Julia Färber, BSc

Biochemical characterization of blue light-regulated LOV-diguanylate cyclases

MASTER'S THESIS

to achieve the university degree of

Master of Science

Master's degree programme: Biochemistry and Molecular Biomedical Sciences

submitted to

Graz University of Technology

Supervisor

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

Institute of Biochemistry

Second supervisor: Ass.Prof. Dipl.-Ing. Dr.techn. Andreas Winkler

Faculty of Technical Chemistry, Chemical and Process Engineering, Biotechnology

Graz, June 2018

AFFIDAVIT

I declare that I have authored this thesis independently, that I have not used other than the declared sources/resources, and that I have explicitly indicated all material which has been quoted either literally or by content from the sources used. The text document uploaded to TUGRAZonline is identical to the present master‘s thesis.

Date	Signature

“Education is not the learning of facts,
but the training of the mind to think.”

Albert Einstein

Acknowledgements

First of all, I would like to thank everybody, who accompanied me throughout the instructive, inspiring but sometimes stressful path of my master's thesis, laboratory practice as well as the written thesis. Thank you very much for all the mental support and even constructive criticism.

Furthermore, I want to sincerely thank my supervisor Peter Macheroux and also my second supervisor Andreas Winkler, who gave me the opportunity to do my master's thesis in his working group at the Institute of Biochemistry, which is also part of Peter Macheroux' research group. Due to his positive attitude and to see things with a certain serenity in a very professional way, the practical work in Andreas' laboratory became a very pleasant time for me. Moreover, he encouraged me, answered my questions and helped me to solve both theoretical and practical problems in a very patient way.

In addition, a special thank goes to Elfriede Zenzmaier, who did not only meet me on a professional but also on a personal level. Thank you for the honesty, advice and support. Elfi often built up my self-confidence, if an experiment did not work and therefore I got frustrated. Furthermore, she was there, if I just needed someone to talk to. Because of her work experience, Elfi helped me to do many laboratory procedures more easy without losing their accuracy.

I would also like to thank Stefan, Geoffrey and several other members of the Institute. They were always accommodating, open-minded and very friendly. I really felt comfortable there.

A very emotional thank goes to my family, especially to my parents. Without their support I would not have been able to come as far as I have. They always believed in me and supported me unconditionally, which is not self-evident. I love you!

Finally, I would like to thank my boyfriend, best friends and fellow students for all the motivation and encouragement up to the end.

Abstract

LOV-diguanylate cyclases are light-sensitive proteins, responding to blue light. Bacteria use this environmental stimulus for the regulation of their cellular functions, like enzymatic activities that correlate with the production of the bacterial second messenger c-di-GMP. At the C-terminus of these photoreceptors, GGDEF domains work as effectors, featuring the catalytic activity of diguanylate cyclases. As output domains, they are responsible for downstream signaling. In addition, such proteins contain an N-terminal LOV sensor as input domain, which uses flavin as cofactor. Upon photon absorption at 445 nm, this results in a light-activated adduct state of the photosensory domain. Sensor and effector domains are covalently linked to each other via specific supercoiled α-helices that are essential for signal transduction within LOV photoreceptors.

In this project, two naturally occurring as well as two engineered LOV-GGDEF versions were structurally and functionally characterized to get knowledge about the modulation of blue light-regulated LOV-diguanylate cyclases. Wild-type *Sf*LadC (light-activated diguanylate cyclase) originates from *Serratia fonticola* and *Ss*LadC from *Salinisphaera shabanensis*. The chimeras LPadC_A and LPadC_B are combinations of the two constructs *Sf*LadC and *Is*PadC (phytochrome-activated diguanylate cyclase) that originates from *Idiomarina species* A28L.

Full-length *Sf*LadC shows aggregation and oligomerization after size-exclusion chromatography, influenced by different light conditions. In contrast, all the other tested systems exhibit light-independent dimeric states, leading to high protein stability. In addition, *Sf*LadC is 100-fold activated upon illumination compared to the dark, whereas LPadC_B is about 10-fold and *Ss*LadC is about 1.5 fold activated by blue light exposure in comparison to darkness. In contrast, LPadC_A offers low enzymatic activity under light conditions and no activity in the dark because of low enzyme concentration or short incubation times. All tested proteins seem to be light-regulated.

Zusammenfassung

LOV-Diguanylatzyklasen sind lichtempfindliche Proteine, die durch blaues Licht aktiviert werden. Bakterien reagieren auf diesen äußeren Umweltreiz und regulieren dadurch ihrer zellulären Funktionen wie beispielsweise enzymatische Aktivitäten, die proportional zur Produktion und teilweise auch an der Herstellung des bakteriellen Signalmoleküls c-di-GMP beteiligt sind. Am C-Terminus dieser Photorezeptoren befinden sich sogenannte GGDEF Domänen als Effektoren, welche die katalytische Aktivität von Diguanylatzyklasen aufweisen. Zusätzlich besitzen solche Proteine LOV Domänen als Sensoren, die Flavin als Kofaktor benutzen. Nach der Absorption von Photonen im Bereich von 445 nm entsteht ein licht-aktivierter Adduktzustand dieser photosensorischen Domänen. Sensor- und Effektordomänen sind über eine superspiralisierte α-Helix kovalent miteinander verbunden, welche essentiell für die Signaltransduktion innerhalb von Photorezeptoren ist.

In dieser Masterarbeit wurden zwei natürlich vorkommende als auch zwei synthetisch hergestellte LOV-GGDEF Systeme hinsichtlich ihrer Struktur und Funktion charakterisiert, um ein besseres Verständnis über die Modulation solcher blaulicht-aktivierten LOV-Diguanylatzyklasen zu erhalten.

Wildtyp *Sf*LadC (licht-aktivierte Diguanylatzyklase) stammt von *Serratia fonticola* und Wildtyp *Ss*LadC von *Salinisphaera shabanensis*. Die Chimären LPadC_A und LPadC_B sind Kombinationen der beiden Konstrukte *Sf*LadC und *Is*PadC (phytochrom-aktivierte Diguanylatzyklase), das von *Idiomarina species* A28L stammt. Nach Größenausschlusschromatographie zeigt das natürlich vorkommende Konstrukt *Sf*LadC sowohl Aggregation als auch Oligomerisierung, beeinflusst durch variierende Lichtverhältnisse. Im Gegensatz dazu weisen alle weiteren, getesteten Varianten licht-unabhängige Dimere auf, was eine hohe Proteinstabilität zur Folge hat. Darüber hinaus ist *Sf*LadC bei Belichtung 100-fach aktiviert im Vergleich zur Dunkelheit, während LPadC_B eine 10-fache und Wildtyp *Ss*LadC eine 1,5 fache Aktivierung durch Belichtung mit Blaulicht verglichen mit der Dunkelheit zeigen. Die Chimäre LPadC_A hingegen hat eine geringe enzymatische Aktivität im Licht auf und ist völlig inaktiv im Dunklen. Abschließend kann man sagen, dass LPadC_B eine deutliche Lichtregulation aufweist, während *Ss*LadC nur minimal durch Licht reguliert wird. Zudem kann man davon ausgehen, dass auch LPadC_A licht-reguliert wird.

Table of Contents

1. Introduction

In nature, light is an essential energy source. Several organisms, like bacteria, utilize this significant environmental stimulus for the molecular regulation of different cellular functions such as enzymatic activities employing light-sensitive protein domains. These photoreceptors react to variable light conditions and for that purpose are linked to various cofactors. Upon illumination, the sensor domains get activated, leading to the start of signaling processes. With optogenetics, such dynamic and complicated biological systems can be visually controlled. This technique is a combination of biochemical and genetic approaches to achieve both spatial and temporal accuracy. The conformation of photo-switchable enzymes in optogenetic tools can be altered by the absorption of photons. There are two different types of photosensitive proteins, naturally occurring and engineered ones [1] [2] [3]. Found in several prokaryotic as well as eukaryotic organisms, blue light-modulated photoreceptors, containing light-oxygen-voltage sensing (LOV) domains, are related to PER–ARNT (aryl hydrocarbon receptor nuclear translocator) –SIM (PAS) proteins. In addition, LOV proteins generally appear in both phototropic and chemotropic bacteria [4].

1.1 Flavin-based light-sensitive proteins

First of all, there are three different types of flavin-based photosensors. They are called LOV domains, sensors of blue light using flavin adenine dinucleotide (BLUF) and cryptochromes (CRYs). The main functions of these proteins are the regulation and modulation of several biological activities in response to blue-light induction [5]. The wavelength of photon absorption is located between 320 and 500 nm (Ultra Violet A (UVA)/blue light). LOV and BLUF domains are able to bind different flavin derivates like flavin mononucleotide (FMN) or flavin adenine dinucleotide (FAD) (Figure 1) as cofactors. These two photoreceptor modules can be combined with several effector domains, which is important for the light regulation of enzyme activity and even for the biomolecular interaction with DNA or other proteins [2].

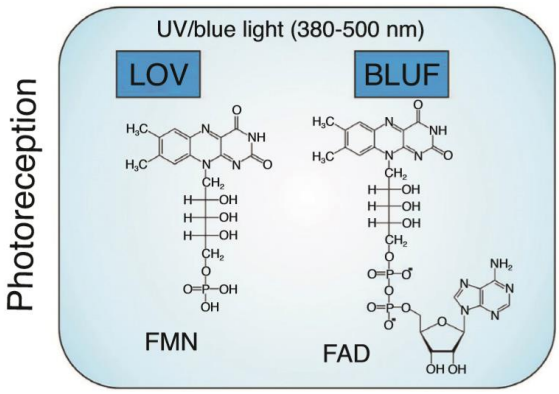

Figure 1. Activation of photosensory domains (LOV/BLUF) upon interactions of their cofactors (FMN/FAD) with photons at a wavelength between 380 and 500 nm (UVA/blue light) [2].

Blue light photoreceptors utilize flavin derivates as chromophores that are bound to proteins' domains or subdomains. Photon absorption of such derivates results in redistributed charges and therefore in altered redox potentials of this molecule, which can be employed to respond to UVA, blue or even green light. These reactions are necessary to start the signal transduction in answer to environmental stimuli [5] [6]. The chromophores of blue light photosensory domains are all derivates of riboflavin (RB) (= 7, 8-dimethyl-10-(1-deoxy-D-ribitol-1-yl) isoalloxazine) also known as vitamin B2 (Figure 2). RB kinase induces the phosphorylation of RB to FMN (riboflavin 5'-phosphate) and FMN adenylyltransferase adds an adenine group to RB, resulting in FAD [6].

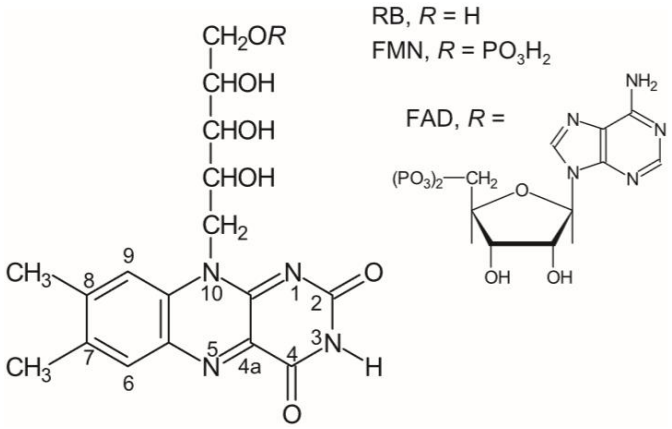

Figure 2. Cofactors of blue light photoreceptors. RB is a cyclic 7,8-dimethyl-10-(1-deoxy-d-ribitol-1-yl)isoalloxazine. FMN and FAD are derivates of RB.
RB = riboflavin; FMN = flavin mononucleotide; FAD = flavin adenine dinucleotide [6]

The spectra of the ground state of a flavin ring has an absorption of approximately 450 nm and about 12.500 $M^{-1}cm^{-1}$ (Figure 3). In water, RB and FMN show fluorescence, whereas FAD features a reduced fluorescence due to internal quenching by the adenosine part. The complete reduction of the ground state of FAD or FMN leads to the interruption of the conjugated system. Because of that, the reduced form of the cofactors is colorless and features only a weak absorption of visible photons, and hence no fluorescence [6].

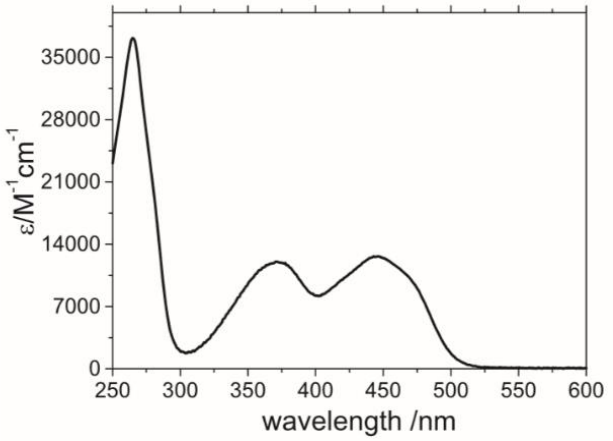

Figure 3. Absorption spectrum of the oxidized ground state of all blue light flavin derivates (RB, FMN and FAD) in solution. RB = riboflavin; FMN = flavin mononucleotide; FAD = flavin adenine dinucleotide [6]

1.1.1 Light–oxygen–voltage sensitive photoreceptors

Phototropin was identified as a photosensory protein in plants that is responsible for reactions and therefore modifications related to UVA and blue light absorption. In that case, the light-sensitive protein is composed of two LOV sensors, which are non-covalently linked to one FMN molecule at the N-termini of these proteins. At the C-termini, kinases are located as effector domains, resulting in the autophosphorylation of these proteins. In contrast, other LOV light-sensitive proteins offer variable effectors like kinases, cyclases or diverse modules for interaction with DNA or other proteins [4] [5] [7]. Moreover, with different sensor/effector combinations, it is also possible to genetically create and develop diverse artificial chimeras [7]. In general, the connection between sensor and effector occurs with one α-helix between the appropriate domains [4].

Turning to the structural composition, there are five different β-strands linked to four different α-helices, building the fundamental β-scaffold of LOV sensors. Two of these helices (Eα and Fα) constitute a pocket, the active site (A-site), which interacts with the cofactor. Helices associated with the N-/C-termini of LOV light-sensitive domains (Ncap and Ccap) are flexible and necessary for signal transduction [5].

Depending on predominant conditions, there might be structural alterations in the whole molecule of the photoreceptor. In darkness, there is a conserved Cys residue close to an isoalloxazine ring system, which are then non-covalently bound to each other. This Cys residue is found on the Eα-helix in the active, hydrophobic center in a conserved motif. In this dark state, the flavin in the LOV photosensory part of the protein is completely oxidized and hence features the classical absorption characteristics described above, with absorption maximum around 370 and 447 nm. Upon illumination with blue light, the photon energy leads to an excited singlet state (Figure 4). Afterwards, an intersystem crossing event generates a short lived triplet state with absorption maxima at 650 and 715 nm [5]. Subsequently, the Cys, close to the cofactor, forms a covalent adduct with the carbon position 4a of the flavin. The result is a reversible flavin-cysteinyl adduct that absorbs photons at 390 nm. Besides, this conversion probably occurs by forming a neutral biradical intermediate after the reduction of the nitrogen 5 position of FMN by Cys. Besides the covalent linkage between the carbon position 4a of FMN and the active Cys residue, a transfer of electrons related to protonation might be necessary for the production of the light state adduct of LOV photoreceptors [5]. Additionally, exchanged amino acid (aa) residues close to the flavin binding pocket can activate or increase the activity of the respective protein and therefore its signal transduction. The chemistry of LOV light-sensitive protein reactions is quite complicated. Due to that complexity, the recovery from light to dark is variable. There is a thermal driven process back to the ground state and this procedure can take a few milliseconds (>500 ms) up to several hours. In this case, the covalent connection between FMN and Cys has to be broken to complete the photocycle. This reaction can occur by deprotonating the nitrogen position 5 of the chromophore and to add this hydrogen to Cys to protonate it again. In addition, this process runs very slowly or is even not reversible. The conserved Cys is not important for LOV-domains to interact with FMN, but it is essential for the light reaction in LOV photoreceptors. Many different factors can influence the rate of dark state recovery,

especially hydrogen bond donors or reducing agents, depending on pH value or salt concentration [4] [5].

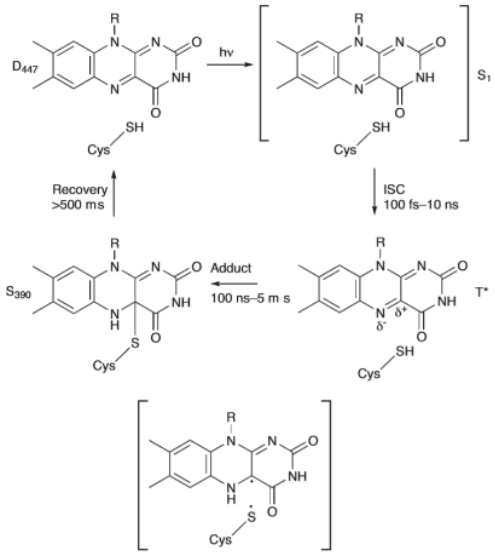

Figure 4. Photocycle in LOV photosensory domains. Photon energy from blue light induction changes the dark state conformation to an excited singlet state and by intersystem crossing to an excited triplet state of FMN. N5 protonation of flavin leads to a neutral, biradical intermediate. This enables formation of a covalent bond between the active Cys residue and C4a of FMN to generate the reversible light state adduct. The lifetimes of the photoadduct range from a few seconds up to several hours. N5 = nitrogen position 5; C4a = carbon position 4a; D447 = dark state, 447 nm absorbance; S1 = singlet state; ISC = intersystem crossing; T* = excited triplet state; S390 = light state, 390 nm absorbance [5]

1.2 Sensor Effector Modularity

In nature, various blue light-sensitive proteins are present. These photoreceptors contain sensor and effector domains, which are covalently linked to each other. Details of signal transduction pathways on a molecular level are mostly unknown, but are essential for signal transduction within cells. Bacterial signaling processes for example feature LOV photosensory domains connected to GGDEF regulators, which enable blue light regulation of the function of diguanylate cylcases (DGCs) [8]. This enzymatic functionality is responsible for the production of the second messenger bis-(3'-5')-cyclic-dimeric-guanosine monophosphate (c-di-GMP) that can only be found in the cytosol of bacterial cells. Such enzymes generate the formation of one c-di-GMP out of two guanosine triphosphate (GTP) molecules, which serve as substrates. The name GGDEF domain originates from a conserved motif for the cellular messenger synthesis in the active center of DGCs. In the active site of each GGDEF domain, the Asp of the consensus sequence to produce c-di-GMP, based on GTP, can be replaced by a Glu in DGCs (GGDEF or even GGEEF motif). Since c-di-GMP levels need to be

tightly controlled in bacterial cells, there are also enzymatic functionalities that degrade c-di-GMP, as for example phosphodiesterases (PDEs) that produce the linear di-GMP (= pGpG) [9] [10].

Generally, blue light modulation of cyclic nucleotides tightly controls the enzymatic activities in flavin-based LOV photoreceptors. [10]. In bacteria, several DGCs as well as PDEs are found, resulting in a complex modulation of cellular c-di-GMP levels in response to various stimuli [11]. High concentrations of the signaling molecule are associated with increased biofilm formation (cytotoxicity). A lower level of the second messenger causes a transition to more motile cell behaviour. C-di-GMP is therefore involved in molecular processes like pathogenesis and the lifestyle of bacterial cells [9] [11].

1.2.1 Well-characterized LOV photosensory proteins in bacteria

In general, 3.5 to 10 % of bacterial genomes, which have been already sequenced, encode for LOV sensor domains. Such photosensory domains are quite important in light-regulated modulation related to signaling processes. The arrangement of such LOV photoreceptors is frequently conserved: N-terminal sensor – linker – effector at the C-terminus [4].

YtvA in gram-positive *Bacillus subtilis*: This photoreceptor consists of a combination of LOV- and sulfate transporter anti-sigma antagonist (STAS) domains as sensor and effector. Its function is the light-dependent modulation of sporulation, including the stress factor σB. In *Bacillus licheniformis*, this pathway can be blocked by blue light illumination. In response to stress, the regulator domain is phosphorylated by a Ser/Thr kinase that modulates the activation of σB via the stressosome. YtvA is linked to GTP-binding *in vitro*, and a modification in the ligand binding site inhibits the light-induced start of σB related transcription *in vivo* [4].

LOV-histidine kinase (HK) in gram-negative *Brucella abortus*: Upon absorption of photons in the blue light range, the corresponding kinase leads to an enhanced rate of autophosphorylation. In addition, the replication in macrophages of mice is also 10 times higher upon illumination of bacterial cells than in darkness. Because of that, this kinase and therefore the conserved Cys residue, which is responsible for the flavin-cysteinyl adduct formation in the photocycle, indicates viral increase in used model organisms. Furthermore, there is almost no recovery of the blue light exposed photosensory protein *in vitro*. Due to the relation of the light exposure to the activity of

the HK *in vivo*, the enzymatic activity will last quite long. Blue light as incoming stimulus might influence cells *in vitro*, and might be proportional to the upregulation of virulence *in vivo*, associated with LOV-HK [4].

LovK (histidine kinase) in *Caulobacter crescentus*: *In vitro*, LovK is phosphorylated by blue light exposure. *In vivo*, LovK is overexpressed upon illumination, resulting in enhanced cell adhesion, which also depends on the conserved Cys residue, forming the flavin-cysteinyl adduct state upon illumination. Moreover, it has been discovered, that both FMN of the LOV-protein and cytosol have closely matching redox potentials. Photosensory proteins require this cofactor oxidized, but its redox state and therefore the acitivity of LovK can be influenced through the cellular redox state. *In vitro*, a reduction of FMN causes a reduction of the light-sensitive modulation of the histidine kinase [4].

LOV-GGDEF-EAL in *Synechococcus elongatus*: The catalytic domains can act as DGC and PDE however the GGDEF domain appears to feature degenerate catalytic motifs because no c-di-GMP formation could be observed. Therefore this photoreceptor contains only blue light-regulated PDE activity *in vitro*, degrading the signaling molecule c-di-GMP. Blue light absorption leads to variable levels of this second messenger, resulting in modulations of the cyanobacterium's functions and physiology. LOV photoreceptors with such a composition are frequently found in bacteria [4].

1.2.2 LOV-diguanylate cyclases

In bacterial LOV photoreceptors, about 20 % are LOV-GGDEF proteins, which are in charge of the modulation and regulation of the unique signaling molecule c-di-GMP [4]. The basic structure of this dynamic system consists of an N-terminal LOV sensor and a C-terminal GGDEF effector, responsible for the catalytic activity of DGCs [8]. The photosensory part acts as the input domain, responding to environmental stimuli, whereas the effector as the output domain is responsible for downstream signaling. The enzymatic activity of DGCs is provided by the GGDEF effector domain [12]. The linker between the photosensory and the output domains is predicted to be composed of supercoiled α- helices that contain a motif of seven amino acid residue repeats (a-g) [13], corresponding to the typical hxxhxxx sequence [8], which can be employed and/or modified in the engineering of light-sensitive proteins. The coiled-coil linker includes both variable hydrophobic amino acids, which are often located at positions

a and d, and hydrophilic amino acids (neutral and charged ones) that can be mostly found at positions e and g. Coiled-coils are very specific structures and any change can cause major, sometimes unpredictable, effects [13] For example, asparagine is responsible for the specificity in supercoiled sequences [10], which can be found at specific positions, affecting the dynamics and the oligomerization behaviour as well as the arrangement and directionality of the involved helices. Therefore, coiled-coil helices that are rich in hydrophobic and Asn residues at the central a and d positions of the heptad repeats specifically build homodimeric conformations [13].

Discovered by X-ray crystallography, one conformation of flavin-based LOV-GGDEF photoreceptors appears to be a tetrameric assembly, composed of two intertwined dimers forming an antiparallel molecule (Figure 5) [8].

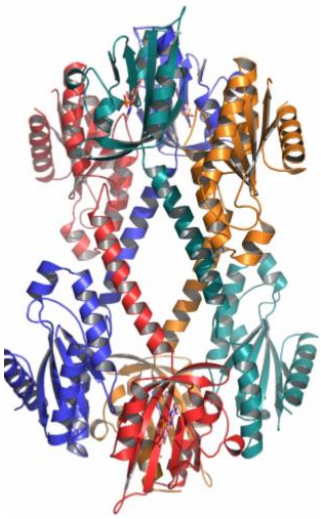

Figure 5. Crystal structure of a tetrameric blue light-sensitive LOV-DGC [8]**.**

Each dimer consists of two LOV-GGDEF protomers. In addition, there is an antiparallel packing of the two LOV dimers [9]. The basic structure of the effector protein is composed of five α-helices (α0- α4), followed by five antiparallel β-strands building one β-sheet. The α1 helix provides residues for phosphate binding. The different secondary elements are linked through a β-hairpin for substrate binding on the side and through a protecting loop (p-loop) on the other side [11].

LOV-diguanylate cyclases need two GTPs as substrate to catalyze the formation of the bacterial signaling molecule c-di-GMP. Each effector binds one single GTP molecule in its binding pocket, inducing the dimerization of two GGDEF domains and

formation of the active site at the dimer interface. This active center positions the GTPs in an optimal orientation for product formation. To sum up, the initiation of the enzymatic activity can occur through the dimeric formation of two GGDEF effector domains. In contrast, product inhibition is a possibility for inactivation of DGCs. In the p-loop of the regulator domain, an RxxD motif is located on the outside of the catalytic GGDEF dimer. This sequence is also called inhibitory site (I-site). If c-di-GMP is bound to this element, the activity of the corresponding DGCs is blocked due to cross-links between individual GGDEF domains that are incompatible with enzymatic turnover [12].

1.3 Project

This study aimed at a better understanding of the regulation of GGDEF activity by LOV photosensory modules. For this purpose, two naturally occurring light-activated diguanylate cyclases (LadCs) were tested. Both of them are composed of a LOV domain as sensor and a GGDEF domain as effector, including the typical function of diguanylate cyclases. These constructs are found in different bacterial organisms. Both wild-type LOV-GGDEF variants were generated with their full-length sequences, derived from sequencing projects of *Serratia fonticola* (*S. fonticola*) and *Salinisphaera shabanensis* (*S. shabanensis*). The corresponding proteins are subsequently abbreviated as *Sf*LadC and *Ss*LadC (Figures 6a and 6b). The accession numbers of the two homologs are WP_021805398 and WP_084623639. In addition, two chimeras of the *S. fonticola* LOV-domain and the GGDEF domain of WP_007419415 from *Idiomarina species* (*Idiomarina sp.)* A28L, which is derived from the red-light activated diguanylate cyclase *Is*PadC [10], were established. In LPadC_A, the LOV domain of *Sf*LadC is connected to the linker region and the GGDEF domain of *Is*PadC (phytochrome-activated diguanylate cyclase) (Figures 6a and 6b). In contrast, LPadC_B is a combination of the LOV domain and the linker region of *Sf*LadC that is linked to the GGDEF domain of *Is*PadC. The difference between these modified variants is therefore the coiled-coil linker, which either belongs to its cognate sensor or effector domain. Moreover, some truncations within the linker heptad as well as specific mutations of *Sf*LadC were produced, but were not characterized in full detail (Figures 6a and 6b).

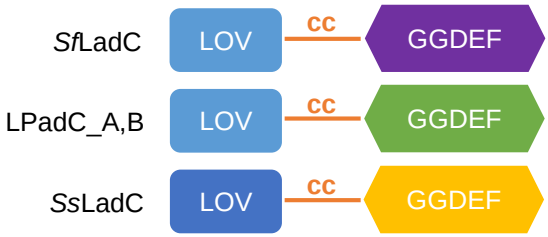

Figure 6a. Domain architecture of generated wild-type and synthetic LadCs. LOV = sensor domain, cc = coiled-coil linker region, GGDEF = effector domain; different domain colors = domains of different organisms

	120-148
SfLadC	RMIIQQQ LRDEHRS LEEMKIH FEQLSIK DG
*Sf*LadC -1	RMIIQQQ LRDEHR LEEMKIH FEQLSIK
*Sf*LadC -3	RMIIQQQ LRDE LEEMKIH FEQLSIK
*Sf*LadC -4	RMIIQQQ LRD LEEMKIH FEQLSIK
*Sf*LadC -7	RMIIQQQ LEEMKIH FEQLSIK
*Sf*LadC H132N	RMIIQQQ LRDENRS LEEMKIH FEQLSIK
*Sf*LadC K139N	RMIIQQQ LRDEHRS LEEMNIH FEQLSIK
*Sf*LadC H132N K139N	RMIIQQQ LRDENRS LEEMNIH FEQLSIK
SsLadC	RLLVDEE LRARNRD LESAKKA LEELVTL
LPadC_B	RMIIQQQ LRDEHRS LEEMKIH FEQLSIK DD
	500-528
LPadC_A	IVADSMQ LNLLNDQ LADANEN LEKLASF

Figure 6b. Sequences of the coiled-coil linker regions of generated wild-type and synthetic LadCs. Blue/violet = differences in truncated versions and specific mutations of *Sf*LadC compared to wild-type *Sf*LadC; red/green = amino acid conservation within linker heptad positions between different generated constructs

The aim of this project was to gain insight into the modularity of such light-sensitive proteins through a detailed functional characterization of these dynamic systems.

After initial synthesis of codon optimized genes and modifications to the expression plasmids by standard tools of molecular biology, individual constructs had to be isolated and identified by sequencing. Then, the protein expression in *Escherichia coli* (*E. coli*) as production organism was performed, followed by the purification of the produced protein. After that, the enzymatic activities of the appropriate diguanylate cyclases were qualitatively and quantitatively measured, under both dark and blue light conditions. Based on the different oligomeric states of LOV proteins, a better understanding of their regulation, depending on functional and structural properties should be established.

At the first place, full-length *Sf*LadC had been partly characterized already. In general, full-length *Sf*LadC shows a typical photocycle, leading to about 100 times more

activation upon illumination compared to dark conditions [8]. After optimization, wild-type *Sf*LadC is quite easy to express and purify, resulting in adequate protein yields that are essential for the crystallization of the desired protein. After x-ray crystallography, the crystal structure of *Sf*LadC appears to be a tetrameric assembly [8]. This construct would be promising for detailed characterization, but its aggregation and oligomerization in solution, identified by size-exclusion chromatography (SEC), might be a problem for further analysis.

Because of that, another similar wild-type construct, *Ss*LadC, was generated for comparison. The isolated protein amount after expression and purification is quite low, but SEC chromatograms show one single peak that corresponds to a dimeric conformation. However, there appears to be no significant light regulation of enzymatic activity present under the conditions of my in vitro assays.

Afterwards, diverse truncated versions and specific mutations of full-length *Sf*LadC were established but not completely characterized. In addition, two chimeras, LPadC_A and LPadC_B, were engineered too. These synthetic proteins are composed of the LOV sensor of *Sf*LadC, but the GGDEF domain as effector originates from *Is*PadC. LPadC_B exhibits significant light-regulation and both chimeras show light-independent dimeric conformation in solution. LPadC_B offers high yields of the desired protein, whereas LPadC_A shows quite low amounts of the protein of interest. Later on, expression and purification procedures of wild-type *Ss*LadC and chimera LPadC_A were optimized, resulting in higher protein yields, which will be important for further functional analyses. Moreover, another quantitative assays associated with c-di-GMP production and therefore enzymatic activities should be performed via high performance liquid chromatography (HPLC) for verification. Finally, crystallization of the chimera LPadC_B should be done by x-ray crystallography, potentially leading to a crystal structure with a dimeric assembly.

2. Results

2.1 Qualitative light regulation of enzymatic activity

After the generation of diverse plasmid constructs, activities of corresponding DGCs were qualitatively determined under light as well as under dark conditions. Before quantification, whole-cell screenings of enzymatic activity in living *E. coli* BL21 (DE3) were performed to get first impressions of light regulation to pre-screen for interesting constructs. In cells featuring DGC activity, c-di-GMP production and therefore cellulose synthesis was detected, leading to red colonies on congo red (CR) plates, while inactive cells remain white.

All tested proteins included pET-GB1_6His_TEV, except for the negative control. This assay was done with wild-type *Sf*LadC as positive control, and also four modified versions of it, showing truncations within one heptad of the linker region (-1, -3, -4, -7) were analyzed. In addition, two specific coiled-coil linker variants of *Sf*LadC were tested. In one of these variants, histidine was replaced by asparagine (H132N), whereas in the other version, lysine was exchanged for asparagine (K139N). Furthermore, enzymatic activities of a double mutant consisting of both specific mutations (H132N K139N) as well as the generated chimeras (LPadC_A and LPadC_B) were screened for light regulation. Two CR plates containing dried drops of all mentioned constructs were incubated at 15 °C overnight. One of these plates was exposed to blue light, while the other one was kept in dark instead.

Generally, there is no significant difference between varying light conditions (Figures 7a and 7b), except for a lower intensity of the whole plate that was illuminated. Maybe, a bleaching effect might be responsible, resulting from the high intensitiy of blue light by the utilized lamp. Under both conditions, the positive control assumes a red colony phenotype, whereas the negative control, using pET-M11 Appa ΔC shows white colouring (Figures 7a and 7b). By illumination, *Sf*LadC -1, -3 and both chimeras seem to be not activated compared to the negative control (Figure 7a). In darkness, truncations -1 and -3 exhibit no DGC activity either (Figure 7b). Moreover, all specific mutations (H132N, K139N and H132N K139N) and truncated *Sf*LadC -4 show intensive red colonies comparable to the positive control under dark conditions. Under light stimulation, truncation -4 and mutant K139N are clearly activated (Figure 7a), followed by modified construct H132N and double mutation H132N K139N, which exhibit lower enzymatic activity. In darkness, LPadC_A as well as LPadC_B feature a

lower level of activity, showing red colonies with low intensity (Figure 7b) comparable to inactivate white phenotypes upon light stimulation (Figure 7a). Truncation -7 appears to be quite active in darkness, but cannot be evaluated by illumination (Figures 7a and 7b) because of a hole in the gel, which complicates interpretation of the result. In addition, full-length *Ss*LadC, cloned into the expression vector pET-M11, was also tested via this DGC activity assay, resulting in almost no activation by blue light stimulation as well as in darkness.

Finally, differences between light and dark conditions associated with enzymatic activities, indicate light-regulated constructs, which are interesting for further characterization. In this case, the *in vivo* screening of DGC activity of several constructs was not significant, because of weak differences between light and dark conditions, even if a light-regulation of about 100-fold activation by blue light illumination of the wild-type *Sf*LadC had been already verified *in vitro*. Because of that, varying conditions related to various temperature and β-D-1-thiogalactopyranoside (IPTG) concentration for induction were tested without any success. Anyway, some of the screened constructs were chosen for continuing analysis (Figures 7a and 7b): full-length *Sf*LadC for verification, truncation -4 because of intensive red colonies under both conditions and the two chimeras (LPadC_A and LPadC_B) because of low but maybe light-regulated activities. Additionally, wild-type *Ss*LadC containing pET-GB1_6His_TEV was also analyzed in comparison with pET-M11 *Ss*LadC.

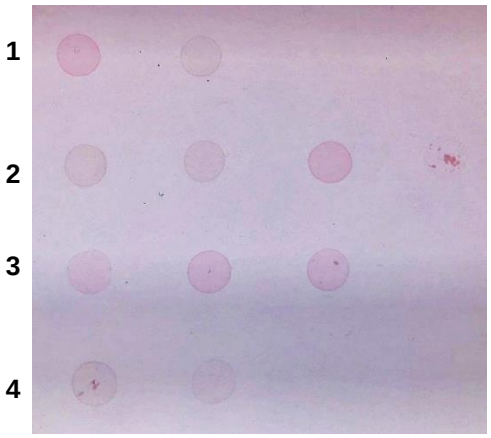

Figure 7a. *In vivo* screening of DGC activity in *E. coli* BL21 (DE3) including different constructs under light conditions. Activation = red colonies; inactivation = white colonies.

1 = wild-type *Sf*LadC**,** Appa ΔC; **2** = truncations -1, -3, -4, -7; **3** = specific mutations H132N, K139N, double mutant H132N K139N; **4** = chimeras LPadC_A and LPadC_B

Figure 7b. *In vivo* screening of DGC activity in *E. coli* BL21 (DE3) including different constructs under dark conditions. Activation = red colonies; inactivation = white colonies.

1 = wild-type *Sf*LadC**,** Appa ΔC; **2** = truncations -1, -3, -4, -7; **3** = specific mutations H132N, K139N, double mutant H132N K139N; **4** = chimeras LPadC_A and LPadC_B

2.2 Expression and contamination

At the beginning of qualitative and quantitative analyses on the molecular level of proteins, it is very important to avoid contaminations of the purified proteins. On the one hand, contaminations lead to falsified concentration estimation. On the other hand, copurified proteins might result in unexpected reaction products. Some established proteins are more highly expressed than others, mostly affecting final protein amounts but also purities. During the expression and purification of desired constructs, several samples were taken for quality control via sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE).

2.2.1 Two wild-type *Ss*LadCs

Generated versions associated with *Ss*LadC vary in the length of the His-Tags, containing six or nine histidine residues. Purified GB1 6His TEV *Ss*LadC is expected to have a molecular weight of about 47.4 kDa, whereas GB1 9His TEV *Ss*LadC has a size of about 47.8 kDa. After removal of the tags, both proteins are about 38.1 kDa.
Before the lysis of the cells, the variant with only six histidines shows a distinct and quite intensive band at the relevant size between 43 kDa and 66 kDa compared to the standard, indicating in an adequate protein expression (Figure 8a). After mechanical sonication, the sample with *E. coli* BL21 (DE3) including the 6His version appears to be overloaded, similar to the appropriate flow through after the first Ni^{2+} column, resulting in a smear within the whole lane. In the case of both constructs, GB1 6His TEV *Ss*LadC and GB1 9His TEV *Ss*LadC, some of the desired protein might get lost after the lysis (Figures 8a and 9), because of the formation of inclusion bodies. In comparison, washing procedure of GB1 6His TEV *Ss*LadC (Figure 8a) with a salt concentration of 60 mM imidazole removes more unspecifically bound proteins but also a certain amount of the expressed protein, resulting in quite a loss of the desired GB1 6His TEV *Ss*LadC. In addition, both cleaning steps show equal results, so washing should have been repeated more often or with an enhanced salt concentration. For increasing the affinity of SsLadC to the Ni^{2+} matrix and allowing more stringent washing conditions, the His-Tag had been extended to nine histidine residues. This allows washing steps of GB1 9His TEV *Ss*LadC (Figure 9), which do not show removal of the protein of interest but effectively removes contaminating species. In both constructs, bands of desired protein are visible at appropriate sizes upon elution with high (250 mM) imidazole concentrations (Figures 8a and 9). Protein

yields seem to be quite high, however GB1 6His TEV *Ss*LadC features much more unspecific bands due to the less stringent washing conditions. Later on, this construct was also purified via SEC to get rid of the contamination, resulting in less unspecific bands (Figure 8b). After overnight dialysis in presence of TEV protease and a second Ni^{2+} chromatography step, *Ss*LadC originating from the 6His construct shows a new protein band between 29 kDa and 43 kDa compared to the standard but with lower intensity (Figure 8a). There was a loss of protein because of insufficient TEV protease or incomplete digestion. For verification, the dialysate of *Ss*LadC coming from the 9His version was also separated via SDS-PAGE (Figure 9), looking similar to the elution fraction and was quantitatively verified by NanoDrop.

In comparison, protein amounts of both generated *Ss*LadC constructs out of 1 l protein expression each, were almost identical after the second Ni^{2+} column purification. The outcome of *Ss*LadC coming from the 6His variant was about 1.9 mg with quite a lot of contamination, whereas the outcome of the other version originating from the 9His construct was about 1.8 mg with less unspecific components (Figures 8a and 9). After ÄKTA purification, the yield of *Ss*LadC coming from the 9His version was about 0.44 mg totally. By the way, 9 mg of 6His originating *Ss*LadC and 5.2 mg of 9His originating *Ss*LadC were present after the first Ni^{2+} purification.

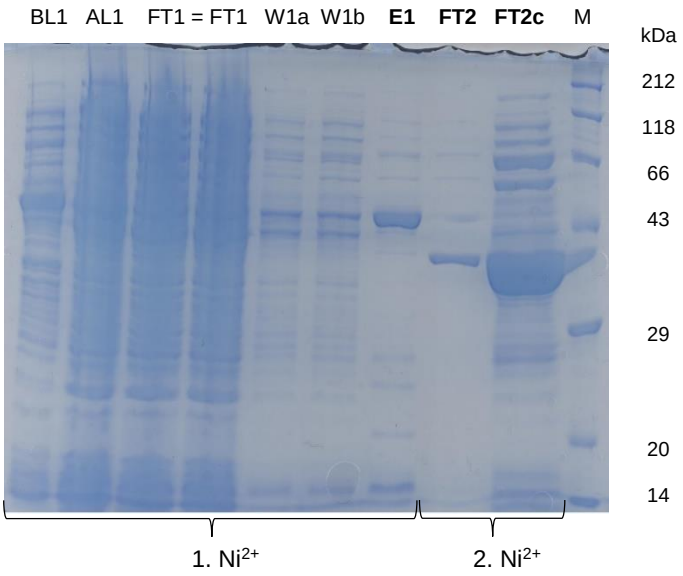

Figure 8a. SDS-Page of full-length GB1 6His TEV *Ss*LadC and *Ss*LadC. Several fractions during different purification steps of the first and second affinity chromatography, using Ni^{2+} columns were analyzed. GB1 6His TEV *Ss*LadC (**E1**) is about 47.4 kDa, whereas *Ss*LadC (**FT2**, **FT2c**) is about 38.1 kDa.
BL1 = before lysis; AL1 = after lysis; FT1 = flow through; W1a, 1b = wash; E1 = elution (protein); FT2 = flow through (protein); FT2c = flow through (concentrated); M = marker/standard; 1./2. Ni^{2+} = 1./2. Ni^{2+} column

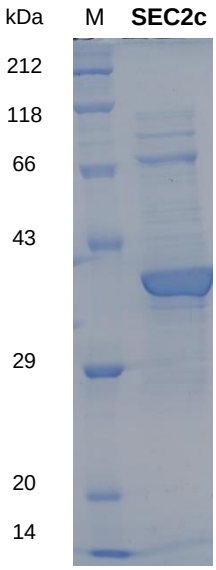

Figure 8b. SDS-Page of full-length *Ss*LadC coming from the 6His variant. Fractions during SEC chromatography containing *Ss*LadC were pooled. *Ss*LadC (**SEC2c**) is about 38.1 kDa.
M = marker/standard; SEC2c = purification via SEC (concentrated)

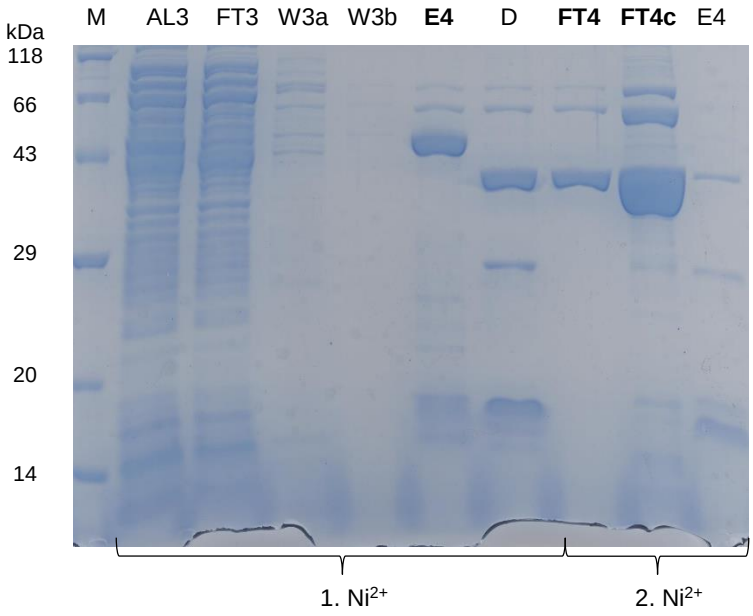

Figure 9. SDS-Page of full-length GB1 9His TEV *Ss*LadC and *Ss*LadC. Several fractions during different purification steps of the first and second affinity chromatography, using Ni^{2+} columns were analyzed. GB1 9His TEV *Ss*LadC (**E1**) is about 47.4 kDa, whereas *Ss*LadC (**FT2**, **FT2c**) is about 38.1 kDa.
BL3 = before lysis; AL3 = after lysis; FT3 = flow through; W3a, 3b = wash; E3 = elution (protein); FT4 = flow through (protein); FT4c = flow through (concentrated); M = marker/standard; 1./2. Ni^{2+} = 1./2. Ni^{2+} column

2.2.2 Full-length *Sf*LadC versus *Sf*LadC -4 truncation

The protein construct *Sf*LadC inclusive GB1, 6His and TEV site is about 44.4 kDa, whereas the protein's size decreases to about 35.1 kDa without the solubility tag after digestion with TEV protease. Another tested variant of *Sf*LadC is a -4 truncation, because it was one modification of *Sf*LadC, which seemed to be light-regulated after a qualitative screening was performed. In this protein, four amino acids within the heptad of the linker between LOV sensor and GGDEF effector had been removed. The version, which contains the TEV site is about 43.9 kDa. In contrast, the variant without TEV site after overnight dialysis exhibits a size of about 34.6 kDa.

In both cases, the protein of interest was expressed (Figures 10 and 11), showing an explicit band around 43 kDa compared to the marker. After mechanical lysis by sonication, bacterial cells should have been destroyed. According to this, an overload was observed in both samples, resulting in a smear and not showing accurate bands. After the first Ni^{2+} chromatography, flow throughs of GB1 6His TEV *Sf*LadC and GB1 6His TEV *Sf*LadC -4 look similar. This was expected, because of same imidazole concentrations (30 mM). Furthermore, only one washing step was done during the first Ni^{2+} purification of GB1 6His TEV *Sf*LadC (Figure 10) in comparison to the engineered variant, where this step was repeated (Figure 11). During this processes, further contaminations should have been removed with an increased 60 mM imidazole concentration. In comparison, the first wash of GB1 6His TEV *Sf*LadC -4 appears to be more efficient (Figure 11), because of the elimination of many unspecific components. This observation could be verified by the purity of following fractions, containing soluble protein each. Elution of the soluble protein occurred by 250 mM imidazole respectively (Figures 10 and 11). The full-length *Sf*LadC exhibits more contamination in comparison to the engineered variant. Furthermore, the protein band of naturally occurring *Sf*LadC is much more intensive compare to *Sf*LadC -4, leading to a higher concentration and therefore higher protein amount of wild-type *Sf*LadC. After digestion with TEV protease upon overnight dialysis, essential bands of both generated variants were shifted to a position between 29 kDa and 43 kDa associated with appropriate marker bands. Moreover, truncation -4 offers differences in the intensities of desired protein bands before and after dialysis (Figure 11), resulting in a dramatic loss of protein between these steps.

At the end, final protein yields of about 9.2 mg of full-length *Sf*LadC and about 1.1 mg of *Sf*LadC -4 out of 12 l protein expression each, were received and compared.

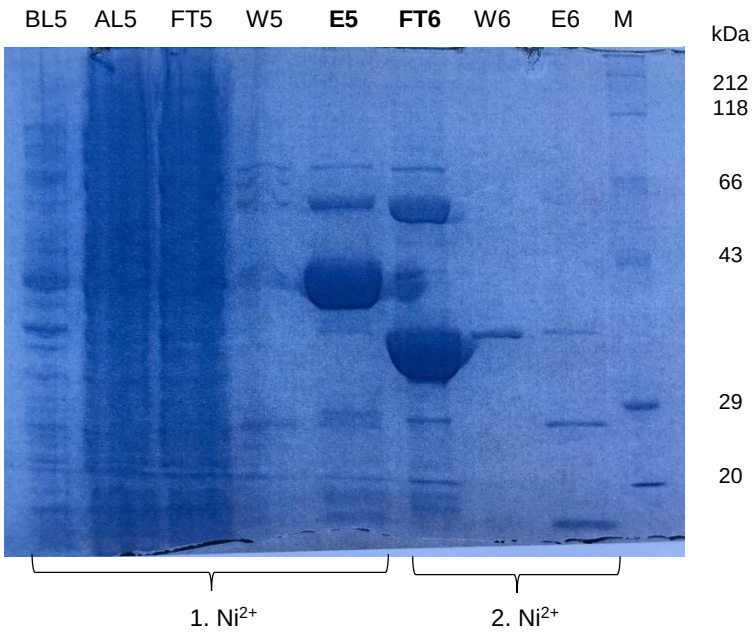

Figure 10. SDS-Page of full-length GB1 6His TEV *Sf*LadC and *Sf*LadC. Several fractions during different purification steps of the first and second affinity chromatography, using Ni^{2+} columns were analyzed. Wild-type GB1 6His TEV *Sf*LadC (**E5**) is about 44 kDa, whereas *Sf*LadC (**FT6**) is about 35 kDa.
BL5 = before lysis; AL5 = after lysis; FT5 = flow through; W5, 6 = Wash; E5 = elution (protein); FT6 = flow through (protein); E2 = elution; M = marker/standard; 1./2 .Ni^{2+} = 1./2. Ni^{2+} column

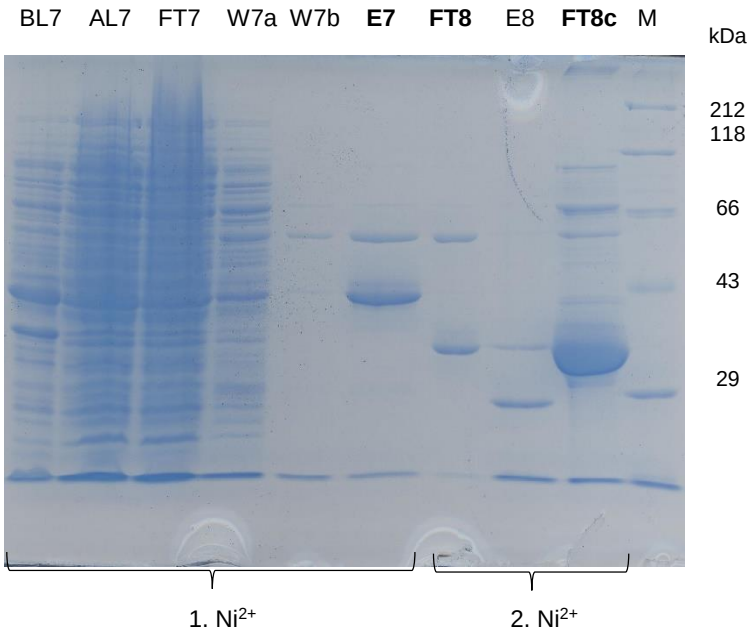

Figure 11. SDS-Page of GB1 6His TEV *Sf*LadC -4 truncation and *Sf*LadC -4 truncation. Several fractions during different purification steps of the first and second affinity chromatography, using Ni^{2+} columns were analyzed. GB1 6His TEV *Sf*LadC -4 (**E7**) is about 44 kDa, whereas *Sf*LadC -4 (**FT8**, **FT8c**) is about 35 kDa.
BL7 = before lysis; AL7 = after lysis; FT7 = flow through; W7a, 7b = wash; E7 = elution (protein); FT8 = flow through (protein); E8 = elution; FT8c = flow through (concentrated); M = marker/standard; 1./2. Ni^{2+} = 1./2. Ni^{2+} column

2.2.3 Chimeras LPadC_A and LPadC_B

Two different synthetic chimeras were purified, LPadC_A and LPadC_B. Both of them differ in sequence lengths with as well as without GB1, 6His and TEV sites. The A variant is about 43.9 kDa prior to TEV digestion, whereas version B has a size of about 44.3 kDa under the same conditions. After the second affinity chromatography, construct A is reduced to 34.6 kDa and chimera LPadC_B exhibits a molecular weight of about 35 kDa.

In variant A and B, the relevant protein is visible at about 43 kDa, which demonstrates protein expression (Figure 12). In version B, the desired band is also available at about 43 kDa but is significantly increased in intensity (Figure 13). After the lysis of *E. coli* BL21 (DE3) cells including chimera B, the sample shows an extreme overload, resulting in a smear within the whole lane. After the first affinity chromatography, flow throughs of GB1 6His TEV LPadC_A and GB1 6His TEV LPadC_B containing the soluble protein, look similar as expected (Figures 12 and 13). Thereafter, two washing steps of both chimeras were performed respectively. These steps had been done by increasing the salt concentration from 30 mM to 60 mM imidazole. The first wash of GB1 6His TEV LPadC_A was very efficient, removing multiple unspecific proteins (Figure 12). In addition, there was a loss of certain amount of the desired protein after washing. In chimera B, cleaning steps did not remove much contamination, but leading to quite a high loss of protein (Figure 13). Both constructs, especially version B, offer elution fractions including an intensive band of the protein of interest (Figures 12 and 13), indicative of a high protein yield. During overnight dialysis, between the first and second affinity chromatography step, protein amounts of both chimeras had been decreased. In LPadC_B, the amount of the desired protein decreases from about 31.5 mg to 17.6 mg out of 6 l expressed protein, when TEV cleavage occurred. In contrast, LPadC_A shows a decrease from about 8.6 mg to about 1.33 mg out of 6 l protein expression, involving TEV restriction. The reasons for this loss could be the above mentioned problems like partial cleavage or not enough TEV protease, because a faint band is visible at 43 kDa (Figure 12). Additionally, precipitation could have happened too.

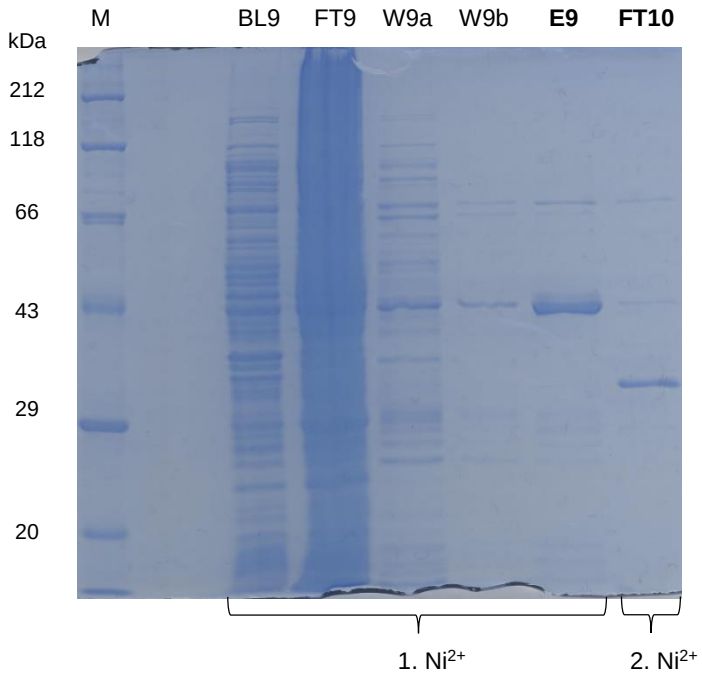

Figure 12. SDS-Page of the chimera LPadC_A. Several fractions during different purification steps of the first and second affinity chromatography, using Ni^{2+} columns were collected. GB1 6His TEV LPadC_A (**E9**) is about 43.9 kDa, whereas LPadC_A (**FT10**) is about 34.6 kDa.
M = Marker/Standard; BL9 = before lysis; FT9 = flow through; W9a, 9b = wash; E9 = elution (protein); FT10 = flow through (protein); 1./2. Ni^{2+} = 1./2. Ni^{2+} column

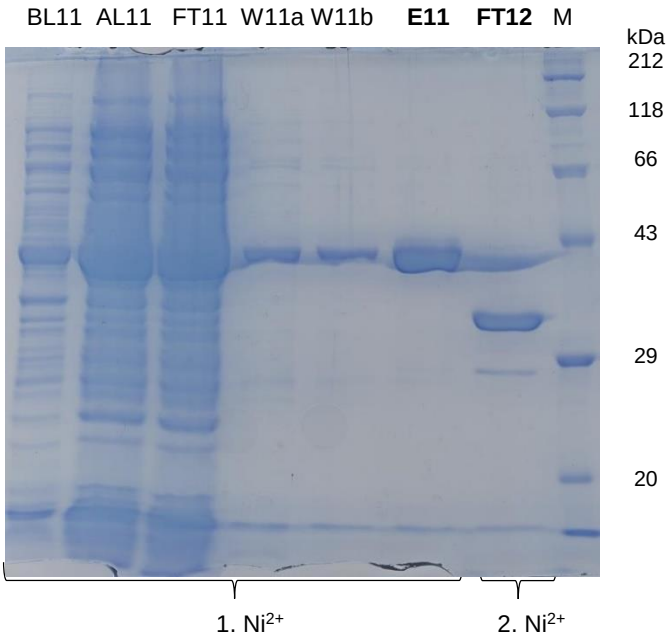

Figure 13. SDS-Page of the chimera GB1 6His TEV LPadC_B and LPadC_B. Several fractions during different purification steps of the first and second affinity chromatography, using Ni^{2+} columns were collected. GB1 6His TEV LPadC_B (**E11**) is about 44.3 kDa, whereas LPadC_B (**FT12**) is about 35 kDa.
BL11 = before lysis; AL11 = after lysis; FT11 = flow through; W11a, 11b = wash; E11 = elution (protein); FT12 = flow through (protein); M = marker/standard; 1./2. Ni^{2+} = 1./2. Ni^{2+} column

2.3 Structural conformations – light regulation

In the last step of protein purification, SEC is utilized to obtain some information about naturally occurring oligomeric states of the generated constructs. Additionally, there is the ability for getting rid of unwanted contaminations or even unspecific aggregates of the proteins of interest by fractionation, leading to almost pure proteins. In addition, diverse illumination tests are performed with this method to characterize the influence of light on the conformational stability of synthesized variants. The molecular weight is related to the elution volume in ml, while the protein amount is proportional to the absorbance in milli absorption units (AU).

2.3.1 Spectral characterization of natural occurring DGCs

In darkness, the absorption maximum of flavin-based light-sensitive proteins is at about 445 nm, showing the typical flavin double peak (“shoulder”). In addition, the absorption maximum of LOV photosensory domains upon blue light illumination is at 390 nm, exhibiting one single peak. Furthermore, the recovery of full-length *Sf*LadC (Figure 14) back to the dark state after exposure to blue light takes about 2 h. In contrast, the recovery of wild-type *Ss*LadC back to the ground state after blue light illumination takes over 7 h (Figure 15).

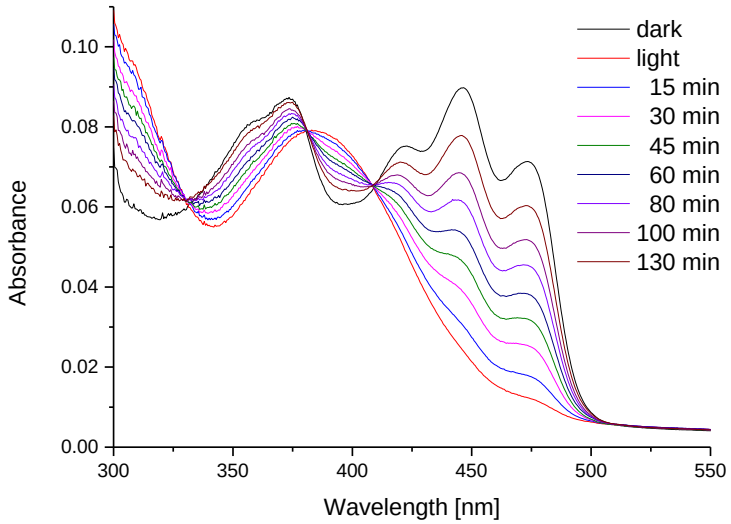

Figure 14. Spectral characterization of wild-type *Sf*LadC originating from GB1 6His TEV *Sf*LadC. After the second affinity chromatography using Ni^{2+} columns, the recovery of *Sf*LadC from light to dark was characterized by photometric measurement and takes about 2 h.

X-axis = wavelength in nm; y-axis = absorbance in AU; 280 nm = absorption maximum of proteins; 445 nm = absorption maximum of flavin-based light-sensitive proteins in darkness; 390 nm = absorption maximum of flavin-based light-sensitive proteins in light

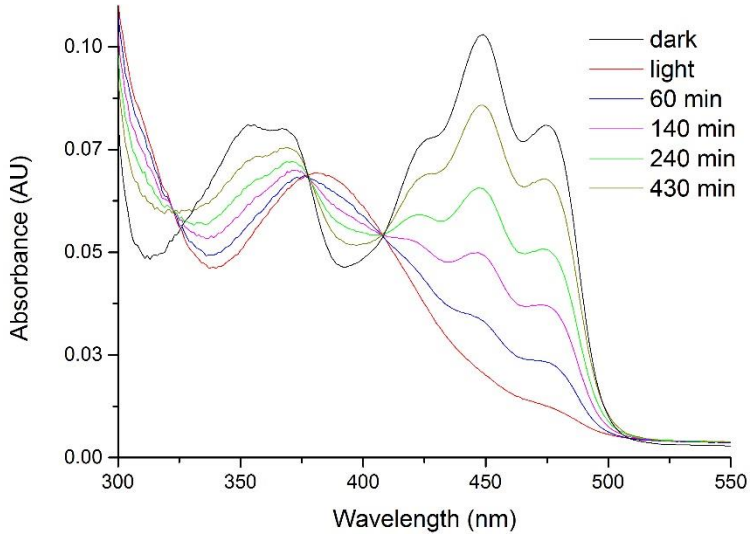

Figure 15. Spectral characterization of wild-type *Ss*LadC originating from GB1 9His TEV *Sf*LadC. After the second affinity chromatography using Ni^{2+} columns, the recovery of *Ss*LadC from light to dark was characterized by photometric measurement and takes over 7 h.
X-axis = wavelength in nm; y-axis = absorbance in AU; 280 nm = absorption maximum of proteins; 445 nm = absorption maximum of flavin-based light-sensitive proteins in darkness; 390 nm = absorption maximum of flavin-based light-sensitive proteins in light

2.3.2 Aggregation of full-length *Sf*LadC

Upon size-exclusion chromatography, wild-type *Sf*LadC was purified using the big Superdex 200 16/600 PG column, revealing pronounced aggregation and oligomerization within the whole sample. The measurement was specifically taken at an absorbance of 280 nm for all proteins and at 445 nm for flavin-based photoreceptors under dark conditions (Figure 14). Compared to the retention time and elution volume of calibration proteins associated with specific molecular weights, the elution volume indicates a mixture of aggregates, different oligomeric, tetrameric and dimeric assemblies (Figure 16). The first peak, beginning at a volume of about 40 ml shows soluble aggregates of *Sf*LadC. Because of that, the yield of protein after SEC is correspondingly reduced. Moreover, octameric conformation of *Sf*LadC seems to be probable between fractions 1 to 10 (Figure 16). Additionally, fractions 11 to 39 possibly contains a combination of both tetramers and dimers one after another.
Finally, the truncated *Sf*LadC -4 was also tested in the dark, resulting in much more aggregation compared to full-length *Sf*LadC.

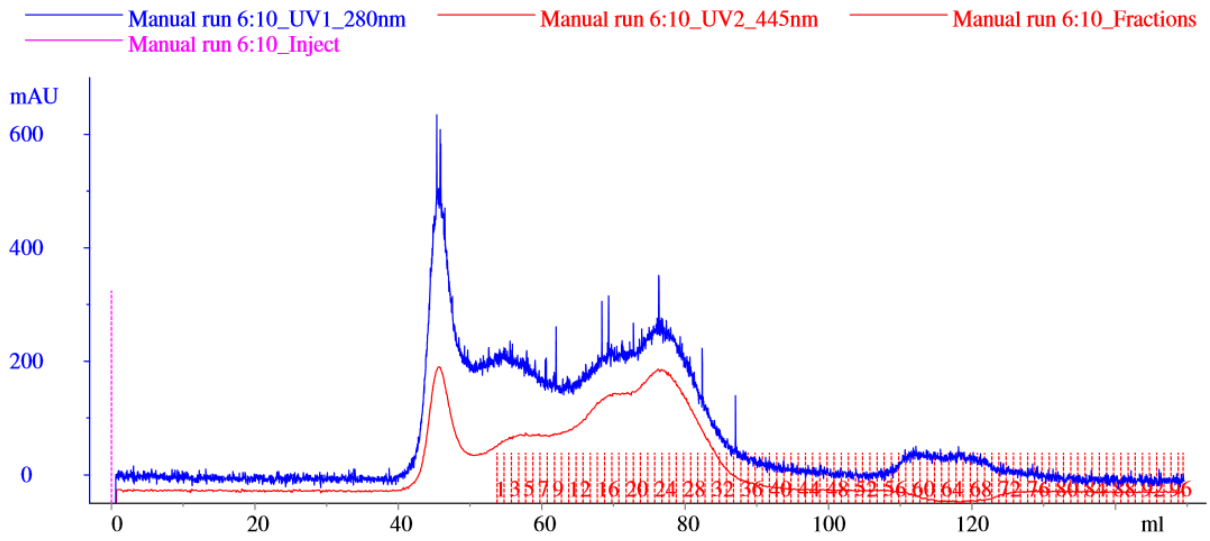

Figure 16. Size-exclusion chromatography of wild-type *Sf*LadC using the 16/600 Superdex 200 PG column under dark conditions. After the second affinity chromatography using Ni^{2+} columns, proteins were purified with SEC to get information about light-regulated formation and also get rid of contamination/oligomerization. Full-length *Sf*LadC includes aggregates as well as oligomeric, tetrameric and dimeric states in the dark.

X-axis = elution volume in ml; y-axis = absorbance in mAU; 280 nm = absorption maximum of all proteins; 445 nm = absorption maximum of flavin-based light-sensitive proteins in darkness

For direct comparison, light-dependent testing of wild-type *Sf*LadC was repeated under both dark and light conditions, to get information about structural conformations that are possibly light-regulated. On the one hand, the analysis of full-length *Sf*LadC in the dark (Figure 17) offers similar results as mentioned and observed above (Figure 16). The only difference is the utilized protein concentration that was not as high as before, and therefore the Superdex 200 increase 10/300 GL column was used here. On the other hand, same concentrated *Sf*LadC, which was measured by SEC upon blue light illumination during the whole run (Figure 18), shows more aggregates and oligomers instead of tetrameric or even dimeric assemblies compared to dark conditions. Because of that, the formation of oligomeric states of wild-type *Sf*LadC are strongly influenced and regulated by blue light exposure. Finally, the absorption maximum at 390 nm is higher under light conditions, while the absorption maximum at 445 nm is higher in the dark (Figure 14).

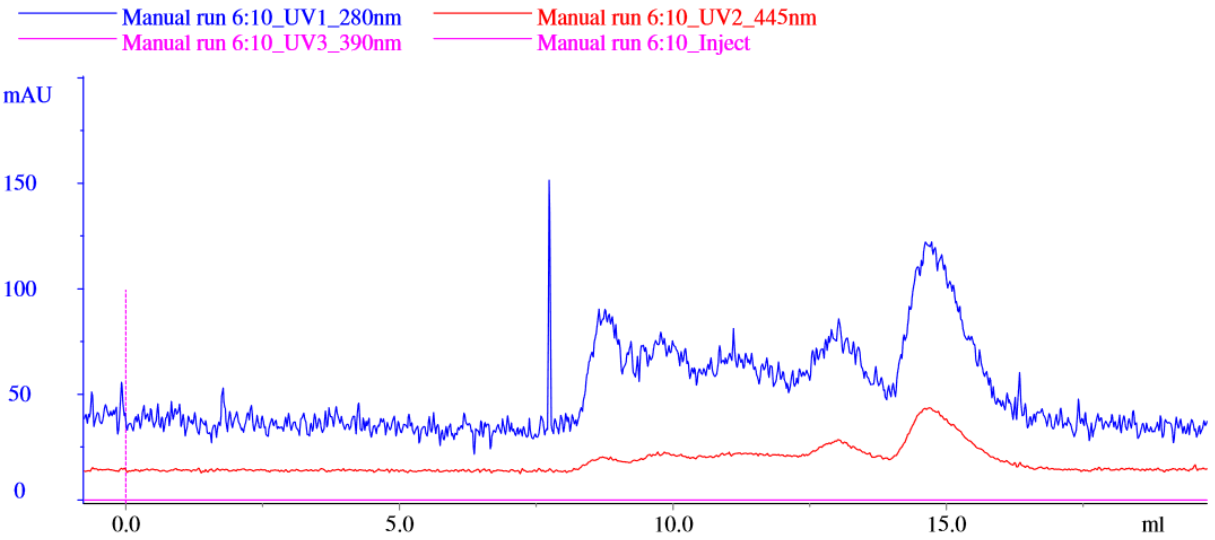

Figure 17. Size-exclusion chromatography of *Sf*LadC originating from GB1 6His TEV *Sf*LadC using the Superdex 200 increase 10/300 GL column under dark conditions. After the second affinity chromatography using Ni^{2+} columns, proteins were purified with SEC to get information about light-regulated formation and also get rid of contamination/oligomerization. *Sf*LadC exhibits dimeric conformation under dark conditions.
X-axis = elution volume in ml; y-axis = absorbance in mAU; 280 nm = absorption maximum of all proteins; 445 nm = absorption maximum of flavin-based light-sensitive proteins in darkness; 390 nm = absorption maximum of flavin-based light-sensitive proteins in light

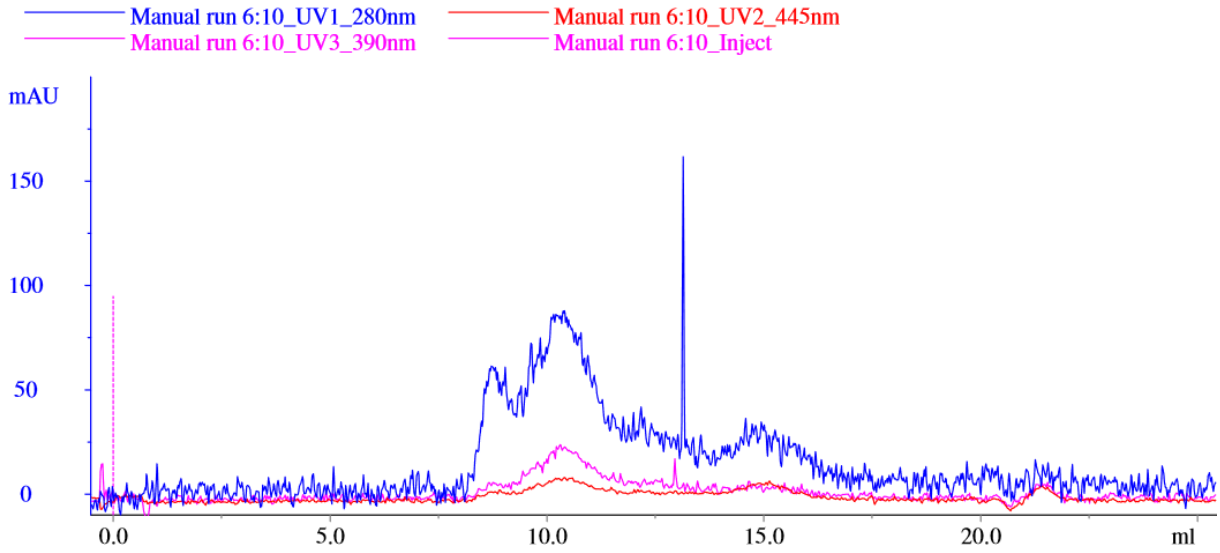

Figure 18. Size-exclusion chromatography of *Sf*LadC originating from GB1 6His TEV *Sf*LadC using the Superdex 200 increase 10/300 GL column under light conditions. After the second affinity chromatography using Ni^{2+} columns, proteins were purified with SEC to get information about light-regulated formation and also get rid of contamination/oligomerization. *Sf*LadC exhibits dimeric conformation under light conditions.
X-axis = elution volume in ml; y-axis = absorbance in mAU; 280 nm = absorption maximum of all proteins; 445 nm = absorption maximum of flavin-based light-sensitive proteins in darkness; 390 nm = absorption maximum of flavin-based light-sensitive proteins in light

At the end, I tested the influence of buffer systems on the aggregation and oligomerization tendency of *Sf*LadC by changing conditions of the SEC runs. Instead of Tris/HCl pH 8.0, MES (2-(N-morpholino) ethanesulfonic acid) pH 6.0 as well as CHES (N-Cyclohexyl-2-aminoethanesulfonic acid) pH 9.0 were used as alternative buffer systems. In addition, I also tested the influence of the salt concentration by including high [1 M] and also low concentrations of NaCl [50 mM]. In conclusion, no effect of pH and salt concentration was seen neither under light nor under dark conditions. Because these observations, we stopped to work with this construct, because the functional interpretation would have been too difficult.

2.3.3 Contamination of wild-type *Ss*LadCs

As already mentioned, two different variants of *Ss*LadC were generated, varying in the length of the N-terminal His-Tag. These two naturally occurring DGCs were produced, because of much aggregation and the formation of several oligomeric states of the similar wild-type *Sf*LadC, identified after SEC.

Because of increased affinity to the Ni^{2+} columns and less visible contamination, *Ss*LadC originating from GB1 9His TEV *Ss*LadC was used for purification by SEC under dark conditions, followed by quantitative light-regulated measurements of enzymatic activities via HPLC. In contrast, SsLadC coming from GB1 6His TEV *Ss*LadC was utilized for light-dependent testing by SEC, to get information about this structural conformation of this construct under dark as well as under light conditions.

First of all, *Ss*LadC coming from the 9His version was purified by SEC at 4 °C in darkness. The resulting chromatogram shows one single protein peak at about 14 ml elution volume (Figure 19) featuring characteristic absorption for a flavoprotein at 445, 390 and 280 nm (Figure 15). Because of dark conditions, the absorbance at 445 nm is typically higher than the absorbance at 390 nm. Compared to previously purified chimeras LPadC_A as well as LPadC_B, where the oligomeric assembly was verified by MALS to be dimeric (Figure S2), *Ss*LadC also features a dimeric assembly in the dark. At the beginning of fractionation coupled measurement, much contamination of irrelevant proteins was detected at 280 nm (Figure 19), even though only few unspecific bands were available in the appropriate SDS gel (Figure 9). Collection tubes 1-5 include only little quantity of the relevant construct, based on the absorbance at 445 nm (Figure 19). However, fractions 6 and 7 that correspond to the eluting of dimers were then collected, concentrated and stored in liquid nitrogen at -80 °C for further

analysis, leading to pure protein with consistent conformation. Because of separation of undesired contaminants, there is a lower concentration and quantitative yield of the construct.

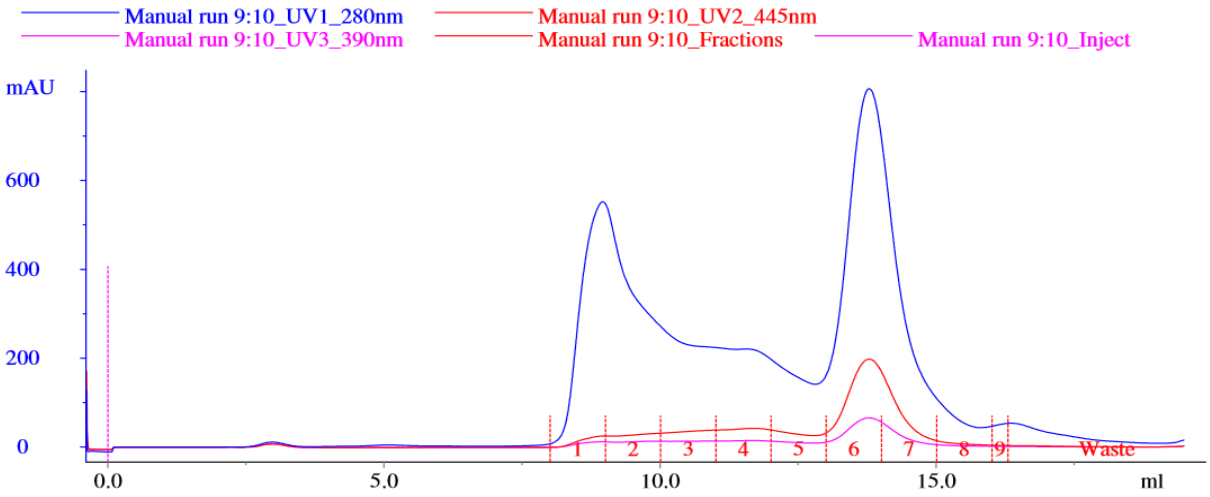

Figure 19. Size-exclusion chromatography of *Ss*LadC originating from GB1 9His TEV *Ss*LadC using the Superdex 200 increase 10/300 GL column under dark conditions. After the second affinity chromatography using Ni^{2+} columns, proteins were purified with SEC to get information about light-regulated formation and also get rid of contamination/oligomerization. *Ss*LadC exhibits dimeric conformation under dark conditions.
X-axis = elution volume in ml; y-axis = absorbance in mAU; 280 nm = absorption maximum of all proteins; 445 nm = absorption maximum of flavin-based light-sensitive proteins in darkness; 390 nm = absorption maximum of flavin-based light-sensitive proteins in light

In addition, possible light regulation associated with structural conformation was verified under both dark and light conditions. These light-dependent tests for further quantitative analyses were performed with wild-type *Ss*LadC, which comes from the 6His variant, because of low protein yield of *Ss*LadC that originates from the 9His version. At first, one aliquot was analysed in darkness, whereas another sample was preilluminated before injection for one minute and then exposed to blue light during the whole measurement (Figures 20 and 21). Both chromatograms represent almost the same curve, featuring one peak at about 14 ml. This indicates a light-independent dimeric conformation. In darkness, the absorption maximum at 445 nm is higher, while it is higher at 390 nm upon illumination (Figures 15, 20 and 21).

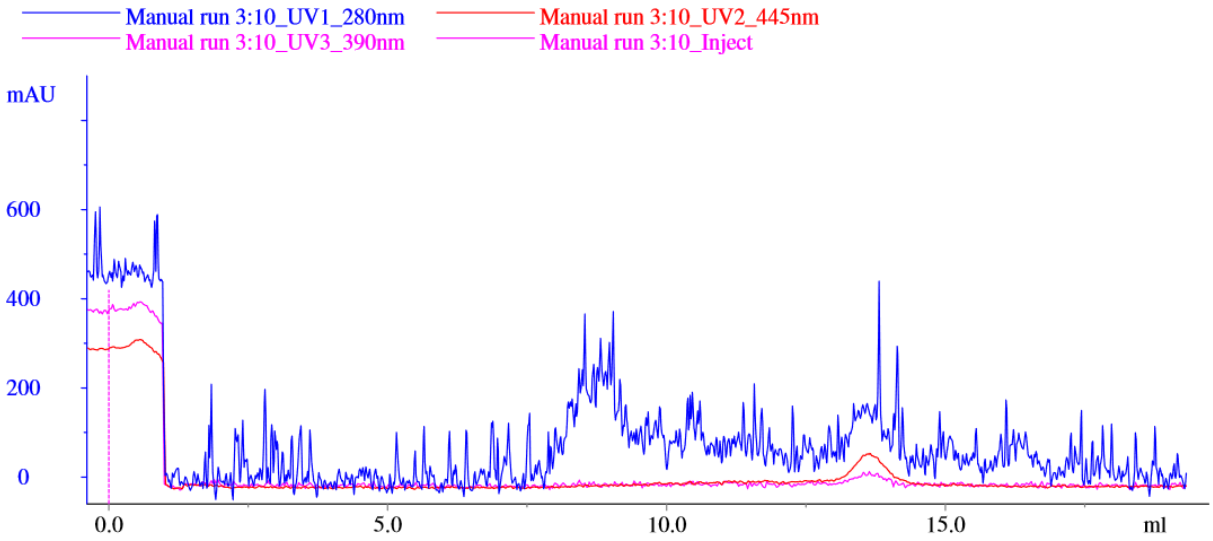

Figure 20. Size-exclusion chromatography of *Ss*LadC originating from GB1 6His TEV *Ss*LadC using the Superdex 200 increase 10/300 GL column under dark conditions. After the second affinity chromatography using Ni^{2+} columns, proteins were purified with SEC to get information about light-regulated formation and also get rid of contamination/oligomerization. *Ss*LadC exhibits dimeric conformation under dark conditions.
X-axis = elution volume in ml; y-axis = absorbance in mAU; 280 nm = absorption maximum of all proteins; 445 nm = absorption maximum of flavin-based light-sensitive proteins in darkness; 390 nm = absorption maximum of flavin-based light-sensitive proteins in light

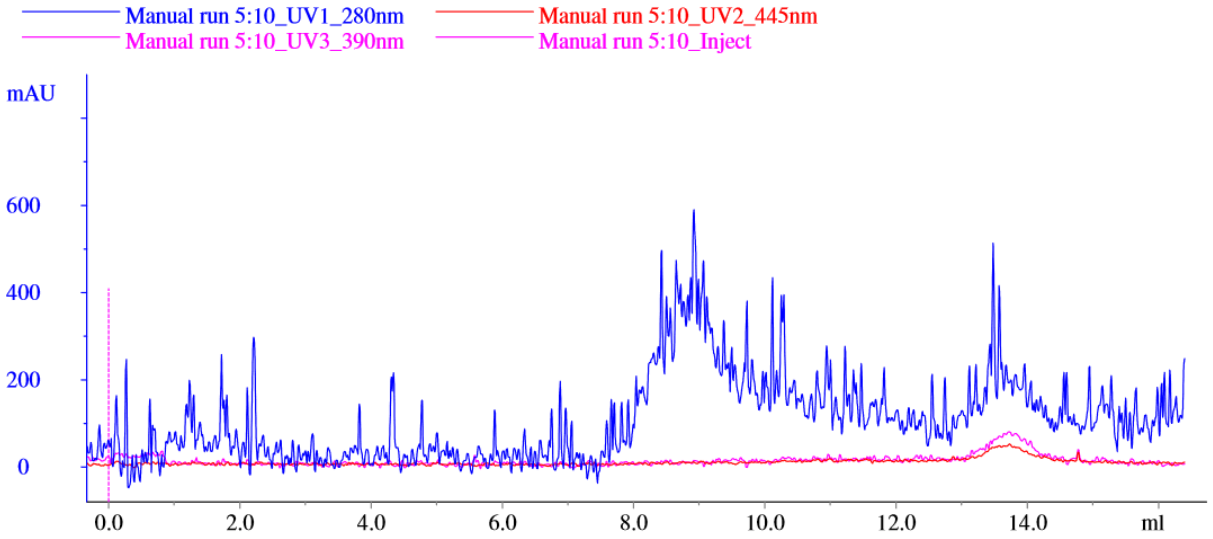

Figure 21. Size-exclusion chromatography of *Ss*LadC originating from GB1 6His TEV *Ss*LadC using the Superdex 200 increase 10/300 GL column under light conditions. After the second affinity chromatography using Ni^{2+} columns, proteins were purified with SEC to get information about light-regulated formation and also get rid of contamination/oligomerization. *Ss*LadC exhibits dimeric conformation under light conditions.
X-axis = elution volume in ml; y-axis = absorbance in mAU; 280 nm = absorption maximum of all proteins; 445 nm = absorption maximum of flavin-based light-sensitive proteins in darkness; 390 nm = absorption maximum of flavin-based light-sensitive proteins in light

For visible verification, collected fractions containing the important peak of *Ss*LadC at about 14 ml were pooled and concentrated. Thereafter, the sample was measured via size-exclusion chromatography by exposure to blue light during the whole SEC run again. The resulting chromatogram of this second run shows the dimeric state without any contamination, measured at 280 nm (Figures 15 and 22). Finally, the absorption maximum at 390 nm is characteristically higher than at 445 nm.

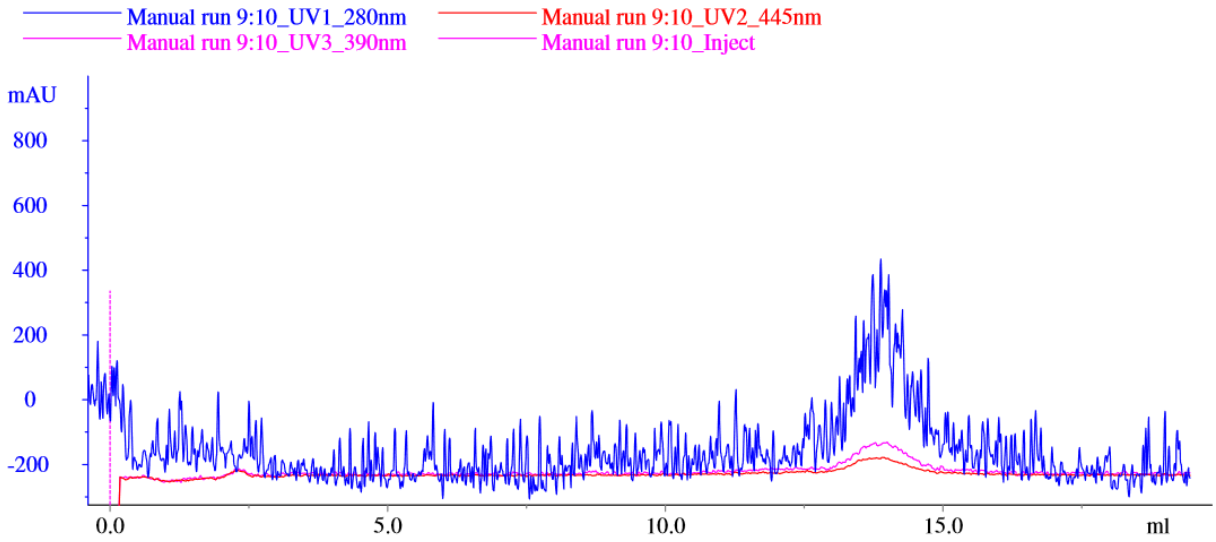

Figure 22. Size-exclusion chromatography of *Ss*LadC originating from GB1 6His TEV *Ss*LadC using the Superdex 200 increase 10/300 GL column under light conditions. After the second affinity chromatography using Ni^{2+} columns, proteins were purified with SEC to get information about light-regulated formation and also get rid of contamination/oligomerization. *Ss*LadC exhibits dimeric conformation under light conditions.
X-axis = elution volume in ml; y-axis = absorbance in mAU; 280 nm = absorption maximum of all proteins; 445 nm = absorption maximum of flavin-based light-sensitive proteins in darkness; 390 nm = absorption maximum of flavin-based light-sensitive proteins in light

2.3.4 Conformation of chimeras LPadC_B and LPadC_A

Both constructs, full-length *Sf*LadC and also truncated *Sf*LadC -4, exhibit much aggregation and a tendency to oligomerize, which might be a problem for further analysis. Anyway, *Sf*LadC had been partly characterized, offering high enzymatic activity under light regulation. Because of that, the chimeras LPadC_A and LPadC_B were generated, composed of combinations of wild-type *Sf*LadC that contain a different GGDEF domain coming from *Is*PadC. These engineered variants were used to test the influence of the coiled-coil linker and the GGDEF domain on the observed oligomerization. After expression and purification via affinity chromatography, the effect of light on both chimeras was qualitatively tested by SEC.

Generally, three aliquots of the variant LPadC_B were measured under different conditions (Figure 23). The first sample was tested in darkness, the second one was exposed to blue light for one minute before injection, and the third aliquot of this construct was illuminated for one minute before injection and also during the whole process of the SEC run. The result was a single peak with comparable elution volume at about 14 ml under all tested conditions (Figure 23). The conclusion of this procedure is, that this chimera exhibits a stable conformation under all light conditions. The only difference between individual peaks is the absorption maximum at 445 nm as well as at 390 nm, depending on different conditions (Figures 14 and 15). The first peak shows a higher absorption maximum at 445 nm than at 390 nm, which is typical for darkness. In contrast, the third peak exhibits a higher absorption maximum at 390 nm than at 445 nm, which is characteristic of illumination. The peak in the middle of the chromatogram is similar to the first one that was measured in the dark (Figure 23), even though there was a short exposure to blue light in this case. Either illumination was not sufficient to reach the adduct state or the recovery from the light state to the ground state is fast, and the protein fully recovers during the SEC run.

In addition, retention times associated with masses of calibration proteins were compared to LPadC_B. Consequently, the LPadC_B peak might be assigned as a tetramer at about 14 ml after injection (Figure 23). However, domains of the tested constructs are arranged in an elongated fashion in contrast to globular structures of the calibration proteins. Due to that, correlations of elution volumes are not perfect, requiring the additional characterisation of LPadC_B by MALS (Figure S2), which demonstrated that the relevant peak corresponds to a dimeric assembly.

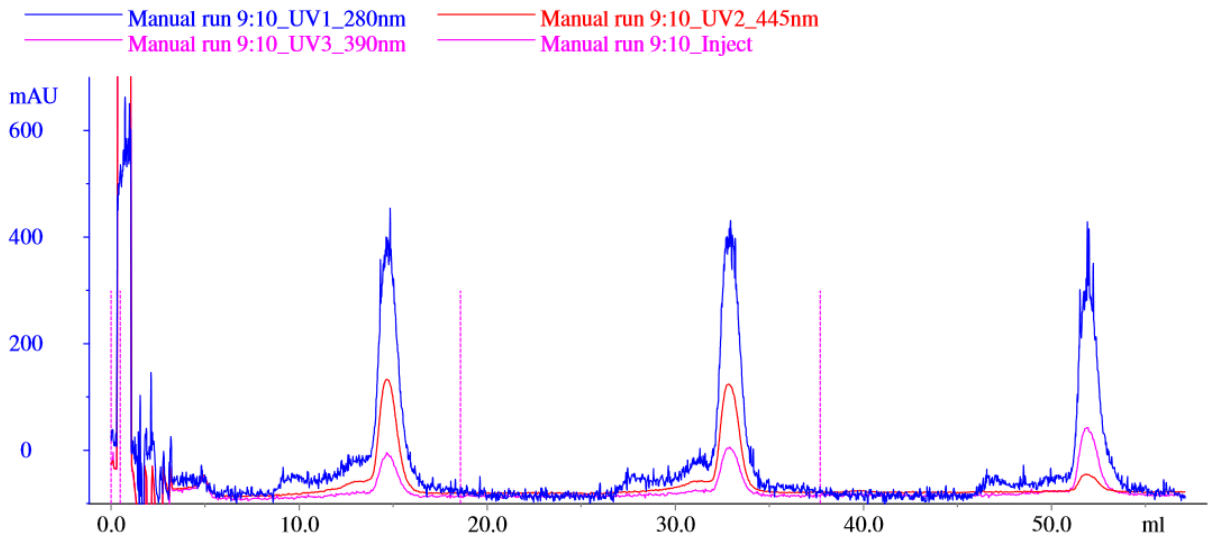

Figure 23. Size-exclusion chromatography of LPadC_B using the Superdex 200 increase 10/300 GL column under dark (1.) and light (2., 3.) conditions. After the second affinity chromatography using Ni^{2+} columns, proteins were purified with SEC to get information about light-regulated formation and also get rid of contamination/oligomerization. Chimera LPadC_B exhibits dimeric conformation under dark as well as under light conditions.

X-axis = elution volume in ml; y-axis = absorbance in mAU; 280 nm = absorption maximum of all proteins; 445 nm = absorption maximum of flavin-based light-sensitive proteins in darkness; 390 nm = absorption maximum of flavin-based light-sensitive proteins in light

The other chimera LPadC_A was also tested under dark as well as under light conditions, resulting in the same relevant peak at about 14 ml under both conditions (Figures 24 and 25). The elution volume of the peak corresponding to LPadC_A, indicates a predominant dimeric state. Additionally, the dark sample shows a specific absorption maximum that is higher at 445 nm and lower at 390 nm (Figures 14, 15 and 24), similar to the aliquot that was only preilluminated before injection (Figure 25). Probably, the stimulation with blue light was not sufficient for reaching the adduct state or the recovery to the ground state was quite fast. Comparably to LPadC_B (Figure 23), all tested samples of both chimeras offer constant quaternary assemblies (Figures 24 and 25), resulting in light-independent dimeric states and therefore high protein stability.

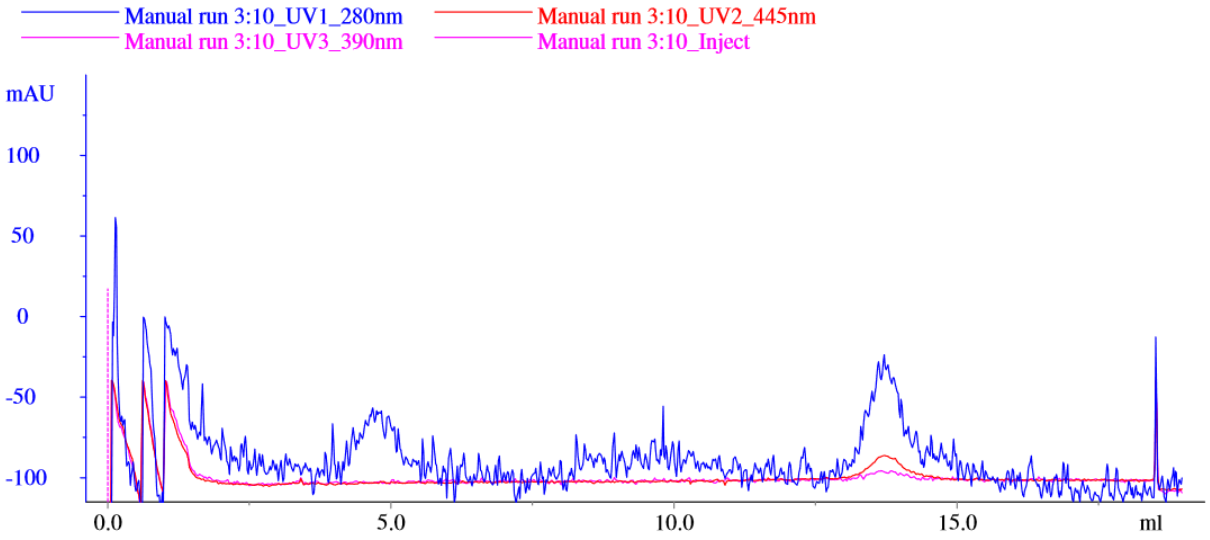

Figure 24. Size-exclusion chromatography of LPadC_A using the Superdex 200 increase 10/300 GL column under dark conditions. After the second affinity chromatography using Ni^{2+} columns, proteins were purified with SEC to get information about light-regulated formation and also get rid of contamination/oligomerization. Chimera LPadC_A exhibits dimeric conformation under dark conditions.
X-axis = elution volume in ml; y-axis = absorbance in mAU; 280 nm = absorption maximum of all proteins; 445 nm = absorption maximum of flavin-based light-sensitive proteins in darkness; 390 nm = absorption maximum of flavin-based light-sensitive proteins in light

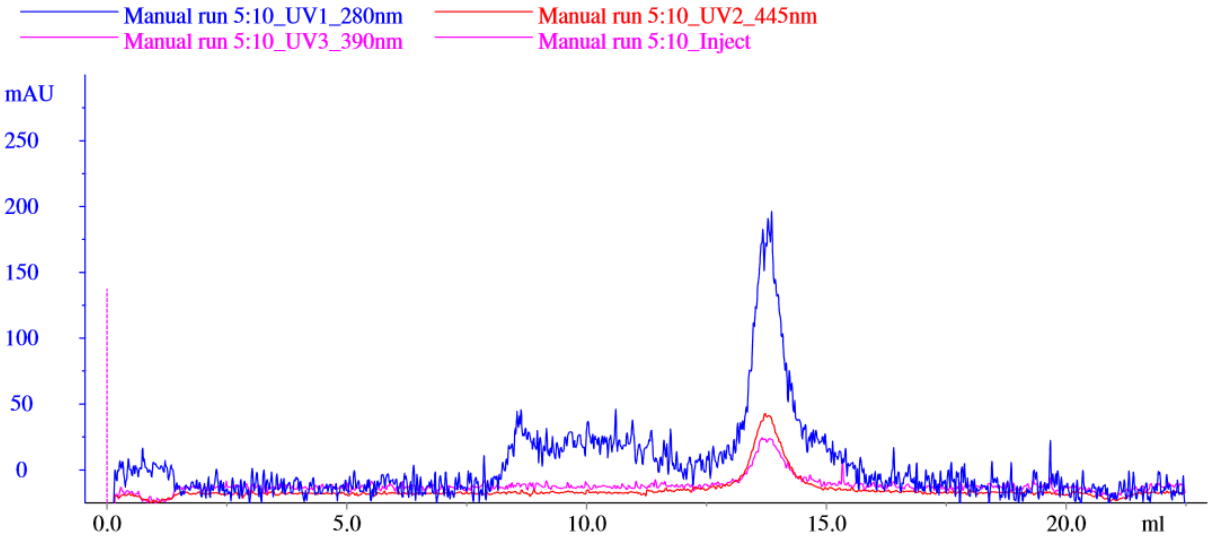

Figure 25. Size-exclusion chromatography of LPadC_A using the Superdex 200 increase 10/300 GL column under light conditions. After the second affinity chromatography, using Ni^{2+} columns, proteins were purified with SEC to get information about light-regulated formation and also get rid of contamination/oligomerization. Chimera LPadC_A exhibits dimeric conformation under light conditions.
X-axis = elution volume in ml; y-axis = absorbance in mAU; 280 nm = absorption maximum of all proteins; 445 nm = absorption maximum of flavin-based light-sensitive proteins in darkness; 390 nm = absorption maximum of flavin-based light-sensitive proteins in light

2.4 Quantitative light regulation of enzymatic activity

After qualitative analysis of structural conformation and light-regulation by size-exclusion chromatography, a quantitative analysis was performed via HPLC, coupled to UV detector for verification. During this process, the conversion of GTP as substrate to its product c-di-GMP was measured. In addition, an intermediate related to c-di-GMP could be formed. Under light as well as under dark conditions, the amount of produced c-di-GMP is proportional to enzymatic activities of the tested constructs. For this method, three different concentrations of GTP were tested [50 µM, 200 µM and 500 µM] with varying enzyme concentrations of the different proteins ([1 µM] for wild-type *Ss*LadC and [0.01 µM] for the chimeras LPadC_B and LPadC_A). Furthermore, all variants were measured in darkness but also by blue light illumination for different incubation times. Afterwards, enzymatic activities of relevant proteins were compared to each other, giving some important information about activation or inactivation of tested DGCs.

Three purified variants of blue light photosensory proteins were quantitatively analyzed: wild-type *Ss*LadC originating from GB1 9His TEV *Ss*LadC and the chimeras LPadC_A and LPadC_B. All tested constructs were purified via affinity chromatography twice. In addition, the naturally occurring variant was purified by SEC. As already mentioned, wild-type *Sf*LadC is 100-fold activated by blue light illumination in comparison with dark conditions [8]. In contrast, *Sf*LadC -4 truncation shows a 4-fold activation by exposure to blue light compared to darkness.

Wild-type *Ss*LadC exhibits enzymatic activities between 0.57 and 0.86 µmol c-di-GMP/ min/ µmol protein in the dark and 0.86 µmol c-di-GMP/ min/ µmol protein under light conditions dependent on the employed substrate concentrations (Figure 26). Therefore, *Ss*LadC offers a 1.5-fold activation by illumination compared to darkness at lower GTP concentrations [50 µM and 200 µM]. However, there is no difference between dark and light conditions at a substrate concentration of 500 µM. This construct exhibits low enzymatic activities under both conditions and shows almost no light regulation. Furthermore, the chimera LPadC_B shows a lower c-di-GMP production from 1.43 to 2.86 µmol c-di-GMP/ min/ µmol protein in darkness (Figure 27), whereas activities between 11.43 and 28.57 µmol c-di-GMP/ min/ µmol protein are present by exposure to blue light. In the dark, the activity of LPadC_B is higher at 200 µM GTP than the c-di-GMP production at 500 µM GTP, which is probably an outlier. This results in a 5-fold activation at 50 µM GTP, 16-fold at 200 µM GTP and

10-fold at 500 µM GPT under light conditions in comparison to dark conditions. The variant LPadC_B is significantly light-regulated, offering high enzymatic activities under light conditions and lower activity in the dark. Moreover, the chimera LPadC_A was only analysed at both lower substrate concentrations [50 µM and 200 µM] (Figure 28). The production of c-di-GMP was below the detection limit in darkness, which means no activity, and offers 0.57 µmol c-di-GMP/ min/ µmol protein upon illumination. Therefore, LPadC_A shows low enzymatic activity by blue light exposure and no activity in the dark. This construct appears to be light-regulated too, but no activation fold could have been calculated. On average, LPadC_B is regulated the most (~10 times), followed by the wild-type SsLadC (~1.5 times) under light conditions in comparison to dark conditions (Figures 27 and 26). According to this, LPadC_B exhibits the highest enzymatic activities, followed by wild-type *Ss*LadC and the chimera LPadC_A (Figures 26, 27 and 28).

Under both tested conditions, ci-di-GMP production of wild-type *Ss*LadC extremely increases at the lowest substrate concentration [50 µM] and exhibits constant activation at 200 and 500 µM GTP under dark as well as under light conditions (Figure 26), similar to the chimera LPadC_A at both lower substrate concentrations (Figure 28). At 500 µM GTP, the illuminated sample of *Ss*LadC ends up in saturation (Figure 26), whereas the dark aliquot reaches the same activity level, resulting in no light regulation of *Ss*LadC anymore at the highest substrate concentration. In addition, there was no activation of LPadC_A in darkness (Figure 28). Moreover, enzymatic activity and therefore c-di-GMP production of the chimera LPadC_B increases with increased substrate concentrations upon illumination, ending up in saturation (Figure 27). In darkness, LPadC_B shows only low activity at a constant level.

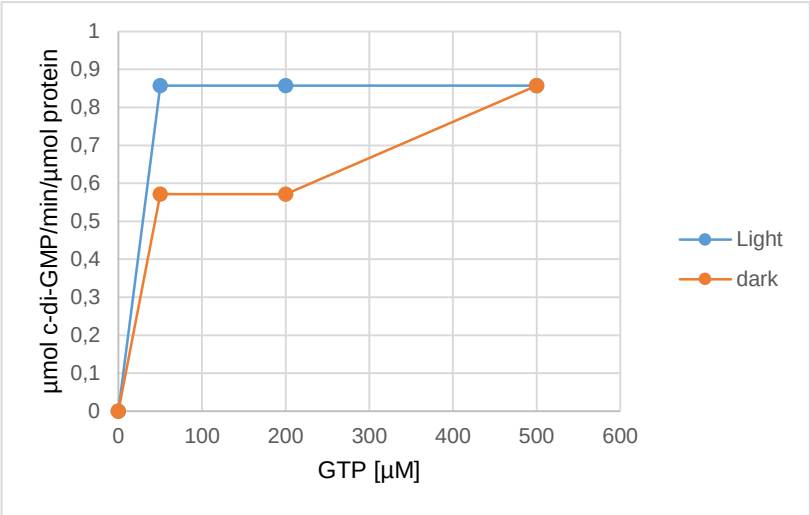

Figure 26. Quantification of 1 µM wild-type *Ss*LadC activity under light as well as under dark conditions and its dependence on GTP concentrations [50 µM, 200 µM and 500 µM]. Increasing enzymatic activitites are proportional to synthesis of product and intermediate (c-di-GMP).

X-axis = substrate concentrations in µM; y-axis = µmol product per minute per µmol dimer protein in dimeric conformation

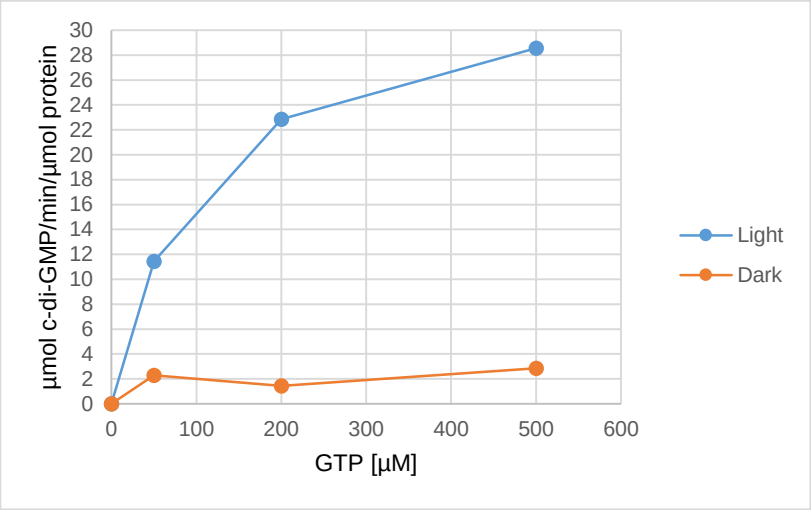

Figure 27. Quantification of 0.01 µM LPadC_B activity under light as well as under dark conditions and its dependence on GTP concentrations [50 µM, 200 µM and 500 µM]. Increasing enzymatic activitites are proportional to synthesis of product and intermediate (c-di-GMP).

X-axis = substrate concentrations in µM; y-axis = µmol product per minute per µmol protein in dimeric conformation

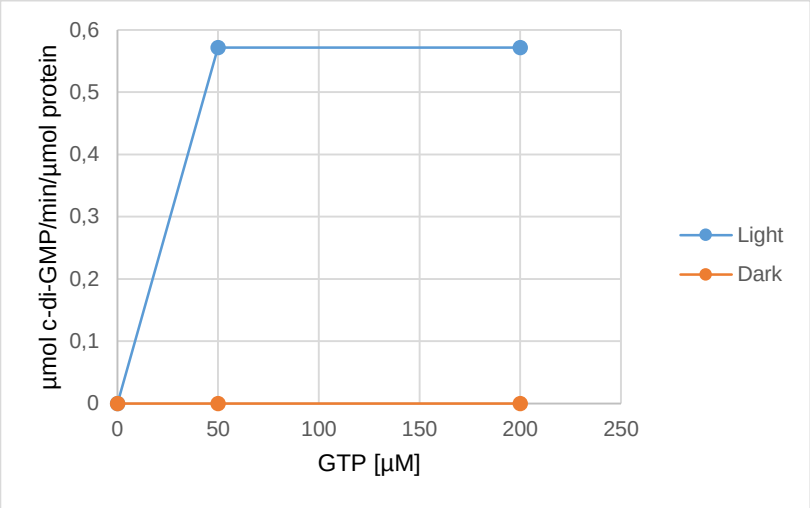

Figure 28. Quantification of 0.01 µM LPadC_A activity under light as well as under dark conditions and its dependence on GTP concentrations [50 µM and 200 µM]. Increasing enzymatic activitites are proportional to synthesis of product and intermediate (c-di-GMP).

X-axis = substrate concentrations in µM; y-axis = µmol product per minute per µmol protein in dimeric conformation

3. Discussion

In general, blue light-regulated photosensitive proteins consist of three different components. At the N-terminus, the LOV domain works as sensor that senses the incoming light, leading to structural and therefore functional changes [8] [12]. LOV domains are coupled to FMN as cofactors, building flavin-cysteinyl adducts under blue light stimulation. The thermal recovery from the light state back to the dark conformation takes seconds up to hours, depending on the LOV system [5]. For example, the recovery of wild-type *Sf*LadC takes about 2 h (Figure 14), while naturally occurring *Ss*LadC recovers over more than 7 h (Figure 15). At the C-terminus, the GGDEF domain works as effector, featuring the enzymatic part of DGCs [8]. This effector is responsible for producing bacterial c-di-GMP out of two GTP substrate molecules [9]. The third important component, the linker region between LOV and GGDEF domain, is composed of a specific supercoiled α-helix, varying in length and also offering various compositions of amino acid residues, resulting in modified signal transduction because of potential conformational alteration [13]. In this project, two naturally occurring LOV-GGDEF systems (*Sf*LadC and *Ss*LadC) as well as two engineered variants (LPadC_A and LPadC_B) had been examined (Figures 6a and 6b). Both qualitative and quantitative analyses were performed, leading to structural

and functional characterization of these mentioned constructs. Discovered by SEC, full-length *Sf*LadC in solution exhibits aggregates and also different oligomeric states in the dark (Figures 16 and 17), while exposure to blue light leads to almost complete but reversible aggregation of this protein (Figure 18). In contrast, wild-type *Ss*LadC and both chimeras (LPadC_A and LPadC_B) show stable dimeric states under dark as well as under light conditions in solution (Figures 19-22 and 23-25), resulting in light-independent oligomerization states, which offer the same relevant peak at 14 ml elution volume under all tested conditions. In the case of LPadC_B, this peak was verified by MALS (Figure S2), corresponding to LPadC_A and wild-type *Ss*LadC, all resulting in a global dimeric structure. Both chimeras contain the LOV sensor of *Sf*LadC from *S. fonticola* and the effector domain of *Is*PadC from *Idiomarina sp.* A28L (Figures 6a and 6b). Due to that, the GGDEF domain of full-length *Sf*LadC seem to be responsible for the formation of inclusion bodies and diverse oligomeric assemblies under dark and especially under light conditions, while the GGDEF domain of *Is*PadC leads to dimeric assemblies under different conditions. In the case of full-length *Ss*LadC, all components originate from *S. shabanensis* (Figures 6a and 6b).

3.1 Activation and inactivation by dimerization

As already mentioned, full-length *Sf*LadC is 100-fold activated by blue light illumination compared to darkness [8]. In contrast, the truncated version *Sf*LadC -4 exhibits activation that is four times higher by blue light exposure in comparison to the dark. Furthermore, c-di-GMP production and therefore enzymatic activities of wild-type *Ss*LadC and the chimeras LPadC_A and LPadC_B were quantitatively analysed via HPLC too. All of the tested proteins show dimeric conformations under dark and light conditions, however they offer various enzymatic activities. Due to that, there are activated as well as inhibited dimers by measuring the c-di-GMP production of these variants. In addition, it has already been discovered, that dimeric conformations of diverse DGCs can be active as well as inactive by GGDEF effector regulation. There can be activation by dimerization but also inactivation by product inhibition, corresponding to different dimeric states [8].

The naturally occurring variant *Ss*LadC exhibits low DGC activities of approximately 0.86 µmol c-di-GMP/ min/ µmol protein under all tested conditions (Figure 26), leading to dimeric assemblies, which are almost not functionally light-regulated. In contrast, LPadC_B shows high enzymatic activity up to 28.57 µmol c-di-GMP/ min/ µmol protein

upon illumination versus lower c-di-GMP production up to 2.86 µmol c-di-GMP/ min/ µmol protein in the darkness (Figure 27), resulting in light-regulated dimers that are active under light conditions and inhibited under dark conditions. For comparison, LPadC_A offers activation by blue light exposure (0.57 µmol c-di-GMP/ min/ µmol protein) but was below the detection limit in the dark (Figure 28). Because of that, LPadC_A exhibits an active dimeric state in light and an inactive dimeric state in darkness. Maybe, the used enzyme concentration of LPadC_A was too low or the incubation times were too short.

The tested chimeras are composed of the sensor domain of *Sf*LadC from *S. fonticola* and the effector domain of *Is*PadC from *Idiomarina sp.* A28L (Figure 6a). In addition, LPadC_A only contains the LOV domain of *Sf*LadC, whereas LPadC_B include the LOV domain and also the linker region of *Sf*LadC (Figures 6a and 6b). Both chimeras show light regulation, varying in the production of c-di-GMP and therefore in DGC activities. These differences in enzymatic activities might occur because of the mentioned structural differences. Developed in diverse studies, LOV-containing light-sensitive proteins are positive modulators in signal transduction during the stress metabolism of *Bacillus subtilis* for example [14]. In general, N-terminal input domains like LOV sensors are involved in the regulation of protein activation or inactivation, because of signal transduction from the incoming signal down to the output effector domain [15]. Subsequently, the sensor domain of full-length *Sf*LadC appear to be responsible for the light regulation in LPadC_A and LPadC_B. In contrast, the combination of LOV domain and linker of *Sf*LadC might be the reason for high enzymatic activities in LPadC_B compared to *Sf*LadC, which is 100-fold activated upon illumination [8]. The DGC activity of LPadC_B is not as high as the enzymatic activity of wild-type *Sf*LadC because of their different effector domain. The C-terminal GGDEF domain of both chimeras comes from *Is*PadC, which exhibits big differences in c-di-GMP production [10]. GGDEF effector domains do not offer significant DGC activities without coupled sensor domains, whereas combinations with several N-terminal input domains lead to the dimerization of GGDEF domains, resulting in increased c-di-GMP concentrations and therefore higher enzymatic activities [11].

In addition, wild-type *Ss*LadC offers low activities and almost no light regulation. Moreover, high concentrations of c-di-GMP are related to quite stable, inactive conformations of analysed constructs [9]. In contrast, generated *Ss*LadC shows very low production of c-di-GM, even though the highest enzyme concentration [1 µM] of

tested proteins were utilized. Probably, there are always changes between both active and inactive dimeric states. Maybe, this results in low activity of the corresponding protein under dark and light conditions.

3.2 Linker length and specificity

Generally, wild-type versions *Sf*LadC and *Ss*LadC and the chimeras LPadC_A and LPadC_B had been biochemically characterized by size-exclusion chromatography followed by quantitative HPLC assays. These different variants feature linker regions from diverse naturally occurring homologs of LOV-GGDEF constructs, influencing enzyme activities and structural conformations.

Linker regions between sensor and effector domains of LOV diguanylate cyclases are composed of highly specific α-helices. These supercoiled helices might influence structural and functional properties [8]. Characteristically, these coiled-coil structures consist of seven different amino acid residues (a-g), which can be repeated [13]. Due to that, the linker of each tested protein exhibit four heptads, resulting in 28 aa residues in total (Figure 6b). Usually, hydrophobic amino acid residues are often located at positions a and d, whereas hydrophilic aa residues (neutral and charged) can be mostly found at positions e and g [13]. Certainly, full-length *Sf*LadC does not offer this typical hxxhxxx arrangement, similar to the other tested variants (Figure 6b) [8]. Wild-type *Ss*LadC and the chimera LPadC_A offer more hydrophobic amino acid residues, than wild-type *Sf*LadC and the chimera LPadC_B, which exhibit distinct higher light-regulated enzymatic activities (Figure 27) [8]. In addition, some of their hydrophilic aa residues are replaced by Asn in wild-type *Ss*LadC and the chimera LPadC_A, which show lower enzymatic activitites (Figures 26 and 28). Therefore, coiled-coils that are rich in hydrophobic and Asn residues at the central a and d positions of the heptad repeats specifically build homodimeric conformations [13], compared to tested LOV-GGDEF systems (Figure 6b). Anyway, asparagine is implicated in the specificity of supercoiled sequences [10], potentially leading to the big differences of c-di-GMP production and DGC activities under light and dark conditions between tested proteins. Maybe, Asn might be responsible for less c-di-GMP production and therefore lower DGC activity of the chimera LPadC_A in comparison to the chimera LPadC_B under both tested conditions, because this is the only big difference in the linker regions of both chimeras (Figure 6b). Furthermore, sensor and effector domains of LPadC_A and LPadC_B are the same (Figure 6a). In addition, wild-type *Ss*LadC offers more

asparagine residues than the chimera LPadC_A in its linker, resulting in the lowest enzymatic activities under all conditions. In contrast, wild-type *Sf*LadC shows no Asn in its linker region, leading to a 100-fold DGC activity under light conditions compared to the dark [8].

4. Conclusion

For summing up, the biochemical characterization of naturally occurring and engineered LOV-GGDEF blue light-sensitive proteins results in different structural and functional properties of the individual systems.

First of all, wild-type *Sf*LadC had been characterized already. The expression and purification were quite easy after optimization, leading to adequate protein yields. Therefore, x-ray crystallography of this protein was performed, showing a tetrameric assembly as crystal structure. However, there was much aggregation and oligomerization after the purification by SEC, probably because of unidentified surface reactions in solution. To solve this problem, different buffer systems with various pH and salt concentration were tested without any effect. Because of that, a similar wild-type variant, *Ss*LadC, and also two chimeras of *Sf*LadC (LPadC_A and LPadC_B), containing the GGDEF domain of IsPadC, were generated. All these variants result in light-independent but therefore high stable dimeric states after size-exclusion chromatography in solution. Expression and purification steps of LPadC_A and LPadC_B were uncomplicated, resulting in high protein amounts of LPadC_B and a lower yield of LPadC_A at the end. In contrast, wild-type *Ss*LadC was hard to express and purify, getting only low amounts of the desired protein.

In addition, full-length *Sf*LadC offers a 100-fold activation upon blue light illumination compared to darkness [8]. In contrast, this work identified the 1.5 times activated wild-type *Ss*LadC and the 10 times activated chimera LPadC_B. Probably, the chimera LPadC_A is also light-regulated, but the c-di-GMP production was below the detection limit in the dark. For further experiments, higher enzyme concentrations of LPadC_A have to be tested, combined with longer incubation times, to detect enzymatic acitivity under dark conditions. Furthermore, it was important to optimize the expression and purification steps of wild-type *Ss*LadC and the chimera LPadC_A to receive higher protein yields. Due to that, another quantitative HPLC assays should be performed to verify functional results of enzymatic activities and the light-regulation of individual

constructs. In addition, crystallization of the chimera LPadC_B should be performed, probably resulting in a dimeric assembly as crystal structure.

During SEC of wild-type *Sf*LadC, fractions of desired oligomeric states, like tetramers or dimers, should be collected and then separately analysed by SEC again and probably via functional HPLC assays too. Furthermore, some different components and conditions during the expression and purification of this protein could be varied and optimized to solve this problem.

5. Materials and Methods

5.1 Synthesis of constructs

First of all, the gene cassette of *Sf*LadC was available in the vector pET-GB1 with kanamycin resistance (Kn^R), including an N-terminal His-tag of six histidine residues, the tobacco etch virus (TEV) polyhistidine tag as cleavage site and five extra Gly residues between the TEV-site and the start of the protein coding sequence to improve TEV cleavage of isolated proteins. The gene cassette of *Ss*LadC had already existed in another plasmid, pET-M11. Because of that, the second homolog was cloned into the preferred expression vector pET-GB1.

To this end, both constructs, *Sf*LadC in pET-GB1 as well as *Ss*LadC in pET-M11, were transformed into *E. coli* XL10 Gold (Gene Art, Life Technologies) using β-mercaptoethanol treated cells for increased transformation efficiency. At first, chemically competent cells were thawed at 4 °C on ice. Afterwards, 1 µl of 1.43 M β-mercaptoethanol was added to each 50 µl *E. coli* XL10 Gold, reaching 25 mM in total, followed by incubations at 4 °C for 10 min. Furthermore, 1 µl per [100 ng μl^{-1}] concentration of each plasmid was added and also incubated for 10-15 min on ice. Then, bacterial cells were shocked at 42 °C for 45 sec, followed by cooling at 4 °C for about one minute. At the end, 150 µl of warm super optimal broth (SOC) medium were added for relaxing respectively. After that, samples were incubated at 37 °C for about one hour, shaking at 900 rpm. Finally, 200 µl of each transformation reaction were plated on Luria-Bertani (LB) plates (peptone [10 g l^{-1}], yeast extract [5 g l^{-1}], NaCl [5 g l^{-1}] and agar-agar [15 g l^{-1}) containing kanamycin [30 mg ml^{-1}; 1:1000] and then incubated overnight at 37 °C.

The next step was the isolation of the DNA, preparing appropriate plasmids (pEQGOLD Plasmid Miniprep Kit I). Cell material from each transformation plate had been resuspended in 250 µl of cold solution I, which was complemented with ribonucelase A (RNaseA). After vortexing, 250 µl of solution II were added to the mixture and then inverted. Last, 350 µl of solution III were added to each sample and also mixed gently. All components united, centrifugation occurred (10,000 g; 10 min; room temperature (RT)). These steps were performed for chemical lysis, followed by neutralization of the cell suspensions. After centrifugation, each supernatant was loaded onto a small column (PerfectBind DNA Column) up to approximately 750 µl. Because of a volume of almost 1 ml each, this step was done twice, followed by

centrifugation (10,000g; 1 min; RT). While the DNA was bound to the matrix of the columns, the filters were cleaned with 750 µl washing buffer including EtOH and following centrifugation for two times (10,000 g; 1 min; RT). Thereafter, the membrane was dried by another centrifugation step, which is a quite important part of this preparation (10,000 g; 2 min; RT). Finally, 50 µl dH_2O were added to the columns, eluting DNA of produced plasmids by one final centrifugation (5,000 g; 1 min; RT).

Subsequently, the digestion of 2 µg DNA of each pET-GB1_6His_TEV_5G_*Sf*LadC and pET-M11_*Ss*LadC was performed at NcoI and NotI restriction sites with 1 µl of cold high-fidelity enzymes NcoI-HF and NotI-HF (New England Biolabs, Inc.) to establish sticky ends. In addition, fresh and vortexed CutSmart Buffer (10x) (New England Biolabs, Inc.) was added to the samples and incubated for 30 min at 37°C. Afterwards, agarose gel electrophoresis (1 %; 1 g agarose, 100 ml tris-acetate-EDTA (TAE) buffer (1x) and 10 µl ethidium bromide (EtBr)) was performed for 1.5 h at 90-100 V for separation of the required fragments. In this method, an appropriate DNA marker was utilized for comparison (O'GeneRuler™ 1kb DNA Ladder, ready-to-use; Thermo Scientific™) and a 6x loading dye (orange; New England Biolabs, Inc.) was used for staining. Under UV light, both the vector band of the first construct (pET-GB1; ~5.5 kb) and the insert band of the second construct (*Ss*LadC; ~1 kb) were cut out of the agarose gel, followed by DNA extraction (Zymoclean™ Gel DNA Recovery Kit). At first, the gel fragments were blended with agarose dissolving buffer (ADB) at a ratio of 1:3 related to their weights. For complete dissolution, the mixtures were incubated at 50 °C for 10-15 min. For DNA binding, the suspensions were then loaded onto a column (Zymo-Spin™ Column; Collection Tube), followed by centrifugation (10,000 g; 2 x 60 sec; RT). Afterwards, the removal of contaminations was performed by adding 200 µl of washing buffer supplemented with EtOH and centrifugation twice (10,000 g; 30 sec; RT). Finally, elution buffer or dH_2O was added to the columns, eluting DNA via another centrifugation (10,000 g; 60 sec; RT).

Thereafter, pET-GB1 got dephosphorylated by treatment with 1 µl of antarctic phosphatase in the presence of the appropriate buffer (10x) (New England Biolabs, Inc.), which was then incubated at 37 °C for 15 min, followed by inactivation at 70°C for 5 min. After that, final DNA extraction was performed (DNA Clean and Concentrator™-5). For this purpose, the expression vector fragment was diluted with 3 volumes of DNA binding buffer and the insert was mixed with 5 volumes of DNA binding buffer. Then, each sample was transferred to a column (Zymo-Spin™ Column;

Collection Tube), followed by a centrifugation step (10,000 g; 30 sec; RT). Moreover, the matrix of each column was washed to remove contaminations twice with 200 µl washing buffer containing EtOH. Vector and insert DNA were eluted with elution buffer via centrifugation (10,000 g; 30 sec; RT). Subsequently, the ligation of pET-GB1 and *Ss*LadC (ratio 1:10) by adding 1 µl of T4 DNA ligase and T4 ligase buffer (10x) (New England Biolabs, Inc.) occurred for about 3.5 h at 16 °C. Then, a transformation was performed overnight. Afterwards, a cultivation of individual colonies in LB medium complemented with kanamycin [30 mg ml^{-1}; 1:1000] was done at 37 °C under shaking overnight. These cultures were used for another plasmid preparation, for which the cells were pelleted by centrifugation (5,000g; 10 min; RT). Finally, sequencing of the generated construct was performed (Microsynth) by adding T7 primers [10 mM] to the DNA fragment *Ss*LadC, which was assembled into the expression vector pET-GB1.
In addition, a modification in full-length *Ss*LadC relating to the length of the His-Tag was performed, to extend the tag from six to nine His residues. This was done, because of a relatively high loss of desired protein after efficient washing steps with 60 mM imidazole during Ni^{2+} purifications. Furthermore, the *Ss*LadC variant was produced to get rid of contaminations, but also to obtain a higher concentration of the protein of interest, by increasing its affinity during chromatography via Ni^{2+} columns. The generation of this modified construct started with mutagenesis (Figure S1) described by Liu and Naismith [16], followed by restriction digestion via the enzyme DpnI (New England Biolabs, Inc.) for about 30 min at 37 °C and about 1.5 h at RT. The same procedure was performed for the establishment of some truncations in the coiled-coil linker as well as specific mutations. In contrast, the previously mentioned chimeras LPadC_A and LPadC_B were produced by two separate polymerase chain reactions (PCRs) each (Figure S1), (NEBuilder HiFi DNA Assembly) because of different combinations between two constructs. The time profile was composed of 30 cycles, including initial denaturation (98 °C; 30 sec), denaturation (98 °C; 10 sec), annealing (62 °C; 20 sec), extension (72 °C; 30 sec kb^{-1}), final extension (72 °C; 5 min) and hold (4 °C) [16]. Each sample consisted of a mix of the appropriate template (10 ng), suitable forward and reverse primers [5 µM] (Sigma-Aldrich), deoxynucleotide triphosphates (dNTPs) [0.2 mM], Q5 buffer (5x) and Q5 HiFi polymerase [5 µM] (New England Biolabs, Inc.) [16]. All PCRs were performed by using the same PCR device (GeneAmp PCR System 9700 Thermocycler (96-well), PE Applied Biosystem).

The purification of individual amplified components of LPadC_A and LPadC_B was done via agarose gel electrophoresis, followed by DNA extraction (Zymoclean^{TM} Gel DNA Recovery Kit). After that, the NEBuilder assembly procedure was performed to generate the final chimeras (NEBuilder HiFi DNA Assembly Reaction Protocol). Totally, 0.1 pmols of *Ss*LadC should have been inserted into the expression vector pET-GB1. On the basis of weight and length of desired fragments, a formula was given for calculation (pmols = (weight in ng) x 1,000 / (base pairs x 650 daltons). In addition, the required ratio of vector to insert was 1:2. Each reaction volume was 5 µl in total, containing 2.5 µl NEBuilder HiFi DNA Assembly Master Mix. The mix was incubated for 15 min at 50 °C.

Subsequent steps were similar to the generation of the initial full-length constructs and included transformation, cultivation, plasmid preparation by DNA isolation (pEQGOLD Plasmid Miniprep Kit I) and sequencing to confirm the desired mutations (Microsynth).

5.2 *In vivo* screening of DGC activity

Before quantification, whole-cell screenings of DGC activity in living *E. coli* BL21 (DE3) were performed to get first impressions of light regulation to pre-screen for interesting constructs. In cells featuring DGC activity, c-di-GMP production and therefore cellulose synthesis was detected, leading to red colonies on CR plates, while inactive cells remain white. Except for *Ss*LadC variants in the pET-GB1 expression vector, all generated constructs were screened by a cell-based DGC activity assay. This procedure has been adapted from the protocol by Antoniani *et al* [17]. By using this method, the enzymatic activities of each construct were qualitatively analyzed under dark as well as light conditions. Expression plasmids were transformed in *E. coli* BL21 (DE3), and bacterial cells were then grown in yeast extract casamino acids (YESCA) medium at 30 °C via shaking cultivation, including kanamycin [30 mg ml^{-1}; 1:1000] as selection marker, yeast extract [1.5 mg ml^{-1}], casamino acids [0.01 g ml^{-1}], $MgSO_4$ [0.05 mg ml^{-1}] and $FeSO_4$ [0.5 µg ml^{-1}]. The cells were grown to a density of about 0.5 at 600 nm, verified by photometric measurements (HITACHI; U-1100 Spectrophotometer). Afterwards, protein expression of each bacterial culture was induced by 0.1 mM IPTG under dark conditions. Then, the samples were incubated at 16 °C for 3-4 h with constant shaking at 900 rpm in the darkness. After centrifugation (16,000 g; 5 min, RT) the cells of each culture were resuspended in YESCA medium, to reach an OD_{600} of 10. Finally, 3 µl of each sample were spotted onto two CR [0.01

mg ml^{-1}]-YESCA agar [0.02 g l^{-1}] plates, supplemented with kanamycin [30 mg ml^{-1}; 1:1000]. After waiting for about 5 min, the culture drops had been dried and then the plates were incubated overnight at 15-20 °C for 16-18 h. One plate was wrapped in aluminium foil to create dark conditions, while the other one was constantly illuminated by blue light. The distance between plate and light source (Spot LED with 48 LEDs SMD, E27, 3.0 W, blue; Luminea™) was about 1 m. In addition, both a positive control (pET-GB1 6His TEV *Sf*LadC) and a negative control (pET-M11 AppA ΔC) were performed. The longer the incubation time, the more intense the spots appeared in case of activation and at some point the signal was saturated.

5.3 Spectral characterization

Each absorbance spectrum of the wild-type constructs *Sf*LadC and *Ss*LadC was measured with a Specord 200 Plus spectrophotometer (Analytik Jena) 300-500 nm. The dark state spectra and individual spectra during the thermal recovery after 1 min illumination with blue light (Spot LED with 48 LEDs SMD, E27, 3.0 W, blue; Luminea™) were measured at 20 °C. In addition, the recovery of each tested protein was followed at 390 nm after exposure to blue light for one minute (Spot LED with 48 LEDs SMD, E27, 3.0 W, blue; Luminea™). Individual samples were diluted in the corresponding buffer, containing 10 mM Tris/HCl pH 8.0, 500 mM NaCl and 2 mM $MgCl_2$.

5.4 Protein expression

All generated plasmid constructs were used for protein expression in *E. coli* BL21 (DE3). For this purpose, the cells were grown in LB medium (peptone [10 g l^{-1}], yeast extract [5 g l^{-1}] and NaCl [5 g l^{-1}]) with a supplementation of kanamycin [30 mg ml^{-1}; 1:1000]. This process took place in the shaking incubator at 37 °C and 120-140 rpm. Once the bacterial cells had reached the exponential phase of growth, the temperature was reduced to 16 °C. At this time, the cells' density was about 0.5 to 0.8 at an absorbance of 600 nm as determined by photometric measurements (U-1100 Spectrophotometer, HITACHI). The cells were adapted to the lower temperature for about 25 min to improve yields of soluble proteins. Afterwards, protein expression was induced with 0.1 mM isopropyl IPTG. At this moment, the work was continued under dark conditions, to limit any effect of light on the sensory part of the protein. The temperature was shifted to 18 °C now. Then, protein expression in the shaking incubator at 120 rpm was continued for 16-18 h.

The collection of bacterial cells occurred via centrifugation (5,000 rpm; ~12 min; 4 °C) followed by resuspension in cold buffer composed of 50 mM HEPES pH 7.0, 500 mM NaCl, 5 % glycerol, 5 mM $MgCl_2$ and 30 mM imidazole. Thereafter, an extra sample of about 50 µl before lysis was taken for later analysis. At the end, harvested cells were frozen at -20 °C.

5.5 Protein purification

All the following steps were performed under safe-light (indirect dim orange light) and cold conditions to protect the protein. For lysis, the frozen *E. coli* cells were quickly thawed by adding resuspension buffer and using a magnetic mixer. Then, lysozyme with a final concentration of 200 µg ml^{-1} as well as protease inhibitory mix (1 tablet of complete protease inhibitor – EDTA free (Roche) per 10 l expressed protein) were added to the suspension. After 20 min incubation at RT, cell disruption was mechanically done by sonication for about 3 x 4 min (100 W; Labsonic L, Sartorius). In the next step, cellular fragments were separated from the soluble material, including the desired protein, via ultracentrifugation (45,000 rpm (Sorvall T-865, 65,000 rpm max; Thermo Scientific) 1 h; 4 °C). Then, the obtained supernatants were collected separately and one additional sample of 40 µl after lysis was collected for further analysis. Thereafter, FMN was added to the protein in solution to fully saturate the expressed LOV domains.

5.5.1 Affinity chromatography

Subsequently, the soluble protein was purified via affinity chromatography. For this purpose, a column (Protino, Macherey-Nagel) containing a nickel (Ni^{2+}) matrix (Ni Sepharose 6 Fast Flow, GE Healthcare) was utilized. After its equilibration with about 15 ml resuspension buffer including 30 mM imidazole, the combined supernatant was loaded onto the column. The sample was flowing through the Ni^{2+} matrix, which was then washed with about 2 x 15 ml washing buffer containing 60 mM imidazole. For the elution of the appropriate protein, an increased concentration of 250 mM imidazole was used. About 10-30 ml elution buffer were necessary to receive the purified protein in solution. At the end of this chromatography, the eluted protein was filled into a dialysis tube. Enzymatic TEV-protease at a ratio of about 1:10 was added, cleaving off the protein at the TEV site. Overnight, the dialysis procedure was run in dialysis buffer (10 mM 4-(2-hydroxyethyl)piperazine-1-ethanesulfonic acid (HEPES) pH 7.0,

500 mM NaCl, 5 mM $MgCl_2$, 5 % glycerol, 1 mM ethylene diamine tetra-acetic acid (EDTA) pH 8.0, 0.5 mM dithioerythritol (DTE) and imidazole) to end up with 30 mM imidazole concentration of the soluble protein at 4 °C in darkness. To get rid of the cleaved GB1 and His-tags as well as the TEV protease, another affinity chromatography with a Ni^{2+} column was done. The flow through containing the protein of interest was collected and then concentrated by centrifugation at 4 °C (Amicon Ultra-15, Merck Millipore). Finally, the sample was aliquoted, flash-frozen in liquid nitrogen, and deep-frozen at -80 °C. Additionally, 20 µl of each purification fraction were collected for experiments later.

5.5.2 Size-exclusion chromatography

Furthermore, proteins were also purified via size-exclusion chromatography (SEC) on a Superdex 200 increase 10/300 GL column (GE Healthcare). The proteins were thawed and about 0.5 to 1 ml were injected into the purification system (ÄKTA, GE Healthcare) under dark conditions at 4 °C. The SEC buffer was composed of 10 mM Tris/HCl pH 8.0, 500 mM NaCl and 2 mM $MgCl_2$ and was filtered and degassed prior to usage. In addition, the system and column had to be equilibrated with the relevant buffer before protein loading. After protein loading, the sample was eluted from the column, and was collected in fractions of 1 ml each. Those including proteins of interest were then pooled and concentrated via centrifugation at 4 °C again (Amicon Ultra-15, Merck Millipore). In addition, some extra samples of 20 µl were taken during this process for quality control. At the end, the final protein samples were aliquoted, flash-frozen in liquid nitrogen, and deep-frozen at -80 °C until needed.

Gel filtration was not only used for purification, but also for testing the light influence on generated constructs. SEC is a method that can be used to estimate the oligomeric states of proteins by the elution volumes of individual eluting peaks in comparison to molecular weights of known standards. Therefore conformational changes associated with light activation during the SEC run can be analyzed. To this end, two aliquots of one construct were diluted with gel filtration buffer to approximately 1 ml in total, reaching a concentration of about 18-20 µM each. Afterwards, one aliquot was measured regularly under dark conditions. The second aliquot was illuminated with blue light for one minute before injection and during SEC, the whole column was exposed to blue light too. After the experiment, it was possible to compare the chromatograms and to identify conformational differences within the same construct

in response to blue light treatment as well as with others. In addition, the synthetic chimera LPadC_B was also measured by multi-angle light scattering (MALS; Wyatt Technologies) in phosphate-buffered saline (PBS) buffer with pH 7.4. This method detects the same parameters like SEC related to conformational states, but also enables an absolute quantification of the molecular masses of tested constructs, using both Light Scattering (LS; miniDAWN TREOS) and UV (Generic UV) detectors.

DNA and protein concentrations were all measured and determined by NanoDrop Microvolume Spectrophotometers and Fluorometer (Thermo Fisher Scientific). DNA concentrations are measured at an absorbance of 260 nm and also the ratio of 260/280 nm can be identified. In contrast, protein amounts are determined at 280 nm generally. For the calculation of final protein concentrations, the molar absorptivity or molar extinction coefficient ε (epsilon) is required, evaluated by the Lambert-Beer Law. In addition, free flavin is shown at an absorbance of about 445 nm, considering the contribution to the total protein amount shown at 280 nm.

5.6 SDS gel electrophoresis

SDS-PAGE was used for quality control of protein purification. It was composed of 12.5 % polyacrylamide in the stacking gel and 5 % polyacrylamide in the separating gel. All proteins of interest were separated by SDS-PAGE, to confirm their molecular weight and also for analyzing expression levels and protein purity. Because of that, several individual samples of different steps during protein expression and purification were taken. Afterwards, these samples were supplemented with 4x sample buffer and then heated for 5 min at 95 °C to denature the protein. Moreover, 10 µl of each sample were put into individual slots of a prepared 12.5 % SDS gel. Additionally, 5 µl of a protein marker (Roti®-Mark STANDARD, ready-to-use, Roth) were also filled in as a standard for comparison. After loading, the gel was run for 15 min at 90 V, followed by another 45 min at 180 V. At the end of the SDS-PAGE run, the gel was put into Coomassie-Brilliant-Blue R (sigmaR) solution for about 15 min and slewing. Thereafter, the gel was destained by using a mixture of acetic acid and ethanol for 3 x 30 min under shaking. Eventually, gels were rinsed with dH_2O and scanned for documentation.

5.7 Kinetic characterization of constructs

The quantitative analysis of enzymatic activities and therefore blue light regulation of naturally occurring and engineered photoreceptors was done via HPLC assays, connected to a UV detector. This method utilizes a reversed-phase column (SunFire C18 4.6 mm x 100 mm, Waters) and follows the protocol described in Gourinchas *et al* [10]. The method allows baseline separation of the substrate GTP to the product c-di-GMP as well as a linear GTP-GMP (pppGpG) intermediate that is sometimes observed for the DGC reaction cycle. The amount of produced c-di-GMP is proportional to the enzymatic activities of tested proteins. At the beginning, the whole HPLC system was purged with 100 % of buffer A (10 mM triethylammonium formate (TEAF) pH 6.0) for 5 min, followed by 5 min with 100 % of buffer B (= buffer A in 90 % MeOH) at 1ml min^{-1}. Afterwards, the equilibration was done by 2 % of buffer A combined with 98 % of buffer B at a flow rate of 1 ml min^{-1} for about 30 min. The same buffers were used for the linear gradient from 2-20 % of MeOH during which the individual compounds eluted from the column. For analysis, three different substrate concentrations were utilized [50 µM, 200 µM and 500 µM] (GTP; Roth). Protein concentrations were adapted to the observed conversion of GTP and were prepared by diluting in the appropriate buffer (50 mM HEPES pH 7.0, 500 mM NaCl and 50 mM $MgCl_2$). All components were kept on ice (~4 °C) and in relative darkness during preparation and before using. Enzymatic reactions were started by the addition of GTP and then stopped at different times by heating to 95° C for about 1.5 min. Similarly, experiments were done under blue light. Samples for the analysis of blue light activation were preilluminated for 1 min prior to the addition of GTP. Blue light exposure was performed with a Thorlabs (M455L3 SM2P50-A) lamp and maximum intensity. After stopping the reactions by heating, denatured protein was separated by centrifugation (20,000 rpm; 5 min; RT). Finally, 50 µl of each supernatant were filled into vials and analysed by HPLC. Each measuring process took 15 min at a flow rate of 1 ml min^{-1}.

At the end, HPLC chromatograms were evaluated via integration of individual peak areas of the corresponding products. This enabled the quantification of GTP, c-di-GMP and possible intermediates. The time dependent increase in c-di-GMP was used to estimate the initial rates of product formation and eventually normalized to the protein concentration to obtain apparent turnover rate constants.

6. Supplemental material

GB1 9His TEV *Ss*LadC:

Fw: 5' GAAACACCATCACCATCA CCATCACCATCACCCCATG 3' (59/66)
Rv: 5' TGATGGTGATGGTGTTTC ATGGTATATCTCCTTCTTAAAGTTAAACAAAATTATTTCTAGAGG 3' (59/66)

LPadC_A: *Sf*LadC 1-119 - *Is*PadC 500-683 (...FIGIQKDVT – LIVADSMQ...)

FwL: 5' TAAGCGGCCGCACTCGAG 3' (71)
RvL: 5' GGTAACATCTTTCTGAATGCCAATAAAGTTG 3' (66)

FwP: 5' GGCATTCAGAAAGATGTTACC CTGATTGTGGCAGATAGCATG 3' (61/63)
RvP: 5' TGCTCGAGTGCGGCCGCTTA CTGGCTACAAACCTGATTAC 3' (75/60)

LPadC_B: *Sf*LadC 1-148 - *Is*PadC 529-683 (...FEQLSIK – DDLTGIF ...)

FwL: 5' TAAGCGGCCGCACTCGAG 3' (71)
RvL: 5' TTTGATGCTCAGCTGTTCAAAATGG 3' (66)

FwP: 5' GAACAGCTGAGCATCAAA GATGATCTGACCGGTATCTTTAATC 3' (60/62)
RvP: 5' TGCTCGAGTGCGGCCGCTTA CTGGCTACAAACCTGATTAC 3' (75/60)

Figure S1. Primer sequences for the generation of wild-type GB1 9His TEV *Ss*LadC and the chimeras LPadC_A and LPadC_B

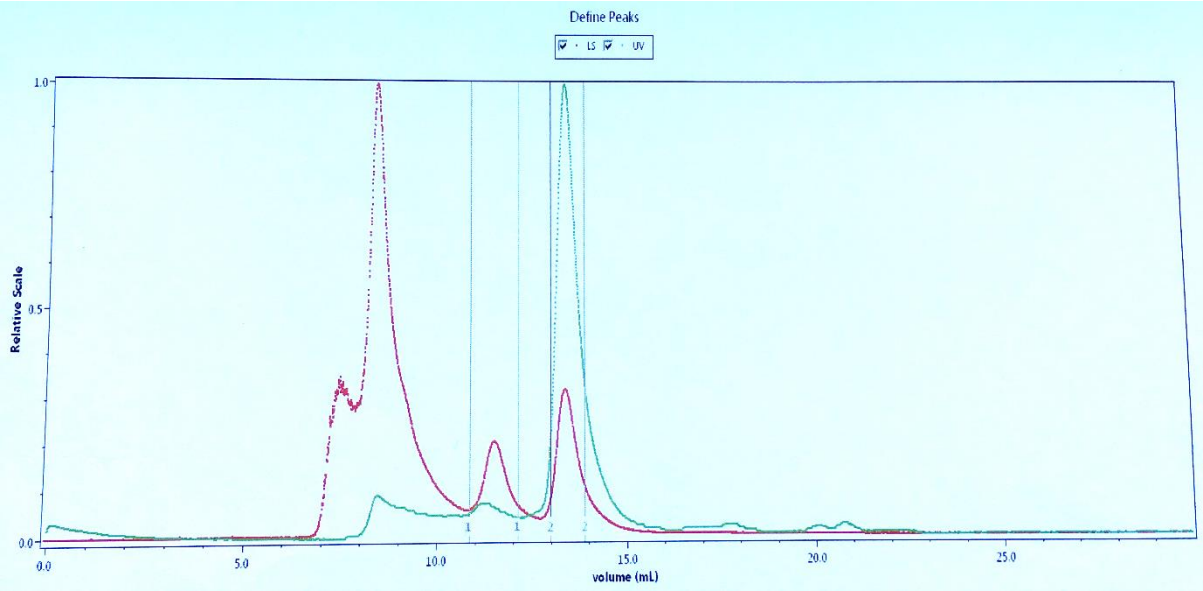

Figure S2. Multi-angle light scattering of the chimera LPadC_B using LS and UV detectors. Detection of the conformational state and the absolute quantification of the molecular masses of this synthetic protein. X-axis = elution volume in ml; y-axis = relative Scale

7. Abbreviations

aa	amino acid(s)
ADB	Agarose Dissolving Buffer
A-site	Active site
ARNT	Aryl hydrocarbon Receptor Nuclear Translocator
AU	Absorption Units
C-di-GMP	Cyclic-dimeric-Guanosine Monophosphate
CHES	N-Cyclohexyl-2-aminoethanesulfonic acid
CR	Congo Red
DGC(s)	Diguanylate Cyclase(s)
dNTPs	deoxynucleotide Triphosphates
DTE	Dithioerythritol
E	Elution
EDTA	Ethylene Diamine Tetra-acetic Acid
EtBr	Ethidium Bromide
FAD	Flavin Adenine Dinucleotide
FMN	Flavin Mononucleotide
FT	Flow Through
GMP	Guanosine Monophosphate
HEPES	4-(2-Hydroxyethyl)Piperazine-1-Ethanesulfonic acid
HF	High Fidelity
HK	Histidine Kinase
HPLC	High Performance Liquid Chromatography
Idiomarina sp.	*Idiomarina species*
I-site	Inhibitory/Inhibition site
Kn^R	Kanamycin Resistance
LadC(s)	Light-activated diguanylate Cyclase(s)
LB	Luria-Bertani
LS	Light Scattering
M	Marker/Standard
MALS	Multi-Angle Light Scattering

MES	2-(N-Morpholino)Ethanesulfonic acid
PadC(s)	Phytochrome-activated diguanylate Cyclase(s)
PAGE	Polyacrylamide Gel Electrophoresis
PBS	Phosphate-Buffered Saline
PDE(s)	Phosphodiesterase(s)
PG	Prep Grade
p-loop	protection loop
Pseudomonas sp.	*Pseudomonas species*
RB	Riboflavin
RNaseA	Ribonucelase A
RT	Room Temperature
SDS	Sodium Dodecyl Sulfate
SEC	Size-Exclusion Chromatography
S. fonticola	*Serratia fonticola*
S. shabanensis	*Salinisphaera shabanensis*
SOC	Super Optimal Broth
STAS	Sulfate Transporter Anti-Sigma antagonist
TAE	Tris-Acetat-EDTA
TEAF	Triethylammonium Formate
Tris	Tris (hydroxymethyl)-aminomethan
UV	Ultra Violet
W	Wash
YESCA	Yeast Extract Casamino Acids

8. References

[1] J. M. Christie, J. Gawthorne, G. Young, N. J. Fraser and A. J. Roe, "LOV to BLUF: flavoprotein contributions to the optogenetic toolkit," *Molecular Plant 5(3),* pp. 533-544, May 2012.

[2] V. Jansen, J. F. Jikeli and D. Wachten, "How to control cyclic nucleotide signaling by light," *Current Opinion in Biotechnology 48,* pp. 15-20, December 2017.

[3] K. E. Brechun, K. M. Arndt and G. A. Woolley, "Strategies for the photo-control of endogenous protein activity," *Current Opinion in Structural Biology 45,* pp. 53-58, August 2017.

[4] J. Herrou and S. Crosson, "Function, structure, and mechanism in bacterial photosensory LOV proteins," *Nature Reviews Microbiology 9(10),* pp. 713-723, 8 August 2011.

[5] K. S. Conrad, C. C. Manahan and B. R. Crane, "Photochemistry of flavoprotein light sensors," *Nature Chemical Biology 10,* pp. 801-809, October 2014.

[6] A. Losi and W. Gärtner, "Old Chromophores, New Photoactivation Paradigms, Trendy Applications: Flavins in Blue Light-Sensing Photoreceptors," *Photochemistry and Photobiology 87(3),* pp. 491-510, May/June 2011.

[7] S. T. Glantz, E. J. Carpenter, M. Melkonian, K. H. Gardner, E. S. Boyden, G. Ka-Shu Wong and B. Y. Chow, "Functional and topological diversity of LOV domain photoreceptors," *PNAS 113(11),* pp. 1442-1451, 15 May 2016.

[8] S. Etzl, G. Gourinchas, E. Zenzmaier and A. Winkler, *Molecular mechanisms of light regulation in LOV-diguanylate cyclases,* 2017.

[9] N. De, M. Pirruccello, P. Violinova Krasteva, N. Bae and R. V. Raghavan, "Phosphorylation-Independent Regulation of the Diguanylate Cyclase WspR," *PLoS Biology 6(3),* p. e67, 25 March 2008.

[10] G. Gourinchas, S. Etzl, C. Göbl, U. Vide, T. Madl and A. Winkler, "Long-range allosteric signaling in red light-regulated diguanylyl cyclases," *Science Advances 3(3),* p. e1602498, 3 March 2017.

[11] T. Schirmer, "C-di-GMP Synthesis: Structural Aspects of Evolution, Catalysis and Regulation," *Journal of Molecular Biology 428(19),* pp. 3683-3701, 25 September 2016.

[12] Y. Chen, S. Liu, C. Liu, Y. Huang, K. Chi, T. Su, D. Zhu, J. Peng, Z. Xia, J. He, S. Xu, W. Hu and L. Gu, "Dcsbis (PA2771) from Pseudomonas aeruginosa is a highly active diguanylate cyclase with unique activity regulation," *Scientific Reports 6(29499),* 8 July 2016.

[13] G. Grigoryan and A. E. Keating, "Structural specificity in coiled-coil interactions," *Current Opinion in Structural Biology 18(4),* pp. 477-483, August 2008.

[14] J. T. Henry and S. Crosson, "Ligand binding PAS domains in a genomic, cellular, and structural context," *Annual Review of Microbiology 65,* pp. 261-286, October 2011.

[15] F. Zähringer, E. Lacanna, U. Jenal and T. Schirmer, "Structure and Signaling Mechanism of a Zinc-Sensory Diguanylate Cyclase," *Structrure 21(7),* pp. 1149-1157, 2 July 2013.

[16] H. Liu and J. H. Naismith, "An efficient one-step site-directed deletion, insertion, single and multiple-site plasmid mutagenesis protocol," *BMC Biotechnology 8(91),* 4 December 2008.

[17] D. Antoniani, P. Bocci, A. Maciąg, N. Raffaelli and P. Landini, "Monitoring of diguanylate cyclase activity and of cyclic-di-GMP biosynthesis by whole-cell assays suitable for high-throughput screening of biofilm inhibitors," *Applied Microbiology and Biotechnology 85(4),* pp. 1095-1104, January 2010.