions (Fig. 55 and Appendix Fig. 6).



$D_{rel}$ $Is$PadC$^{Reg2}$ Dark - $D_{rel}$ $Is$PadC Dark        $D_{rel}$ $Is$PadC$^{Reg2}$ Light - $D_{rel}$ $Is$PadC Light
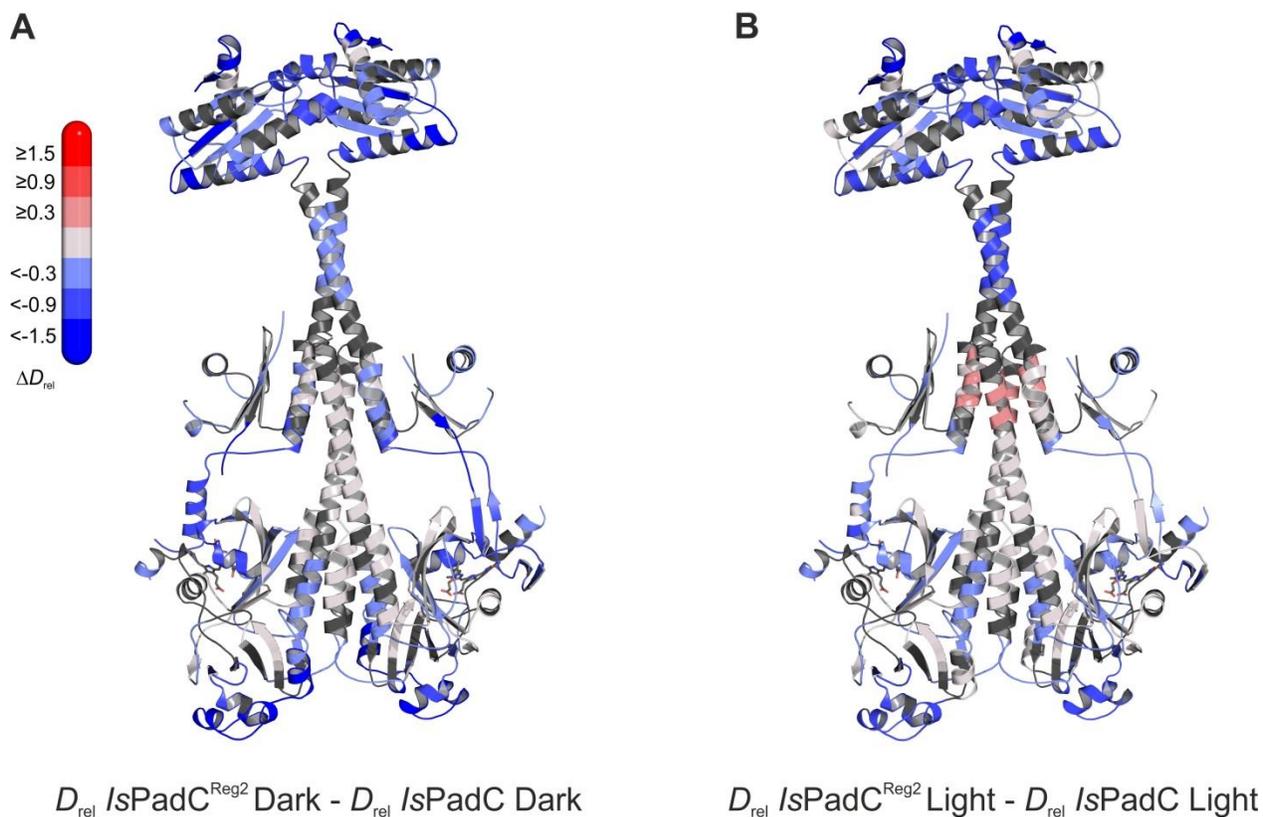
**Figure 55. Comparison of changes in conformational dynamics of $IsPadC^{Reg2}$ and $IsPadC$ evaluated by HDX-MS.** (**A**), The presented structures are colored according to the observed changes in relative deuterium incorporation ($\Delta D_{rel}$) between the dark-adapted states of $IsPadC$ and $IsPadC^{Reg2}$ ($D_{rel}$ Dark $IsPadC^{Reg2}$ – $D_{rel}$ Dark $IsPadC$) (**A**), and between the light illuminated states of $IsPadC$ and $IsPadC^{Reg2}$ ($D_{rel}$ Light $IsPadC^{Reg2}$ – $D_{rel}$ Light $IsPadC$) (**B**), after 3 min of deuteration. The changes in $\Delta D_{rel}$ are indicated by the scale in the top left corner with blue corresponding to reduce deuterium incorporation and red reflecting increased exchange of amide protons in the $IsPadC^{Reg2}$ variant stabilized in the stimulating register. The biliverdin chromophore is represented as stick model and colored in gray.

Remarkably, the pronounced decrease of deuterium incorporation observed at the DGC interface notably at peptides constituting of the helices α1/α2 and α4 is rather supporting the

DGC dimer configuration observed in $Is$PadC$^{Reg2}$ structure (*cf.* Fig. 51). This stabilization is also transmitted to the GGEEF region around the GTP binding site that appears more stabilized in $Is$PadC$^{Reg2}$. This stabilization of the DGC dimer is obviously promoted by the stabilization of the coiled-coil stimulating register observable by the subtle decrease of deuterium exchange at this region in $Is$PadC$^{Reg2}$. Since the increase of conformational dynamics in the coiled-coil linker region was previously correlated with a stimulation of GTP turnover, the fact that $Is$PadC$^{Reg2}$ is overall more active than $Is$PadC suggests that this variant might more or less constantly populate the stimulating register in its coiled-coil linker. The observed increase of deuterium incorporation upon light illumination for $Is$PadC$^{Reg2}$ would therefore reflect an increase of global dynamics of the system rather than a switching between two conformational registers as described for $Is$PadC. Because $Is$PadC$^{Reg2}$ features similar Pr-state spectral characteristics as $Is$PadC, the coiled-coil conformational register in this variant might behave independently of the PSM dynamics of activation. Since $Is$PadC$^{Reg2}$ shows only a residual two fold difference of activation between dark and light conditions (Appendix Table 2), we can conclude that the structural dynamics of activation of the phytochrome PSM and its coordination to the conformational dynamics of the coiled-coil linker are involved in a complex molecular cross-talk crucial for the enzymatic activity modulation of the DGC dimer. The residues composition and number of successive heptad repeats in the coiled-coil linker profoundly influence the equilibrium between the register 1 and 2. However, as seen with our coiled-coil register variants of $Is$PadC, a fully stabilized register 2 do not correlate with the highest activity because a certain conformational dynamics of switching between the two registers in the linker is necessary for the GTP binding and catalytic conversion to c-di-GMP (Fig. 56A). Because the linker register switching is inherently linked to the refolding event of the PHY-tongue element, stabilization of the coiled-coil linker promotes the stabilization of the downstream sensor domains as observed by HDX-MS (*cf.* Fig. 55). Moreover, as observed from the $Is$PadC$^{Reg2}$ GTP soaked structure, conformational changes in the DGC dimer might also influence the coiled-coil linker and downstream sensor conformations highlighting the complex molecular cross-talk between phytochrome sensor, coiled-coil linker, and DGC effector (Fig. 56B).
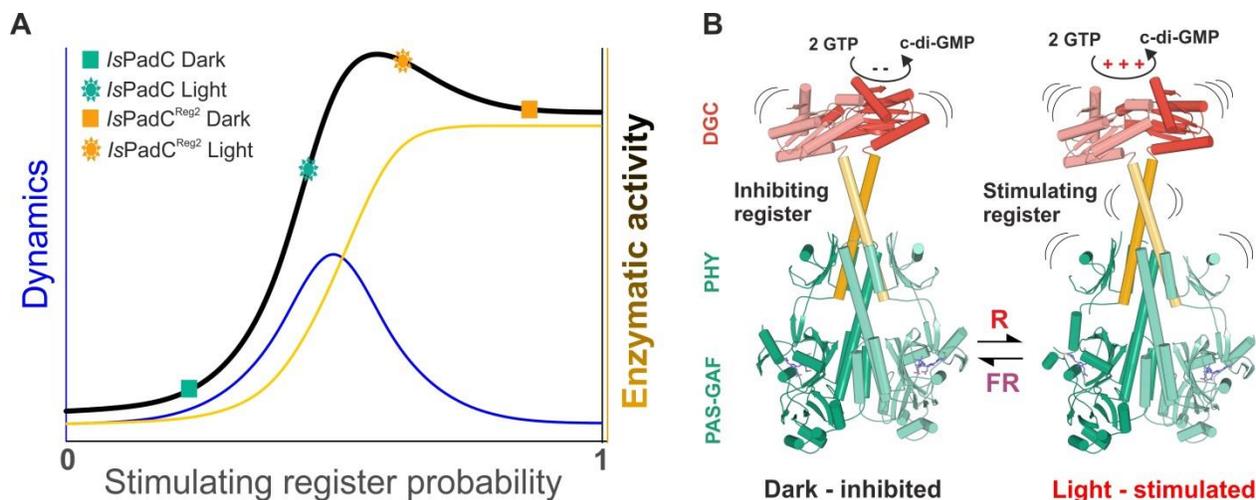
**Figure 56. A toggle mechanism between linker conformations modulates DGC activity.** (**A**) Schematic representation of how DGC activity is modulated by the coiled-coil linker register and its intrinsic conformational dynamics. The transition between the inhibiting and the stimulating coiled-coil conformations is represented on the x-axis as the probability to be in the stimulating register. Under conditions where a switching between coiled-coil registers is possible the dynamics of the system are increased (blue curve). In combination with the structural contribution of the stimulating register for DGC activity (orange curve), the overall enzymatic activity is approximated as the sum of both contributions (dark curve). A large photodynamic range can only be obtained by a fine-tuned balance of conformational flexibility and structural differences of the coiled coil conformations. (**B**) Schematic model of PadC activation. In the dark, a stable β-hairpin conformation of the PHY tongue elements maintains the coiled-coil linker in its inhibiting conformational register. Upon red light illumination and isomerization of the biliverdin of one protomer, the dimeric interface of the PSM rearranges leading to more conformational freedom at the PHY interface allowing the population of the stimulating linker register that in turn facilitates GTP conversion at the GGDEF domains.

## 2.2.5.    Discussion

In our previous characterization of PadC systems, we solved the full-length structure of *Is*PadC in its dark-adapted state, thereby highlighting the tight coupling between phytochrome and DGC dimers (*cf.* Chapter 1. and ref.(32)). We suggested that fine-tuned conformational dynamics of the coiled-coil linker dimer enable the stimulation or inhibition of GTP turnover by the DGC dimer. Finally, the structure of *Is*PadC^Reg2 allows us to confirm the intricate relation between PSM activation and coiled-coil linker conformational changes allowing signal transduction. Indeed, the Pfr-state structural reorganization observed for the biliverdin environment of one protomer in this structure promotes a switching of the coiled-coil linker

conformation to the expected stimulating register (*cf.* Fig. 46 & 50). This observation directly confirmed that the switching between the inhibiting and stimulating conformational registers of the coiled-coil linker is required for DGC activity modulation. Furthermore, the fact that only one biliverdin environment is activated to the Pfr-state within the homodimer was compatible with the coiled-coil linker presents in its stimulating register rather correlates with the observation that *Is*PadC features a composite illuminated spectrum of Pfr and meta-R-like contributions that differs from the observations made for other phytochrome systems featuring highly stable Pfr-states and enabling both PHY-tongue elements to be refolded (39, 80, 85). Considering the observation that the other protomer shows a classical Pr-state, the fact that both chromophores are present in 15*E* form under illumination in solution suggests that the Pfr/Pr heterodimer structure obtained for *Is*PadC[Reg2] is one possible conformation of the homodimer during its transition from a Pr/Pr homodimer to an asymmetric Pfr/meta-R-like heterodimer. The highly stable photoactivated conformation of *Is*PadC[Reg2] and the stabilized stimulating conformation of the coiled-coil linker might always retain a subtle population of the Pfr/Pr heterodimer which might favor crystallization of this conformation further amplified by crystal seeding of the crystallization drops. Importantly, *Is*PadC[Reg2.a] which shows a highly stabilized photoactivated state does not present more Pfr contribution in its steady-state illuminated spectrum compared to *Is*PadC (*cf.* Fig. 42), supporting the fact that an asymmetric heterodimer Pfr/meta-R-like is the signaling conformation of *Is*PadC under illumination. In contrary to phytochrome PSM homodimers enabling both PHY-tongues to refold to a typical Pfr-state, *Is*PadC features differences at the dimer interface and/or within the local environment of the chromophore that avoids a symmetric conversion to the Pfr-state. Interestingly, *D. radiodurans* phytochrome (*Dr*BphP) variants containing mutations of local residues around the biliverdin feature spectral characteristics similar to *Is*PadC and have been described as potential mixtures between Pfr and meta-R species (97). Considering the relevant signaling properties of Pfr/Pr heterodimer in plant phytochromes (155), Pfr/Pr and/or Pfr/meta-R-like asymmetric heterodimers in bacteriophytochromes might as well be relevant for molecular signal transduction especially in the case of the modulation of dimeric output effectors. As previously observed, structures of phytochrome PSMs from different species revealed an apparent plasticity at the dimer interfaces (32, 79). These differences of global assembly can potentially differently influencing the photoconversion route of each chromophore environment. Inter-protomeric allosteric effects of the PSM dimer interface on the chromophore environments have been notably described in

engineered fluorescent proteins (96). Also in the bacteriophytochrome *Dr*BphP, the disruption of the PAS-GAF interface dramatically slows down the thermal recovery to the ground state further supporting the major influence of the dimer interface on the chromophores environments (88, 156).

By featuring pronounced differences in their thermal recovery rates, the differently stabilized coiled-coil variants of *Is*PadC also show the fundamental impact of the sensor-effector linker dimer interface on the global dynamics of photoconversion. For all of these variants however, the relative amplitude and rate constant of the first phase of the recovery process is only barely affected by the different assemblies that allowed us to propose the hypothesis that this phase of the recovery reflects the thermal reversion of the chromophore environment stabilized in an intermediate meta-R-like state. Similarly to other described phytochromes, this intermediate meta-R-like species could be part of an unproductive shunt reaction pathway (144). Potentially, the evolutionary adaptation of phytochrome dimerization and tuning of the coordination of important structural elements to control the biliverdin environments might have been naturally selected for the modulation of specific output modules. Notably in PadCs, the composition and length of the coiled-coil linker is obviously finely tuned to be coordinated with the asymmetric photoactivation behavior of the PSM assembly, remarkably observed by the importance of the asparagine residues in the *a* position of the central heptad repeat of the stimulating register that provides enough dynamics to the linker to conserve a fold of activation in *Is*PadC$^{Reg2}$ compared to *Is*PadC$^{Reg2.a}$. Considering only the field of photoreceptors, asymmetric activation mechanisms are rather sporadically described (157, 158) even though the asymmetric integration of allosteric modulation emerges as a critical evolutionary aspect (159–161). Based on our characterization of *Is*PadC systems we can now say that the asymmetric activation of dimeric photoreceptors might be a conserved mechanism of allosteric regulation of specific effector domains active by concerted catalytic mechanisms. The origin of the asymmetric behavior observed in *Is*PadC PSM remains at the moment rather unexplained. Potentially, specific rotamer differences of important residues in the biliverdin environment of one protomer might prevent the population of a typical Pfr-state by retaining specific interactions of the PHY-tongue element with the GAF domain. Equally, the establishment of a typical Pfr-state in one protomer might inhibit the other protomer to form a full Pfr-state by avoiding the unfolded PHY-tongue element to refolds to the α-helical conformation due to rearrangement of the PSM dimer interface.

Obviously the photoresponse adaptations of phytochromes rely on a complex equilibrium of intra- and inter-protomeric interactions. In order to better understand the structural interactions involved in the functional cross-talks between phytochrome sensors and DGC effectors, we looked at the effect of domain replacements between full-length PadC homologs.

## 2.3. Functional inter- and intra-protomeric molecular crosstalks involved in phytochrome photoresponses.

The results presented in this chapter are mainly based on the following publication currently in preparation:

# Functional inter- and intra-protomeric molecular crosstalks involved in phytochrome photoresponses

Gourinchas G[1], Vide U[1], Winkler A[1*]

[1]*Institute of Biochemistry, Graz University of Technology, petersgasse 12/II, 8010 Graz, Austria*

*Corresponding author Email: andreas.winkler@tugraz.at*

**2018** (In preparation)

## 2.3.1. The stability of specific PSM intra-protomeric interactions influences the phototransduction

The precise modulation of nucleotidyl cyclase activity catalyzing cyclic 3',5'-guanylyl and adenylyl nucleotides involved in the regulation of many metabolic and behavioral responses in various organisms recently appeared as a promising pharmacological target (16). However, even though some photoreceptors have been already successfully coupled to nucleotidyl cyclases and used as optogenetics tools (76, 109–111, 162) the rational design of new photoreceptor – nucleotidyl cyclase couples still needs a substantial amount of screening efforts with little control on the output photomodulation capacity depending on the mechanistic complexity of the used photoreceptor. Due to its role of bacterial second messenger involved in many biological pathways modulation, notably to bacterial biofilm formation (72), the precise *in vivo* regulation of the c-di-GMP level by red light sensors appears promising (163). However, the rational fusion of specific effectors to phytochrome sensors with desired photomodulation capacities remains difficult due to the fine-tuned allosteric cross-talks between sensors and effectors.

To improve the understanding about the molecular communication between phytochrome sensors and DGC effectors, we generated synthetic chimeras between two naturally occurring PadC homologs featuring overall 37 % sequence identity and distinct photoresponse to red light (Fig. 57). *Is*PadC and *Ts*PadC respectively found in *Idiomarina species* A28L and *Thioalkalivibrio species* ALMg3 feature the same arrangement of functional domains (32), however a closer look at the primary sequence revealed that one major difference between the two homologs occurs at the coiled-coil linker region previously described as crucial modulator of DGC activity (32, 77). Indeed, the *Ts*PadC coiled-coil linker is two helical turns shorter (7 residues) and presents a Gly residue in position *f* of its first heptad repeat which is rather uncommon in helical patterns as it might potentially break the helical structure and change the coiled-coil dimer assembly (120, 164) (Fig. 57 and Appendix Fig. 1).
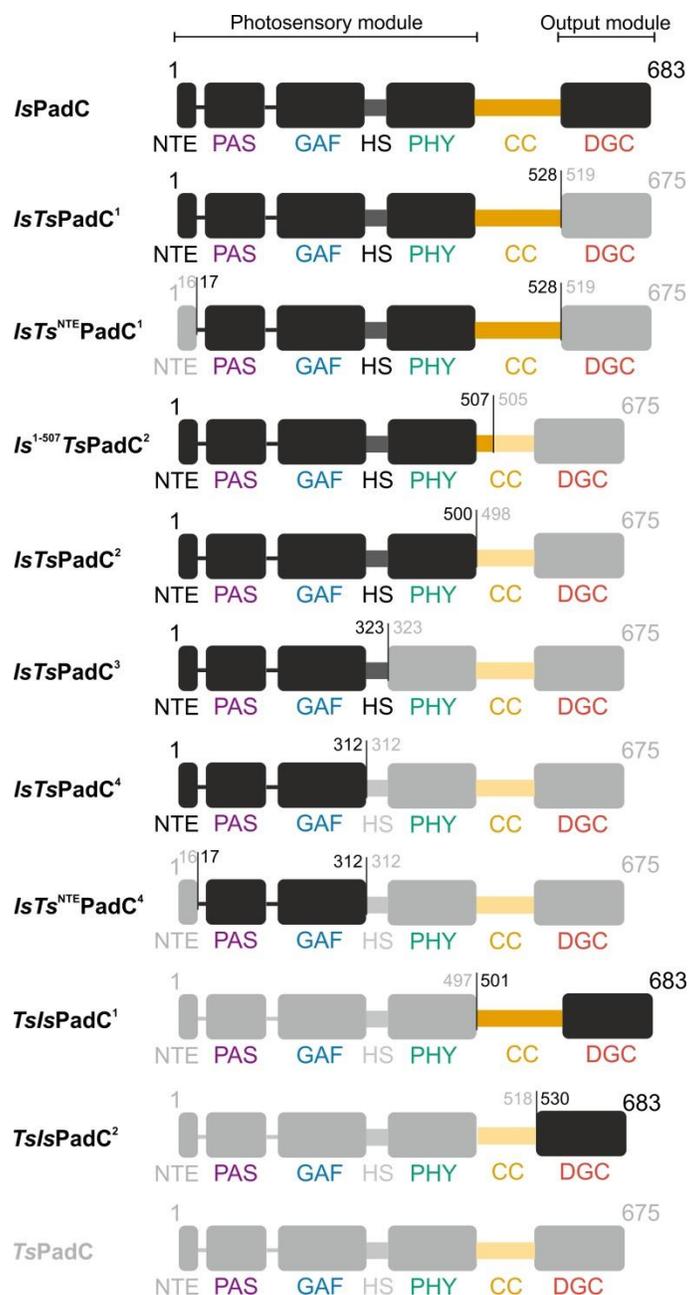
**Figure 57. Overview of the synthetic chimeras generated between *Is*PadC and *Ts*PadC.** Schematic representation of the synthetic chimeras generated between *Is*PadC and *Ts*PadC. Domains of *Is*PadC and *Ts*PadC are colored in black and light gray, respectively. The positions of the fusions points are indicated on each side of a vertical black line. The sequence numeration is based on the sequence of the *Is*PadC sequence in black and *Ts*PadC sequence in gray.

Nevertheless, the representation of the succession of heptad repeats of the *Ts*PadC coiled-coil linker shows a rather similar dark state inhibiting register with hydrophobic residues in

positions *a* and *d* as well as one destabilizing residue in position *d* (Fig. 58). Considering a similar linker translation mechanism as observed for *Is*PadC (77), Asn509 would be found at the position *a* in the stimulating register along with other destabilizing residues and stabilizing hydrophobic residues at the position *d*. This observation supports potential similarities in the modulation mechanism of signal transduction for *Ts*PadC and *Is*PadC coiled-coil linkers since the *Is*PadC linker also conserves Asn residues in position *a* of its stimulating register (Fig. 58). However, the difference of 7 residues in their sensor-effector linker regions might potentially lead to a slightly different DGC dimer orientation respective to the phytochrome sensor due to the supercoiling of the linker region (122).
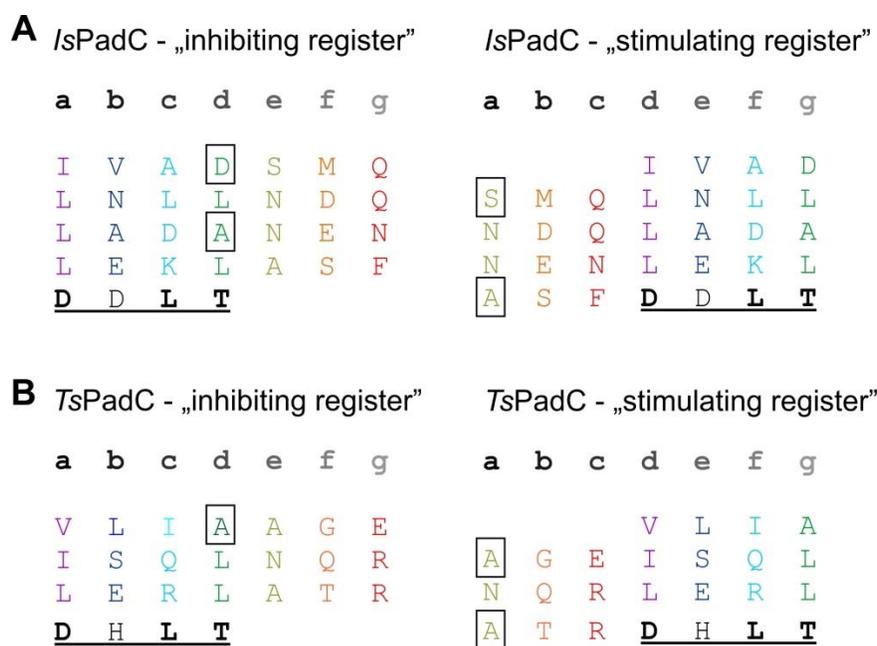


**Figure 58. Comparison of sequence registers between *Is*PadC and *Ts*PadC coiled-coil linkers.** Schematic representation of the inhibiting and stimulating coiled-coil sequence registers for *Is*PadC (**A**), *Ts*PadC (**B**), respectively. Destabilizing residues are boxed.

Due to the absence of crystallization of *Ts*PadC full-length construct, structural information on the global assembly was obtained by homology modelling using the I-TASSER server (165, 166) and the structure of *Is*PadC as template (32). The best generated model with a high confidence score (C-score=1.55) and a correct topology (TM-score=0.94) was kept. Due to the difference of 7 residues in coiled-coil linker length between *Ts*PadC and *Is*PadC, *Ts*PadC PSM-linker (residues 1-518) and *Ts*PadC DGC (residues 519-675) were modelled separately

using *Is*PadC PSM (residues 1-521) and *Is*PadC DGC (residues 529-683) as templates, respectively. The modelled *Ts*PadC DGC dimer was aligned on one *Ts*PadC PSM protomer by structural alignment of *Ts*PadC PSM and *Is*PadC PSM. The *Ts*PadC DGC dimer was then manually rotated to accommodate the positioning of the DGC wide turn containing the DXLT motif in respect of the end of the *Ts*PadC coiled-coil linker. The global assembly obtained showed that a rotation of at least ~20° of the *Ts*PadC DGC dimer can be expected in comparison to *Is*PadC DGC dimer that highlights the fact that the supercoiling of the coiled-coil linker induces slightly different sensor-effector structural alignments even in the case of an even number of helical turns difference (Fig. 59). Irrespective of the length, the amino acid composition of the coiled-coil linker element might have a stronger impact on the stability of the full-length system and therefore on the DGC regulation.
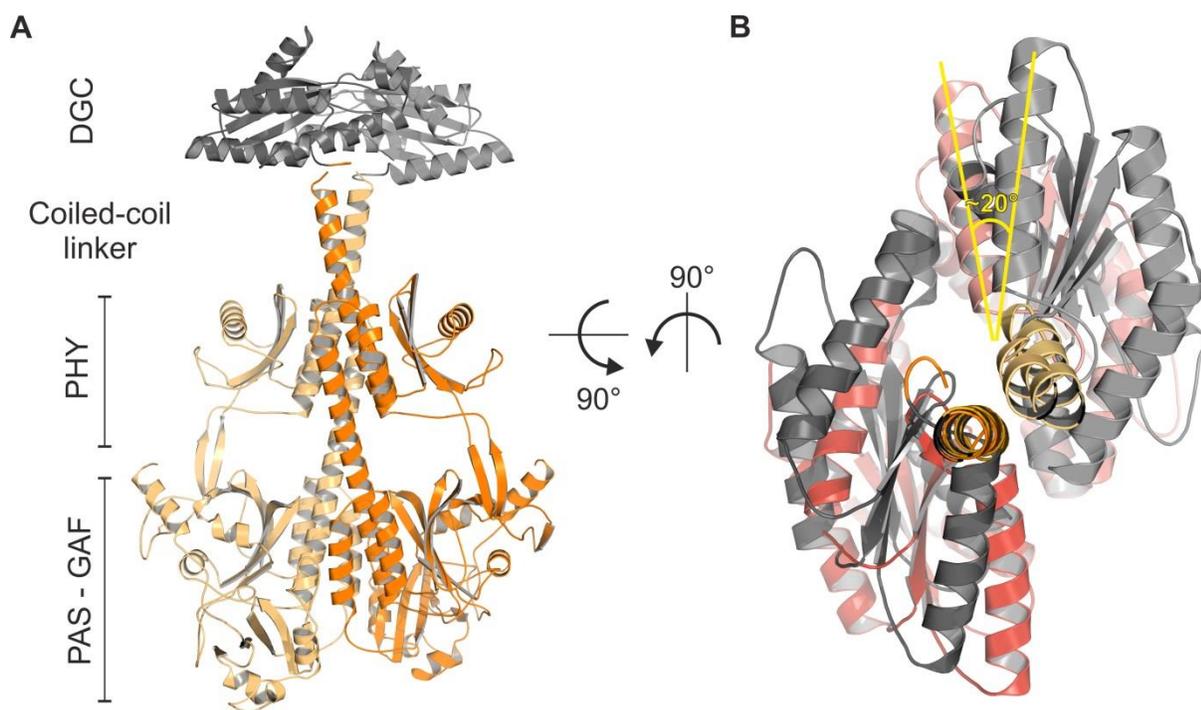


**Figure 59. Homology modelling of *Ts*PadC.** (**A**) Homology model of the full-length Pr-state structure of *Ts*PadC generated by the I-TASSER server (165) using parts of *Is*PadC full-length structure as templates. (**B**) The superposition of *Ts*PadC and *Is*PadC PSMs (residues 1-312, chains A) shows that the *Ts*PadC DGC dimer could be rotated of ~20° in the full-length construct due to the supercoiling of the coiled-coil linker.

A closer look to the interactions involved in the stability of the PHY-tongue Pr-state β-hairpin conformation in *Is*PadC structure suggests that important interactions are overall well

conserved between *Is*PadC and *Ts*PadC (Fig. 60). However, the extremity of the β-hairpin PHY-tongue in *Ts*PadC is most probably different than in *Is*PadC notably due to the Arg460 that is potentially involved in a salt bridge with the Glu5 located in the NTE. Due to the uncertainty of the homology model it is difficult to interpret any functional implication of this supplementary salt bridge in *Ts*PadC, however, an arginine at this position in the PHY-tongue element might also be involved in some interactions with the GAF domain upon refolding of the PHY-tongue that could potentially stabilize the photoactivated state in *Ts*PadC.
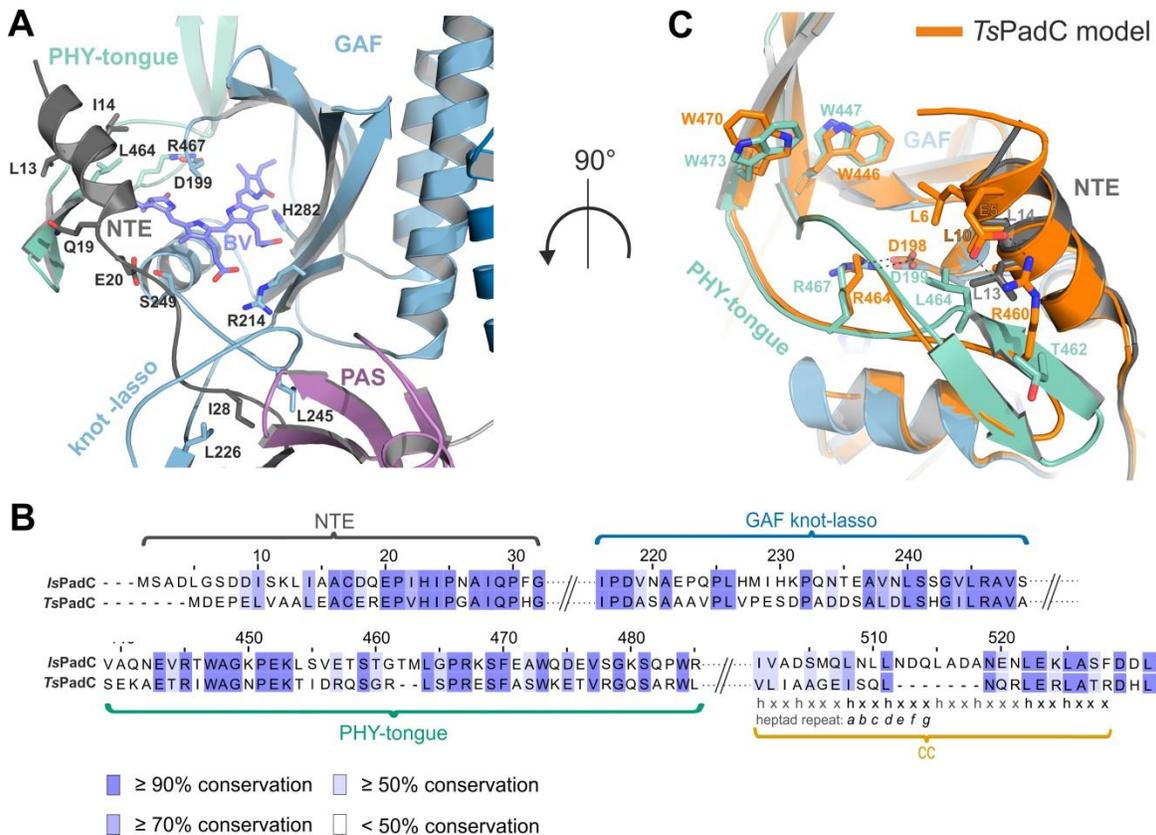


**Figure 60. Details of sequence conservation between *Is*PadC and *Ts*PadC.** (**A**) Close-up view on the NTE interactions with the GAF domain and the PHY-tongue element in Pr-state. BV chromophore and important residues involved in the interactions are highlighted in sticks representation. (**B**) Sequence alignement of important sructural features of the *Is*PadC and *Ts*PadC photosensory modules and the sensor effector linker regions. (**C**) Close up view on the NTE – PHY-tongue interactions in *Ts*PadC homology model in comparison to *Is*PadC. *Is*PadC domains are colored accordingly to panel A and *Ts*PadC model is colored in orange.

A preliminary HDX-MS analysis of *Ts*PadC between dark and illuminated conditions revealed that peptides constituting of the PHY-tongue element show two different hydrogen-

deuterium exchange profiles similarly to *Is*PadC (Fig. 61). Based on the homology modeling of *Ts*PadC and in comparison with the *Is*PadC β-hairpin PHY-tongue and refolded PHY-tongue from *Is*PadC[Reg2] we observed that the fast exchanging peptides would correspond to the unfolded region of the PHY-tongue while the slower exchanging peptides would correspond to the α-helical part of the PHY-tongue that refolds upon illumination (Fig. 61). This hydrogen-deuterium exchange distribution supports the same asymmetric activation of the PSM in *Ts*PadC as observed for *Is*PadC (77). The fast deuterium uptake observed for the NTE region in *Ts*PadC would suggest a rather instable structure of the NTE and/or instable interactions between the NTE and the PHY-tongue elements.
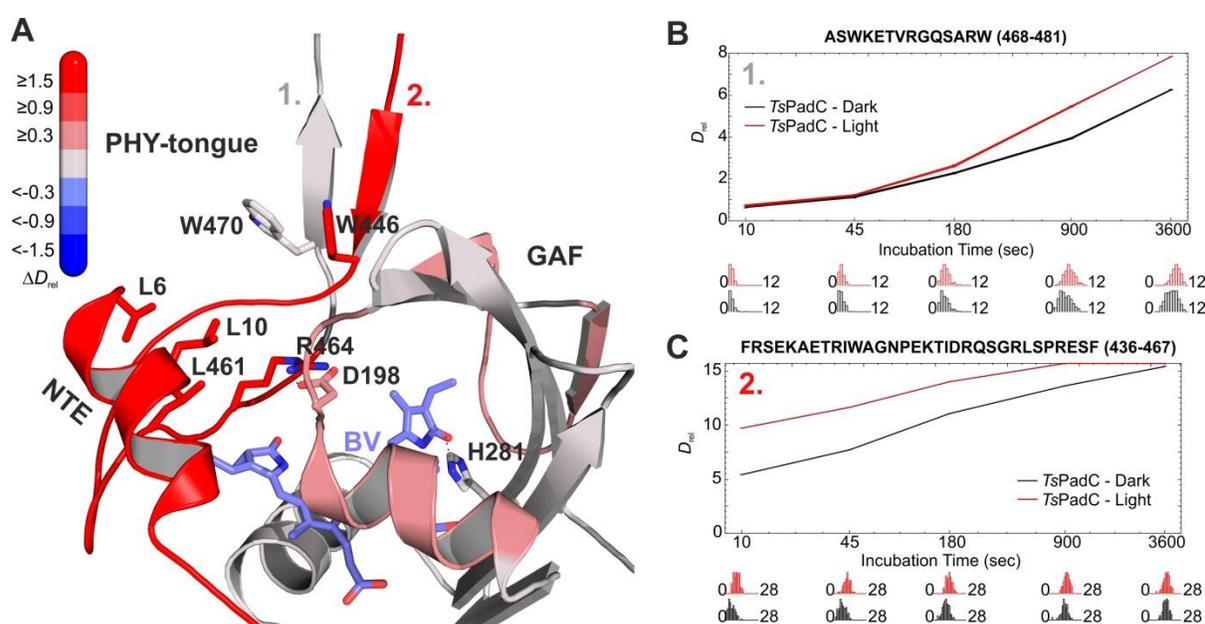


**Figure 61. Initial characterization of *Ts*PadC changes in conformational dynamics upon light illumination.** (**A**) Close-up view on the chromophore environment after 45 sec incubation in $D_2O$ that shows a similar deuterium exchange pattern as observed for *Is*PadC (*cf.* Fig 36). (**B-C**) Deuterium uptake profiles of peptides situated on each β-strand of the PHY-tongue β-hairpin support a potential similar refolding of part of the PHY-tongue as observed for *Is*PadC (77).

Obviously, subtle differences of interaction between the NTE, the PHY-tongue, and the GAF domain might influence the positioning of the biliverdin chromophore and therefore the absorption properties of the full-length holoprotein.

The characterization of the *Ts*PadC absorption spectrum revealed a same partial Pfr photoconversion upon illumination as observed for all the so far characterized PadCs (32, 77). In

the case of *Is*PadC, as described in the previous chapter, this incomplete Pfr photoconversion was attributed to a functional asymmetric Pfr/meta-R heterodimer presents in the steady-state illuminated spectrum. The possibility of heterogeneous chromophore environments in *Ts*PadC is also in some extend represented by the non-isosbestic behavior observed for the thermal recovery to the ground state indicating more than one colored species involved in the steady-state illuminated spectra (Fig. 62).
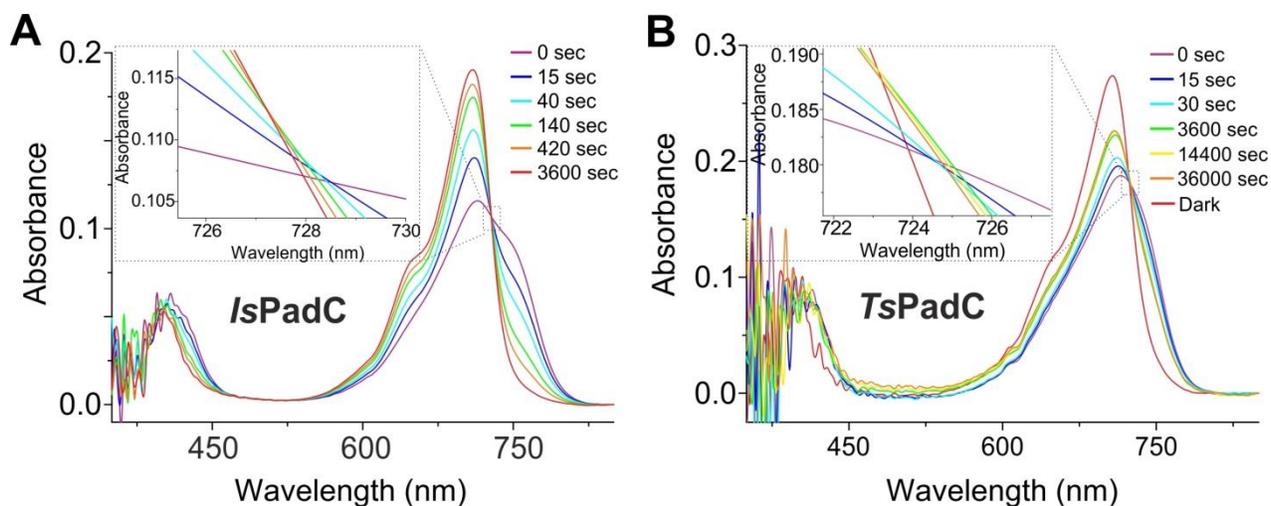


**Figure 62. Close-up view on the non-isosbestic region of the *Is*PadC spectrum.** Spectral recovery of *Is*PadC (**A**) and *Ts*PadC (**B**) recorded using a diode-array spectrophotometer with a sampling time of 5 s. The insets show a close up view of the non-isosbestic recovery.

Interestingly, this spectroscopic behavior seems to be conserved also by other phytochrome species not regulating DGC effectors, which supports the potential relevance of asymmetric activation in phytochrome homodimers (30, 66, 141, 142). Both *Is*PadC and *Ts*PadC are canonical bacteriophytochromes with the same Pr-state as a ground state featuring an absorption maximum of the Q-band respectively at 710 and 708 nm (Fig. 63A). Upon red light illumination, both phytochromes show a steady-state spectrum shifted to the far-red light region. However, it is interesting to observe that the Pfr-state contribution of *Ts*PadC is slightly shifted to 744 nm in comparison to *Is*PadC found at 750 nm. Remarkably, both phytochromes feature more than 90 % of the chromophore population in the 15*E* isomer upon illumination (Fig. 63B).
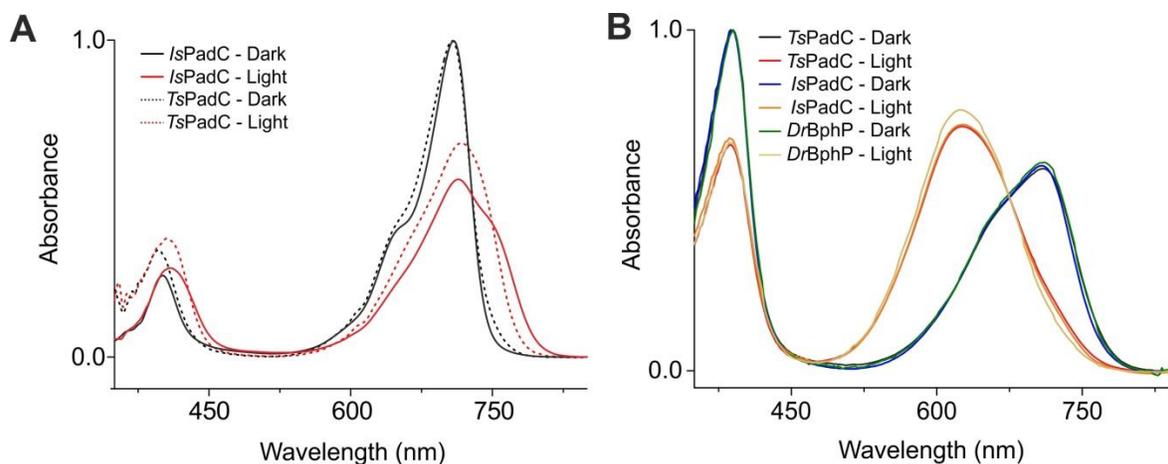
**Figure 63. Spectroscopic characterization of *Ts*PadC.** (**A**) UV/Vis absorption spectra of *Is*PadC and *Ts*PadC scaled to their Pr-state Q-band maximum absorption. (**B**) Denaturation of *Ts*PadC in methanol/TCA 0.1 % under dark and light conditions in comparison to *Is*PadC and *Dr*BphP.

Although both bacteriophytochromes have similar spectroscopic characteristics, their global dynamics of signal transduction are rather different. Indeed, *Is*PadC is able to recover thermally to its ground state on a minute time scale while *Ts*PadC needs days to recover fully to its ground state (Fig. 64). Importantly, the fitting of the time courses of 15*E* to 15*Z* chromophore isomerization measured by acidic denaturation for both *Is*PadC and *Ts*PadC produced similar recovery rates constants suggesting that both isomerized chromophores are involved in the illuminated state absorption (Fig. 64B). Obviously, some specific interactions between the protomers or intra-protomeric interactions retain the photoactivated conformation longer in *Ts*PadC. Considering the violin model of signal transduction (134), a different conformation of the DGC dimer respective to a shorter coiled-coil linker might also influences the downstream sensor and stabilizing the photoactivated state in *Ts*PadC. All together, these observations highlight the adaptation of related phytochromes to various photomodulations of specific effector by tuning the coordination of structural elements influencing the overall stability of the photoactivated state.
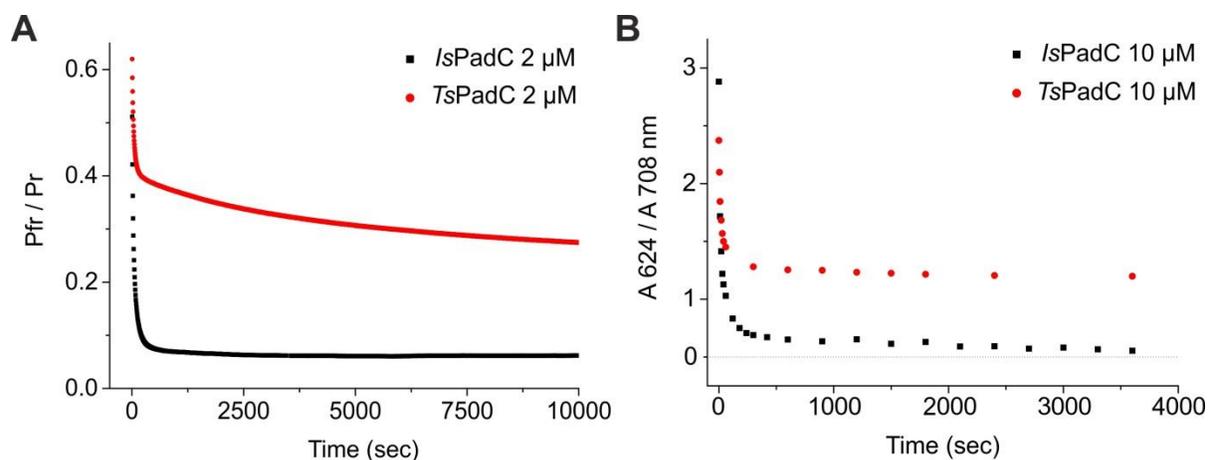
**Figure 64. Difference of photoactivated state stability between *Is*PadC and *Ts*PadC. (A)** Thermal recovery of *Is*PadC (black) and *Ts*PadC (red) recorded for identical protein concentrations of 2 μM and represented as the ratio Pfr-state absorption over Pr-state absorption. (**B**) Thermal reversion of the 15*E* to the 15*Z* isomer for *Is*PadC and *Ts*PadC at a concentration of 10 μM represented as the ratio of 15*E* absorption (624 nm) over 15*Z* absorption (708 nm).

## 2.3.2. NTE coordination to the PHY-tongue dominates the properties of phytochromes photoactivation

From the spectroscopic characterization of the generated synthetic chimeras, we observed that all the constructs show the same spectral signature as *Is*PadC as long as the PAS-GAF core of *Is*PadC is present. Even for *IsTs*PadC[3] and *IsTs*PadC[4] which feature the PHY-tongue of *Ts*PadC in front of the *Is*PadC PAS-GAF core and NTE region, the same spectral signature as in *Is*PadC is observed. Strikingly, as soon as we replaced the NTE region of *Is*PadC by the corresponding one of *Ts*PadC (*IsTs*[NTE]PadC[4]), we completely reestablished the spectral signature of *Ts*PadC (Fig. 65). Importantly, the exchange of only the NTE region between the two bacteriophytochrome PSMs (*IsTs*[NTE]PadC[1]) does not change the spectral signature of the illuminated state but slightly shifts the Pr-state Q-band maximum of absorption that highlights the importance of the NTE region on the positioning of the biliverdin chromophore in its protein environment but also suggests that specific interactions between the NTE and the PHY-tongue element are crucial for the spectroscopic properties of the full-length holoprotein (Fig. 65). Obviously, there is a coordinated interaction between the PHY-tongue and the NTE region that appears to dominate the spectroscopic properties of bacteriophytochromes. Structures of

phytochromes in dark-adapted and light activated states (39, 80, 85), as well as the comparison of the *Is*PadC dark–adapted state and the *Is*PadC[Reg2] structures (77) confirm the pronounced rearrangement of both the PHY-tongue and the NTE regions upon light activation that affect the molecular interactions between these two elements. Studies on plant phytochromes also revealed the importance of the NTE region on the photoactivated state stability (167).
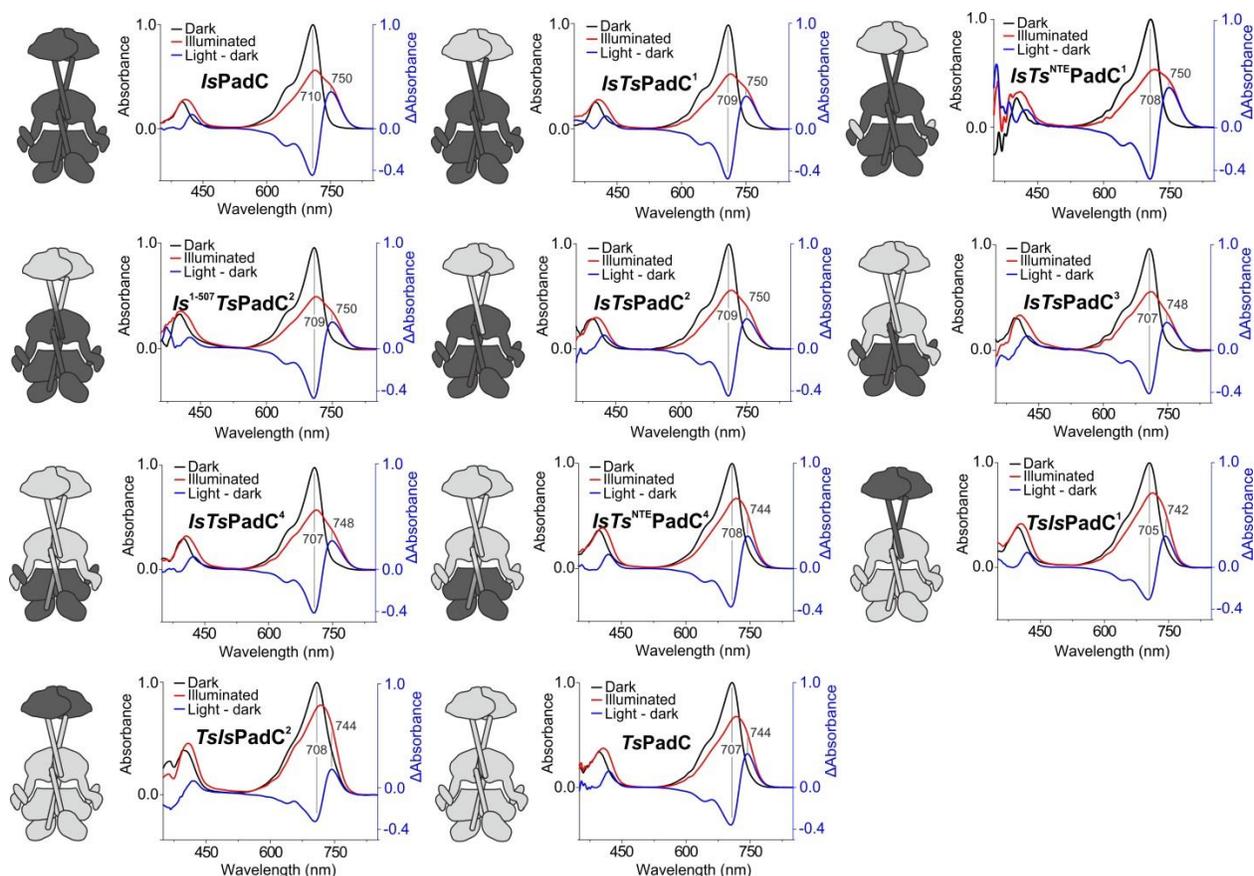


**Figure 65. Biochemical characterization of the synthetic chimeras.** UV/Vis spectrum of the various chimeras. Schematic representations of the chimera dimers are represented next to each spectrum. Domains corresponding to *Is*PadC are colored in black while domains corresponding to *Ts*PadC are colored in light gray.

Even though the spectroscopic properties of the chimeras are barely affected by exchanging functional domains other than GAF or NTE, striking differences of photoactivated state stability and light regulation of DGC activity are observed (Appendix Tables 1 and 2). Interestingly, the thermal recovery of *Is*PadC from the illuminated state spectrum to the Pr/Pr homodimer can be fitted by the sum of two exponential decay functions at a protein concentration of 2 μM whereas the *Ts*PadC thermal recovery at the same protein concentration necessitates the

use of three exponential decay functions to describe the data accurately (Fig. 64A and Appendix Table 1). Remarkably, the inclusion of a third exponential decay contribution is also required for *Is*PadC variants featuring a highly stabilized photoactivated state (Appendix Table 1) (77). As described in the preceding chapter, we observed that the relative amplitude of the fast phase of the recovery is almost identical between *Is*PadC wildtype and *Is*PadC variants stabilized in the stimulating register of the coiled coil linker (77) (Appendix Table 1). This first phase is obviously barely affected by the variation of global dynamics within the dimer and was previously attributed to the recovery of the protomer featuring a meta-R like intermediate species. Notably, due to the low percentage of the relative amplitude over the overall amplitude of recovery, the fast phase can only represent a molecular species of lower extinction coefficient than a classical Pfr-state. Strikingly, *Ts*PadC recovery also presents a fast phase with rate-constants and relative amplitudes rather in the range of the *Is*PadC recovery (Appendix Table 1). This observation would potentially support a similar Pfr/meta-R-like asymmetric dimer in *Ts*PadC. The introduction of the *Ts*PadC GGDEF domain or the coiled-coil plus GGDEF domain to *Is*PadC (*IsTs*PadC[1] and *IsTs*PadC[2], respectively) results in a strong increase of the relative amplitude of the fast phase for the overall amplitude of recovery that suggests a destabilization of the Pfr-state in those variants potentially due to an impaired molecular coordination between the coiled-coil linker and the PHY-tongue regions leading to a higher population of meta-R like species (Appendix Table 1). Already the exchange of the DGC effector between the two PadCs highlights the finely tuned coordination between the DGC effector, the coiled-coil linker, and the phytochrome PSM. Due to its direct link to the coiled-coil linker by the structural wide turn including the conserved DXLT motif, DGC dynamics are intricately linked to the coiled-coil linker dynamics and vice-versa. The construct *TsIs*PadC[1] that features the coiled-coil linker and DGC dimer of *Is*PadC fused to *Ts*PadC PSM monomerized upon dark state recovery and reversibly re-dimerized upon light illumination that further highlights the complexity of the interactions of all functional elements and consequently also the dimerization interface of the PSM (Fig. 66).
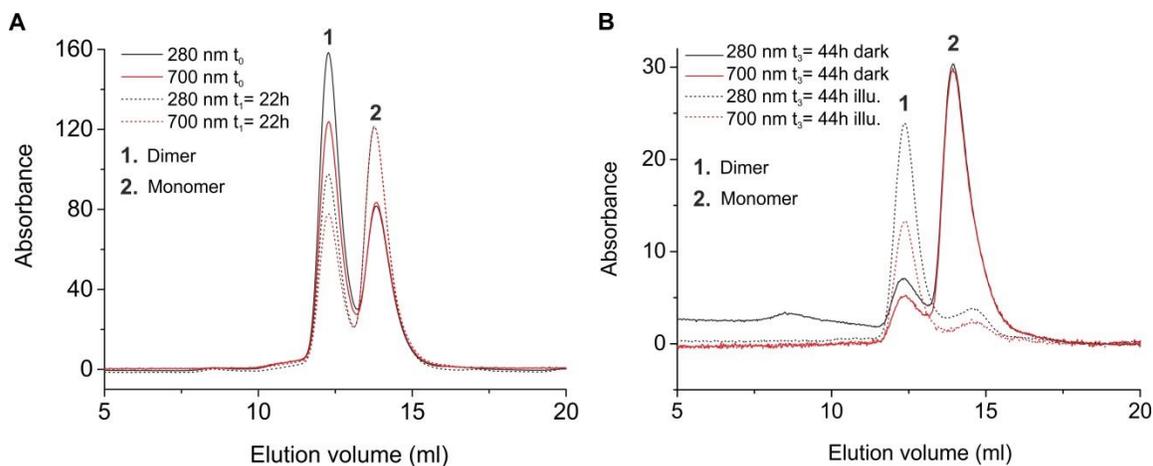
**Figure 66.** *TsIs*PadC[1] **monomerization during ground state recovery.** (**A**) Gel filtration analysis of the time dependence monomerization of *TsIs*PadC[1] upon Pr-state recovery using a Superdex 200 10/300 analytical grade column. (**B**) After 44h in the dark a sample of *TsIs*PadC[1] was illuminated 1 min with red light and analyzed by gel filtration under constant illumination that shows the reformation of the dimer interface upon illumination.

As previously characterized, the conformational dynamics of the coiled-coil linker strongly depends on the PHY-tongue refolding event and therefore, the introduction of the *Ts*PadC PHY-domain in *IsTs*PadC[2] (*IsTs*PadC[3]) reestablishes the coordination between DGC, coiled-coil, and PHY-tongue elements decreasing the frequency of meta-R like species population and reestablishing a proper Pfr-state as conclude from the decrease of the first recovery phase amplitude contribution in the overall amplitude of the recovery (Appendix Table 1). In line with the structural rearrangements observed between *Is*PadC and *Is*PadC[Reg2] structures (77), this observations further confirms that the Pfr-state recovery is represented by the slow phase of recovery (Appendix Table 1). Interestingly, the introduction of the *Ts*PadC central helical spine into *IsTs*PadC[3] (*IsTs*PadC[4]) only barely affects the rate constants and relative amplitudes of the two first phases of the recovery while massively affecting the rate constant of the third phase highlighting the importance of the PSM inter-protomeric interactions on the thermal reversion (Appendix Table 1). The following introduction of the *Ts*PadC NTE region in *IsTs*PadC[4] led to almost identical thermal recovery to *Ts*PadC highlighting the fundamental coordination of the NTE region with the PHY-tongue in the control of the overall dynamics of signal transduction as well as spectroscopic properties of the biliverdin cofactor.

## 2.3.3. Impact of the various domain replacements on DGC activity regulation

Although the chimeras generated feature only subtle differences in their UV/Vis spectra, their strong differences in thermal recoveries to the ground state indicate a strong impact on the signal integration mechanism that is also reflected in a different photomodulation of the DGC effector (Fig. 67). Both *Is*PadC and *Ts*PadC feature a photoactivation of DGC, with respectively ~70 and ~6-fold activation under illumination (660 nm, 45 mW cm$^{-2}$). Strikingly, *Ts*PadC presents much more GTP conversion already in dark condition (Fig. 67 and Appendix Table 2). We assume that spontaneous formation of a catalytically active DGC dimer in *Ts*PadC might be promoted due to the shorter coiled-coil linker length and hence weaker stabilization of the inhibiting coiled-coil register. In line with the violin model (134), the productive DGC dimer might also stabilize the coiled-coil linker in its stimulating register potentially retaining a productive dimer over longer times and resulting in substantial GTP conversion already under dark conditions. Interestingly, the swapping of GGDEF domains between *Is*PadC and *Ts*PadC, (*IsTs*PadC$^1$ and *TsIs*PadC$^1$) maintains a rather high fold of DGC photoactivation (Appendix Table 2). However, when both coiled-coil linker and DGC dimers are exchanged (*IsTs*PadC$^2$ and *TsIs*PadC$^2$), the GTP turnover rate is strongly affected highlighting the crucial coordination between the PSM and the coiled-coil linker conformational dynamics to properly position the productive DGC interface. Strikingly, the replacement of the first *Ts*PadC heptad repeat by the corresponding one of *Is*PadC in *IsTs*PadC$^2$ (*Is*$^{1-507}$*Ts*PadC$^2$) reestablishes a similar photoactivation property as in *IsTs*PadC$^1$. This highlights the importance of the structural interactions between the PHY domain and the coiled-coil linker, and potentially indicates that the Gly residue contained in the first *Ts*PadC heptad repeat introduces potentially a different conformational dynamics in the linker compared to *Is*PadC. Similarly, the inclusion of the *Ts*PadC PHY domain in *IsTs*PadC$^2$ (*IsTs*PadC$^3$) reestablishes a decent GTP turnover rate and a fold activation similar to *IsTs*PadC$^1$ further supporting the importance of the PHY interface on the dynamics of the coiled-coil linker. The inclusion of the *Ts*PadC central helical spine in *IsTs*PadC$^3$ (*IsTs*PadC$^4$) only reduces the GTP turnover rate that together with the strong impact on the thermal recovery highlights the tuning of the central helical spine in respect of the PHY domain dynamics. Remarkably, the inclusion of the *Ts*PadC NTE in *IsTs*PadC$^4$ that resulted in

identical UV/Vis spectra and thermal recovery behavior as for the *Ts*PadC wildtype shows a rather weak GTP turnover capacity. Obviously, reestablishing the crucial interactions between the NTE and the PHY-tongue are enough to properly integrate and transduce the light signal, however, efficient GTP turnover can only be observed if the DGC dimer interface is well coordinated to the coiled-coil linker and PSM conformational changes (Fig. 67).
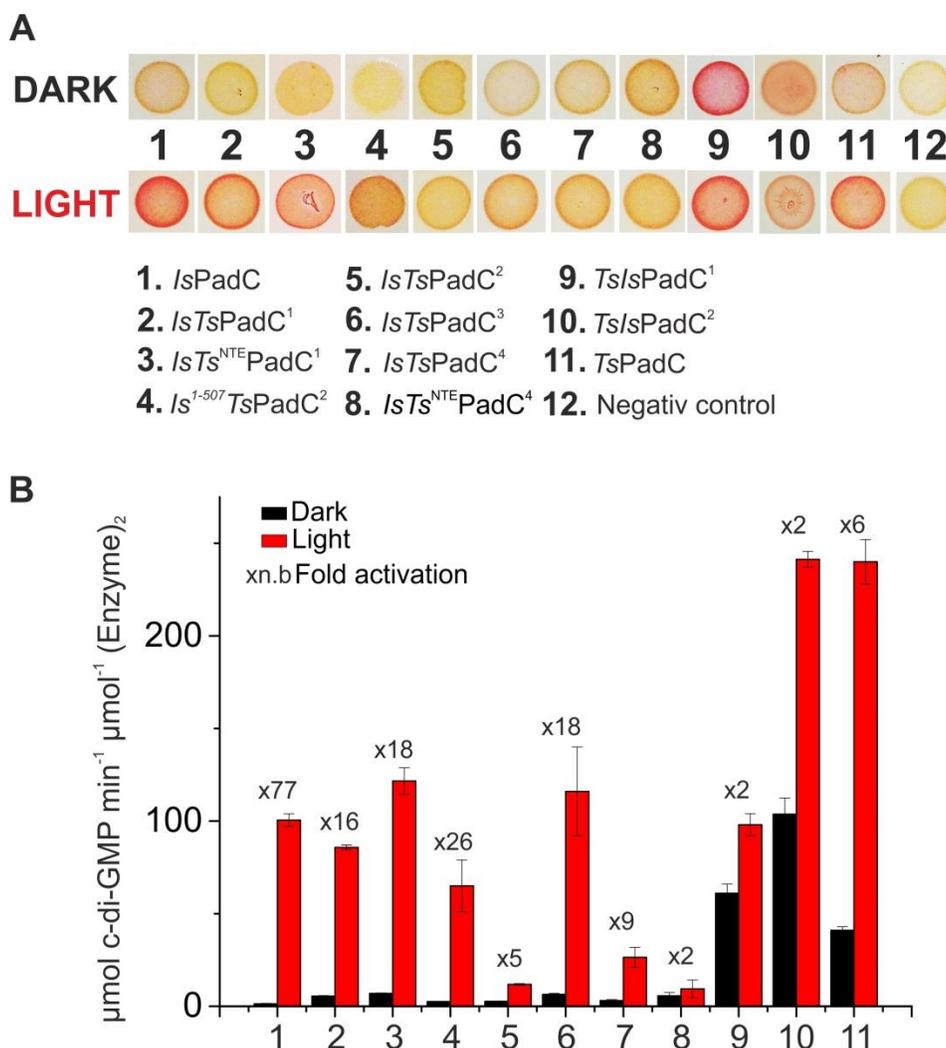


**Figure 67. Photoactivation capacity of the synthetic chimeras.** (**A**) Congo-red based *in vivo* screening of DGC activity. *Is*$^{1-507}$*Ts*PadC$^2$, *IsTs*$^{NTE}$PadC$^1$ and *TsIs*PadC$^1$ were part of another screening plate and have been cut out to show the screening result on a same panel for all the constructs. All the other constructs results of a same screening plate but have been cut out to conserve the same order of construct as described in Fig. 57. (**B**) Comparison of reaction rates between dark and light at 200 μM GTP. The fold activation is indicated above the columns. Initial rates are quantified from experimental triplicates for three time points, and the sample standard deviation of individual points contributed to the error estimation of the linear fit that is used to calculate the initial rate of product formation. The SE of the estimate from the linear regression is used as error indicator.

## 2.3.4. The asymmetric behavior of *Is*PadC is controlled by the PSM dimer interface

In an attempt to characterize the origin of the asymmetric *Is*PadC behavior, we attempted to monomerize the *Is*PadC PSM by introducing a total of 6 amino acid substitutions along the PSM dimer interface (Phe132Glu; Phe136Glu; Phe303Glu; Leu310Glu; Glu272Ala; and Asp494Ala) (Appendix Table 4). Some of the mutations are based on previously reported mutations to generate *Dr*BphP PSM monomer (41). The purification of the variant featuring all 6 substitutions in high salt buffer led to the desired monomeric species with only a subtle residual population of dimers under dark conditions. However upon light illumination the dimer interface is further destabilized and only PSM monomers are observable based on gel filtration analysis (Fig. 68).
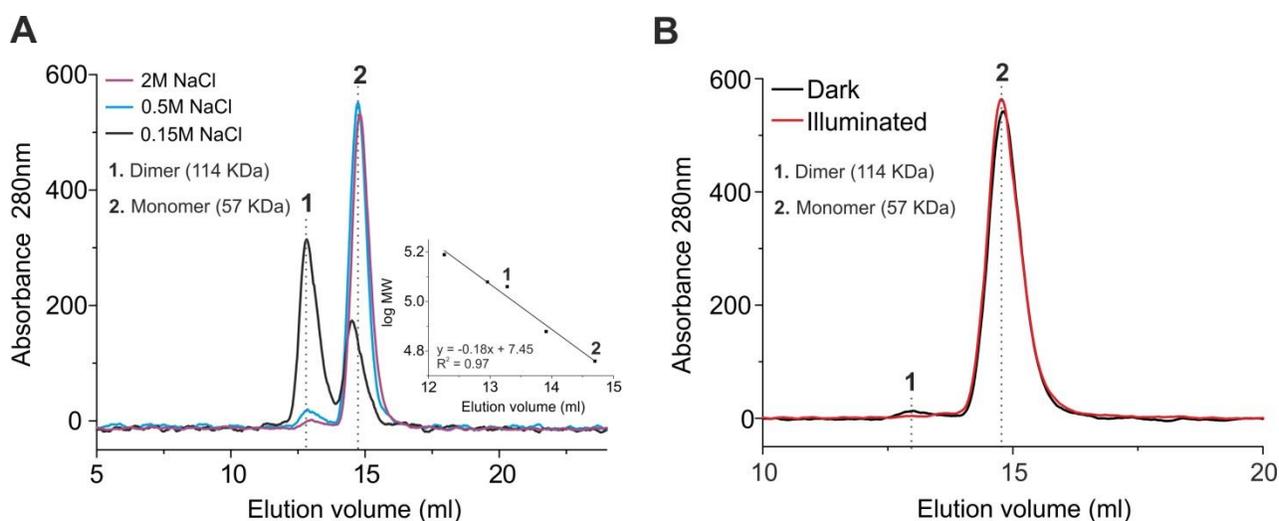


**Figure 68. Characterization of *Is*PadC^PSM monomer.** (**A**) Gel filtration analysis of *Is*PadC^PSM monomerized eluted in dark conditions under different NaCl concentration. (**B**) Illumination of the protein in 2 M NaCl buffer destabilizes further the dimer interface. Monomer and dimer species are determined by internal calibration of the gel filtration column as represented in the inset included in panel A.

Unexpectedly, the UV/Vis spectra of this variant do not indicate the expected homogeneous Pfr-state population. However, the illuminated spectrum shows an increased Pfr-state contribution in equilibrium with other intermediate contributions (Fig. 69A). Potentially the dimer interface of *Is*PadC might also be needed to stabilize the Pfr-state leading to fast

spontaneous dark state reversion in monomerized PSM. Remarkably, the dimer interface is reestablished when lowering the salt concentration, and the UV/Vis spectrum resembles the spectrum of *Is*PadC wildtype in the same salt condition. Denaturation of *Is*PadC^PSM monomerized in low salt condition shows a similar quantity of 15*E* isomer as for the wildtype protein (Fig. 69B). Obviously, lower salt concentration led to dimerization of the PSM and increased the frequency of meta-R like species due to inter-protomeric allosteric effects of the PSM dimer interface on the chromophore environments. Moreover, the thermal recovery of monomerized *Is*PadC^PSM in low salt conditions shows a fast phase with relative amplitude of ~36% similar to *Is*PadC wildtype and indicative of meta-R like species recovery (Appendix Table 1).
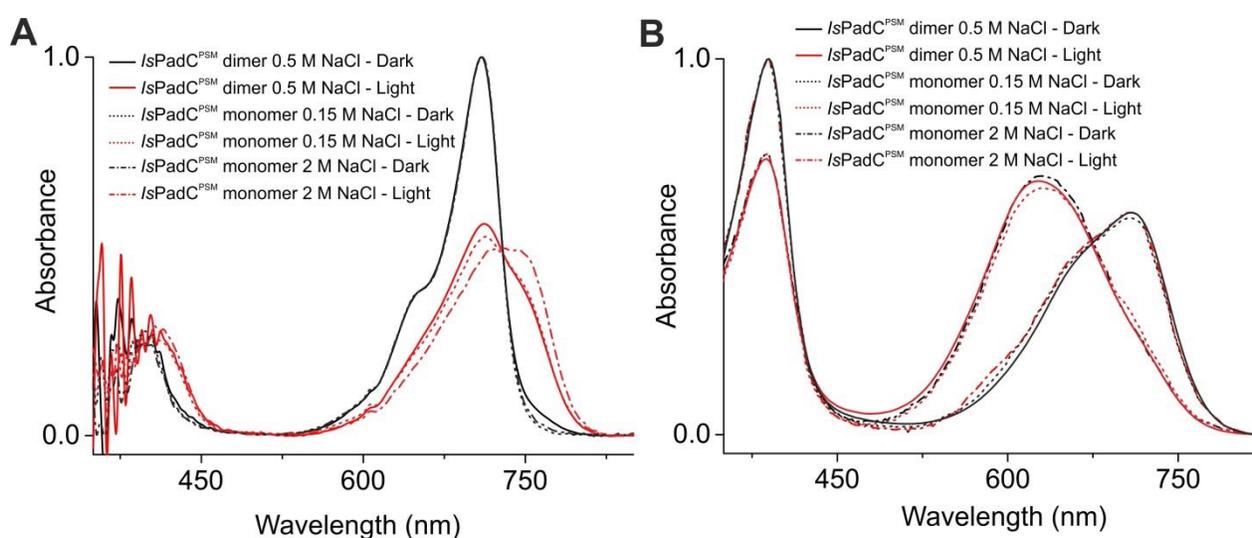


**Figure 69. Spectroscopic characterization of *Is*PadC^PSM monomer.** (**A**) Gel filtration analysis of *Is*PadC^PSM monomerized eluted under dark conditions with different NaCl concentrations. Illumination of the protein further destabilizes the dimer interface as represented in the upper inset. Monomer and dimer species are determined by internal calibration of the gel filtration column as represented in the lower inset. (**B**) UV/Vis spectra of *Is*PadC^PSM monomerized (dashed lines) at different NaCl concentrations in comparison to *Is*PadC^PSM dimer spectra (solid lines). (**C**) Denaturation in Methanol/TCA 0.1 % of *Is*PadC^PSM monomerized (dashed lines) at different NaCl concentrations in comparison to *Is*PadC^PSM dimer spectra (solid lines)

As previously described (32, 77), the deletion of the PHY-tongue in *Is*PadC (*Is*PadC^{Δ442-477::SG}) led to an asymmetry in chromophore environments accompanied by an increase in fluorescence (*cf.* Fig. 43A-B). Interestingly, the PAS-GAF truncation in both *Is*PadC and *Ts*PadC led to a similar asymmetric behavior indicating that the PAS-GAF core dimer interface might be at the origin of the asymmetry in the *Is*PadC PSM homodimer (Fig. 70).
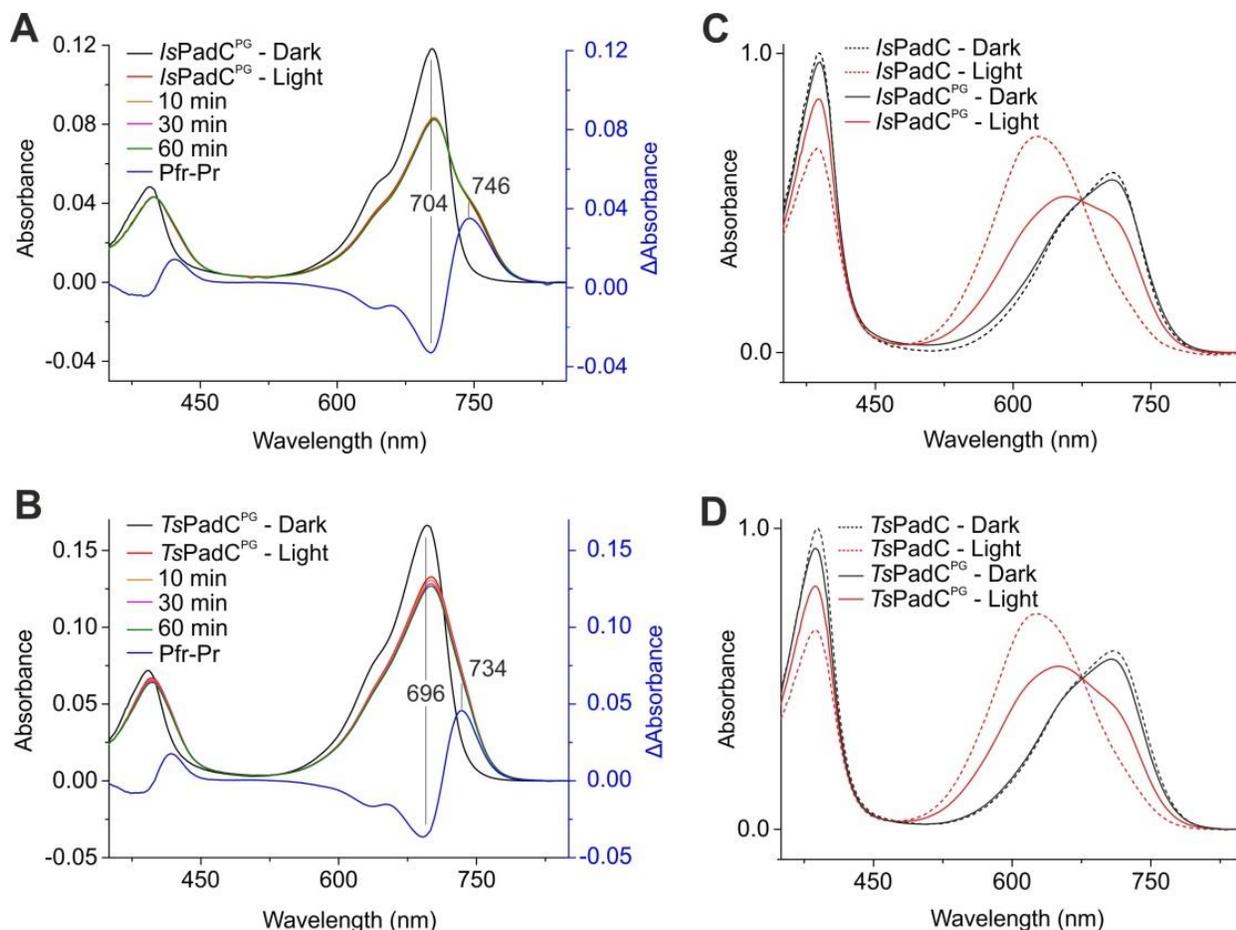
**Figure 70. Spectroscopic characterization of *Is*PadC and *Ts*PadC PAS-GAF truncation variants.** (**A-B**) UV/Vis spectra of *Is*PadC$^{PG}$ (**A**) and *Ts*PadC$^{PG}$ (**B**). After light illumination and population of the illuminated steady state spectrum, several spectra were recorded over 1 hour to show that the PAS-GAF truncation is not able to recover to the Pr-ground state. For *Ts*PadC$^{PG}$, a reproducible time-dependent subtle decrease of the absorption at 696 nm is observed after red-light illumination but the origin of this phenomenon is not understood yet. The Pfr-Pr difference spectrum is shown in blue. (**C-D**) Denaturation in methanol/TCA 0.1 % of *Is*PadC$^{PG}$ in comparison to *Is*PadC (**C**), and denaturation of *Ts*PadC$^{PG}$ in comparison to *Ts*PadC (**D**).

Strikingly, the monomerization of *Is*PadC$^{\Delta442\text{-}477::SG}$ PSM showed a fair increase of chromophore 15*E* population under illumination in comparison to the dimeric version of the variant in identical NaCl concentration (Fig. 71 and 72). Moreover, in high NaCl conditions the dimer interface is further destabilized that led to an even higher increase of the 15*E* chromophore population (Fig. 72). These observations confirmed that the PSM dimer interface mostly dominated by the PAS-GAF dimer interface, affect both chromophores environments and might be at the origin of the asymmetric activation in full-length *Is*PadC.
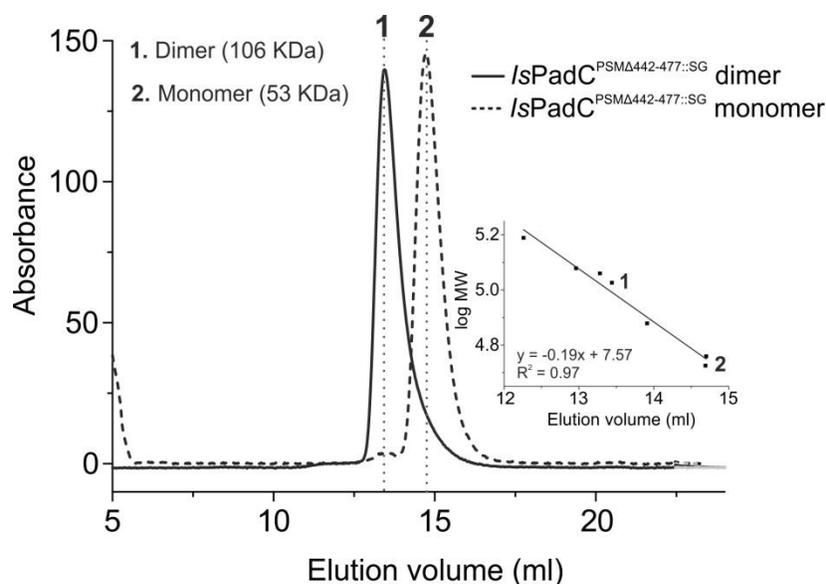
**Figure 71. Monomerization of *Is*PadC^{Δ442-477::SG} PSM.** Gel filtration analysis of *Is*PadC^{Δ442-477::SG} PSM dimer and monomerized variant performed in 500 mM NaCl buffer condition and non-actinic light condition. Monomer and dimer species are determined by internal calibration of the gel filtration column as represented in the inset.
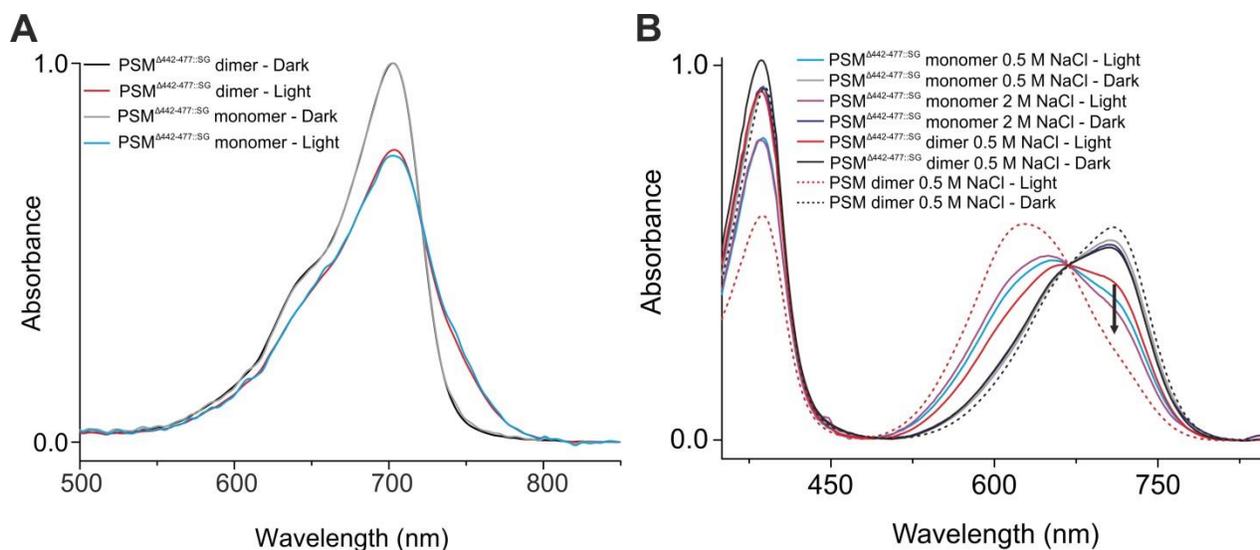


**Figure 72. Inter-protomer interactions induce heterogeneous chromophore environments.** (**A**) UV/Vis spectra of *Is*PadC^{Δ442-477::SG} PSM dimer and monomerized variants recorded in buffer containing 500 mM NaCl. (**B**) Denaturation in methanol/TCA 0.1 % of *Is*PadC^{Δ442-477::SG} PSM dimer and monomerized variants in comparison to *Is*PadC^{PSM} dimer. The increase of 15*E* chromophore population upon *Is*PadC^{Δ442-477::SG} PSM monomerization supports the allosteric effect of inter-monomeric contacts on the chromophore environments.

## 2.3.5.    Discussion

Photoreceptors have naturally evolved to precisely sense various light qualities and intensities to modulate biological output functions. Nowadays, this property is extensively targeted for the development of innovative applications in the field of optogenetics (18, 19, 94, 96). Usually, those photoreceptors are used as building blocks directly fused to biological functions that require a precise spatiotemporal regulation (16, 76, 109, 111). However, since the molecular mechanism of light signal integration and transduction by the light sensor to achieve precise modulation of effectors domains remains poorly understood, such fusions still require a strong effort of screening various fusions points with little control on the final photomodulation capacity. To address this issue, we need to better understand the conformational coordination between the various functional domains and their impact on the molecular mechanism of photoactivation in full-length sensor-effector couples. To this end, guided by the full-length structures of *Is*PadC in dark-adapted (32) and signaling active conditions (77), we designed synthetic chimeras by exchanging functional *Is*PadC domains by the corresponding regions of the homolog *Ts*PadC. Both bacteriophytochromes possesses the same arrangement of functional domains that modulate a DGC effector and form dimer in solution. However, they feature important primary sequences divergences along the dimer interface, within the NTE region, the local environment of the biliverdin, and the PHY-tongue element, that lead to different spectral signatures. The main sequence divergence is however found within the sensor-effector linker region that is 7 residues shorter in *Ts*PadC. Due to the supercoiling character of the coiled-coil linker the *Ts*PadC full-length model generated by homology modelling suggests a subtle reorientation of the DGC dimer in respect to the PSM sensor. The net effect of these sequences differences is that both full-length PadCs present different dynamics of signal transduction with striking difference of photoactivated state stability. In *Ts*PadC, the shorter linker leads most probably to a more flexible dimer interface destabilizing the inhibiting coiled-coil conformation in the dark with the result of substantially more GTP conversion already in the Pr resting state. Even though the DGC effectors are well conserved, their respective exchange between *Is*PadC and *Ts*PadC (*IsTs*PadC[1] and *TsIs*PadC[2]) already slightly impacts the GTP turnover rates and fold of photoactivation. Upon inclusion of the coiled-coil linker from *Ts*PadC and *Is*PadC respectively, (*IsTs*PadC[2] and *TsIs*PadC[1]) the global dynamics of signal transduction are dramatically affected by either decreasing the GTP turnover rate in *IsTs*PadC[2] or the oligomeric

state of the full-length assembly in *TsIs*PadC[1]. Notably, the replacement of the first *Ts*PadC heptad repeat in *IsTs*PadC[2] by the corresponding *Is*PadC heptad repeat reestablished a photomodulation similar to *IsTs*PadC[1]. A similar effect is observed by the inclusion of *Ts*PadC PHY domain in *IsTs*PadC[2] that reestablished a photoactivation capacity by enabling a proper dynamics of signal transduction between the PHY domain and the coiled-coil linker. In line with the structural rearrangements observed between the *Is*PadC and *Is*PadC[Reg2] structures (77) phytochrome PSMs and coiled-coil linkers have co-evolved to provide coordinated conformational changes to achieve the desired photomodulation of specific effector domains. The fusion point between phytochrome PSMs and sensor-effector linkers is therefore an important aspect of PadC engineering as well as of the rational design of any optogenetic tool based on phytochrome systems.

Another conformational interaction that appeared to be crucial for phytochrome photoactivation is the coupling between the NTE and the PHY-tongue elements. Indeed, the insertion of the *Ts*PadC NTE region in *IsTs*PadC[1] slightly shifts the Pr peak maximum absorption towards the *Ts*PadC Pr-state absorption. Strikingly, the *Ts*PadC NTE in coordination with the *Ts*PadC PHY-tongue in *IsTs*[NTE]PadC[4] reestablished a UV/Vis spectrum and thermal recovery properties almost identical to *Ts*PadC wildtype highlighting the crucial role of the molecular interactions between the NTE and PHY-tongue regions on phytochromes photoactivation properties. Identical signal integration properties, however, do not obviously lead to identical signal transduction capacities as observed for *IsTs*[NTE]PadC[4] that features a much lower GTP turnover compared to the *Ts*PadC wildtype. As previously shown for plant phytochromes (168, 169), the NTE region in bacteriophytochromes, which extends through the typical figure-of-eight knot and includes the covalent attachment site of the biliverdin chromophore, has a primordial role in the photochromic diversity of bacteriophytochromes. Even though conserved hydrophobic interactions are present between the NTE and the PHY-tongue, the NTE region remains rather variable in residues composition and length between various phytochrome species even among closely related subfamilies (Appendix Fig. 1). A similar variability in length is observed for the looped tip of the PHY-tongue elements. This might generate different interactions between the NTE and the PHY-tongue in both dark and photoactivated conformations leading to different photoresponses of the full-length holoproteins. Due to the intricate relation between PHY-tongue refolding and the coiled-coil linker conformational dynamics, the stability of the PHY-tongue

interactions with the NTE and the GAF domain might be crucial for the inhibition or stimulation of the coiled-coil linker element.

In respect of the chromophore local environment, the GAF domain features by far most of the direct interactions with the biliverdin chromophore and is therefore directly involved in the chromophore photochemistry. Minimalist phytochromes like cyanobacteriophytochrome described as stand-alone GAF-fold are able to support both covalent binding and proper photoisomerization of the bilin chromophore highlighting the fundamental implication of the GAF domain in the signal integration (170). Indeed, even for the *IsTs*PadC[4] variant that features the *Is*PadC PAS-GAF core and the *Ts*PadC PHY domain, the full-length protein showed the spectral signature of *Is*PadC. Moreover, the fact that the swapping of the GAF domains between *Agrobacterium tumefaciens* canonical and bathy phytochromes species lead to the adoption of a canonical character by the previously bathy phytochrome (168), indicates that subtle sequence conservation in the GAF domain might favor different energetically favorable configurations of the bilin chromophore in the ground state in various phytochromes species. In the case of our two homologs canonical bacteriophytochromes, the high sequence conservation in the GAF domain allows the maintain of the biliverdin in 15Z isomer in the ground state and the molecular interactions between the GAF domain, the NTE, and PHY-tongue elements tune the spectral properties of the full-length holoproteins.

As far as the mechanism of molecular signal transduction is concerned, the high sequence similarity between the local chromophore environments suggests similar initial local changes upon D-ring rotation in *Is*PadC and *Ts*PadC. Indeed, both full-length PadCs exhibit the property of partial Pfr-state conversion within the homodimer population, a property that is conserved by all the so far characterized PadC homologs (32). The denaturation of both bacteriophytochromes in acidic conditions reveals that both chromophore environments feature a 15*E* biliverdin upon constant red light illumination, meaning that an asymmetric Pfr/meta-R-like homodimer might also be a relevant signaling state for *Ts*PadC. However, Pfr contributions in the *Ts*PadC spectrum are slightly blue-shifted (744 nm). Since the local residues around the chromophore are strongly conserved, the molecular basis for this Pfr absorption shift is so far not understood. One should however mention the Tyr189 in *Ts*PadC GAF domain that replaces the Met190 in *Is*PadC and is in relative proximity to the D-ring and potentially affects the absorption spectrum. However the fact that *IsTs*[NTE]PadC[4] containing the PAS-GAF core of *Is*PadC features the same spectral

signature as *Ts*PadC rather suggests that the NTE – PHY-tongue interactions dominate the spectroscopic properties over the PAS-GAF core.

Both PHY-tongue deleted and PAS-GAF truncation variants of *Is*PadC and *Ts*PadC show an asymmetric isomerization of the chromophore population correlating with an increase in fluorescence intensity. Indeed, upon illumination only roughly half of the chromophore population isomerizes to the 15*E* configuration while the remaining chromophore population radiatively decays into fluorescence. This observation suggests that the PAS-GAF dimer interface, which is mainly composed of the central helix bundle, affects the chromophores environments allosterically by inter-protomeric interactions. Indeed, upon monomerization of *Is*PadC$^{\Delta442-477::SG}$ PSM, the population of 15*E* chromophore is increased. Potentially, the molecular cross-talk between the two PSM protomers also applies in the case of the full-length holoprotein. Due to the PAS-GAF dimer interface, when one protomer photoconverts to the final Pfr photoproduct, the other protomer is trapped in an intermediate state, most probably an early meta-R state that is not able to branch to the final Pfr photoproduct thereby forming a functional Pfr/Pr heterodimer.

# 3. Conclusion & Outlook

## 3.1.    Results summary

Phytochromes have sparked much excitement over the last years. Since the first low resolution SAXS envelope of a full-length bacteriophytochrome has been solved in 2006 (136) where only partial functional domains were able to be placed, striking new pictures of phytochromes in dark-adapted and light activated states revealed fundamentally conserved trends in phytochrome activation mechanisms (39, 77, 80, 85, 137). More recently, the structure and functional characterization of full-length phytochrome - effector couples brought a clearer view on the intricate conformational changes involved in modulation of phytochrome sensors and output effectors. Notably, the functional characterization of the sensor-effector linker region, often not resolved in phytochrome 3D structures, reveals the fundamental role of this region in signal transduction and allosteric modulation of the effector domains (32, 56, 76, 77). In this thesis, the focus was set on diguanylyl cyclases as effectors which are regulated by bacteriophytochromes and conclude on new aspects of bacteriophytochrome activation and diguanylyl cyclase regulation. Due to the global structural conservation between diguanylyl cyclases and other nucleotidyl cyclases, which represent interesting pharmacological targets, like guanylyl and adenylate cyclases (171), our results might also be used to better understand the complex modulation of nucleotidyl cyclases by photoreceptors with the ultimate aim to efficiently couple photoreceptors  and nucleotidyl cyclases for optogenetics purposes. Furthermore, the appreciation of molecular details of the sensor-effector linker element might benefit the long-term objective of coupling photoreceptors with unrelated effector systems.

Even though phytochromes have evolved to specifically bind different bilin chromophores, in cyanobacterial-, bacterial-, and plant phytochromes the bilin chromophore is usually constrained in a 15$Z$ isomer form (172) within the GAF binding pocket under the Pr-resting state, and isomerizes at the C15=C16 methine bridge to its 15$E$ form (173) upon irradiation with red light. The only exception comes from the bathyphytochromes that feature a stable 15$E$ bilin isomer under dark condition (173) that can be switched back to the 15$Z$ form upon far-red light illumination. In all of these phytochrome families, photoisomerization of the chromophore triggers local structural rearrangements around the chromophore to accommodate

the rotation of ~180° of the D-ring that ultimately destabilizes the interactions of the PHY-tongue in its β-hairpin conformation leading to a refolding of part of the PHY-tongue to an α-helical conformation (34, 39, 77, 83, 85, 86). Due to the diversity of output effectors naturally modulated by phytochromes, this common early activation mechanism is then differently transduced along the PHY domain and the sensor-effector linker to finally modulate an output function (32, 56, 76, 77, 84). Since many regulated effector domains work in a concerted manner (127, 171, 174), phytochromes often dimerize to form both a stable inhibiting conformation in dark condition and an activated conformation under illumination. Differences of dimer interfaces in full-length phytochrome – effector couples lead to striking differences of photoactivated state stabilities and thermal recoveries to the ground state (32, 77, 79). Even within a phytochrome subfamily modulating the same effector domain, variations in global assembly trigger striking differences in effector modulation (32, 77). Comparative sequence analysis of various full-length PadCs showed that the various homologs feature different sensor-effector linker lengths and composition although retaining a common coiled-coil character (*cf.* Appendix 1). Apparently subtle sequence variations around the chromophore binding site led to different UV/Vis spectral signatures and photomodulation of the GGDEF effectors (32). The elucidation of the *Is*PadC dark-adapted state full-length structure and a closer look at the sensor-effector linker region revealed that this region features a conserved coiled-coil character among PadCs and is able to populate at least two different conformational registers either inhibiting or stimulating the catalytic DGC dimer (32, 77). Subtle deletions in *Is*PadC coiled-coil linker strongly affect the dynamic range of photoactivation although barely affecting the UV/Vis spectral properties (32). These results highlighted the fine-tuning of the coiled-coil sequence composition and length of successive heptad repeats to achieve different stabilities of the inhibiting and stimulating registers that allow high dynamic range of DGC photoactivation. The generation of *Is*PadC coiled-coil linker variants with increased inhibiting or stimulating register stability confirmed the conformational flexibility of the linker to modulate the GGDEF interface, as well as the importance of the coiled-coil sequence composition to satisfy the balanced equilibrium between stabilization and destabilization of each conformational register (77). The full-length structure of an almost constitutively active variant stabilized in the coiled-coil stimulating register (*Is*PadC$^{Reg2}$) confirmed at the molecular level the stimulating register conformation of the coiled-coil linker and revealed that this conformation is intricately linked to the PSM photoactivation structural changes. A single protomer conversion to the Pfr-state is compatible with the coiled-

coil linker in the stimulating register leading to the formation of an asymmetric Pfr/Pr heterodimer (77). The UV/Vis characterization of several PadCs homologs revealed a common illuminated steady-state spectrum with only partial Pr- to Pfr-state photoconversion under constant light illumination independent of the light intensity. However, the denaturation of two naturally occurring PadCs, *Is*PadC and *Ts*PadC, in acidic condition revealed that both chromophores within the homodimer are in 15*E* isomer under illumination. We concluded that the protomer that do not form a classical Pfr-state might most probably be trapped in an early meta-R-like conformation forming an asymmetric Pfr/meta-R-like heterodimer. Considering that Pfr/Pr heterodimers exist in plant phytochromes to sense various ratios of red to far-red light within the ambient light (175–177), similar asymmetric activation of bacteriophytochrome homodimers might also have evolved to adapt to different light qualities (Fig. 73).

The asymmetric activation in *Is*PadC might have been conserved, to specifically activate the dimeric GGDEF effector to produce sequentially the pppGpG intermediate and then the final product c-di-GMP, and might thereby be relevant for many other PadC homologs. While the asymmetric activation of *Is*PadC and its variants leads to similar UV/Vis spectral properties, domain truncation or single residue mutations affecting the global dynamics of signal transduction result in diverse regulation capacities of DGC activity. Coordinated molecular interactions between the various functional domains during signal transduction must be understood in order to rationally design new phytochrome – DGC couples with specific DGC modulation. Especially, the fusion point between the phytochrome PSM and the coiled-coil linker region revealed to be crucial for the GTP turnover rate and dynamic range of photoactivation. Indeed, the swapping of coiled-coil linker – DGC region between the two homologs *Is*PadC and *Ts*PadC led in both cases to an inefficient photoactivation of DGC, and for the fusion of *Ts*PadC PSM with the coiled-coil – DGC region of *Is*PadC a time dependent change in oligomeric state was observed upon thermal recovery to the Pr ground state. However upon replacement of the *Is*PadC PHY domain by the *Ts*PadC PHY-domain, or replacing the first heptad repeat of the coiled-coil linker by the first heptad repeat of *Is*PadC coiled-coil linker, reestablished a high photoactivation capacity in this variant. These results highlight the importance of the fusion point between phytochrome PSMs and coiled-coil linkers and the coevolution of coiled-coil linker compositions and lengths to respond to subtle differences in phytochrome PSMs dynamics.

With respect to the *Is*PadC^Reg2 structure, the conformational dynamics of the coiled-coil linker are intricately linked to the dynamics of the PHY-tongue refolding and the flexibility of the PSM dimer interface. Impairment of the communication between the PHY-tongue, PSM and coiled-coil linker have profound impacts on the global dynamics of signal transduction and eventually on the UV/Vis properties since the PHY-tongue refolding is so far considered as determinant event in the photoconversion to the finale Pfr-state. PHY-tongue refolding triggers the destabilization of interactions between the PHY-tongue, the GAF domain, and the NTE region as supported by the *Is*PadC and *Is*PadC^Reg2 structures. Indeed, in the *Is*PadC^Reg2 structure the NTE region of the protomer adapting a Pfr-state presents a pronounced structural rearrangement. Replacement of both the *Is*PadC PHY-tongue and the NTE regions by the one of *Ts*PadC led to UV/Vis properties almost identical to *Ts*PadC. The NTE region that features the covalent attachment site of the chromophore in bacteriophytochrome, and extends through the typical figure-of-eight knot of the GAF domain, has obviously a crucial effect on bacteriophytochrome spectral properties, as already observed for plant phytochromes (167). In dark conditions its interactions with the PHY-tongue might further stabilize the PHY-tongue in its β-hairpin conformation maintaining the PSM, the coiled-coil linker and the DGC dimer in inhibiting conformations whereas under red light illumination and chromophore isomerization, the NTE undergoes a structural rearrangement further destabilizing the β-hairpin conformation of the PHY-tongue. Altogether, it appears that the NTE region in the context of its interactions with the PHY-tongue element might be an important scaffold for engineering spectroscopic properties of bacteriophytochromes.

During this thesis, interesting new structural insights have been obtained on full-length bacteriophytochromes coupled to diguanylyl cyclase effectors. Together with a functional characterization of naturally occurring full-length homologs and derived variants, the fine-tuned conformational dynamics between the functional domains constituting of phytochrome PSM was highlighted. It was shown that the sensor-effector coiled-coil linker can adopt different conformational registers to modulate the enzymatic activity of DGC effectors. Although important structural rearrangements within the PSM are conserved in respect of previously characterized phytochrome species (137), PadC phytochromes seem to conserve an asymmetric activation of the PSM homodimer. This feature is potentially shared by other phytochrome species based on their UV/Vis spectrum characteristics and might have evolved in

bacteriophytochromes to modulate specific effectors requiring asymmetry for their catalytic activity (178). In the case of *Is*PadC, inter-protomeric interactions have allosteric effects on the PSM dimer interface influencing the local environment of the chromophores. Due to the high flexibility of the PSM dimer interface observed for several phytochrome species this region might also represent an important scaffold for phytochrome design and engineering.
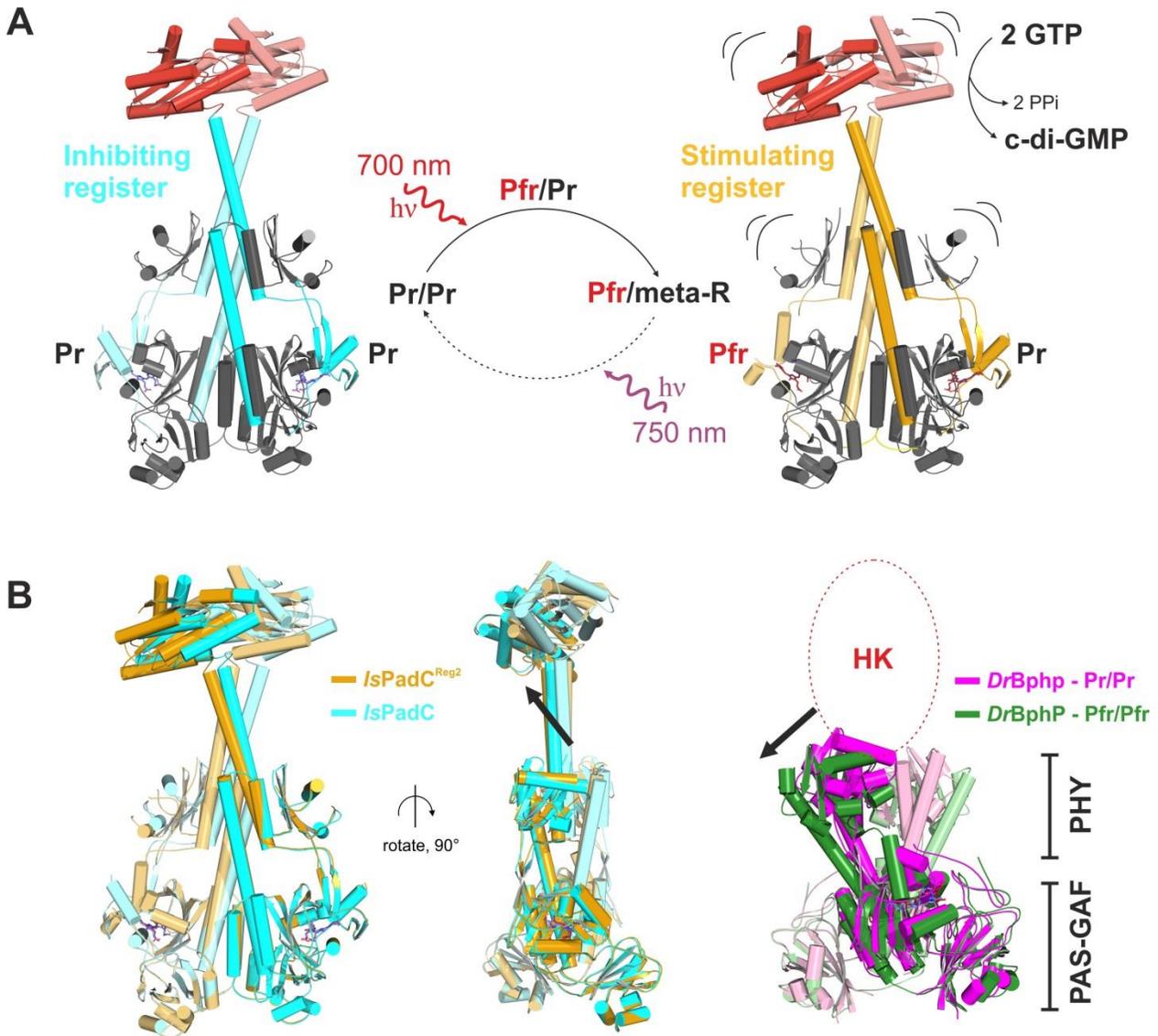


**Figure 73. Global mechanism of light signal transduction in *Is*PadC.** (**A**) Following illumination with red light, the stable *Is*PadC Pr/Pr homodimer in dark conditions absorbs photons from the red light that results in isomerization of both chromophores. Upon rotation of the D-ring, the conformation of the chromophore environments changes triggering the transition from the Pr- to Pfr-state. In *Is*PadC, the conversion of one protomer to the final Pfr photoproduct affects the chromophore environment of the other protomer allosterically, avoiding a complete transition to the Pfr-state in this protomer. The resulting PSM asymmetry allows the translation of the coiled-coil

linker helices to populate the stimulating register that activates the conversion of 2 GTP into c-di-GMP. (**B**) The superposition of *Is*PadC Pr/Pr homodimer and *Is*PadC$^{Reg2}$ Pfr/Pr homodimer structures reveals that activation of the DGC effector only requires a subtle translational movement of the coiled-coil helices (left side). In comparison, the symmetric Pfr/Pfr activation of *Dr*BphP – PSM (green, PDB 5C5K (85)) indicates much larger structural rearrangements of the PHY dimer interface in comparison to the Pr/Pr homodimer (magenta, PDB 4Q0J (34)) (right side). This highlights the fine-tuning of phytochromes dimeric assembly to modulate specific effectors. However the activated structure represents only the PSM truncation of the full-length assembly and recent SAXS based studies on full-length *Dr*BphP proposed more subtle conformational rearrangements upon illumination (88). HK: histidine kinase.

## 3.2.    Outlook

### 3.2.1.    Time resolved vibrational IR- and Raman-spectroscopy of *Is*PadC

All the so far characterized PadC homologs present a common behavior of partial Pfr conversion at the steady-state illuminated spectrum. In the case of *Is*PadC we were able to show that the PSM dimer interface features inter-protomeric contacts affecting the chromophore environments within the homodimer resulting in an asymmetric Pfr/meta-R-like dimer upon illumination. In order to further characterize the conformation of the chromophore environments in full-length *Is*PadC, we started collaboration with the research group of Peter Hildebrandt from the Technical University of Berlin, Germany. His research group is specialized in time-resolved vibrational spectroscopy techniques that are advantageous techniques to precisely attribute a defined spectroscopic state to both chromophore environments and potentially to better understand the unexpected photoconversion cycle of *Is*PadC.

In this respect, the full-length *Is*PadC and the PAS-GAF truncation were analyzed by pre-resonance FT-Raman (RR) with a 1064 nm excitation wavelength and difference Fourier Transform Infrared Spectroscopy (FTIR) in $H_2O$ and $D_2O$ buffers. The full-range RR spectra in dark and under illumination (Fig. 74A-B, left panels) show a similar intensity distribution between *Is*PadC, *Is*PadC$^{PG}$, and Agp1 (used as reference for a symmetric Pfr/Pfr homodimer). Having a closer look at the hydrogen atom out-of-plane rocking (HOOP) region located at the CD

methine bridge of the BV chromophore, the RR spectra shows the expected out-of-plane modes of the D-ring (Fig. 74A-B, right panels). However, due to the mixture Pfr/meta-R of *Is*PadC under constant light illumination, we expected to see a splitting and upshifted HOOP band for *Is*PadC full-length as it is observed for *Is*PadC$^{PG}$. Nevertheless, the HOOP band is relatively broad in comparison to the same region in Agp1 which features a typical Pfr/Pfr homodimer under illumination that could potentially suggests a mixture of spectroscopic states. While this is indicative of meta-R species, a true mixture between Pfr and meta-R species would result of a sharp HOOP band at ~810 cm$^{-1}$ corresponding to Pfr species with a broad shoulder at ~824 cm$^{-1}$ corresponding to meta-R species, but in the case of full-length *Is*PadC this behavior is not observed thereby suggesting that both chromophores are in meta-like configuration under illumination (Fig. 74).
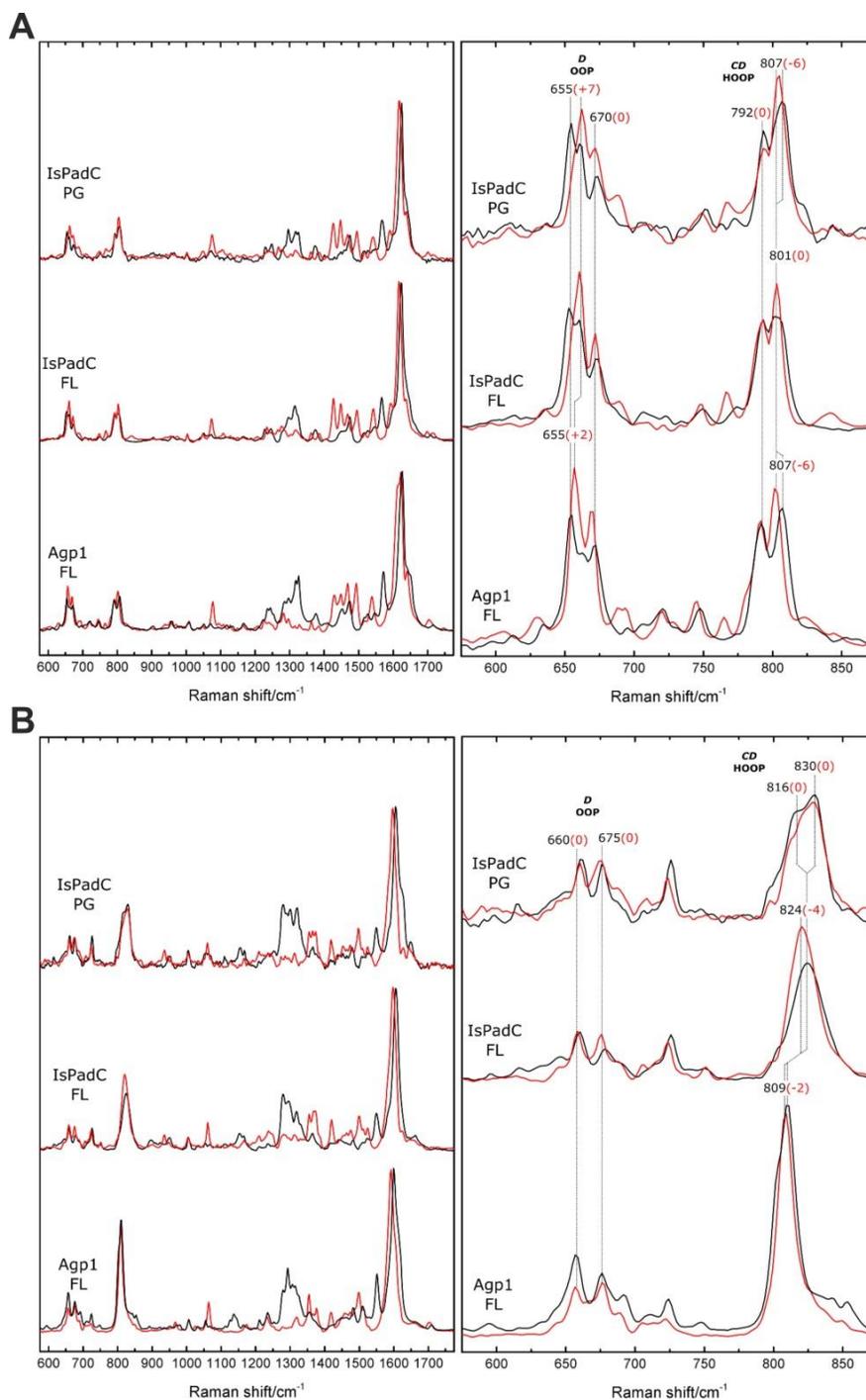
**Figure 74. Resonance raman spectroscopy analysis.** (**A-B**) Full-range RR spectra (left side) and close-up view on the HOOP region (right side) of *Is*PadC[PG], *Is*PadC wildtype, and Agp1 recorded in dark conditions (**A**), and under light illumination (**B**). Black traces are recorded in $H_2O$ and red traces are recorded in $D_2O$. Measurements performed by David Buhrke, TU Berlin, and reproduced with his permission.

The RR region corresponding to the D-ring carbonyl stretching band (Fig. 75) suggests that the C=O band is not visible for *Is*PadC compared to Agp1 under illumination. This observation could suggest an enolic form of the D-ring carbonyl group under illumination that could be in line with the fact that *Is*PadC has a rather quick thermal recovery to the ground state. An enolic form of the D-ring carbonyl would promote the rotation of the D-ring to convert back to the 15Z isomer. However, C=O bands are little Raman-active and show low intensity in RR spectrum. Therefore C=O band might also potentially be present under illumination but at lower intensity requiring higher protein concentration to observe the signal.
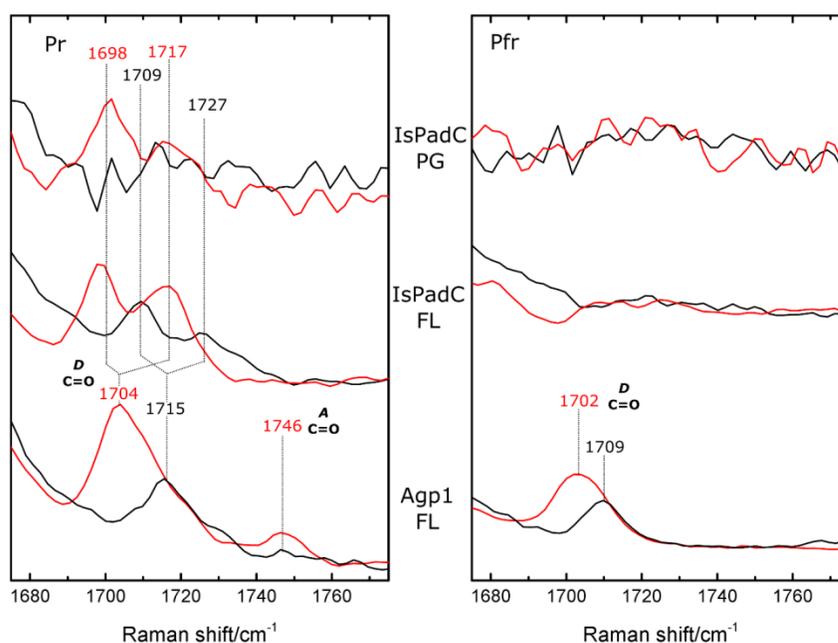


**Figure 75. Resonance raman spectra in the C=O stretching region.** Close-up view on the C=O stretching region of the RR spectroscopy analysis of *Is*PadC$^{PG}$, *Is*PadC and Agp1 in dark conditions (left side) and under light illumination (right side). Black traces are recorded in $H_2O$ and red traces are recorded in $D_2O$. Measurements performed by David Buhrke, TU Berlin, and reproduced with his permission.

From the IR analysis of *Is*PadC we observed that the IR signal of the amide I located at 1636 cm that is characteristic of the PHY-tongue conformational changes has a relatively low intensity in comparison to the same signal for Agp1. This observation would support the existence of an asymmetric Pfr/meta-R heterodimer in *Is*PadC because less structural changes are observed in the PHY-tongue region upon illumination in comparison to the symmetric Agp1 Pfr/Pfr homodimer (Fig. 76A). Interestingly, a time dependent change in the amide I band

intensity is observed for *Is*PadC. While the origin of the time dependent changes so far remains unclear, these changes in the signal could originate from transient tetramerization of activated *Is*PadC dimers due to the very high protein concentration used during the measurements (Fig. 76B), as also observed during our SAXS measurements performed under illumination (*cf.* Appendix Fig. 2).
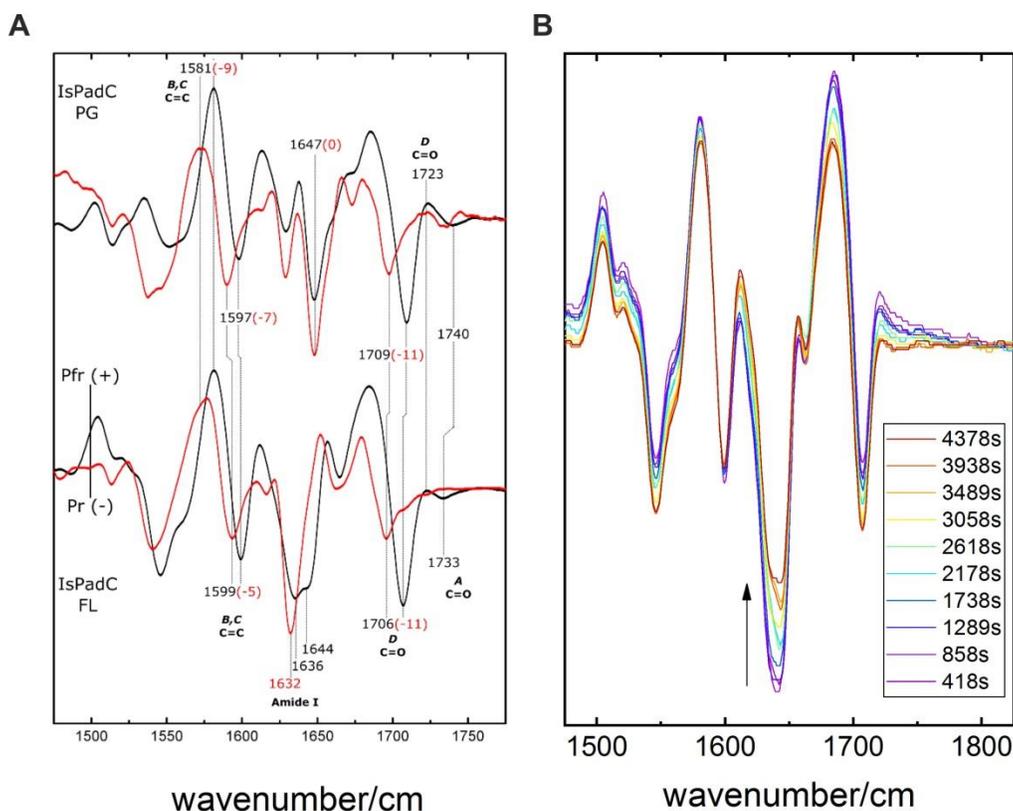


**Figure 76. IR spectroscopy analysis of *Is*PadC and *Is*PadC^{PG}.** (**A**) Full-range IR spectra of *Is*PadC and *Is*PadC^{PG}. (**B**) Time dependent changes in the amide I band intensity. Measurements performed by David Buhrke, TU Berlin, and reproduced with his permission.

## 3.2.2.    NMR analysis of full-length PadC

In another attempt to obtain structural information on PadC photoactivation dynamics, we started to collaborate with the research group of Tobias Madl of the Medical University of Graz to analyze the full-length *Is*PadC and *Is*PadC^{Reg2} variant by NMR methyl-selective isotope labeling. After successful expression of *Is*PadC and *Is*PadC^{Reg2} in minimal media containing ^{13}C isotope labelled amino acids precursors of Ile, Leu/Val, and Met (50 mg, 100 mg, and 50 mg,

respectively), the proteins were then purified using our traditional procedure (*cf.* materials and methods part). A methyl-TROSY NMR spectrum was then recorded successively on a 600 MHz and 700 MHz NMR magnet under dark conditions.

Both proteins were stable over the measurement time and presented a rather good signal to noise ratio at a standard protein concentration of 5 mg ml$^{-1}$. Promising chemical shifts are observed for some peaks corresponding to Leu or Val residues by comparing *Is*PadC with the variant *Is*PadC$^{Reg2}$ (Fig. 77). Further investigations need to be done to assign the various peaks to their corresponding residues. To this extend, the measurements will be repeated from more concentrated samples prepared from bacterial cells cultivated in D$_2$O minimal medium in order to increase the signal to noise ratio. Furthermore, individual modules of *Is*PadC (PSM and DGC) will be analyzed in order to assign specifically the NMR signals coming from the residues in the coiled-coil linker region.
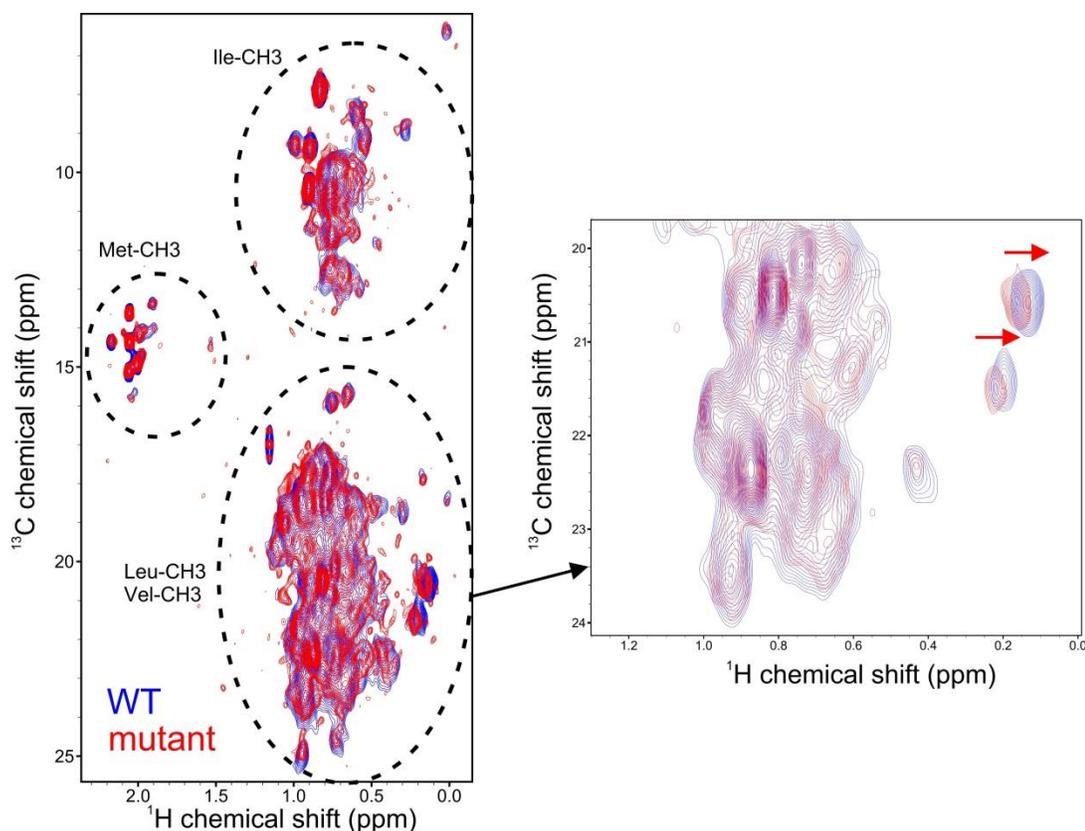


**Figure 77. NMR Methyl-TROSY spectra of *Is*PadC and *Is*PadC$^{Reg2}$.** 2D Methyl-TROSY spectra of $^{13}$C isotope specific labeling of methyl groups in Ile, Leu/Val, and Met residues were recorded with at 700 MHz NMR spectrometer. *Is*PadC$^{Reg2}$ variant featuring preferentially the stimulating register in its coiled-coil linker region show

interesting chemical shifts of Leu/Val residues. Measurements performed by Tobias Madl at the Medical University of Graz, and reproduced with his permission.

### 3.2.3.     Spectroscopic diversity of phytochromes

So far in this thesis, the focus was set on PadC homologs diverging by multiples of 7 residues in their sensor-effector linker regions. However an extended analysis or PadC homologs revealed other PadC homologs with untypical linker lengths (Appendix Fig. 1). We initially expressed, purified, and biochemically characterized 3 selected homologs, *As*PadC, *Th*PadC, and *Tc*PadC, with linker length differences of 11, 5, and 14 residues in comparison to the *Is*PadC linker, respectively. Strikingly, two homologs (*Th*PadC, and *Tc*PadC) present a strongly blue shifted maximum of the Q-band absorption in the Pr-state (665 nm) while their Pfr-state absorption remains at 750 nm. Such a Pr-state absorption is usually typical for plant and cyanobacterial phytochromes binding phytochromobilin or phycocyanobilin, and was also observed for an eukaryotic algal phytochrome binding phycocyanobilin (179) but so far has not observed for bacteriophytochromes binding biliverdin. *Tc*PadC and *Th*PadC spectra feature a smaller cross-section of Pr and Pfr absorptions peaks compared to traditional bacteriophytochrome spectra and thereby are very promising systems to better understand the origin of the photochromic diversity in phytochromes. So far relatively low amounts of pure full-length proteins for those 3 homologs (*Tc*PadC, *Th*PadC, *As*PadC) have been isolated and some optimization of expression and/or purification protocols will be necessary to purify higher amounts for their further biochemical characterization. Since *Th*PadC presents a light inhibition of the DGC dimer upon screening in *E. coli* (Fig. 78), it would be interesting to solve the full-length structure of *Th*PadC in dark conditions to potentially observe a catalytically active DGC dimer linked to the Pr conformation and comparing it to the *Is*PadC dark-adapted and activated state structures.
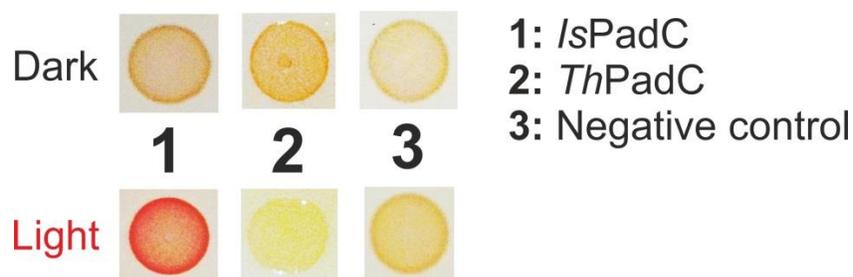
**Figure 78. *In vivo* screening of *Th*PadC activity.** In comparison to the well characterized *Is*PadC, *Th*PadC shows a red light inhibition of *in vivo* DGC activity.

## 3.2.4.    Diguanylyl cyclase catalytic reaction and regulation

The dynamics involved in diguanylyl cyclase catalytic activity so far remain relatively poorly understood. The conversion of GTP to c-di-GMP passes by the formation of the linear pppGpG intermediate, however, it remains unclear how the substrate reorients within the active site following the first phosphodiester bond to form the second phosphodiester bond, or if the two phosphodiesters bonds are form successively with unchanged orientation of the substrate in the active site. Since no structure of the DGC dimer in complex with the pppGpG was so far obtained, the conformational change of the DGC dimer upon GTP catalysis remain overall poorly understood. Moreover expression and purification of isolated DGC dimers frequently results in complexes with the substrate GTP or with the product c-di-GMP on the inhibitory I-site, which complicates the functional interpretation of the various catalytic states of DGC dimers (127, 151). Furthermore, the rather flexible dimer interface of the DGC observed from SAXS reconstruction of full-length *Is*PadC, suggested that the modulation of DGC activity involves the sampling of various DGC dimer conformations. Although the structure of *Is*PadC<sup>Reg2</sup> shows an interesting opened DGC dimer never observed so far, it is difficult to attribute a specific functional state to this dimer considering the absence of substrate during crystallization.

We observed that the mixing of the unproductive analogue GTP-α-S with classical GTP led to the formation of an artificial intermediate ppp-αS-GpG that cannot be further processed by the DGC due to the S atom on the interacting phosphate group. Taking advantage of this property we purified some of this artificial ppp-αS-GpG via HPLC. However, the relatively low probability of the presence of a GTP and a GTP-α-S in the DGC active sites and the inhibiting

effect of the presence of two GTP-α-S led to overall a very low production yield of the artificial ppp-αS-GpG. Furthermore, with the polarity of this compound being rather similar to the product c-di-GMP, the two molecules elute at relatively similar elution volume on the HPLC column with the methanol gradient used, thereby limiting the collection of the artificial ppp-αS-GpG to the non-overlapping part of the elution peaks. Therefore, reactions conditions (ratio GTP:GTP-α-S) and the HPLC methanol gradient need to be optimized to efficiently purify larger quantities of the artificial ppp-αS-GpG intermediates. Nevertheless, from the so far ppp-αS-GpG purified, an initial soaking experiment was tried on *Is*PadC^Reg2 crystal to see if the opened DGC dimer might represent the pppGpG binding state of the DGC dimer. However, the relatively low concentration used during soaking (100 μM) did not lead to any ppp-αS-GpG electron density within the DGC active site and the solved structure is identical to *Is*PadC^Reg2 structure. Only weak electron density blobs were observed on the DGC dimer that could potentially correspond to pyrophosphate.

To further understand the DGC dimer catalytic mechanism, we attempted to isolate and purify the coiled-coil linker – DGC dimer from *Is*PadC (ccDGC) with the expectation to obtain high resolution structure of the DGC dimer coupled to its coiled-coil linker. However, expression and purification of ccDGC led to the DGC dimer complexed with c-di-GMP at the inhibitory site. Crystallization of the variant only led to a 3Å resolution structure with c-di-GMP in the active and inhibitory sites and an unfolded coiled-coil linker (Fig. 79). Removing of the c-di-GMP bound by incubation of the protein with phosphodiesterase led to the free protein which however was not able to crystallize so far (*cf.* materials and methods part).
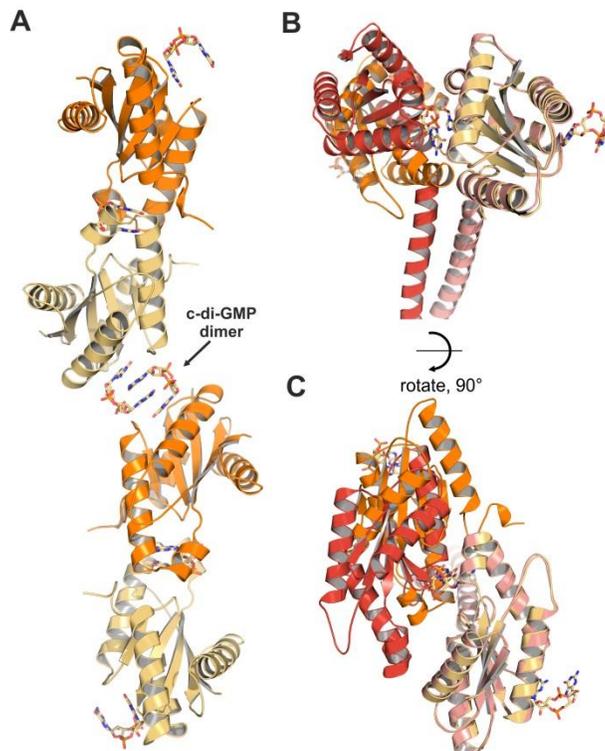
**Figure 79. Structure of the *Is*PadC coiled-coil - DGC truncation (ccDGC) in complex with c-di-GMP.** (**A**) Two symmetry related dimers of ccDGC crosslinked by a c-di-GMP dimer bound to the DGC I-site. The two protomers of a ccDGC dimer are colored in orange and light-orange. C-di-GMP bound at the I-site and active center is represented in sticks. (**B-C**) Superposition of the ccDGC structure and the *Is*PadC DGC dimer based on residues 528-683 of chains A visualized in front view (**B**) and top view (**C**). The coiled-coil linker of the ccDGC structure is unstructured and no electron density for this region was observed.

# 4. Materials and methods used

The methods and material used to generate the results presented in this thesis are briefly summarized in the following paragraphs.

## 4.1.    Protein preparation

The coding sequences of naturally occurring PadC homologs from *Idiomarina sp.* A28L (WP_007419415), *Marinimicrobium agarilyticum* (WP_027329460), *Thioalkalivibrio sp.* ALMg3 (WP_026331574), *Thauera chlorobenzoica* (WP_075148475.1), *Thiohalospira halophila* (WP_093427120.1), and *Agromyces sp. Leaf222* (WP_082462169.1) corresponding respectively to *Is*PadC, *Ma*PadC, *Ts*PadC, *Tc*PadC, *Th*PadC, and *As*PadC were synthetized with codon optimization for expression in *Escherichia coli* bacterial host (GeneArt, Life Technologies). Gene cassettes were then cloned into the pETM-11 vector by means of restriction ligation using the NcoI and NotI restriction sites in order to express the proteins with a Tobacco etch virus (TEV)-cleavable polyhistidine tag. All the sequence protein variants described in this thesis are derived from the PadC homologs by site-directed mutagenesis following the protocol described by Liu and Naismith (180) (primers used can be found in Appendix Table 4).

All the protein variants generated were expressed as $(His)_6$-tagged holoproteins complexed with the biliverdin chromophore in *E. coli* BL21 (DE3) co-expressing the heme oxygenase (HO-1) from *Synechocystis* sp. PCC6803. Pre-cultures were diluted 1/50 in a main culture volume grown until mid-log phase at 37 °C in LB medium supplemented with 10 mM $MgCl_2$ and 0.2 % glucose. Cultures were then shifted to 16 °C prior to induction with 0.25 mM isopropyl-β-D-thiogalactopyranoside (IPTG) for ~15 hours in the presence of δ-aminolevulinic acid (10 mg $l^{-1}$). After harvesting, cells were lysed in lysis buffer (*cf.* Appendix Table 4) by combination of lysozyme (100 µg $ml^{-1}$) and sonication (2 x 5 min, 100 W, continuous mode). Lysates were clarified by ultracentrifugation (206,000 g) and the soluble holoprotein purified by affinity chromatography on a $Ni^{2+}$-Sepharose matrix by gravity flow. After loading of the column matrix with the lysate, contaminant proteins were washed out by lysis buffer containing 30 mM imidazole, and the holoproteins were finally eluted by lysis buffer containing 250 mM imidazole.

To decrease the imidazole concentration and cleave the polyhistidine tag, the eluted proteins were subjected to overnight dialysis in the presence of a ratio ~1:16 for TEV/substrate. The Histidine-tagged TEV protease and the cleaved (His)6-tag were separated from the holoproteins by reloading the dialysate onto the $Ni^{2+}$-Sepharose matrix and collecting the flow-through. After concentration, the holoproteins were finally purified by size-exclusion chromatography on a 16/60 Superdex 200 prep grade column in buffer according to Appendix Table 4. The purified proteins were concentrated, flash-frozen in liquid nitrogen and conserved at -80 °C until needed. All the steps mentioned above were performed under non-actinic light conditions.

To express *Is*PadC loaded with selenomethionine, the protein was expressed in *E. coli* BL21 (DE3) co-expressing the heme oxygenase grown in minimal medium containing increased concentrations of certain amino acids for metabolic inhibition of methionine biosynthesis (181) and supplemented with 50 mg $ml^{-1}$ of selenomethionine. The expressed protein was purified as described above with the only modification that 1 mM DTE was added to the finale storage buffer to prevent oxidation of selenomethionine.

*Degradation of the c-di-GMP bound to the ccDGC truncation of IsPadC* – The ccDGC truncation variant of *Is*PadC can only be co-purified with the product c-di-GMP bound to the allosteric I-site and to the GGDEF active site. Due to the inhibitory effect of c-di-GMP bound to the I-site (151), it was required to remove the c-di-GMP from the protein before measuring the activity of GTP conversion. To this end, ccDGC was incubated 40 min at 30 °C in presence of the phosphodiesterase RocR, previously purified following the method described in Rao *et al,.* (182), and stored in the following buffer [Tris/HCl pH 8.0 20 mM, NaCl 500 mM, glycerol 5 % (w/v), KCl 25 mM, $MgCl_2$ 10 mM, imidazole 10 mM]. RocR and ccDGC - c-di-GMP were mixed with a ratio ccDGC to RocR of 1:1. Because the RocR carries a His-tag, it is possible to separate RocR from the ccDGC variant on a $Ni^{2+}$ affinity chromatography column after the incubation time. The flow-through was collected and loaded on a Superdex 200 10/300 analytical grade column to fully purify the treated ccDGC.

## 4.2. Congo red *in vivo* based assay

The previously described method of Antoniani *et al,.* (125) was adapted to efficiently estimate *in vivo* DGC activity induced in *E. coli* by the various constructs generated. To this aim, *E. coli* BL21 (DE3) cells co-expressing the heme oxygenase and transformed with pETM-11 based PadC expression plasmids were grown at 30 °C in YESCA medium [casamino acids (10 mg ml$^{-1}$) and yeast extract (1.5 mg ml$^{-1}$)] supplemented with MgSO$_4$ (0.05 mg ml$^{-1}$), FeSO$_4$ (0.5 mg ml$^{-1}$), kanamycin (30 mg ml$^{-1}$), and chloramphenicol (34 mg ml$^{-1}$) to mid-log phase. After normalization of each culture to an OD$^{600}$ of 0.5, each culture was induced with 0.25 mM IPTG and δ-aminolevulinic acid (10 mg l$^{-1}$) for 4 hours at 16 °C. Afterwards, 3 µl of the induced culture was spotted on YESCA agar plates containing kanamycin (30 mg ml$^{-1}$), chloramphenicol (34 mg ml$^{-1}$), and Congo red (0.01 mg ml$^{-1}$) and incubated at 20 °C for 16 hours in the dark or under constant red light illumination (75 mW cm$^{-2}$ at 630 nm). As a negative control, we used the pETM-11 AppA construct (183) that does not show any DGC activity.

## 4.3. UV/Vis characterization

Ultraviolet (UV)‑visible (Vis) absorption spectra were acquired with a CCD-based Specord S300 Vis spectrophotometer using diluted samples in their appropriate storage buffer (*cf.* Appendix Table 4) and equilibrated at 20 °C prior to measurement. For protein variants with highly stabilized photoactivated states and partially purified in the activated state due to ambient light, a prolonged incubation in dark or a short illumination with far-red light (720 nm) was performed prior to measurements. Dark-adapted Pr-state absorption spectra were measured from samples equilibrated under non-actinic light conditions and minimized time contact with the measuring light of the spectrophotometer using a neutral density filter (ND = 1.0 or 2.0) between the light source and the sample cuvette. Pfr-enriched spectra were recorded under constant red light illumination (660 nm, 45 mW cm$^{-2}$) in the presence of the same neutral density filter for attenuating the far-red light contributions of the measuring light.

Pr-state recovery kinetics were followed at the Q-band maximum absorption of respectively a homogeneous Pr-state sample (recovery of Pr-state contributions), and illuminated

sample after 1 min red light illumination (depletion of Pfr and meta-R-like states contributions) using a Specord 200 spectrophotometer with 10 ms integration time sampled every 5 sec. Integration times and time intervals of the time scans were adjusted for the individual variants to minimize the effect of the actinic measuring lights. Reported time constants of dark-state recovery differ by less than 10 % between the Pr- and Pfr-state Q-band absorption wavelengths used.

The UV/Vis absorption of 15$Z$ and 15$E$ isomeric forms of biliverdin were measured based on the previously described protocol of Thümmler *et al,*. (143). The protein sample sufficiently concentrated to observe enough amplitude in the signal ($\geq$ 10 µM) is diluted at least 1:10 in quenching buffer (0.1 % trichloroacetic acid in methanol) in dark condition or directly after 1 min illumination (660 nm, 45 mW cm$^{-2}$) and directly measured with a Specord 200 spectrophotometer. Re-illumination of the denatured Pfr-state led to the conversion from 15$E$ biliverdin isomer to the 15$Z$ isomer. Therefore Pfr-peptides were denatured directly after switching off the red light source and kept in the dark until their measurements.

## 4.4. Fluorescence measurements

Fluorescence measurements were performed using an RF-5301 PC spectrofluorophotometer equipped with a concave blazed holographic grating excitation and emission monochromator, and containing a 150 W Xenon lamp as light source. Samples diluted to 2 µM were equilibrated at room temperature and complete darkness for the dark state measurement or 1 min under saturating red light (660 nm, 45 mW cm$^{-2}$) for the light state measurement prior to data acquisition. A sampling interval of 0.2 nm was used with a slit opening of 1.5 nm for the excitation slit and 5 nm for the emission slit.

## 4.5. HPLC-based enzyme kinetics measurements

To record the conversion of GTP to c-di-GMP for all the variants, we adapted a protocol for high performance liquid chromatography (HPLC) from a previously described method (184).

Briefly, purified protein samples were mixed with GTP at various concentrations in reaction buffer (50 mM Hepes pH 7.0, 500 mM NaCl, and 50 mM $MgCl_2$) at 20 °C under non-actinic light for the dark state measurement, or under constant red light illumination (630 nm, 0.7 mW $cm^{-2}$ or 660 nm, 45 mW $cm^{-2}$) following a 1 min pre-illumination of the sample for the light state measurements. Samples were then thermally inactivated by 1 min incubation at 95 °C. After separating the substrate and products from denatured protein by centrifugation, the nucleotides were separated by a linear 7 min gradient from 2 % to 20 % methanol using a reversed phase HPLC column (SunFire C18 4.6 mm x 100 mm) equilibrated in 10 mM triethylammonium formate (pH 6.0). For protein variants isolated in a partially activated state and featuring a slow thermal recovery a prolonged incubation of the sample under far-red light (730 ± 20 nm) coupled to a far-red light bandpass filter (750 nm) at room temperature and complete darkness was performed to fully populate the Pr-state of the sample prior to the dark state measurements. All kinetic data were normalized to the concentration of the dimeric protein, and all samples were corrected for the amount of c-di-GMP formed during the inactivation step.

## 4.6.     Structure data collection and refinement

*Crystallization setup* – Full-length *Is*PadC (PDB 5LLW) has been crystallized under non-actinic light conditions by a hanging-drop vapor diffusion setup at 289 K. 2 µl drops containing equal volumes of *Is*PadC (7 mg $ml^{-1}$) in the Pr-state, and reservoir solution (0.1 M Bis-tris pH 5.5, 0.1 M ammonium acetate, 17 % (w/v) polyethylene glycol 10,000) were equilibrated against 500 µl of reservoir solution. Already after overnight incubation, elongated thin plate-like crystals were observed that became thicker during one week of incubation. To obtain selenomethionine-labeled *Is*PadC crystals, the same procedure was performed with 1 mM dithioerythritol (DTE) in the protein solution and identically looking crystals were harvested after 15 days of growth. Crystals were harvested under low-intensity non-actinic light conditions (green LED, 520 ± 20 nm) by transferring the crystal in a cryoprotectant solution (reservoir solution containing 25 % glycerol) and immediately flash-frozen in liquid nitrogen. The GTP-bound crystal structure of *Is*PadC dark-adapted state (PDB 5LLX) was obtained by soaking *Is*PadC crystals with 10 mM GTP in cryoprotectant solution for 5 min under non-actinic light conditions. The structure of *Is*PadC PSMcc (PDB 5LLY) resulted from the proteolytic degradation of *Is*PadC crystals after

prolonged incubation (> 1 month) forming plate-like crystal with a slightly different morphology compared to full-length crystals.

The variant $IsPadC^{Reg2}$ (PDB 6ET7) was crystallized at 293 K under non-actinic light conditions using a sitting-drop vapor diffusion setup. The best diffracting crystals were obtained by seeding, using streak seeding method, 2 µl drops containing equal volumes of $IsPadC^{Reg2}$ (5 mg ml$^{-1}$) and reservoir solution (0.1 M magnesium formate, 12 % (w/v) polyethylene glycol 3,350) equilibrated against 35 µl of reservoir solution. Plate-like crystal morphologically very similar to $Is$PadC crystals was observed after overnight incubation and reached final dimension within 10 days of crystal growth. Crystals were harvested under non-actinic light conditions and transferred into a cryoprotectant solution of reservoir containing 30 % glycerol and immediately flash-frozen in liquid nitrogen. A GTP-bound structure of the variant was obtained by soaking $IsPadC^{Reg2}$ crystal in 10 mM GTP in cryoprotectant solution for different times ranging from 5 min to overnight incubation, however, a rather fragility of the crystals is observed upon soaking resulting to poor quality diffracting crystals. Nevertheless, diffraction around ~3 Å has been collected from an overnight soaked crystal but the refinement procedure is not yet complete and the structure has not been deposited.

The Coiled-coil – DGC variant of $Is$PadC (ccDGC) was crystallized at 289 K in its purified form with c-di-GMP bound by sitting-drop vapor diffusion setup. To obtain the best diffracting crystals 2 µl drops of ccDGC (10 mg ml$^{-1}$) were seeded by streak seeding method and equilibrated against 35 µl of reservoir solution (0.1 M potassium thiocyanate, 30 % (w/v) polyethylene glycol 2,000 MME). Crystals were harvested and transferred in cryoprotectant solution containing 25 % glycerol and immediately flash-frozen in liquid nitrogen. The refinement of the structure is not complete and the structure has not been yet deposited.

*Diffraction data collection and refinement - Diffraction* data for native and Se-Met labeled $Is$PadC were collected at beamline ID29-10 and ID23-1, respectively, of the European Synchrotron Radiation Facility (ESRF). Diffraction data for $IsPadC^{Reg2}$ were collected at beamline P11 of the Deutsches Elektronen-Synchrotron (DESY). Diffraction data for ccDGC were collected at the beamline XRD1 of the Elettra-Sincrotrone Trieste.

Diffraction data were processed using the XDS program package (185). For Selenomethionine labeled $Is$PadC, several data sets collected at multiple positions were merged

for successful phasing by selenium single-wavelength anomalous dispersion. Due to the moderate quality of the initial anomalous electron density map calculated with PHENIX AutoSol (186) obtained from the 3Å Selenomethionine *Is*PadC diffraction data, we combined the phase information obtained from the molecular replacement performed with individuals search models of the PAS-GAF core, PHY, and GGDEF domains adapted to *Is*PadC sequence using PHENIX Sculptor (128). The PAS-GAF core of *R. Palustris* bacteriophytochrome *Rp*A3015 (PDB 4S21) (79), the PHY domain of *R. palustris* bacteriophytochrome *Rp*BphP3 (PDB 4R70) (79), and the GGDEF domain of WpsR from *Pseudomonas aeruginosa* (PDB 3I5B) (127) were used as search models. Two molecules of each search model were successfully places in the asymmetric unit using PHENIX Phaser (187). The initial model of PAS-GAF-PHY-GGDEF dimer was subsequently used for rigid-body fitting in the anomalous density map and finally refined in several rounds of maximum likelihood target (MLHL). In addition, torsion non-crystallographic symmetry (NCS) restraints and secondary structures restraints were included together with Translation/Libration/Screw (TLS) groups for the individual domains of the protein. During the final rounds of refinement, reference model restraints from the higher-resolution PSMcc structure were included and optimization of X-ray and ADP weights was performed.

The crystal structure of *Is*PadC PSMcc fragment was solved by molecular replacement using PHENIX Phaser with an intermediate model of the PAS-GAF-PHY fragment from the full-length structure of *Is*PadC as search model. Refinement included an initial simulated annealing (torsion) step followed by several rounds of maximum-likelihood least-squares refinement of models modified with Coot, as described above. TLS and secondary structure restraints were applied during the refinement.

The crystal structure of *Is*PadC soaked with GTP was solved by molecular replacement using PHENIX Phaser with models of the dimeric *Is*PadC PSM containing the coiled-coil linker and the dimer of the *Is*PadC DGC. Refinement included an initial rigid-body fit of individual domains because the linker element revealed a more pronounced kink in the soak crystal, followed by several rounds of maximum-likelihood least-squares refinement of models modified with Coot, as described above. TLS, NCS, and secondary restraints were applied during the refinement.

The crystal structure of *Is*PadC$^{Reg2}$ variant was solved by molecular replacement using PHENIX Phaser with the PAS-GAF fragment and the PHY-domain lacking the tongue extensions of the full-length structure of IsPadC (PDB 5LLW), as well as the DGC fragment of the IsPadC GTP soaked structure (PDB 5LLX). Two molecules for the PAS-GAF, PHY, and DGC search models were successfully placed in the asymmetric unit. The initial model of a PAS-GAF-PHY-GGDEF dimer was then manually extended to build the PHY-tongue and coiled-coil linker regions in several rounds of maximum-likelihood refinement of models modified with Coot using σA-weighted $2mFo – DFc$ and $Fo – Fc$ electron density maps. In addition, torsion-NCS restraints and secondary structure restraints were included together with TLS groups for the individual domains of the protein. During the final rounds of refinement reference model restraints were applied on PAS-GAF-PHY excluding the PHY β-hairpin extension using a previously solved high resolution *Is*PadC PSMcc structure (32), and on GGDEF domains using the *Is*PadC dark state structure (32) and optimization of X-ray and ADP weights was performed.

Spectral properties of the *Is*PadC and *Is*PadC$^{Reg2}$ crystal were recorded at 100 K under a cryostream and at room temperature using a micro-spectrophotometer at the cryobench ID29S of the ESRF facilities (188).

## 4.7. Small-angle x-ray scattering

SAXS data for solutions of *Is*PadC in the light and dark states were recorded by Christoph Göbl and Tobias Madl at the Technical University of Munich on an in-house SAXS instrument (SAXSess mc2, Anton Paar) equipped with a Kratky camera, a sealed x-ray tube source, and a two-dimensional Princeton Instruments PI·SCX:4300 CCD detector. For dark- and light-state measurements, the protein solutions were either kept in the dark or exposed to red light (1.8 mW cm$^{-2}$ at 630 nm) for 90 min before the measurements. The scattering patterns were measured with a 180 min exposure time (1,080 frames, 10 sec each) for several solute concentrations ranging from 0.8 to 13.0 mg ml$^{-1}$. Radiation damage was excluded on the basis of a comparison of individual frames of the 180 min exposures, where no changes were detected. A range of momentum transfer of 0.012 Å$^{-1}$ < s < 0.63 Å$^{-1}$ was covered [s = 4π sin(θ)/λ, where 2θ is the scattering angle and λ = 1.542 Å is the x-ray wavelength]. All SAXS data were analyzed with the

package ATSAS (version 2.5). The data were processed with the SAXS Quant software (version 3.9) and desmeared using the programs GNOM and GIFT (189, 190). The forward scattering I(0), the radius of gyration Rg, the maximum dimension Dmax, and the interatomic distance distribution functions [P(R)] were computed with the program GNOM (189). The masses of the solutes were evaluated by comparison of the forward scattering intensity with that of a human serum albumin reference solution (molecular mass, 69 kDa) and using Porod's law. To generate *ab initio* shape models, a total of 50 models were calculated using the program DAMMIF (191) and aligned and averaged using the program DAMCLUST. C2 symmetry was defined. The ab initio shape models were aligned with the crystal structure of *Is*PadC using the program SUPCOMB (192). The structures of *Is*PadC were modeled using the program CORAL (193). Input was the crystal structure of *Is*PadC determined here and SAXS data. The orientation of the DGC domains was kept flexible during the calculations, and no dimerization interface was restrained to account for dynamics. A total of 50 structures were calculated, and the best structures based on the fit to the experimental data were selected to prepare the figures.

## 4.8. Hydrogen-Deuterium exchange coupled to mass spectrometry

Deuterium labeling experiments were performed to address the effect of illumination on the conformational dynamics of various variants in order to compare them and observing the effect of specific mutation or truncations on the overall conformational dynamics. Briefly, 200 µM protein sample were equilibrated at 20 °C in the dark or under red light illumination (630 nm, 0.7 mW cm$^{-2}$) for 1 min prior to starting the deuterium labelling reaction by diluting the sample 20-fold in $D_2O$ containing 10 mM Hepes (pD 7.0), 150 mM NaCl, and 10 mM $MgCl_2$. Reactions were quenched with ice-cold 200 mM ammonium formic acid (pH 2.5) after 10 sec, 45 sec, 3 min, 15 min, and 60 min and injected into a cooled HPLC setup as described previously (132). To sum up, deuterated samples were digested on an immobilized pepsin column operated at 10 °C. Resulting peptides were desalted on a 2 cm C18 guard column, and separated during a 7 min acetonitrile gradient (15–50 %) in the presence of 0.6 % (v/v) formic acid on a reversed phase column (XR ODS 75 x 3 mm, 2.2 mM). Peptides were then infused into a maXis electrospray

ionization-ultra high resolution-time-of-flight mass spectrometer. Deuterium incorporation was analyzed and quantified using the Hexicon 2 software package ([http://hx2.mpimf-heidelberg.mpg.de](http://hx2.mpimf-heidelberg.mpg.de)) (194).

## 4.9.    Limited proteolysis

Trypsin digestion patterns of PadC homologs and variants were obtained by following the proteolytic degradation over 60 min in the dark or with constant red light illumination (0.7 mW cm$^{-2}$ at 630 nm) at 20 °C. A 1:100 (w/w) ratio of protease to enzyme was used for *Is*PadC and derived variants, 1:750 (w/w) ratio for *Ma*PadC, and 1:1000 (w/w) ratio for *Ts*PadC with PadC concentrations varying between 6 and 8 mM. Tryptic digests were performed in reaction buffer containing 10 mM tris-HCl (pH 8.0), 500 mM NaCl, and 10 mM MgCl$_2$. PadC (5 mg) was quenched per time point upon mixing with SDS sample buffer [4× stock: 10 % (w/v) glycerol, 0.6 % (w/v) tris-HCl (pH 6.8), 2 % (w/v) SDS, 0.02 % (w/v) bromophenol blue, 1.5 % (w/v) dithiothreitol] and heated at 95 °C for 5 min. All samples were loaded onto a 12 % SDS‑polyacrylamide gel for electrophoretic separation of proteolytic fragments.

# References

1.  Kühne W (1878) *On the Photochemistry of the Retina and Visual Purple* (Macmillan, New York).

2.  Wald G (1933) Vitamin A in the retina. *Nature* 132(3330):316–317.

3.  Sakmar TP, Menon ST, Marin EP, Awad ES (2002) Rhodopsin: Insights from Recent Structural Studies. *Annu Rev Biophys Biomol Struct* 31(1):443–484.

4.  van der Horst MA, Hellingwerf KJ (2004) Photoreceptor Proteins, "Star Actors of Modern Times": A Review of the Functional Dynamics in the Structure of Representative Members of Six Different Photoreceptor Families. *Acc Chem Res* 37(1):13–20.

5.  Davis SJ, Vener AV, Vierstra RD (1999) Bacteriophytochromes: Phytochrome-Like Photoreceptors from Nonphotosynthetic Eubacteria. *Science* 286(5449):2517.

6.  Losi A, Polverini E, Quest B, Gärtner W (2002) First evidence for phototropin-related blue-light receptors in prokaryotes. *Biophys J* 82(5):2627–2634.

7.  Foster K, et al. (1984) A rhodopsin is the functional photoreceptor for phototaxis in the unicellular eukaryote Chlamydomonas. *Nature*:756–9.

8.  Gomelsky M, Hoff WD (2011) Light helps bacteria make important lifestyle decisions. *Trends Microbiol* 19(9):441–448.

9.  Fraser DP, Hayes S, Franklin KA (2016) Photoreceptor crosstalk in shade avoidance. *Curr Opin Plant Biol* 33:1–7.

10. Cashmore AR (2003) Cryptochromes: enabling plants and animals to determine circadian time. *Cell* 114(5):537–543.

11. Heintzen C (2012) Plant and fungal photopigments. *Wiley Interdiscip Rev Membr Transp Signal* 1(4):411–432.

12. Jenkins GI (2014) The UV-B Photoreceptor UVR8: From Structure to Physiology. *Plant Cell* 26(1):21–37.

13. Rockwell NC, Su Y-S, Lagarias JC (2006) PHYTOCHROME STRUCTURE AND SIGNALING MECHANISMS. *Annu Rev Plant Biol* 57(1):837–858.

14. Mathes T, Kennis JTM (2016) Editorial: Optogenetic Tools in the Molecular Spotlight. *Front Mol Biosci* 3.

15. Ziegler T, Möglich A (2015) Photoreceptor engineering. *Front Mol Biosci* 2.

16. Ryu M-H, Moskvin OV, Siltberg-Liberles J, Gomelsky M (2010) Natural and Engineered Photoactivated Nucleotidyl Cyclases for Optogenetic Applications. *J Biol Chem* 285(53):41501–41508.

17. Pham VN, Kathare PK, Huq E (2018) Phytochromes and Phytochrome Interacting Factors. *Plant Physiol* 176(2):1025–1038.

18. Isomura A, Kageyama R (2017) Illuminating information transfer in signaling dynamics by optogenetics. *Curr Opin Cell Biol* 49:9–15.

19. Toettcher JE, Weiner OD, Lim WA (2013) Using Optogenetics to Interrogate the Dynamic Control of Signal Transmission by the Ras/Erk Module. *Cell* 155(6):1422–1434.

20. Opoku-Temeng C, Zhou J, Zheng Y, Su J, Sintim HO (2016) Cyclic dinucleotide (c-di-GMP, c-di-AMP, and cGAMP) signalings have come of age to be inhibited by small molecules. *Chem Commun* 52(60):9327–9342.

21. Montgomery BL, Lagarias JC (2002) Phytochrome ancestry: sensors of bilins and light. *Trends Plant Sci* 7(8):357–366.

22. Ho Y-SJ, Burden LM, Hurley JH Structure of the GAF domain, a ubiquitous signaling motif and a new class of cyclic GMP receptor. 12.

23. Wu S-H, Lagarias JC (2000) Defining the Bilin Lyase Domain: Lessons from the Extended Phytochrome Superfamily [†]. *Biochemistry (Mosc)* 39(44):13487–13495.

24. Lamparter T, et al. (2004) The Biliverdin Chromophore Binds Covalently to a Conserved Cysteine Residue in the N-Terminus of *Agrobacterium* Phytochrome Agp1 [†]. *Biochemistry (Mosc)* 43(12):3659–3669.

25. Park C-M, et al. (2000) A Second Photochromic Bacteriophytochrome from *Synechocystis* sp. PCC 6803: Spectral Analysis and Down-Regulation by Light [†]. *Biochemistry (Mosc)* 39(35):10840–10847.

26. Vierstra RD, Davis SJ (2000) Bacteriophytochromes: new tools for understanding phytochrome signal transduction. *Semin Cell Dev Biol* 11(6):511–521.

27. Lamparter T, Michael N, Mittmann F, Esteban B (2002) Phytochrome from Agrobacterium tumefaciens has unusual spectral properties and reveals an N-terminal chromophore attachment site. *Proc Natl Acad Sci* 99(18):11628–11633.

28. Bhoo S-H, Davis SJ, Walker J, Karniol B, Vierstra RD (2001) Bacteriophytochromes are photochromic histidine kinases using a biliverdin chromophore. 414:4.

29. Nagano S, et al. (2016) The Crystal Structures of the N-terminal Photosensory Core Module of *Agrobacterium* Phytochrome Agp1 as Parallel and Anti-parallel Dimers. *J Biol Chem* 291(39):20674–20691.

30. Essen L-O, Mailliet J, Hughes J (2008) The structure of a complete phytochrome sensory module in the Pr ground state. *Proc Natl Acad Sci* 105(38):14709–14714.

31. Wagner JR, Brunzelle JS, Forest KT, Vierstra RD (2005) A light-sensing knot revealed by the structure of the chromophore-binding domain of phytochrome. *Nature* 438(7066):325–331.

32. Gourinchas G, et al. (2017) Long-range allosteric signaling in red light–regulated diguanylyl cyclases. *Sci Adv* 3(3):e1602498.

---

33. Wagner JR, Zhang J, Brunzelle JS, Vierstra RD, Forest KT (2007) High Resolution Structure of *Deinococcus* Bacteriophytochrome Yields New Insights into Phytochrome Architecture and Evolution. *J Biol Chem* 282(16):12298–12309.

34. Burgie ES, et al. (2014) Crystallographic and Electron Microscopic Analyses of a Bacterial Phytochrome Reveal Local and Global Rearrangements during Photoconversion. *J Biol Chem* 289(35):24573–24587.

35. Lamparter T (2004) Evolution of cyanobacterial and plant phytochromes. *FEBS Lett* 573(1–3):1–5.

36. von Stetten D, et al. (2007) Highly Conserved Residues Asp-197 and His-250 in Agp1 Phytochrome Control the Proton Affinity of the Chromophore and Pfr Formation. *J Biol Chem* 282(3):2116–2123.

37. Rudiger W, Thummler F, Cmiel E, Schneider S (1983) Chromophore structure of the physiologically active form (Pfr) of phytochrome. *Proc Natl Acad Sci* 80(20):6244–6248.

38. Yang X, Ren Z, Kuk J, Moffat K (2011) Temperature-scan cryocrystallography reveals reaction intermediates in bacteriophytochrome. *Nature* 479(7373):428–432.

39. Takala H, et al. (2014) Signal amplification and transduction in phytochrome photosensors. *Nature* 509(7499):245–248.

40. Hahn J, et al. (2006) Probing protein-chromophore interactions in Cph1 phytochrome by mutagenesis. *FEBS J* 273(7):1415–1429.

41. Auldridge ME, Satyshur KA, Anstrom DM, Forest KT (2012) Structure-guided Engineering Enhances a Phytochrome-based Infrared Fluorescent Protein. *J Biol Chem* 287(10):7000–7009.

42. Andel FI, Lagarias JC, Mathies RA Resonance Raman Analysis of Chromophore Structure in the Lumi-R Photoproduct of Phytochrome†. 12.

43. Borucki B, et al. (2005) Light-induced Proton Release of Phytochrome Is Coupled to the Transient Deprotonation of the Tetrapyrrole Chromophore. *J Biol Chem* 280(40):34358–34364.

44. Ulijasz AT, et al. (2010) Structural basis for the photoconversion of a phytochrome to the activated Pfr form. *Nature* 463(7278):250–254.

45. Butler WL, Norris KH, Siegelman HW, Hendricks SB (1959) Detection, assay, and preliminary purification of the pigment controlling photoresponsive development of plants. *Proc Natl Acad Sci U S A* 45(12):1703–1708.

46. Rausenberger J, et al. (2010) An Integrative Model for Phytochrome B Mediated Photomorphogenesis: From Protein Dynamics to Physiology. *PLoS ONE* 5(5):e10721.

47. Franklin KA, Quail PH (2010) Phytochrome functions in Arabidopsis development. *J Exp Bot* 61(1):11–24.

48. Purcell EB, Crosson S (2008) Photoregulation in prokaryotes. *Curr Opin Microbiol* 11(2):168–178.

49. van der Horst MA, Key J, Hellingwerf KJ (2007) Photosensing in chemotrophic, non-phototrophic bacteria: let there be light sensing too. *Trends Microbiol* 15(12):554–562.

50. Memmi S, et al. (2008) Photoactive Yellow Protein from the Halophilic Bacterium *Salinibacter ruber*. *Biochemistry (Mosc)* 47(7):2014–2024.

51. Gomelsky M, Klug G (2002) BLUF: a novel FAD-binding domain involved in sensory transduction in microorganisms. *Trends Biochem Sci* 27(10):497–500.

52. Sinha RP, Häder D-P (2002) UV-induced DNA damage and repair: a review. *Photochem Photobiol Sci* 1(4):225–236.

53. Khorchid A, Ikura M (2006) Bacterial histidine kinase as signal sensor and transducer. *Int J Biochem Cell Biol* 38(3):307–312.

54. Bijlsma JJE, Groisman EA (2003) Making informed decisions: regulatory interactions between two-component systems. *Trends Microbiol* 11(8):359–366.

55. Giraud E, et al. (2002) Bacteriophytochrome controls photosystem synthesis in anoxygenic bacteria. *Nature* 417(6885):202–205.

56. Otero LH, et al. (2016) Structure of the Full-Length Bacteriophytochrome from the Plant Pathogen Xanthomonas campestris Provides Clues to its Long-Range Signaling Mechanism. *J Mol Biol* 428(19):3702–3720.

57. Bonomi HR, et al. (2016) *Xanthomonas campestris* attenuates virulence by sensing light through a bacteriophytochrome photoreceptor. *EMBO Rep* 17(11):1565–1577.

58. Bhaya D (2004) Light matters: phototaxis and signal transduction in unicellular cyanobacteria: Phototaxis and signal transduction. *Mol Microbiol* 53(3):745–754.

59. Narikawa R, Fukushima Y, Ishizuka T, Itoh S, Ikeuchi M (2008) A Novel Photoactive GAF Domain of Cyanobacteriochrome AnPixJ That Shows Reversible Green/Red Photoconversion. *J Mol Biol* 380(5):844–855.

60. Yoshihara S, Katayama M, Geng X, Ikeuchi M (2004) Cyanobacterial Phytochrome-like PixJ1 Holoprotein Shows Novel Reversible Photoconversion Between Blue- and Green-absorbing Forms. *Plant Cell Physiol* 45(12):1729–1737.

61. Ikeuchi M, Ishizuka T (2008) Cyanobacteriochromes: a new superfamily of tetrapyrrole-binding photoreceptors in cyanobacteria. *Photochem Photobiol Sci* 7(10):1159.

62. Ishizuka T, et al. (2006) Characterization of Cyanobacteriochrome TePixJ from a Thermophilic Cyanobacterium Thermosynechococcus elongatus Strain BP-1. *Plant Cell Physiol* 47(9):1251–1261.

63. Ishizuka T, Narikawa R, Kohchi T, Katayama M, Ikeuchi M (2007) Cyanobacteriochrome TePixJ of Thermosynechococcus elongatus Harbors Phycoviolobilin as a Chromophore. *Plant Cell Physiol* 48(9):1385–1390.

64. Ulijasz AT, et al. (2009) Cyanochromes Are Blue/Green Light Photoreversible Photoreceptors Defined by a Stable Double Cysteine Linkage to a Phycoviolobilin-type Chromophore. *J Biol Chem* 284(43):29757–29772.

65. Bhaya D, Bianco NR, Bryant D, Grossman A (2000) Type IV pilus biogenesis and motility in the cyanobacterium *Synechocystis* sp. PCC6803. *Mol Microbiol* 37(4):941–951.

66. Tarutina M, Ryjenkov DA, Gomelsky M (2006) An Unorthodox Bacteriophytochrome from *Rhodobacter sphaeroides* Involved in Turnover of the Second Messenger c-di-GMP. *J Biol Chem* 281(46):34751–34758.

67. Fiedler B, Börner T, Wilde A (2005) Phototaxis in the Cyanobacterium Synechocystis sp. PCC 6803: Role of Different Photoreceptors. *Photochem Photobiol* 81(6):1481.

68. Wilde A, Fiedler B, Börner T (2002) The cyanobacterial phytochrome Cph2 inhibits phototaxis towards blue light: Cyanobacterial phytochrome Cph2. *Mol Microbiol* 44(4):981–988.

69. Kyndt JA, Meyer TE, Cusanovich MA (2004) Photoactive yellow protein, bacteriophytochrome, and sensory rhodopsin in purple phototrophic bacteria. *Photochem Photobiol Sci* 3(6):519.

70. Römling U, Gomelsky M, Galperin MY (2005) C-di-GMP: the dawning of a novel bacterial signalling system: C-di-GMP signalling in bacteria. *Mol Microbiol* 57(3):629–639.

71. Römling U, Galperin MY, Gomelsky M (2013) Cyclic di-GMP: the First 25 Years of a Universal Bacterial Second Messenger. *Microbiol Mol Biol Rev* 77(1):1–52.

72. Lopez D, Vlamakis H, Kolter R (2010) Biofilms. *Cold Spring Harb Perspect Biol* 2(7):a000398–a000398.

73. Hall-Stoodley L, Costerton JW, Stoodley P (2004) Bacterial biofilms: from the Natural environment to infectious diseases. *Nat Rev Microbiol* 2:95.

74. Seifert R, Beste KY (2012) Allosteric Regulation of Nucleotidyl Cyclases: An Emerging Pharmacological Target. *Sci Signal* 5(240):pe37.

75. Krasteva PV, Giglio KM, Sondermann H (2012) Sensing the messenger: The diverse ways that bacteria signal through c-di-GMP. *Protein Sci* 21(7):929–948.

76. Etzl S, Lindner R, Nelson MD, Winkler A (2018) Structure-guided design and functional characterization of an artificial red light-regulated guanylate/adenylate cyclase for optogenetic applications. *J Biol Chem*:jbc.RA118.003069.

77. Gourinchas G, Heintz U, Winkler A (2018) Asymmetric activation mechanism of a homodimeric red light-regulated photoreceptor. *eLife* 7:e34815.

78. Nagano S (2016) From photon to signal in phytochromes: similarities and differences between prokaryotic and plant phytochromes. *J Plant Res* 129(2):123–135.

79. Yang X, et al. (2015) Light Signaling Mechanism of Two Tandem Bacteriophytochromes. *Structure* 23(7):1179–1189.

80. Yang X, Kuk J, Moffat K (2009) Conformational differences between the Pfr and Pr states in *Pseudomonas aeruginosa* bacteriophytochrome. *Proc Natl Acad Sci* 106(37):15639–15644.

81. Burgie ES, Bussell AN, Walker JM, Dubiel K, Vierstra RD (2014) Crystal structure of the photosensing module from a red/far-red light-absorbing plant phytochrome. *Proc Natl Acad Sci* 111(28):10179–10184.

82. Yang X, Kuk J, Moffat K (2008) Crystal structure of Pseudomonas aeruginosa bacteriophytochrome: Photoconversion and signal transduction. *Proc Natl Acad Sci* 105(38):14715–14720.

83. Anders K, Daminelli-Widany G, Mroginski MA, von Stetten D, Essen L-O (2013) Structure of the Cyanobacterial Phytochrome 2 Photosensor Implies a Tryptophan Switch for Phytochrome Signaling. *J Biol Chem* 288(50):35714–35725.

84. Bellini D, Papiz MZ (2012) Structure of a Bacteriophytochrome and Light-Stimulated Protomer Swapping with a Gene Repressor. *Structure* 20(8):1436–1446.

85. Burgie ES, Zhang J, Vierstra RD (2016) Crystal Structure of *Deinococcus* Phytochrome in the Photoactivated State Reveals a Cascade of Structural Rearrangements during Photoconversion. *Structure* 24(3):448–457.

86. Song C, et al. (2018) 3D Structures of Plant Phytochrome A as Pr and Pfr From Solid-State NMR: Implications for Molecular Function. *Front Plant Sci* 9.

87. Li H, Zhang J, Vierstra RD, Li H (2010) Quaternary organization of a phytochrome dimer as revealed by cryoelectron microscopy. *Proc Natl Acad Sci* 107(24):10872–10877.

88. Björling A, et al. (2016) Structural photoactivation of a full-length bacterial phytochrome. *Sci Adv* 2(8):e1600920–e1600920.

89. Inomata K, et al. (2005) Sterically Locked Synthetic Bilin Derivatives and Phytochrome Agp1 from *Agrobacterium tumefaciens* Form Photoinsensitive Pr- and Pfr-like Adducts. *J Biol Chem* 280(26):24491–24497.

90. Inomata K, et al. (2006) Assembly of Synthetic Locked Chromophores with *Agrobacterium* Phytochromes Agp1 and Agp2. *J Biol Chem* 281(38):28162–28173.

91. Rohmer T, et al. (2008) Light-induced chromophore activity and signal transduction in phytochromes observed by 13C and 15N magic-angle spinning NMR. *Proc Natl Acad Sci* 105(40):15229–15234.

92. Song C, et al. (2011) Two ground state isoforms and a chromophore D-ring photoflip triggering extensive intramolecular changes in a canonical phytochrome. *Proc Natl Acad Sci* 108(10):3842–3847.

93. Ulijasz AT, et al. (2008) Characterization of Two Thermostable Cyanobacterial Phytochromes Reveals Global Movements in the Chromophore-binding Domain during Photoconversion. *J Biol Chem* 283(30):21251–21266.

94. Tischer D, Weiner OD (2014) Illuminating cell signalling with optogenetic tools. *Nat Rev Mol Cell Biol* 15(8):551–558.

95. Pathak GP, Vrana JD, Tucker CL (2013) Optogenetic control of cell function using engineered photoreceptors. *Biol Cell* 105(2):59–72.

96. Chernov KG, Redchuk TA, Omelina ES, Verkhusha VV (2017) Near-Infrared Fluorescent Proteins, Biosensors, and Optogenetic Tools Engineered from Phytochromes. *Chem Rev* 117(9):6423–6446.

97. Wagner JR, et al. (2008) Mutational Analysis of *Deinococcus radiodurans* Bacteriophytochrome Reveals Key Amino Acids Necessary for the Photochromicity and Proton Exchange Cycle of Phytochromes. *J Biol Chem* 283(18):12212–12226.

98. Fischer AJ, Lagarias JC (2004) Harnessing phytochrome's glowing potential. *Proc Natl Acad Sci* 101(50):17334–17339.

99. Shu X, et al. (2009) Mammalian Expression of Infrared Fluorescent Proteins Engineered from a Bacterial Phytochrome. *Science* 324(5928):804–807.

100. Tran MTN, et al. (2014) In Vivo image Analysis Using iRFP Transgenic Mice. *Exp Anim* 63(3):311–319.

101. Filonov GS, et al. (2011) Bright and stable near-infrared fluorescent protein for in vivo imaging. *Nat Biotechnol* 29(8):757–761.

102. Wehr MC, Rossner MJ (2016) Split protein biosensor assays in molecular pharmacological studies. *Drug Discov Today* 21(3):415–429.

103. Shcherbakova DM, et al. (2016) Bright monomeric near-infrared fluorescent proteins as tags and biosensors for multiscale imaging. *Nat Commun* 7:12405.

104. Filonov GS, Verkhusha VV (2013) A Near-Infrared BiFC Reporter for In Vivo Imaging of Protein-Protein Interactions. *Chem Biol* 20(8):1078–1086.

105. Chen M, et al. (2015) Novel near-infrared BiFC systems from a bacterial phytochrome for imaging protein interactions and drug evaluation under physiological conditions. *Biomaterials* 48:97–107.

106. Pandey N, Nobles CL, Zechiedrich L, Maresso AW, Silberg JJ (2015) Combining Random Gene Fission and Rational Gene Fusion To Discover Near-Infrared Fluorescent Protein Fragments That Report on Protein–Protein Interactions. *ACS Synth Biol* 4(5):615–624.

107. Rausenberger J, et al. (2011) Photoconversion and Nuclear Trafficking Cycles Determine Phytochrome A's Response Profile to Far-Red Light. *Cell* 146(5):813–825.

108. Shimizu-Sato S, Huq E, Tepperman JM, Quail PH (2002) A light-switchable gene promoter system. *Nat Biotechnol* 20(10):1041–1044.

109. Ryu M-H, et al. (2014) Engineering adenylate cyclases regulated by near-infrared window light. *Proc Natl Acad Sci* 111(28):10167–10172.

110. Ryu M-H, Gomelsky M (2014) Near-infrared Light Responsive Synthetic c-di-GMP Module for Optogenetic Applications. *ACS Synth Biol* 3(11):802–810.

111. Gasser C, et al. (2014) Engineering of a red-light-activated human cAMP/cGMP-specific phosphodiesterase. *Proc Natl Acad Sci* 111(24):8803–8808.

112. Engelhard C, Diensthuber RP, Möglich A, Bittl R (2017) Blue-light reception through quaternary transitions. *Sci Rep* 7(1).

113. Scapin G (2013) Molecular replacement then and now. *Acta Crystallogr D Biol Crystallogr* 69(11):2266–2275.

114. Terwilliger TC, et al. (2016) Can I solve my structure by SAD phasing? Anomalous signal in SAD phasing. *Acta Crystallogr Sect Struct Biol* 72(3):346–358.

115. Taylor GL (2010) Introduction to phasing. *Acta Crystallogr D Biol Crystallogr* 66(4):325–338.

116. Schirmer T (2016) C-di-GMP Synthesis: Structural Aspects of Evolution, Catalysis and Regulation. *J Mol Biol* 428(19):3683–3701.

117. Römling U, Amikam D (2006) Cyclic di-GMP as a second messenger. *Curr Opin Microbiol* 9(2):218–228.

118. Tamayo R, Pratt JT, Camilli A (2007) Roles of Cyclic Diguanylate in the Regulation of Bacterial Pathogenesis. *Annu Rev Microbiol* 61(1):131–148.

119. Schirmer T, Jenal U (2009) Structural and mechanistic determinants of c-di-GMP signalling. *Nat Rev Microbiol* 7:724.

120. Lupas AN, Bassler J (2017) Coiled Coils – A Model System for the 21st Century. *Trends Biochem Sci* 42(2):130–140.

121. Indelicato G, et al. (2016) Principles Governing the Self-Assembly of Coiled-Coil Protein Nanoparticles. *Biophys J* 110(3):646–660.

122. Truebestein L, Leonard TA (2016) Coiled-coils: The long and short of it. *BioEssays* 38(9):903–916.

123. Lupas A, Van Dyke M, Stock J (1991) Predicting coiled coils from protein sequences. *Science* 252(5009):1162–1164.

124. Guruprasad K, Reddy BVB, Pandit MW (1990) Correlation between stability of a protein and its dipeptide composition: a novel approach for predicting *in vivo* stability of a protein from its primary sequence. *Protein Eng Des Sel* 4(2):155–161.

125. Antoniani D, Bocci P, Maciąg A, Raffaelli N, Landini P (2010) Monitoring of diguanylate cyclase activity and of cyclic-di-GMP biosynthesis by whole-cell assays suitable for high-throughput screening of biofilm inhibitors. *Appl Microbiol Biotechnol* 85(4):1095–1104.

126. Krissinel E, Henrick K (2007) Inference of Macromolecular Assemblies from Crystalline State. *J Mol Biol* 372(3):774–797.

127. De N, Navarro MVAS, Raghavan RV, Sondermann H (2009) Determinants for the Activation and Autoinhibition of the Diguanylate Cyclase Response Regulator WspR. *J Mol Biol* 393(3):619–633.

128. Bunkóczi G, Read RJ (2011) Improvement of molecular-replacement models with *Sculptor*. *Acta Crystallogr D Biol Crystallogr* 67(4):303–312.

129. Vierstra RD, Zhang J (2011) Phytochrome signaling: solving the Gordian knot with microbial relatives. *Trends Plant Sci* 16(8):417–426.

130. Zähringer F, Lacanna E, Jenal U, Schirmer T, Boehm A (2013) Structure and Signaling Mechanism of a Zinc-Sensory Diguanylate Cyclase. *Structure* 21(7):1149–1157.

131. Konermann L, Pan J, Liu Y-H (2011) Hydrogen exchange mass spectrometry for studying protein structure and dynamics. *Chem Soc Rev* 40(3):1224–1234.

132. Lindner R, Heintz U, Winkler A (2015) Applications of hydrogen deuterium exchange (HDX) for the characterization of conformational dynamics in light-activated photoreceptors. *Front Mol Biosci* 2.

133. Fletcher JM, et al. (2017) N@ *a* and N@ *d* : Oligomer and Partner Specification by Asparagine in Coiled-Coil Interfaces. *ACS Chem Biol* 12(2):528–538.

134. Kornev AP, Taylor SS (2015) Dynamics-Driven Allostery in Protein Kinases. *Trends Biochem Sci* 40(11):628–647.

135. Motlagh HN, Wrabl JO, Li J, Hilser VJ (2014) The ensemble nature of allostery. *Nature* 508(7496):331–339.

136. Evans K, Grossmann JG, Fordham-Skelton AP, Papiz MZ (2006) Small-Angle X-ray Scattering Reveals the Solution Structure of a Bacteriophytochrome in the Catalytically Active Pr State. *J Mol Biol* 364(4):655–666.

137. Burgie ES, Vierstra RD (2014) Phytochromes: An Atomic Perspective on Photoactivation and Signaling. *Plant Cell Online* 26(12):4568–4583.

138. Liu Y, et al. (2013) Bicaudal-D uses a parallel, homodimeric coiled coil with heterotypic registry to coordinate recruitment of cargos to dynein. *Genes Dev* 27(11):1233–1246.

139. Croasdale R, et al. (2011) An Undecided Coiled Coil: the leucine zipper of nek2 kinase exhibits atypical conformational exchange dynamics. *J Biol Chem* 286(31):27537–27547.

140. Hulko M, et al. (2006) The HAMP Domain Structure Implies Helix Rotation in Transmembrane Signaling. *Cell* 126(5):929–940.

141. Loughlin PC, et al. (2016) Spectral properties of bacteriophytochrome AM1_5894 in the chlorophyll d-containing *cyanobacterium Acaryochloris marina*. *Sci Rep* 6(1).

142. Yang X, Stojković EA, Kuk J, Moffat K (2007) Crystal structure of the chromophore binding domain of an unusual bacteriophytochrome, RpBphP3, reveals residues that modulate photoconversion. *Proc Natl Acad Sci* 104(30):12571–12576.

143. Thümmler F, Brandlmeier T, Rüdiger W (1981) Preparation and Properties of Chromopeptides from the Pfr Form of Phytochrome. *Z Für Naturforschung C* 36 c:440–449.

144. Hildebrandt P, Buhrke D, Michael N, Kuhlmann U (2018) The photoconversion of phytochrome includes an unproductive shunt reaction pathway. *ChemPhysChem*.

145. Dasgupta J, Frontiera RR, Taylor KC, Lagarias JC, Mathies RA (2009) Ultrafast excited-state isomerization in phytochrome revealed by femtosecond stimulated Raman spectroscopy. *Proc Natl Acad Sci* 106(6):1784–1789.

146. Eilfeld P Absorption Spectra of Phytochrome Intermediates. 6.

147. Rudiger W, Thummler F (1984) Low temperature spectroscopy of phytochrome: Pr, Pfr and intermediates. *Physiol Plant* 60(3):383–388.

148. D'Arcy A, Bergfors T, Cowan-Jacob SW, Marsh M (2014) Microseed matrix screening for optimization in protein crystallization: what have we learned? *Acta Crystallogr Sect F Struct Biol Commun* 70(9):1117–1126.

149. Liebschner D, et al. (2017) Polder maps: improving OMIT maps by excluding bulk solvent. *Acta Crystallogr Sect Struct Biol* 73(2):148–157.

150. Römling U, Liang Z-X, Dow JM (2017) Progress in Understanding the Molecular Basis Underlying Functional Diversification of Cyclic Dinucleotide Turnover Proteins. *J Bacteriol* 199(5):e00790-16.

151. Chan C, et al. (2004) Structural basis of activity and allosteric control of diguanylate cyclase. *Proc Natl Acad Sci* 101(49):17084–17089.

152. Wassmann P, et al. (2007) Structure of BeF3−-Modified Response Regulator PleD: Implications for Diguanylate Cyclase Activation, Catalysis, and Feedback Inhibition. *Structure* 15(8):915–927.

153. De N, et al. (2008) Phosphorylation-Independent Regulation of the Diguanylate Cyclase WspR. *PLoS Biol* 6(3):e67.

154. Navarro MVAS, De N, Bae N, Wang Q, Sondermann H (2009) Structural Analysis of the GGDEF-EAL Domain-Containing c-di-GMP Receptor FimX. *Structure* 17(8):1104–1116.

155. Klose C, et al. (2015) Systematic analysis of how phytochrome B dimerization determines its specificity. *Nat Plants* 1(7):15090.

156. Takala H, Björling A, Linna M, Westenhoff S, Ihalainen JA (2015) Light-induced Changes in the Dimerization Interface of Bacteriophytochromes. *J Biol Chem* 290(26):16383–16392.

157. Stojković EA, et al. (2014) FTIR Spectroscopy Revealing Light-Dependent Refolding of the Conserved Tongue Region of Bacteriophytochrome. *J Phys Chem Lett* 5(15):2512–2515.

158. Anders K, Gutt A, Gärtner W, Essen L-O (2014) Phototransformation of the Red Light Sensor Cyanobacterial Phytochrome 2 from *Synechocystis* Species Depends on Its Tongue Motifs. *J Biol Chem* 289(37):25590–25600.

159. Neiditch MB, et al. (2006) Ligand-Induced Asymmetry in Histidine Sensor Kinase Complex Regulates Quorum Sensing. *Cell* 126(6):1095–1108.

160. Narayanan A, Kumar S, Evrard AN, Paul LN, Yernool DA (2014) An asymmetric heterodomain interface stabilizes a response regulator–DNA complex. *Nat Commun* 5.

161. Chervitz SA, Falke JJ (1996) Molecular mechanism of transmembrane signaling by the aspartate receptor: a model. *Proc Natl Acad Sci* 93(6):2545–2550.

162. Stierl M, Penzkofer A, Kennis JTM, Hegemann P, Mathes T (2014) Key Residues for the Light Regulation of the Blue Light-Activated Adenylyl Cyclase from *Beggiatoa* sp. *Biochemistry (Mosc)* 53(31):5121–5130.

163. Ryu M-H, Fomicheva A, Moskvin OV, Gomelsky M (2017) Optogenetic Module for Dichromatic Control of c-di-GMP Signaling. *J Bacteriol* 199(18):e00014-17.

164. Pedetta A, Massazza DA, Herrera Seitz MK, Studdert CA (2017) Mutational Replacements at the "Glycine Hinge" of the *Escherichia coli* Chemoreceptor Tsr Support a Signaling Role for the C-Helix Residue. *Biochemistry (Mosc)* 56(29):3850–3862.

165. Zhang Y (2008) I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics* 9(1):40.

166. Yang J, et al. (2015) The I-TASSER Suite: protein structure and function prediction. *Nat Methods* 12(1):7–8.

167. Burgie ES, et al. (2017) Photosensing and Thermosensing by Phytochrome B Require Both Proximal and Distal Allosteric Features within the Dimeric Photoreceptor. *Sci Rep* 7(1).

168. Zienicke B, et al. (2013) Unusual Spectral Properties of Bacteriophytochrome Agp2 Result from a Deprotonation of the Chromophore in the Red-absorbing Form Pr. *J Biol Chem* 288(44):31738–31751.

169. Cherry JR, Hondred D, Walker JM, Vierstra RD (1992) Phytochrome requires the 6-kDa N-terminal domain for full biological activity. *Proc Natl Acad Sci* 89(11):5039–5043.

170. Rockwell NC, Ohlendorf R, Moglich A (2013) Cyanobacteriochromes in full color and three dimensions. *Proc Natl Acad Sci* 110(3):806–807.

171. Linder JU (2006) Class III adenylyl cyclases: molecular mechanisms of catalysis and regulation. *Cell Mol Life Sci* 63(15):1736–1751.

172. Murgida DH, et al. (2007) The Chromophore Structures of the Pr States in Plant and Bacterial Phytochromes. *Biophys J* 93(7):2410–2417.

173. Salewski J, et al. (2013) Structure of the Biliverdin Cofactor in the Pfr State of Bathy and Prototypical Phytochromes. *J Biol Chem* 288(23):16800–16814.

174. Abriata LA, Albanesi D, Dal Peraro M, de Mendoza D (2017) Signal Sensing and Transduction by Histidine Kinases as Unveiled through Studies on a Temperature Sensor. *Acc Chem Res* 50(6):1359–1366.

175. Sharrock RA, Clack T (2004) Heterodimerization of type II phytochromes in *Arabidopsis*. *Proc Natl Acad Sci U S A* 101(31):11500–11505.

176. Clack T, et al. (2009) Obligate Heterodimerization of *Arabidopsis* Phytochromes C and E and Interaction with the PIF3 Basic Helix-Loop-Helix Transcription Factor. *PLANT CELL ONLINE* 21(3):786–799.

177. Xie X, Kagawa T, Takano M (2014) The Phytochrome B/Phytochrome C Heterodimer Is Necessary for Phytochrome C-Mediated Responses in Rice Seedlings. *PLoS ONE* 9(5):e97264.

178. Kim TH, et al. (2017) The role of dimer asymmetry and protomer dynamics in enzyme catalysis. *Science* 355(6322):eaag2355.

179. Rockwell NC, et al. (2014) Eukaryotic algal phytochromes span the visible spectrum. *Proc Natl Acad Sci* 111(10):3871–3876.

180. Liu H, Naismith JH (2008) An efficient one-step site-directed deletion, insertion, single and multiple-site plasmid mutagenesis protocol. *BMC Biotechnol* 8(1):91.

181. Walden H (2010) Selenium incorporation using recombinant techniques. *Acta Crystallogr D Biol Crystallogr* 66(4):352–357.

182. Rao F, Yang Y, Qi Y, Liang Z-X (2008) Catalytic Mechanism of Cyclic Di-GMP-Specific Phosphodiesterase: a Study of the EAL Domain-Containing RocR from Pseudomonas aeruginosa. *J Bacteriol* 190(10):3622–3631.

183. Winkler A, et al. (2013) A ternary AppA-PpsR-DNA complex mediates light regulation of photosynthesis-related gene expression. *Nat Struct Mol Biol* 20(7):859–867.

184. Enomoto G, et al. (2014) Cyanobacteriochrome SesA Is a Diguanylate Cyclase That Induces Cell Aggregation in *Thermosynechococcus*. *J Biol Chem* 289(36):24801–24809.

185. Kabsch W (2010) *XDS*. *Acta Crystallogr D Biol Crystallogr* 66(2):125–132.

186. Adams PD, et al. (2010) *PHENIX*: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* 66(2):213–221.

187. McCoy AJ, et al. (2007) *Phaser* crystallographic software. *J Appl Crystallogr* 40(4):658–674.

188. von Stetten D, et al. (2015) *In crystallo* optical spectroscopy ( *ic* OS) as a complementary tool on the macromolecular crystallography beamlines of the ESRF. *Acta Crystallogr D Biol Crystallogr* 71(1):15–26.

189. Svergun DI (1992) Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *J Appl Crystallogr* 25(4):495–503.

190. Bergmann A, Fritz G, Glatter O (2000) Solving the generalized indirect Fourier transformation (GIFT) by Boltzmann simplex simulated annealing (BSSA). *J Appl Crystallogr* 33(5):1212–1216.

191. Franke D, Svergun DI (2009) *DAMMIF*, a program for rapid *ab-initio* shape determination in small-angle scattering. *J Appl Crystallogr* 42(2):342–346.

192. Kozin MB, Svergun DI (2001) Automated matching of high- and low-resolution structural models. *J Appl Crystallogr* 34(1):33–41.

193. Petoukhov MV, et al. (2012) New developments in the *ATSAS* program package for small-angle scattering data analysis. *J Appl Crystallogr* 45(2):342–350.

194. Lindner R, et al. (2014) Hexicon 2: Automated Processing of Hydrogen-Deuterium Exchange Mass Spectrometry Data with Improved Deuteration Distribution Estimation. *J Am Soc Mass Spectrom* 25(6):1018–1028.

195. Takala H, Lehtivuori H, Hammarén H, Hytönen VP, Ihalainen JA (2014) Connection between Absorption Properties and Conformational Changes in *Deinococcus radiodurans* Phytochrome. *Biochemistry (Mosc)* 53(45):7076–7085.

196. Svergun D, Barberato C, Koch MHJ (1995) *CRYSOL* – a Program to Evaluate X-ray Solution Scattering of Biological Macromolecules from Atomic Coordinates. *J Appl Crystallogr* 28(6):768–773.

197. Buchan DWA, Minneci F, Nugent TCO, Bryson K, Jones DT (2013) Scalable web services for the PSIPRED Protein Analysis Workbench. *Nucleic Acids Res* 41(W1):W349–W357.

198.  Jones DT Protein Secondary Structure Prediction Based on Position-specific Scoring Matrices. 8.

# Appendix

**Appendix Table 1. Overview of thermal dark state recoveries for the PadC constructs used in this study.** * For comparison of individual variants all measurements were done at a protein concentration of 2 µM. Since we observed suboptimal fits of dark-state recoveries for the variants that feature extremely slow recovery contributions using the sum of two exponentials, we used the sum of three exponentials to fit the recovery of these variants. For variants featuring excessively long recoveries the fit included a fixed known endpoint (y0), which was obtained from the corresponding dark state spectra. After red light illumination of 1 min, changes in absorption at wavelength of the Pr-state Q-band maximum absorption were followed, with automatic sampling every 5 s and an integration time of 0.01 s. The contribution of each phase to the thermal recovery is represented as relative amplitude. The SE of the estimate from the nonlinear curve fit corresponding to $y = A1*exp(-x/\tau1) + A2*exp(-x/\tau2) + y0$ for a sum of two exponentials, and to $y = A1*exp(-x/\tau1) + A2*exp(-x/\tau2) + A3*exp(-x/\tau3) + y0$ for a sum of three exponentials, were used as error indicators. [#1] Recovery time constants measured as described above for *Is*PadC and *Ts*PadC at 10 µM. Due to a dependence of the recovery kinetics on protein concentration, *Is*PadC at higher protein concentrations required a fit based on the sum of three exponentials to fit the dark-state recovery. [#2] After red light illumination, the 10 µM samples were kept in the dark and aliquots were removed at different time points to follow the changes in the acid denatured spectra of biliverdin linked peptides by diluting the sample 1:10 in methanol/TCA 0.1%. $ For the monomerized variant of *Is*PadC PSM, recovery rates are indicated for three different NaCl concentrations; $1 for 2 M NaCl; $2 for 500 mM NaCl; and $3 for 150 mM NaCl. § Indicates samples for which the measurements were recorded after illumination with a weaker red LED intensity of 0.7 mW cm$^{-2}$ and fitted over a restrained period of time. Therefore the values reported for these variants need to be cautiously compared with the other measurements.

| | Dark state recovery recorded at the maximum Pr Q-band absorption wavelength | | | | | |
|---|---|---|---|---|---|---|
| **Constructs** | $\tau_1$ (sec) | **Relative A1 (%)** | $\tau_2$ (sec) | **Relative A2 (%)** | $\tau_3$ (sec) | **Relative A3 (%)** |
| *Is*PadC | $18.5 \pm 0.4$ | $34 \pm 0.5$ | $100.0 \pm 0.7$ | $66 \pm 0.5$ | - | - |
| *Ts*PadC | $33.4 \pm 0.1$ | $27.4 \pm 0.4$ | $1,603 \pm 18$ | $7.7 \pm 0.2$ | $63,000 \pm 500$ | $64.9 \pm 0.3$ |
| § *Ma*PadC | $24.9 \pm 0.4$ | $50.0 \pm 0.5$ | $458 \pm 56$ | $50 \pm 3$ | - | - |
| § *Is*PadC$^{\Delta514-516}$ | $13.5 \pm 0.2$ | $40.0 \pm 0.3$ | $308 \pm 12$ | $60 \pm 1$ | - | - |
| § *Is*PadC$^{\Delta514-517}$ | $13.1 \pm 0.7$ | $11.5 \pm 0.3$ | $115 \pm 1$ | $89 \pm 2$ | - | - |
| § *Is*PadC$^{\Delta514-520}$ | $11.9 \pm 0.3$ | $86 \pm 1$ | $393 \pm 9$ | $14.3 \pm 0.1$ | - | - |
| *Is*PadC$^{Reg1}$ | $18.2 \pm 0.3$ | $37 \pm 0.5$ | $141.5 \pm 0.8$ | $48 \pm 0.4$ | $6,930 \pm 50$ | $15 \pm 0.5$ |
| *Is*PadC$^{Reg2}$ | $28.5 \pm 0.3$ | $37 \pm 0.9$ | $3,480 \pm 4$ | $63 \pm 0.8$ | - | - |

| | | | | | | |
|---|---|---|---|---|---|---|
| *Is*PadC[Reg2.a] | 34.9 ± 0.2 | 36 ± 0.8 | 1,300 ± 20 | 3 ± 0.2 | 270,000 ± 1000 | 61 ± 0.7 |
| *TsIs*PadC[1] | 26.5 ± 0.4 | 28.7 ± 0.4 | 96.4 ± 7.3 | 3.3 ± 0.4 | 4,200 ± 18 | 68 ± 0.2 |
| *TsIs*PadC[2] | 576 ± 7 | 17.4 ± 0.5 | 1,900 ± 4 | 82.6 ± 0.5 | - | - |
| *IsTs*PadC[1] | 31.2 ± 0.2 | 67.6 ± 0.5 | 207 ± 3 | 15.5 ± 0.4 | 5,400 ± 33 | 16.8 ± 0.1 |
| *IsTs*[NTE]PadC[1] | 550 ± 3 | 38 ± 0.5 | 3,400 ± 7 | 62 ± 0.5 | - | - |
| *IsTs*PadC[2] | 15.2 ± 0.2 | 40.9 ± 0.7 | 354.8 ± 1.3 | 48.3 ± 0.2 | 6,800 ± 110 | 10.8 ± 0.1 |
| § *Is*[1-507]*Ts*PadC[2] | 21.5 ± 0.1 | 45.8 ± 0.4 | 112 ± 1 | 22.4 ± 0.4 | 2,800 ± 13 | 31.7 ± 0.1 |
| *IsTs*PadC[3] | 21.5 ± 0.2 | 38.0 ± 0.5 | 169.4 ± 0.5 | 50.8 ± 0.4 | 2,200 ± 7 | 11.1 ± 0.1 |
| *IsTs*PadC[4] | 16.9 ± 0.2 | 38.9 ± 0.7 | 423.4 ± 1.7 | 48.6 ± 0.2 | 110,000 ± 25,000 | 12.5 ± 0.1 |
| *IsTs*[NTE]PadC[4] | 17.3 ± 0.5 | 16.0 ± 0.5 | 84.8 ± 0.8 | 23.0 ± 0.5 | 48,600 ± 200 | 60.9 ± 0.1 |
| *Is*PadC[PSM] | 19.4 ± 0.2 | 35.7 ± 0.2 | 956 ± 34 | 12.1 ± 0.4 | 6,300 ± 90 | 52.1 ± 0.4 |
| *Is*PadC[#1] | 18.2 ± 0.2 | 36.7 ± 0.4 | 110.7 ± 0.6 | 48.1 ± 0.4 | 2,500 ± 50 | 15.2 ± 0.1 |
| *Is*PadC[#2] | 10.2 ± 0.8 | 33.0 ± 0.1 | 89 ± 6 | 42.6 ± 0.1 | 6,500 ± 5,000 | 24.4 ± 0.8 |
| *Ts*PadC[#1] | 32.6 ± 0.2 | 24.6 ± 0.6 | 1,176 ± 20 | 3.5 ± 0.2 | 75,000 ± 400 | 71.7 ± 0.2 |
| *Ts*PadC[#2] | 18 ± 2 | 27.5 ± 0.4 | 144 ± 56 | 12.4 ± 0.4 | 50,000 ± 15,000 | 60.1 ± 0.1 |
| *Is*PadC[PSM]mono[$1] | 257 ± 3 | 57.6 ± 0.5 | 652 ± 9 | 42.4 ± 0.5 | - | - |
| *Is*PadC[PSM]mono[$2] | 20.3 ± 0.7 | 28.4 ± 0.5 | 62 ± 2 | 62.2 ± 0.2 | 279 ± 39 | 9.4 ± 0.3 |
| *Is*PadC[PSM]mono[$3] | 19.3 ± 0.4 | 35.5 ± 0.4 | 102 ± 2 | 30.2 ± 0.4 | 650 ± 6 | 34.3 ± 0.2 |

**Appendix Table 2. Comparison of PadC kinetics of substrate conversion.** *Comparison of product formation between the various constructs was performed for initial reaction rates at 200 µM GTP. Initial rates are quantified from experimental triplicates for three time points, and the sample standard deviation of individual points contributed to the error estimation of the linear fit that is used to calculate the initial rate of product formation. The SE of the estimate from the linear regression is used as error indicator. *§* indicates constructs for which the kinetics of GTP conversion to c-di-GMP have been measured using a weaker red LED intensity of 0.7 mW cm$^{-2}$ compared to an intensity of 45 mW cm$^{-2}$ used for all the other measurements. **For the wildtype *Is*PadC the kinetics values were also measured with the weaker LED intensity of 0.7 mW cm$^{-2}$, at 630 nm and are also reported in brackets to compare with the variants analyzed only with this LED intensity which are indicated with the symbol §.

| | initial rates at 200 µM GTP (µmol product min$^{-1}$ µmol$^{-1}$ enzyme$_2$) | | |
|---|---|---|---|
| **Constructs** | **Dark** | **Light** | **Fold activation** |
| **Is*PadC** | (0.8 ± 0.1) / 1.3 ± 0.3 | (34.2 ± 0.4) / 100.5 ± 3.4 | (43 x) / 77 x |
| *Ts*PadC | 41 ± 2 | 240 ± 12 | 6 x |
| § *Ma*PadC | 2 ± 0.2 | 15.4 ± 0.7 | 8 x |
| § *Is*PadC$^{\Delta514\text{-}516}$ | 11 ± 1 | 10 ± 2 | - |
| § *Is*PadC$^{\Delta514\text{-}517}$ | 25 ± 7 | 49 ± 3 | 2 x |
| § *Is*PadC$^{\Delta514\text{-}520}$ | 27.2 ± 0.6 | 36 ± 4 | 1.3 x |
| § *Is*PadC$^{Reg1}$ | 2.13 ± 0.05 | 2.12 ± 0.01 | - |
| § *Is*PadC$^{Reg2}$ | 52 ± 4 | 117 ± 5 | 2.2 x |
| § *Is*PadC$^{Reg2.a}$ | 53.2 ± 0.4 | 53 ± 3 | - |
| *TsIs*PadC$^1$ | 61 ± 5 | 98 ± 6 | 2 x |
| *TsIs*PadC$^2$ | 103.7 ± 8.8 | 241.3 ± 4.3 | 2.3 x |
| *IsTs*PadC$^1$ | 5.5 ± 0.3 | 85.8 ± 1.3 | 16 x |
| *IsTs*$^{NTE}$PadC$^1$ | 6.9 ± 0.3 | 121.6 ± 7.2 | 17 x |
| *IsTs*PadC$^2$ | 2.6 ± 0.1 | 11.8 ± 0.5 | 5 x |
| *Is*$^{1\text{-}507}$*Ts*PadC$^2$ | 2.5 ± 0.2 | 65 ± 14 | 26 x |
| *IsTs*PadC$^3$ | 6.4 ± 0.6 | 116 ± 24 | 18 x |
| *IsTs*PadC$^4$ | 3 ± 0.6 | 26.4 ± 5.4 | 9 x |
| *IsTs*$^{NTE}$PadC$^4$ | 5.6 ± 1.8 | 9.4 ± 4.8 | 2 x |
| § *Is*PadC$^{\Delta442\text{-}477::SG}$ | 21 ± 3 | 13.0 ± 0.5 (78) | -1.6 x |
| § *Is*PadC - ccDGC | | 0.2 ± 0.1 | -4 x (to *Is*PadC) |

**Appendix Table 3. Data collection, phasing and refinement statistics.** [a] Values in parentheses are for highest-resolution shell. [b] Anomalous dataset: Friedel pairs are considered as different reflections.

| | *Is*PadC Se-Met (PDB code: 5LLW) | *Is*PadC GTP soaked native (PDB code: 5LLX) | *Is*PadC PSMcc native (PDB code: 5LLY) | *Is*PadC[Reg2] native (PDB code : 6ET7) | ccDGC c-di-GMP native (non deposited) |
|---|---|---|---|---|---|
| **Data collection** | | | | | |
| Space group | $P\,2_1\,2_1\,2_1$ | $P\,2_1\,2_1\,2_1$ | $P\,1\,2_1\,1$ | $P\,2_1\,2_1\,2_1$ | $P\,2_1\,2_1\,2_1$ |
| Cell dimensions | | | | | |
| $a$, $b$, $c$ (Å) | 50.87, 77.80, 439.98 | 50.73, 78.64, 452.04 | 83.70, 129.30, 122.28 | 49.43, 78.36, 443.56 | 45.58, 63.29, 125.74 |
| $\alpha$, $\beta$, $\gamma$ (°) | 90.0, 90.0, 90.0 | 90.0, 90.0, 90.0 | 90.0, 96.43, 90.0 | 90.00, 90.00, 90.00 | 90.00, 90.00, 90.00 |
| | | | | | |
| Wavelength | 0.979 | 0.966 | 0.966 | 1.008 | 1.039 |
| Resolution (Å)[a] | 68.7 – 3.0 (3.1 – 3.0) | 48.0 – 2.8 (2.9 – 2.8) | 48.1 – 2.4 (2.5 – 2.4) | 58.77 – 2.85 (2.95 – 2.85) | 70.0 – 3.0 (3.1 – 3.0) |
| $R_{meas\,(\%)}$ | 13.2 (162.4) | 18.4 (261.3) | 7.1 (65) | 22.1 (156.4) | 21.8 (220.2) |
| $I/\sigma(I)$ | 15.85 (0.83) | 7.81 (0.61) | 10.31 (1.47) | 9.28 (1.55) | 7.09 (0.88) |
| $CC_{1/2}$ | 99.8 (29.6) | 99.4 (14.2) | 99.7 (62.7) | 99.6 (55.6) | 98.4 (58.2) |
| Completeness (%) | 99.2 (99.7) | 97.5 (99.3) | 93.0 (94.1) | 99.9 (99.8) | 97.9 (96.3) |
| Redundancy | 14.94 (3.93) | 4.37 (4.55) | 2.85 (2.80) | 10.23 (10.76) | 3.42 (3.26) |
| | | | | | |
| **Refinement** | | | | | |
| Resolution (Å) | 68.7 – 3.0 | 48.0 – 2.8 | 48.1 – 2.4 | 58.77 – 2.85 | 62.87 – 3.0 |
| No. reflections | 67,329[b] | 44,894 | 93,313 | 41,744 | 7,637 |
| $R_{work}$ / $R_{free}$ | 0.205 / 0.259 | 0.223 / 0.270 | 0.177 / 0.230 | 0.21 / 0.27 | 0.20 / 0.33 |
| No. atoms | | | | | |
| Protein | 10,850 | 10,647 | 16,287 | 10,424 | 2,582 |
| Ligand/ion | 86 / 2 | 150 / 4 | 184 / 8 | 86 / 2 | 139 / 0 |
| Water | 0 | 4 | 354 | 6 | 0 |
| $B$ factors | | | | | |
| Protein | 98.0 | 92.7 | 59.0 | 84.0 | 78.98 |
| Ligand/ion | 73.9 | 80.2 | 48.0 | 58.4 | 52.94 |
| Water | - | 56.0 | 48.6 | 31.3 | - |
| r.m.s deviations | | | | | |
| Bond lengths (Å) | 0.010 | 0.009 | 0.008 | 0.008 | 0.016 |
| Bond angles (°) | 1.121 | 0.797 | 0.891 | 1.141 | 1.770 |

**Appendix Table 4. Overview of oligonucleotides and buffers**. (**A**) Oligonucleotides used in this thesis. (**B**) Buffers used for purification and storage of the different PadC homologs.

| Panel **A** - oligonucleotides | | |
|---|---|---|
| **Desired construct** | **Oligonucleotide (5'-3')** | **Template** |
| *Is*PadC $\Delta^{514–516}$ | fw: gcagctgaatctgctgaatgatgatgcaaatgaaaatctggaaaaactggccag<br>rv: atcattcagcagattcagctgcatgctatctgccacaatcagcagatcac | *Is*PadC |
| *Is*PadC $\Delta^{514–517}$ | fw: gcagctgaatctgctgaatgatgcaaatgaaaatctggaaaaactggccagc<br>rv: atcattcagcagattcagctgcatgctatctgccacaatcagcagatcac | *Is*PadC |
| *Is*PadC $\Delta^{514–520}$ | fw: gcagctgaatctgctgaatgataatctggaaaaactggccagctttgatgatc<br>rv: atcattcagcagattcagctgcatgctatctgccacaatcagcagatcac | *Is*PadC |
| *Is*PadC PSMcc | fw: ctggaaaaactggccagctaagcggccgcactcgagc<br>rv: gctggccagttttttccagattttcatttgcatctgccagctgatc | *Is*PadC |
| *Is*PadC register 1 | fw: ctgaatctgctgaatgatcagctggcagatctaaatgaaaatctggaaaaactgg<br>rv: ctgatcattcagcagattcagctgcatgctaagtgccacaatcagc | *Is*PadC |
| *Is*PadC register 2 | fw: aatgatcagctggcagatgcaaatgaaaatctggaaaaactggtcagctttgatgatctg<br>rv: catctgccagctgatcattcagcagattcagctgcatgacatctgccacaatc | *Is*PadC |
| *Is*PadC Reg2.a | fw: atcagctggcagatgcagttgaaaatctggaaaaactggtcagc<br>rv: gcatctgccagctgatcaaccagcagattcagctgcatg | *Is*PadCReg2 |
| *Is*PadC ccDGC | fw: cagctgaatctgctgaatgatcagctggcagatgcaaatgaaaatctg<br>rv: gatcattcagcagattcagctgggcgccctgaaaataaagattctcag | *Is*PadC |
| *Is*PadC $\Delta^{442-477}$::SG | fw: acagagcggcagcggtaaaagccagccgtggcgtac<br>rv: cgctgccgctctgtgcaacacgaaacagcagcagatagc | *Is*PadC |
| *Ts*PadC $\Delta^{441-474}$::SG | fw: ggcttttaccgctaacttcatcagtccttgcaacacgaaacagcagcag<br>rv: gatgaagttagcggtaaaagccagccgtggcgtaccg | *Ts*PadC |
| *TsIs*PadC[1] | fw: tggaacgtctggcaacccgtgatgatctgaccggtatc<br>rv: tgctcgagtgcggccgcttactggctacaaacctgattac | *Is*PadC |
| | fw: taagcggccgcactcgag<br>rv: acgggttgccagacgttc | *Ts*PadC |
| *TsIs*PadC[2] | fw: atctggcaagcgatctggccattgtggcagatagcatg<br>rv: tgctcgagtgcggccgcttactggctacaaacctgattac | *Is*PadC |
| | fw: taagcggccgcactcgag<br>rv: ggccagatcgcttgccag | *Ts*PadC |
| *IsTs*PadC[1] | fw: taagcggccgcactcgag<br>rv: atcaaagctggccagttttttcc | *Is*PadC |
| | fw: tggaaaaactggccagctttgatcatctgaccggtctg<br>rv: tgctcgagtgcggccgcttaatctgcgctaacttcaac | *Ts*PadC |
| *IsTs*[NTE]PadC[1] | fw: tggttgcagcactggaagcatgtgaacgtgaaccgattcatattccgaatgcaattcagc<br>rv: cttccagtgctgcaaccagttccggttcatccatggcgccctgaaaataaagattctcagtagtg | *IsTs*PadC[1] |
| *IsTs*PadC[2] | fw: taagcggccgcactcgag<br>rv: cagcagatcacgtgcaatatcac | *Is*PadC |
| | fw: atattgcacgtgatctgctggttctgattgcagccggt<br>rv: tgctcgagtgcggccgcttaatctgcgctaacttcaacg | *Ts*PadC |
| *Is*[1-507]*Ts*PadC[2] | fw: attgtggcagatagcatgcagattagccagctgaatcagc<br>rv: ctgcatgctatctgccacaatcagcagatcacgtgc | *IsTs*PadC2 |
| *IsTs*PadC[3] | fw: taagcggccgcactcgag<br>rv: aacggtaaccatataacgttcaacattacg | *Is*PadC |
| | fw: gaacgttatatggttaccgttaatgaaagccgtgatgttctgg<br>rv: tgctcgagtgcggccgcttaatctgcgctaacttcaacg | *Ts*PadC |
| *IsTs*PadC[4] | fw: taagcggccgcactcgag<br>rv: atgaatcagccacagacgttct | *Is*PadC |
| | fw: cgtctgtggctgattcatgcacgtgatgatgcagttctgat<br>rv: tgctcgagtgcggccgcttaatctgcgctaacttcaacg | *Ts*PadC |
| *IsTs*[NTE]PadC[4] | fw: tggttgcagcactggaagcatgtgaacgtgaaccgattcatattccgaatgcaattcagc<br>rv: cttccagtgctgcaaccagttccggttcatccatggcgccctgaaaataaagattctcagtagtg | *IsTs*PadC[4] |
| *Is*PadC[PSM] | fw: gcacgtgatctgctgtaagcggccgcactcg<br>rv: cagcagatcacgtgcaatatcacgtgctgcatacagctg | *Is*PadC |
| *Is*PadC[PSM] F132E F136E | fw: gaacagcgtgccgaacagaaactgcgtaattgcaaaac<br>rv: ggcacgctgttcctgaatttcaaaccaattgctggtc | *Is*PadC |

| | | |
|---|---|---|
| *Is*PadC<sup>PSM</sup> F132E F136E F303E F310E | fw: gaa gcagcagaacgtctgtgggagattcatagccgtaatgttg<br>rv: gacgttctgctgcttcttcaacggtacgaaccagcagac | *Is*PadC<sup>PSM</sup> F132E F136E |
| *Is*PadC<sup>PSM</sup> F132E F136E F303E F310E E272A | fw: ccagcattggcatttttaacgcagatgaactgtggggtattgttgc<br>rv: gttaaaaatgccaatgctggtgcttgcgctaacaccaaaattacg | *Is*PadC<sup>PSM</sup> F132E F136E F303E F310E |
| *Is*PadC<sup>PSM</sup> F132E F136E F303E F310E E272A D494A | fw: ctgtatgcagcacgtgctattgcacgtgatctgctg<br>rv: acgtgctgcatacagctgtgcggtacgcc | *Is*PadC<sup>PSM</sup> F132E F136E F303E F310E E272A |
| *Is*PadC<sup>PG</sup> | fw: cagaacgtctgtggctgattcattaagcggccgcac<br>rv: atgaatcagccacagacgttctgctgcaaattcaacggtacg | *Is*PadC |
| *Ts*PadC<sup>PG</sup> | fw: gcatgttttcagctggaactgattcagtaagcggccgcac<br>rv: ctgaatcagttccagctgaaaacatgcggtctgaaccattgcac | *Ts*PadC |

| Panel **B** – buffer systems | | |
|---|---|---|
| **Use** | **Buffer composition** | |
| Storage buffer | 10 mM HEPES pH 7, 0.5 M NaCl, 2 mM MgCl$_2$ | *Is*PadC; *Ts*PadC and related variants |
| | 10 mM Tris/Cl pH 8, 0.5 M NaCl, 2 mM MgCl$_2$ | *Ma*PadC |
| Lysis buffer | 50 mM HEPES pH 7, 0.5 M NaCl, 2 mM MgCl$_2$, 10 mM imidazole, 1 mM EDTA | *Is*PadC; *Ts*PadC and related variants |
| | 50 mM Tris/Cl pH 8, 0.5 M NaCl, 2 mM MgCl$_2$, 10 mM imidazole, 1 mM EDTA | *Ma*PadC |
| Dialysis buffer | 50 mM HEPES pH 7, 0.5 M NaCl, 2 mM MgCl$_2$, 1 mM EDTA, 1 mM DTE | *Is*PadC; *Ts*PadC and related variants |
| | 50 mM Tris/Cl pH 8, 0.5 M NaCl, 2 mM MgCl$_2$, 1 mM EDTA, 1 mM DTE | *Ma*PadC |
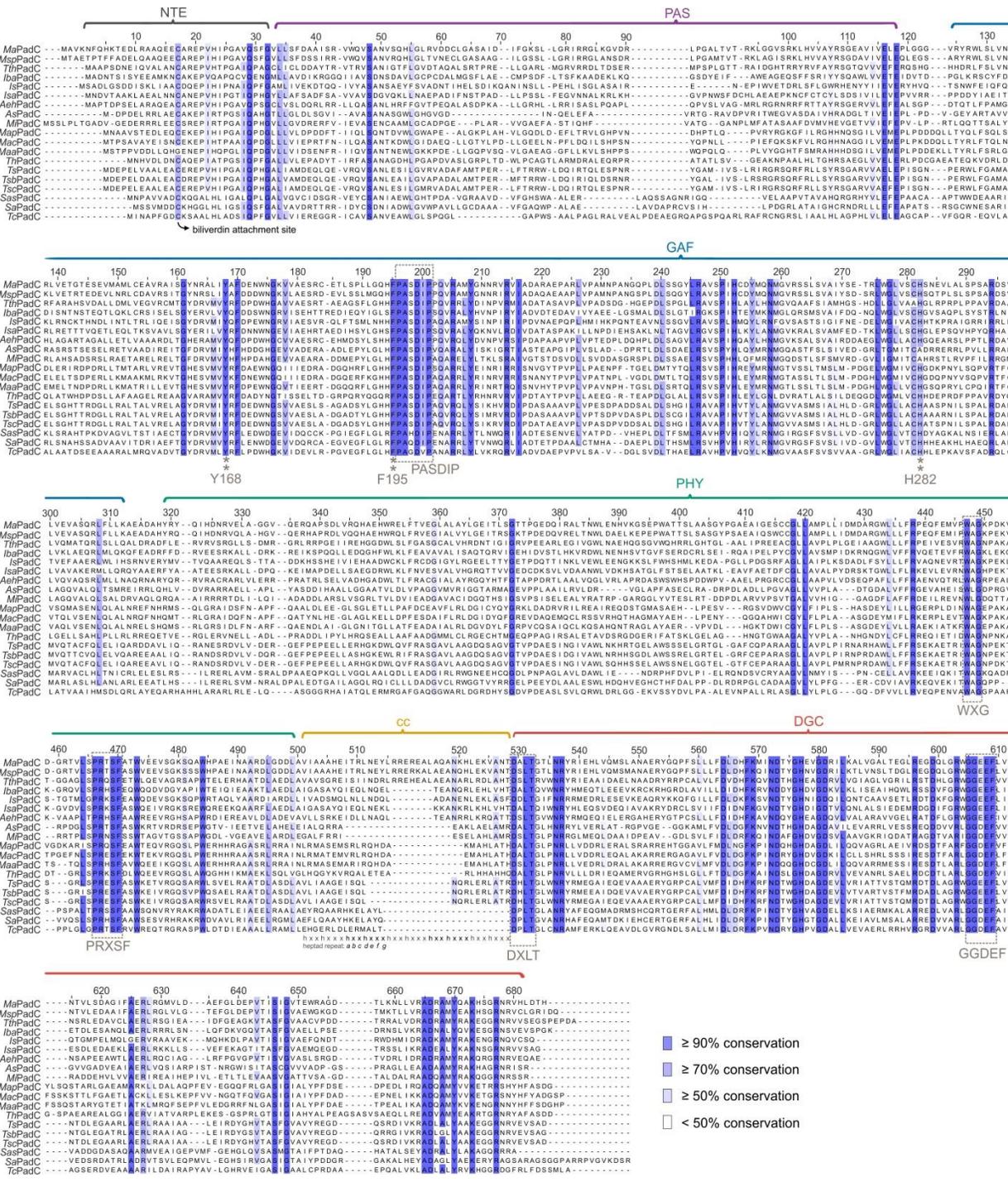
**Appendix Figure 1. Multiple sequence alignment of PadC homologs.** Residues numbering is based on *Is*PadC sequence. *Is – Idiomarina species A28L* and WP_007419415; *Ma – Marinimicrobium agarilyticum* and WP_027329460; *Ts – Thioalkalivibrio species ALMg3* and WP_026331574; *Th – Thiohalospira halophila* and WP_093427120.1; *Maa – Marinospirillum alkaliphilum* and WP_072324728.1 ; *Mac – Marinospirillum celere* and WP_091958617.1 ; *Map – Marinobacter persicus* and WP_091706258.1 ; *Tc – Thauera chlorobenzoica* and WP_075148475.1 ; *Sa – Salinisphaera hydrothermalis* and WP_ 084188582.1; *Sas – Salinisphaera shabanensis* and WP_084623642.1 ; *As – Agromyces sp. Leaf222* and WP_082462169.1 ; *Mi – Microbacterium pygmaeum* and

WP_091492622.1; *Msp – Marinimicrobium species* LS-A18 and WP_ 036160751 ; *Tth – Thioalkalivibrio thiocyanodenitrificans* and WP_051079936; *Iba – diomarina baltica* and WP_006955969; *Isa – Idiomarina salinarum* and WP_034774458; *Aeh – Alkalilimnicola ehrlichii* and WP_011629241; *Tsb – Thioalkalivibrio species* AKL8 and WP_026304848; *Tsc – Thioalkalivibrio species* ALMg11 and WP_018949990. Highly conserved motif are indicated and boxed in dashed lines. Highly conserved residues changing rotamer upon 15$E$ isomerization of the BV are indicated by two asterix (\*\*) and numbered based on IsPadC sequence numeration.
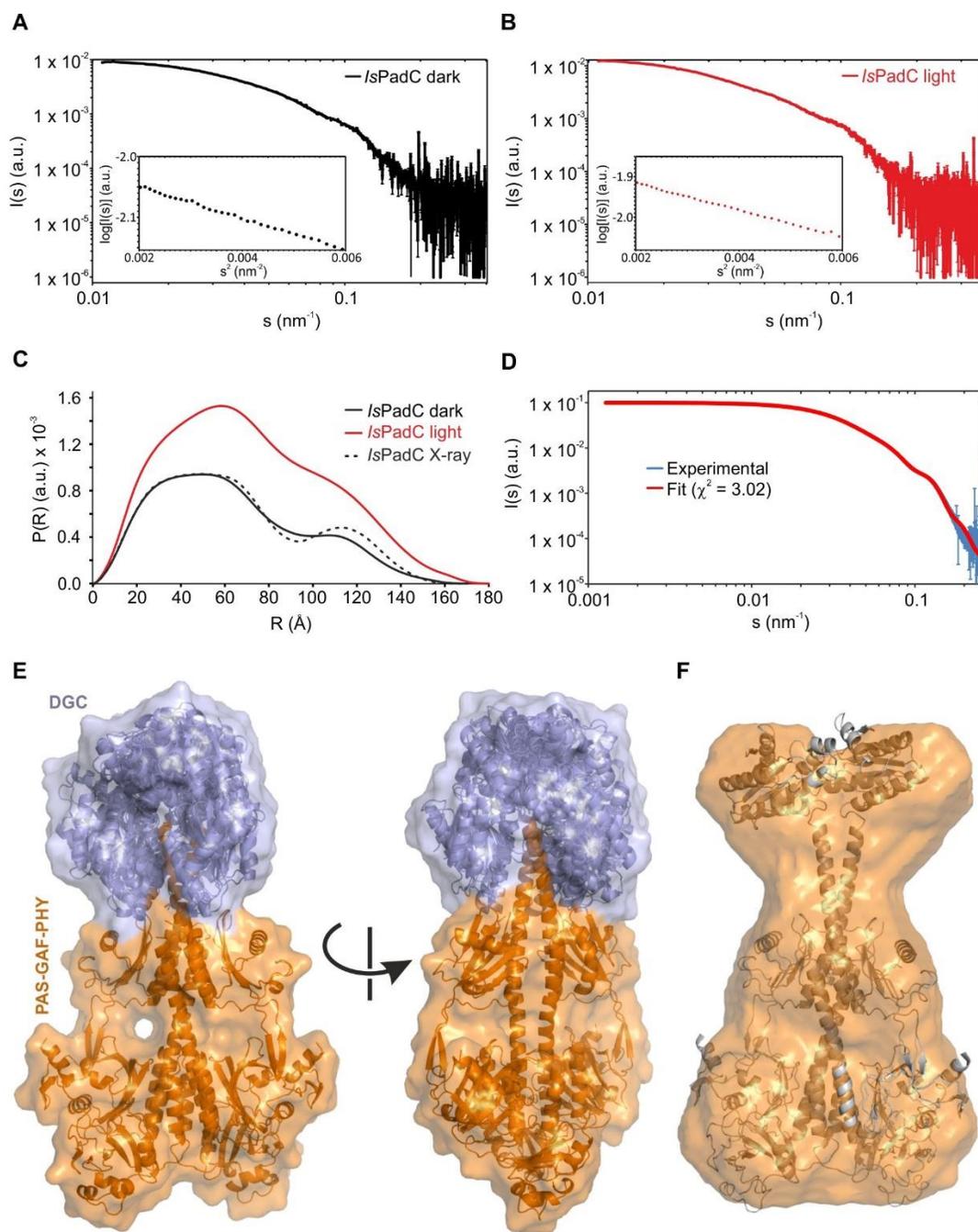
**Appendix Figure 2. Details of SAXS measurements.** (**A** and **B**) Raw SAXS data of dark and light-adapted *Is*PadC states, respectively. The insets show the corresponding Guinier plots, which demonstrate the absence of protein aggregation. Both, the s, and I(s) axes are shown in a logarithmic representation. (**C**) SAXS data showing a comparison of the experimental radial density distributions of *Is*PadC in the dark- and light-state with a back-calculated radial density distribution from the crystal structure. SAXS-derived radius of gyration ($R_g$), maximal dimension ($D_{max}$), and molecular mass are 48.9 Å/160 Å /138 kDa for the dark and 55.9 Å/170 Å/237 kDa for the light-state, respectively. Transient oligomerization is frequently observed for phytochrome modules (195) and not directly linked to the DGC activation mechanism. For comparison SAXS data was back-calculated from the dark-

state crystal structure using the program CRYSOL (196). (**D**) Comparison of experimental *Is*PadC dark-state SAXS data with SAXS data back-calculated from the CORAL model. Both, the s, and I(s) axes are shown in a logarithmic representation. The angular ranges from 0.0012–0.3 $nm^{-1}$ are compared. (**E**) Cartoon and surface representation of the bundle of 7 lowest energy structures from the SAXS-based rigid body modeling calculations of *Is*PadC in the dark-state. PAS-GAF-PHY and DGC domains are shown in orange and blue respectively. (**F**) SAXS-based *ab initio* low-resolution model of *Is*PadC superimposed with the crystal structure determined here.

A  $D_{rel}$ IsPadC(light) − $D_{rel}$ IsPadC(dark)

B  $D_{rel}$ IsPadC PSMcc(light) − $D_{rel}$ IsPadC PSMcc(dark)

C  $D_{rel}$ IsPadC(dark) − $D_{rel}$ IsPadC PSMcc(dark)

D  $D_{rel}$ IsPadC(light) − $D_{rel}$ IsPadC PSMcc(light)
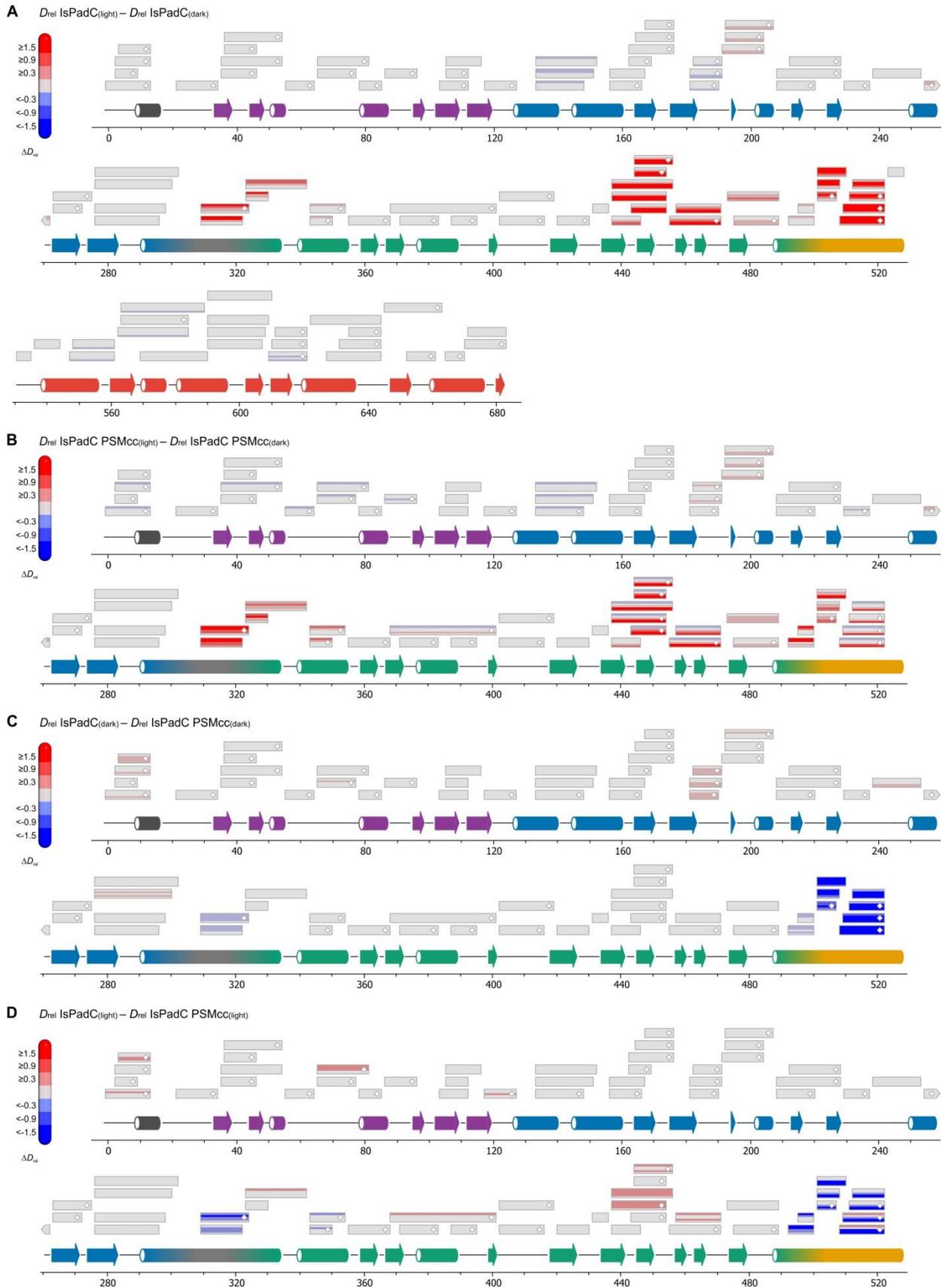
**Appendix Figure 3. Summary of HDX experiments.** A complete illustration of all *Is*PadC peptides evaluated by HDX is presented in panels A-D. Each box represents one peptide and contains up to five different colors that, from the bottom up, correspond to deuteration times of 10, 45, 180, 900 and 3,600 s, respectively. Individual colors correspond to the change in relative deuteration ($\Delta D_{rel}$) of the two compared states according to the legend in the top left corner. $MS^2$ confirmed peptides are marked with diamonds. Terminal arrows at the end of a box indicate continuation of the peptide in the previous or following line. Secondary structure elements are taken from DSSP analysis of the corresponding structures and are colored according to the domain representation of Fig. 2B. A compilation of all individual deuteration plots for *Is*PadC and *Is*PadC PSMcc can be found in Figure S7. (**A**) Changes in *Is*PadC deuteration upon red-light illumination. (**B**) Changes in *Is*PadC PSMcc deuteration upon red-light illumination (also see Movie S2). (**C**) Changes in deuteration of *Is*PadC PSMcc compared to full length *Is*PadC in their dark-adapted states (also see Movie S3). (**D**) Changes in deuterium uptake of the light-adapted states of *Is*PadC PSMcc compared to full length *Is*PadC.
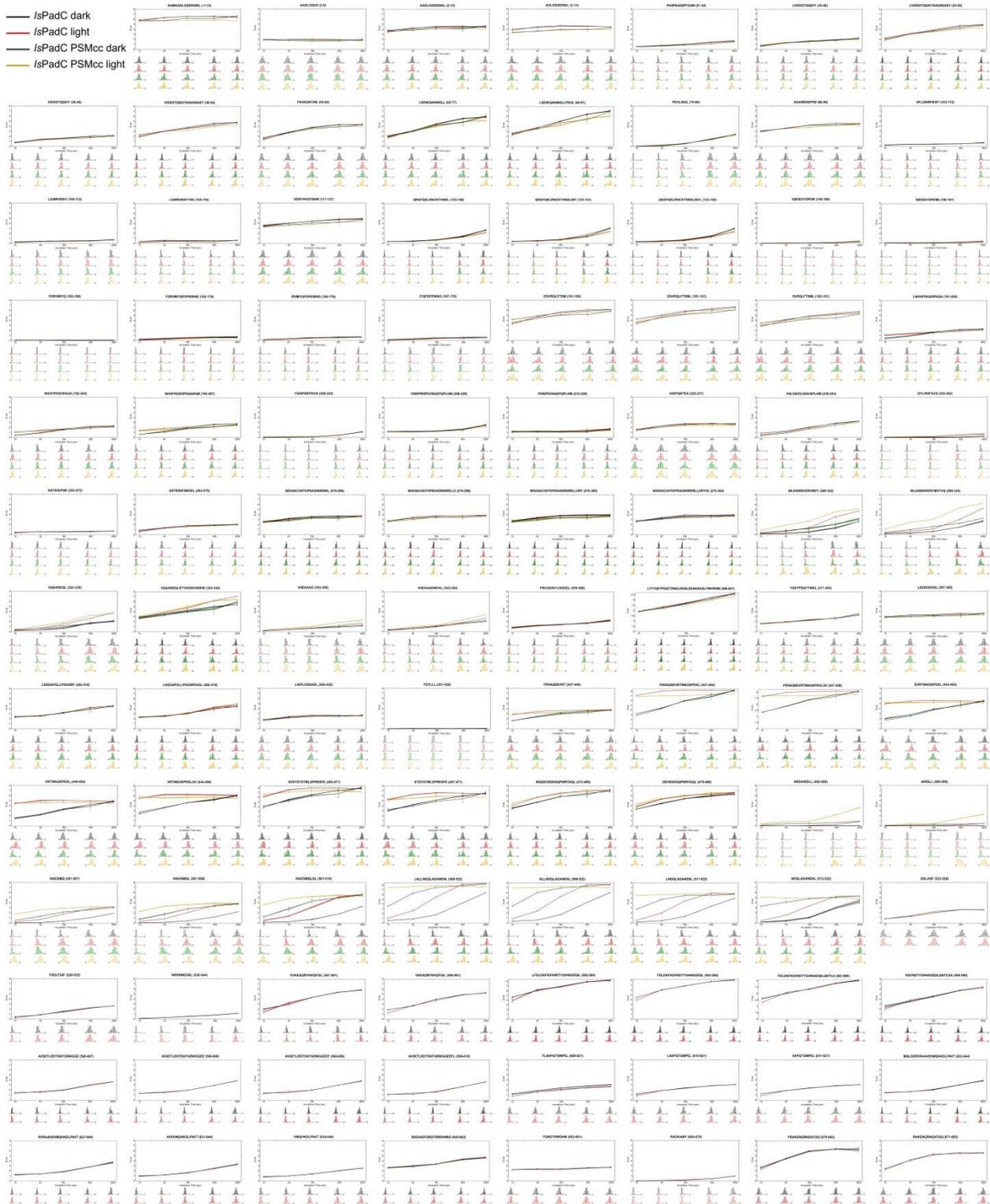
**Appendix Figure 4. Individual deuterium incorporation plots of all evaluated peptides.** Peptide sequences and their corresponding position according to the *Is*PadC sequence are shown on top of each sub-panel. Ordinates of each plot correspond to the relative deuterium incorporation and abscissae to the labeling time in seconds. *Is*PadC traces are colored in black and red corresponding to *Is*PadC in the dark- and light-state, Similarly, *Is*PadC PSMcc traces are

colored in green and orange for the dark- and light-state, respectively. $D$rel values are shown as the mean of three independent measurements and error bars correspond to the standard deviation. A software-estimated abundance distribution of deuterated species is presented in the lower subpanel on a scale from undeuterated to all exchangeable amides deuterated.



**A**

**B**
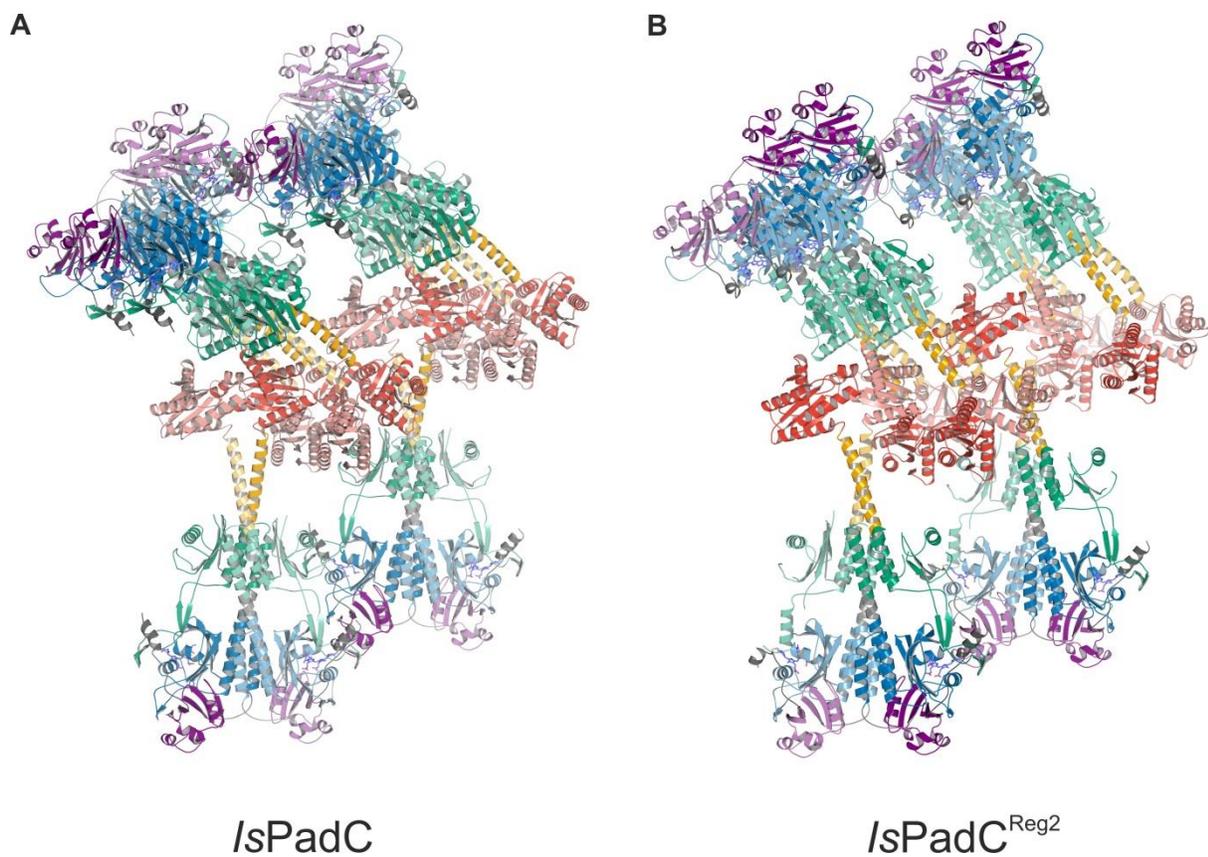
*Is*PadC

*Is*PadC<sup>Reg2</sup>

**Appendix Figure 5. Crystal packing comparison between *Is*PadC and *Is*PadC<sup>Reg2</sup>.** Crystal packing in the *Is*PadC crystals (**A**), and in the *Is*PadC<sup>Reg2</sup> crystals (**B**). Overall similar protein contacts within the crystals are observed.
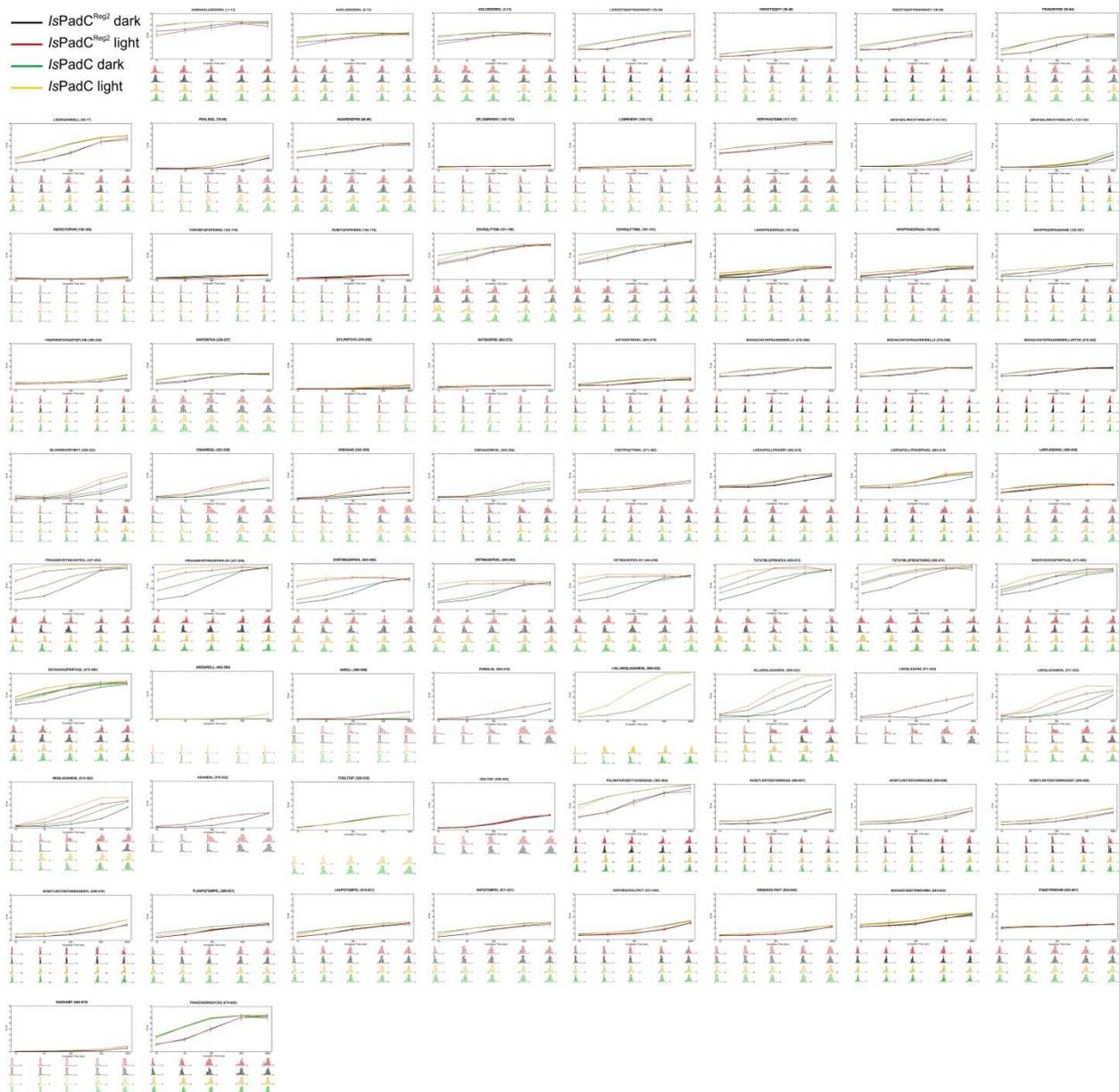
**Appendix Figure 6. Individual deuterium incorporation plots of all evaluated peptides and comparison of common *Is*PadC^Reg2 and *Is*PadC peptides.** Peptide sequences and their corresponding positions are shown on top of each sub-panel. The relative deuterium incorporation is plotted against the labeling time in seconds. *Is*PadC traces are colored in green and orange for the dark- and light-state, respectively, while *Is*PadC^Reg2 traces are colored in black and red for the dark- and light-state, respectively. $D_{rel}$ values are shown as the mean of three independent measurements and error bars correspond to the standard deviation. The lower part of each sub-panel shows software-estimated abundance distributions of individual deuterated species on a scale from undeuterated to all exchangeable amides deuterated.
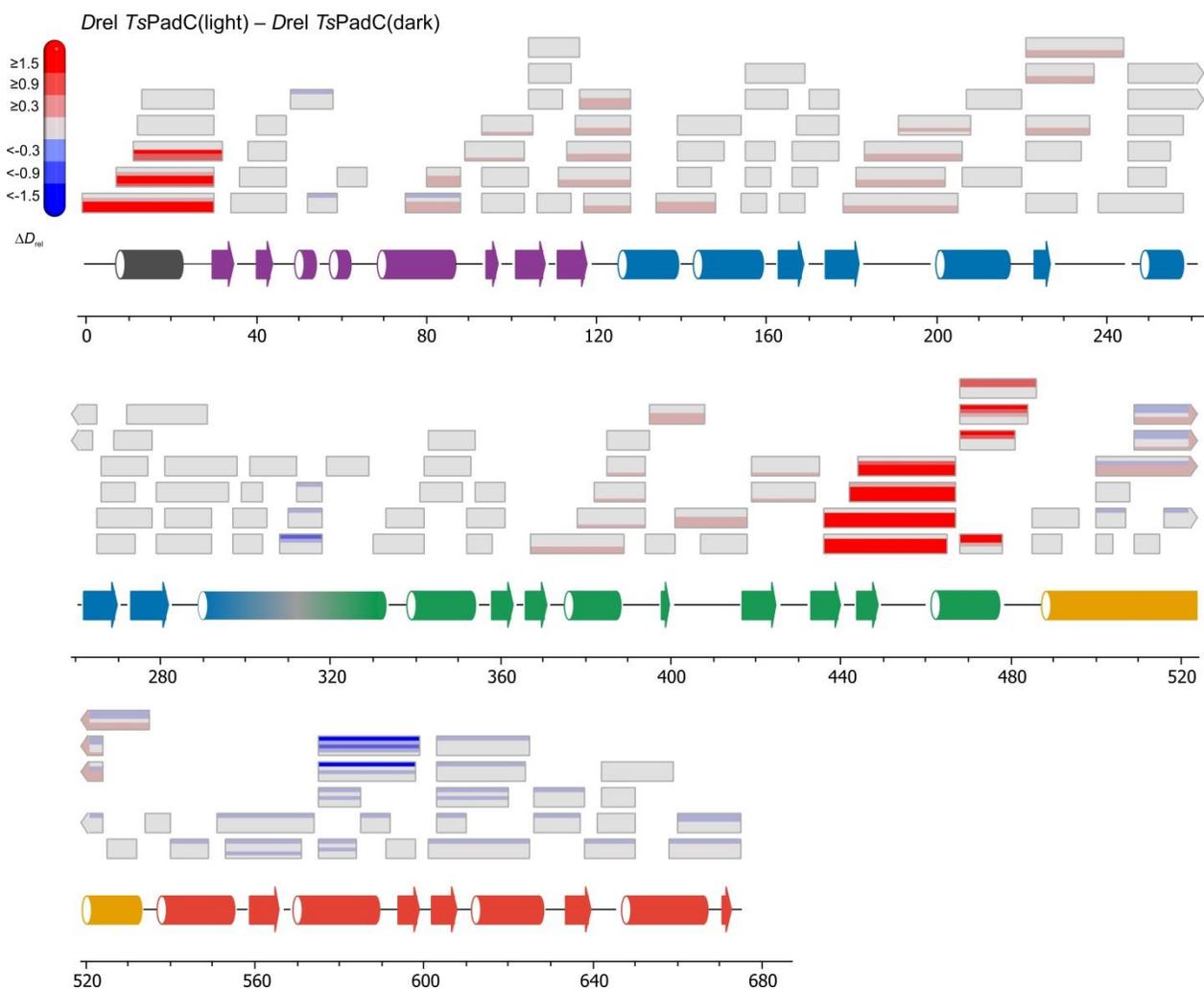
**Appendix Figure 7. Summary of *Ts*PadC HDX analysis.** Changes in *Ts*PadC deuteration upon red-light illumination. A complete illustration of all *Ts*PadC peptides evaluated by HDX is presented. Each box represents one peptide and contains up to five different colors that, from the bottom up, correspond to deuteration times of 10, 45, 180, 900 and 3,600 s, respectively. Individual colors correspond to the change in relative deuteration ($\Delta D_{rel}$) of the two compared states according to the legend in the top left corner. Secondary structure elements are taken from a secondary structure prediction using the PSIPRED server (197, 198). Due to the very slow ground state recovery of TsPadC, the deuterium uptake measurements in dark conditions do not reflect an homogeneous Pr/Pr homodimer population and might be overestimated for certain regions.