Master's Thesis

Purification and characterization of the RfpA phytochrome of a thermophilic cyanobacterium

Syeda Josna Ahmed

University of Jyväskylä

Department of Biological and Environmental Science

31 MAY 2024

UNIVERSITY OF JYVÄSKYLÄ, Faculty of Mathematics and Science
Department of Biological and Environmental Science
Master's Degree Programme in Biological and Environmental Science

Ahmed Syeda J.

Purification and characterization of the RfpA phytochrome of a thermophilic cyanobacterium.

MSci Thesis
32 p., 1 appendix

Supervisors:
Doc. Heikki Takala and
Dr. Amit Srivastava

Tarkastajat:

[May 2024]

Keywords: Bathy phytochrome, photoreceptor, RfpA, UV-Vis spectroscopy.

Perceiving light is a crucial step for all photosynthetic organisms. This perception of light is mediated by a group of photoreceptor proteins called phytochrome that exhibit high responsiveness to light and govern other biological processes like growth and development in plants, bacteria, fungi, algae, and cyanobacteria. They utilize red light and far-red light to regulate different light-dependent cellular processes. Certain cyanobacteria contain a distinct phytochrome called RfpA, which enables them to carry out photosynthesis by harnessing far-red light. So far there is no structural-functional relationship is known about this protein.

The aim of this study was to purify RfpA to explore its signaling mechanism and spectral characterization to observe the photocycle of the phytochrome. Due to the difficulties in the purification of full-length protein, we constructed three different versions of the protein varying in length and containing different domains (GAF-PHY-PAS, GAF-PHY, and GAF only) using site-directed mutagenesis. Although our primary goal was the purification of the full-length protein, we attempted to obtain the longest possible fragment of the protein. We successfully managed to purify GAF-PHY and GAF mutants by affinity and size exclusion chromatography. Furthermore, UV-Vis spectroscopy was performed to study the photocycle of the protein.

Spectroscopy study highlights that RfpA protein exists predominantly in Pfr (Far-red light absorbing form) as its resting state, denoting the characteristics of a bathy phytochrome. These findings provide remarkable insight into the unique spectral properties of RfpA, which can be beneficial in developing optogenetic tools and biotechnological applications.

TABLE OF CONTENTS

TERMS AND ABBREVIATIONS

Terms

Bathy phytochromes	In contrast to prototypical phytochromes, the resting (stable) state of bathy phytochromes is Pfr. They are named after the bathochromic shifted absorption maximum of the Pfr state.
Chromophore	A chemical moiety that is attached to a photoreceptor through a covalent bond. It can change its conformation and absorb different wavelengths of light, undergoing reversible photoisomerization.
DpnI	It is a restriction endonuclease which is purified from *E. coli* but contains gene from *Diplococcus pneumoniae*, and cuts methylated plasmid.
DrBphP	A bacteriophytochrome that derives from *Deinococcus radiodurans*, is prototypical in nature.
EcoRI	Restriction enzyme that derives from *E. coli* strain RY13, cuts the DNA at specific restriction site.
Holoprotein	It is a complete, functional protein consisting of both the apoprotein (e.g., phytochrome) and associated non-protein component (e.g., chromophore).
RfpA	Histidine kinase associated photoreceptor in Rfp signaling components.
RfpB & RfpC	Two response regulators in Rfp signaling components.

Abbreviations

ARNT	Aryl hydrocarbon receptor nuclear transporter
CA	Chromatic acclimation
cGMP	Cyclic guanosine monophosphate
Cph2	Cyanobacterial phytochrome2
FaRLiP	Far-Red Light Photoacclimation response
FRL	Far-Red Light
GAF	cGMP phosphodiesterase/adenylate cyclase/FhlA
Ho1	Heme oxygenase1
kD	Kilodalton
Ni-NTA	Nickel nitriloacetic acid
OPM	Output module
PAS	Period/ARNT/single-minded
PBS	Phycobilisome
PCB	Phycocyanobilin
PCR	Polymerase chain reaction
PcyA	Phycocyanobilin:ferredoxin oxidoreductase
Pfr	Far-red light absorbing form of phytochrome
Pr	Red light absorbing form of phytochrome
PHY	Phytochrome-specific
PSI	Photosystem I
PSII	Photosystem II
PSM	Photo sensory module
Rfp	Regulator for far-red photoacclimation
SDS-PAGE	Sodium dodecyl-sulfate polyacrylamide gel electrophoresis

1 INTRODUCTION

1.1 Phytochromes

Phytochromes, a class of photoreceptor proteins exhibiting photosensitivity, are present in plants, algae, bacteria, and fungi. They are highly responsive to red and far-red light wavelength regions of visible spectrum (Smith 2000). They play a vital role in various aspects of development in plants and regulate seed germination, chlorophyll synthesis, seedling elongation, leaf morphology and movement, and the timing of flowering. Phytochromes are considered responsible for controlling the activity of enzymes primarily involved in metabolic pathways, particularly those associated with the biosynthesis of antioxidants and photosynthetic pigments. They also regulate the expression level of the phytochrome-based gene expression in downstream signaling pathways (Davis *et al.* 1999; Kreslavski *et al.* 2018). Generally, Phytochromes contain two identical chains, each containing an N-terminal core photosensory module (PSM) and a more variable C-terminal output module (OPM). The PSM consists of PAS (period/ARNT/single-minded), GAF (cGMP phosphodiesterase/adenylate cyclase/FhlA), and PHY (phytochrome-specific) domains (Rockwell *et al.* 2006). Phytochromes exist in two forms: the Pr form, which absorbs red light, and the Pfr form, absorbing far-red light and can be photoswitched between these two states (Murgida *et al.* 2007). Based on the form of the resting state, phytochromes are either canonical or bathy. In canonical phytochrome, the metastable Pr form resides in the ground state and is converted into Pfr form when exposed to red light, whereas, regarding bathy phytochrome, Pfr is the dominant form in the dark or ground state and undergoes a conversion of Pfr to Pr upon absorbing far-red light (Auldridge and Forest 2011; Velázquez Escobar *et al.* 2017). Typically, In the case of the Pr form, it absorbs red photons, and the absorbance shows a sharp maximum peak within the range of 650–670 nm and leads to photoconversion to the Pfr state. The "fr" indicates that the phytochrome is able to absorb far-red wavelengths of light, and the maximum absorbance peak of Pfr form exhibits at 705-730nm (Hughes 2010). Phytochromes also contain a bilin chromophore, a linear tetrapyrrolic structure which facilitates the photoresponse. The chromophore varies depending on the organism. For example, plants use phytochromobilin, while bacteria and cyanobacteria use biliverdin and phycocyanobilin respectively (Bhoo *et al.* 2001; Lamparter 2004). Bilin covalently binds with a conserved cysteine residue found in the bilin

binding pocket of the GAF domain in PSM (Ikeuchi and Ishizuka 2008; Burgie and Vierstra 2014). The conversion between Pr and Pfr form is mediated by the bilin chromophore upon absorption of light. The Z⇄E isomerization of the C15=C16 bond in the methine bridge connecting the C and D rings of the chromophore (Figure 1) initiates this photochromism (Takala *et al.* 2018).

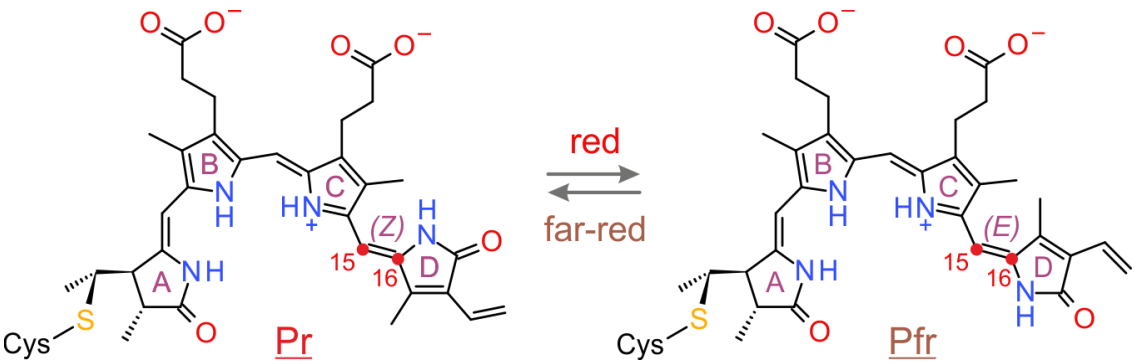

Figure 1. Structural changes in bilin chromophore of canonical phytochrome during photoconversion. Tetrapyrrole rings of chromophore are labeled as A-D, and other chromophore atoms are shown in the picture for clarity. Red and far-red light induces reversible isomerization between the C and D ring of the chromophore, particularly in the C15=C16 double bond, resulting in the flip of the D pyrrole ring. This rearrangement drives the photoconversion of phytochrome by transferring the signal into the downstream cascade (Möglich *et al.* 2010).

In environments rich in red light, such as full sunlight, cause isomerization in the D pyrrole ring of the bilin chromophore. This small change amplified throughout the rest of the protein that triggers rapid conformational alteration in the bilin-binding domain of phytochrome, which in turn drives the change in shape (Figure 2) of phytochrome. The GAF-PHY helix becomes straight, and the tongue of the PHY domain forms a helix structure and converts from Pr to Pfr form. The transition into Pfr hence transmits the signal to histidine kinase and response regulator proteins, ultimately leading to another transcriptional response and facilitating gene transcription involved in biological function (Takala *et al.* 2014; Pham *et al.* 2018).

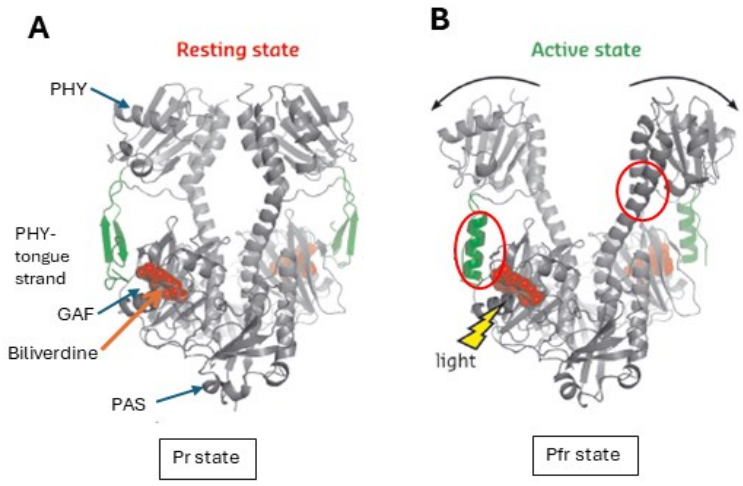

Figure 2. The architectural changes of the photosensory module (PSM) of phytochrome of *Deinococcus radiodurans* concurrent light exposure. The structural characteristics of the PSM in (A) Pr state (PDB id:4O0P) and (B) Pfr state (PDB id:4O01). GAF, PHY, PAS domain, position of bilin chromophore (biliverdin) are shown and marked with arrow. The red loop in figure B illustrates, that upon light exposure PHY tongue turns into a helix, meanwhile GAF-PHY helix shifts to straight form causing alteration in shape and transformation to Pfr form. (Takala *et al.* 2020).

1.2 Importance of cyanobacterial phytochromes

Cyanobacteria are prokaryotic organisms that are able to produce oxygen through photosynthesis. They can survive in widespread environments by utilizing different light wavelengths, including visible light to red light and far-red light in photosynthesis (Rockwell and Lagarias 2017). They also contain a complex set of photoreceptors that control their physiological and metabolic responses in exposure to light, hence regulating all the biological processes, for example, growth and adaptation in distinct environmental conditions (Wiltbank and Kehoe 2019). Cyanobacterial phytochromes along with response regulators, regulate several adaptation strategies that enable them to withstand different ambient conditions of light, named chromatic acclimation (CA) processes. Cyanobacteria are also designated as "green microbial factories" because they can make valuable products such as biofuels and bioplastics (Knoot *et al.* 2018; Zahra *et al.* 2020). These features make them more promising for solving environmental issues like the depletion of fossil fuels and environmental contamination of non-biodegradable plastics. One way to improve their ability to produce beneficial biochemicals is metabolic engineering (Lin and Pakrasi 2019; Pooja *et al.* 2023). It is necessary to obtain the right amount of enzymes for biofuel and bioplastic synthesis. Insufficient enzyme can result in a low yield of bioproduct, whereas

excess target enzyme can compromise the housekeeping function (Li *et al.* 2016; Jeschek *et al.* 2017; Srivastava *et al.* 2021). Therefore, we need precise ways to control enzyme production. Optogenetic systems, which use light to control gene expression, offer a solution (Lindner and Diepold 2022; Reshetnikov *et al.* 2022). By using these tools, we can improve the ability of cyanobacteria to make valuable products and make our biochemical production more sustainable. Cyanobacteria have a variety of light-sensitive proteins, such as phytochromes and cyanobacteriochromes, which make them good candidates for optogenetic tools. But to utilize the metabolic potential of cyanobacteria, a good understanding of the diverse phytochromes is essential (Sivaprakasam *et al.* 2021; Chia *et al.* 2022). Understanding their functional mechanism will help us design better tools that can reveal the full potential of cyanobacteria for biotechnology.

1.3 FaRLiP mechanism and RfpA phytochrome

Cyanobacteria are adaptable microorganism that can acclimate to changing light conditions (Ikeuchi and Ishizuka 2008; Fushimi and Narikawa 2019) . One important adaptation is their ability to remodel their photosynthetic machinery and produce specialized pigments like chlorophyll d (Chl *d*) and chlorophyll f (Chl *f*), which absorb far-red light (FRL; λ = 700–750 nm) (Trinugroho *et al.* 2020). Cyanobacteria proceed photosynthesis by using three major light-harvesting complexes: Photosystem I (PSI), Photosystem II (PSII), and Phycobilisomes (PBS). Some special cyanobacteria have the ability to modify their photosynthetic apparatuses and synthesize Chl *f* and/or Chl *d* in response to FRL, a phenomenon termed Far-Red Light Photoacclimation (FaRLiP) (Gan *et al.* 2014a; Gan and Bryant 2015). These adaptable cyanobacterial strains have specific genomic clusters containing paralogous genes encoding far-red light absorbing counterparts of PSI, PSII, and PBS. This adaptation gives an advantage to these cyanobacteria in environments rich in FRL but limited in white light, such as shaded habitats and canopies (Gan and Bryant 2015). *Leptolyngbya* sp. strain JSC-1 is a FaRLiP cyanobacteria, that contains a gene cluster containing 21 genes crucial for this acclimation (Gan *et al.* 2014a). In this cluster, a single gene encodes Chl f synthase (*chlF/psbA4*), alongside a conserved hypothetical protein, while the other genes code FRL-type PBS and core subunits of PSI and PSII (Ho *et al.* 2017a, b). Also, an operon having three genes—RfpA, RfpB, and RfpC—regulates far-red photoacclimation and is conserved across FaRLiP-capable strains (Zhao *et al.* 2015; Wolf and Blankenship 2019). This operon codes a knotless phytochrome (RfpA) and two response regulators (RfpB and RfpC). The functional role of this regulatory system is obtained from research on terrestrial cyanobacteria like *Chlorogloeopsis fritschii* PCC 9212, *Chroococcidiopsis thermalis* PCC 7203 (hereafter

PCC 7203), and marine cyanobacteria *Synechococcus* sp. PCC 7335 (Ho *et al.* 2017c, b; Shen *et al.* 2019). Recent mutational analyses of the *rfp* genes in these strains have shown their necessity in regulating FaRLiP-related genes, highlighting their role in the synthesis of Chl *f* and Chl *d* (Zhao *et al.* 2015, 2019). So far, the structure and function of Rfp signaling components of FaRLiP response is not known. The output histidine kinase (HK) domain of phytochromes contain two sub-domains: a histidine phosphotransfer (dimerization) domain and an ATP-binding catalytic domain (Möglich *et al.* 2009a). Typically, HKs catalyze autophosphorylation and transfer phosphate to their response regulators. However, HK domains can also contain phosphatase activity and they can hydrolyze the phospho-aspartyl bond in phosphorylated response regulators (Möglich *et al.* 2009b; Wahlgren *et al.* 2022). Moreover, Whether the RfpA histidine kinase has an autophosphorylation or phosphatase activity is also not clear (Figure 3). Besides, the process of signal transmission (phosphorelay) among Rfp proteins is anonymous yet (Zhao *et al.* 2015). In addition, like the knotless phytochrome Cph2, the RfpA phytochrome has a PAS domain at the interface of the PHY and histidine kinase (HK) domains (Figure 4)(Moon *et al.* 2011). The functional significance of this unique PAS domain location in RfpA is also still unidentified.

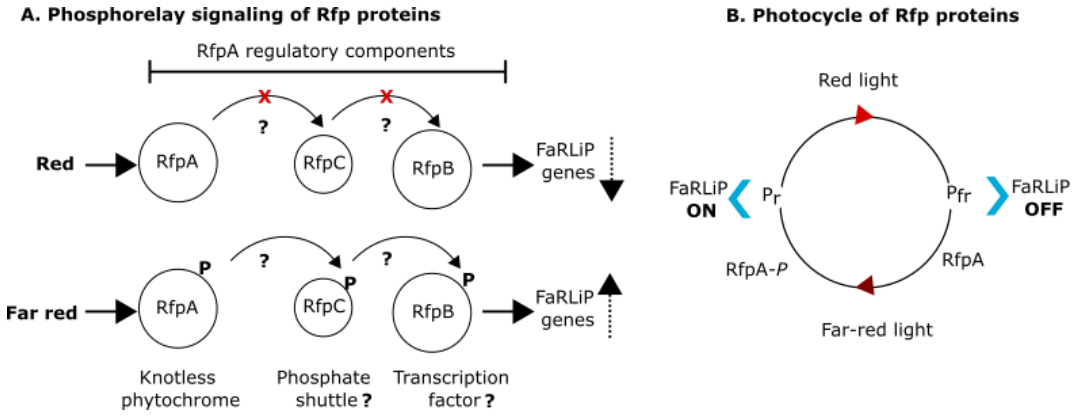

Figure 3. Phosphorelay Signaling of Rfp Proteins and the Photocycle of RfpA. Hypothesis suggests that Far-Red Light (FRL) either activates or deactivates the Histidine Kinase (HK) activity of RfpA. RfpC may act as a mediator for phosphate transfer between RfpA and RfpB. RfpB is presumed to regulate gene transcription within the Far-Red Light Photoacclimation (FaRLiP) cluster, potentially serving as a transcription factor. **(B)** RfpA undergoes photoconversion upon exposure to a wide spectrum of light wavelengths, transitioning into the FRL absorbing form (Pfr). However, only FRL induces the reverse photoconversion, returning RfpA to its red light absorbing form (Pr)(Gan *et al.* 2014a).

Understanding these mechanisms will improve our knowledge of cyanobacterial adaptation strategies. This deeper knowledge will help to promote advanced

techniques in the engineering of the Rfp system for biotechnological and optogenetic applications. In previous study, only RfpA GAF domain was purified, and the analysis revealed that RfpA is a phytochrome (Gan *et al.* 2014b), but the purification of the full-length protein and photocycle of the phytochrome was not accomplished which paves the way for the topic of this thesis. In the lab before, the purification of RfpA protein from *Synechococcus* sp. PCC 7335 was challenging as it was going into to the insoluble fraction after expression in *E. coli* BL21 cells. To solve this problem, it was decided to purify the RfpA protein from the thermophilic cyanobacterium *Chroococcidiopsis thermalis* PCC 7203. It is known that proteins from extremophiles are generally more stable in vitro (Zhao *et al.* 2015). While the full-length RfpA protein is necessary for studying such signaling mechanisms, it is hypothesized that any variant containing the GAF domain could be sufficient for studying the fundamental spectral characteristics and photocycle of RfpA (Rockwell and Lagarias 2024). Therefore, our approach was to generate various RfpA protein versions, and we focused on purifying the possible longest fragment.

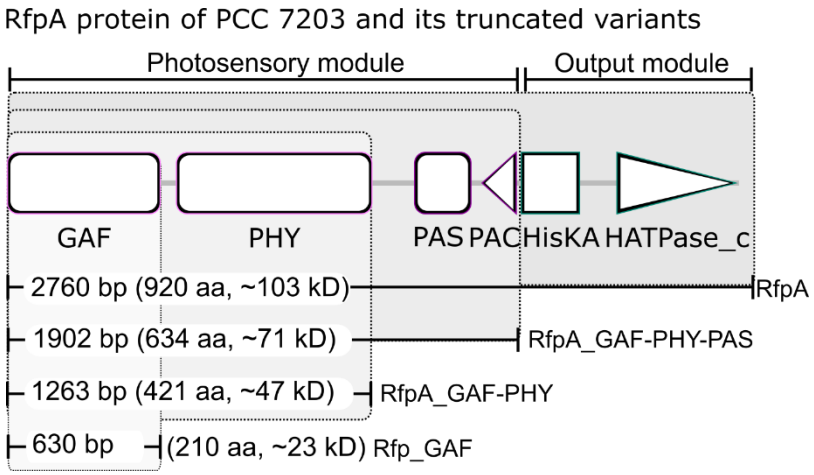

Figure 4. RfpA and its variants. The figure shows the size of RfpA proteins and the length of their nucleotide and polypeptide sequence, that are planned in this work. The protein size is mentioned in kDA (kilodaltons).

1.4 Aims of the study

This study aimed to perform the spectral characterization of the RfpA phytochrome and to find out the characteristics of the phytochrome, whether it is prototypical or bathy. The main objectives of the thesis were to construct different

variants of the RfpA protein (Figure 4) of the thermophilic cyanobacterium *Chroococcidiopsis thermalis* PCC 7203, and purify different variants of the protein, particularly the longest possible variants, and then conduct spectral characterization under light conditions, such as far-red, red light and darkness. The research questions that this work aimed to answer were:

(a) Is the RfpA protein derived from PCC 7203 soluble and stable in vitro? If not, what strategies can be used to solubilize the RfpA protein?

(b) Is RfpA a prototypical or bathy phytochrome?

2 MATERIALS AND METHODS

Obtaining a substantial volume of RfpA holoproteins by expressing the chromophore and RfpA phytochrome from two vector plasmids was challenging. To overcome this problem, it opted to co-express RfpA with its chromophore-phycocyanobilin (PCB) in a single expression vector because holoproteins are known to be more stable than apoprotein (Rockwell *et al.* 2023). To prepare this expression vector, used a plasmid, pDG282, containing the PCB operon with the *ho1* (heme oxygenase 1) and *pcyA* (phycocyanobilin) genes of *Synechocystis* sp. PCC 6803 and *phyB* gene of *Arabidopsis thaliana* (Ge *et al.* 2013; Golonka *et al.* 2019) was used. This plasmid was gift from Prof. Andreas Moglich from University of Bayreuth, Germany. Afterwards, the *phyB* gene was replaced with the *rfpA* gene from PCC 7203 by Dr. Amit Srivastava in our lab, that resulted in a single expression vector named pDG-RfpA, containing both *rfpA* and the PCB-expressing cassette. After construction the vector was confirmed by sequencing (Eurofins).

2.1 Plasmids and primers design

In order to create the variants of the protein, three EcoRI sites between the GAF-PHY, PHY-PAS, and PAS-HK domain junction were inserted (Figure 4). Insertion of EcoRI sites required the design of three specific mutagenic primers designated 58, 59, and 60 (Appendix Table A.1) for the construction of GAF, GAF-PHY, and GAF-PHY-PAS variants, respectively. The other EcoRI site was already present at the end of the C-terminal HK domain of the RfpA protein-encoding gene (Figure 5). Insertion of an additional EcoRI site in the domain junction allowed precise DNA cleavage at these sites using EcoRI restriction enzyme (Fermentas) and resulted in our desired plasmid. Primer designing was accomplished

according to the guidelines of the QuikChange Lightening Multi Site-Directed Mutagenesis kit (Agilent technology). Benchling, a cloud-based platform for biotechnology research, was used to design the primers. Benchling facilitates visualizing and tracking primer binding sites on a target sequence, streamlining the process of attaching single primers or linking them in pairs. After completing designing of the primers, they were synthesized by the company Eurofins Scientific. Also, sequencing primers were fabricated and later used to confirm the final constructs.

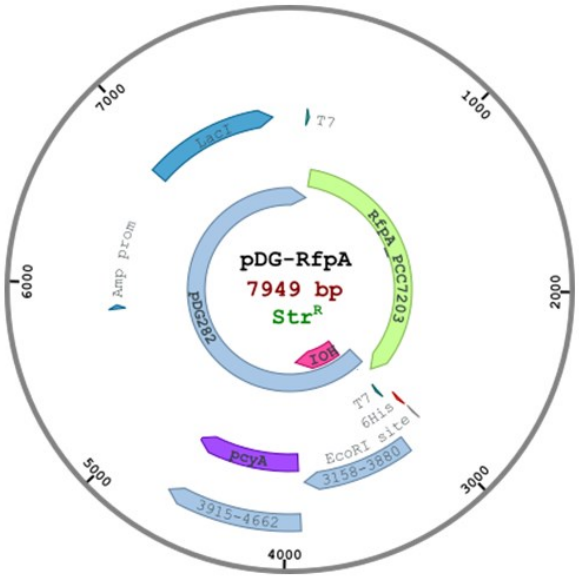

Figure 5. Vector map of pDG-RfpA plasmid. The size of the plasmid is 7949 bp or 7.9 kb. The vector map shows the presence of the EcoRI site in the C-terminal region of the RfpA gene, and the plasmid is streptomycin streptomycin-resistant gene as a marker. HoI and pcyA are the phycocyanobilin chromophore encoding sequences. 6 His gene produces a polyhistidine tag in the protein that works as a remark of the protein in purification. The positions 3158-3880 and 3915-4662 marked in the figure are the positions of CDS (Coding sequence) that contain start and stop codons for protein synthesis. Other markings, such as T7 promoter, Ampicillin promoter (Amp prom), and Lactose suppressor (LacI) marked for clarity.

2.2 Site-directed mutagenesis

2.2.1 PCR amplification and transformation of mutated plasmids

The QuikChange Lighting Multi Site-Directed Mutagenesis kit was used to perform site-directed mutagenesis. Three PCR reaction was incubated using pDG-RfpA as a template and mutagenic primers (Appendix 1) as oligomer by

following the protocol provided with the mutagenesis kit. The thermal cycling parameters (Table 1) was adjusted in a PCR thermocycler (Bio-Rad) and the reactions were set for 3 hours.

TABLE 1. PCR cycle conditions and parameters (30 cycle).

	Temperature (°c)	Time
Initial denaturation	95	2 min
Denaturation	95	20 sec
Annealing	55	30 sec
Extension	65	4 min
Final extention	65	5 min

PCR cloning resulted in recombinant plasmids containing the EcoRI site (GAATTC, 5'-3'direction) at the domain junction (Figure 6). After PCR amplification, 1µl of DpnI restriction enzyme (Thermofisher scientific) was added to each PCR reaction mixture to eliminate the original template DNA to ensure the pure mutated plasmid. DpnI cuts exclusively the methylated plasmid. The previous templates were already methylated as it was a part of *E coli.* cell, but the newly cloned plasmid was not, and therefore, did not cleave with DpnI. The mutated plasmids were further transformed into *E. coli* DH5α using the heat shock method. Apart from the mutated plasmids, one sample with the pDG-RfpA vector plasmid was transformed as a positive control. 5µl of each mutated plasmid and control sample was added in each appropriately leveled tube containing 100µl of E. coli DH5α competent cells. These samples were then subjected to heat shock for 90 seconds at 42°C, following incubation on ice for 30 minutes to facilitate plasmid DNA uptake. Following this procedure, samples were placed on ice for 5 minutes immediately. 900µl of fresh LB medium was added to each sample and then further incubated in a shaker incubator (New Brunswick Excella E24 Incubator Shaker Series) for 30 minutes at 37°C while shaking at 200 rpm. Antibiotic-resistance genes were expressed during this incubation period, and cells recovered after the heat shock. The samples were centrifuged subsequently for 10 minutes at 10000 rpm, and 900µl of supernatant was discarded. The pellet of each sample was resuspended in the remaining 100µl of supernatant inside a laminar hood and plated onto a pre-leveled selective LB plate with 50µg/ml streptomycin. The plates were further incubated at 37°C overnight.

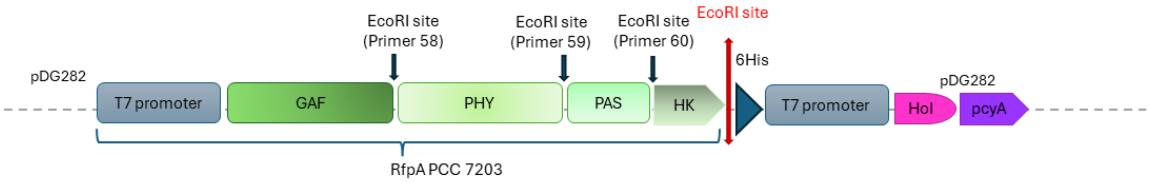

Figure 6. Site-directed mutagenesis of RfpA to construct recombinant plasmid. The figure illustrates the position of the T7 promoters and the domain architecture of RfpA PCC 7203 in the plasmid. The EcoRI site with red color was present previously in the vector expression plasmid. Other EcoRI sites (colored in black) were inserted into the plasmid by performing mutagenesis using specifically designed primers, constructing three recombinant plasmids. When treated with the EcoRI restriction enzyme, it cleaved the DNA at these specific sites.

2.2.2 Restriction digestion of mutated plasmids and ligation

The existence of desired mutations in the resulting plasmids was confirmed through screening assays by restriction digestion (EcoRI digestion). The grown mutated plasmids from the transformed plates were isolated using QIAGEN Plasmid kits (QIAprep spin miniprep kit) to perform the restriction digestion reaction. In the case of isolation, for every construct, three colonies were picked from one plate (a total of 9 colonies). For isolating the plasmids, fresh cultures were prepared separately for each colony, containing 5ml of LB medium, 5µl of streptomycin stock solution (50mg/ml), and the picked colony from transformed plates. The cultures of each plasmid colony were centrifuged at 5000 rpm for 5 minutes at 4°C. After discarding the supernatant, different resuspension and lysis buffers (250µl of p1, 250µl of p2, and 350µl N3) that were provided with the isolation kit were then added to the pellet according to the instruction in the kit protocol and centrifuged after resuspending the pellet at 13000 rpm for 10 minutes. Centrifugation was performed again after transferring the supernatant into the QIAGEN column tube at 10000 rpm for 1 minute. Then, the other two buffer solutions (500µl of pB, and 750µl of pE) of the kit, were added respectively, and subsequent centrifugation was executed by following the protocol. The final plasmid solution was collected in a new microcentrifuge tube by adding 50µl of water to the column, performing centrifugation further 2 to 3 minutes at 10000 rpm, and discarding the column.

After the isolation, restriction digestion reactions following quantifying protein concentration with nanodrop quantification (Thermo scientific) were set for each sample (total 9) by adding 1µl of EcoRI restriction enzyme (Fermentas) along with 2µl of 10x fast digestion buffer and 12µl of water into the 5µl of isolated mutated plasmids and incubated at 37°c for 30 minutes. The EcoRI

enzyme cuts the DNA by recognizing specific sites (Figure 6) that are already present in the plasmid. Gel electrophoresis was performed on the plasmids that were digested. 1% agarose gel, prepared using agarose (2gm) and TAE buffer (200ml) containing ethidium bromide, was used for the gel electrophoresis. The gel images were recorded by Bio-Rad ChemiDoc MP Imaging System. Following the gel electrophoresis process, the target DNA bands were carefully excised and subsequently purified by employing the GeneJET Plasmid Miniprep kit and following the protocols of the kit. The DNA gel bands were measured and then dissolved with the same amount of binding buffer from the GeneJET kit for each band at 60°C in heat block (Thermo scientific). Then, the dissolved gel samples were transferred to the column tubes and centrifuged at 10000 rpm for 1 minute. Following further centrifugation with another 100µl of binding buffer, the columns were washed with 700µl of wash buffer, and finally, the plasmid was collected by adding 50µl of aliquot distilled water to the column and again centrifuged with the same speed for 2 minutes.

The purified plasmid solutions were then assembled for self-ligation for 1 hours at 20°C using 1µl of T4 DNA ligase (ThermoFisher) enzyme along with 10x T4 DNA ligase buffer and nuclease free water according to the protocol provided with the enzyme. After the incubation period, the T4 DNA ligase enzyme was heat inactivated at 65°C for 10 minutes and again the plasmids were transformed into *E. coli* DH5α using LB agar plates containing streptomycin (5µg/ml, stock 50mg/ml). Sequencing of the grown colonies was performed to confirm the final constructs.

2.3 Protein expression

The expression vectors carrying specific *rfpA* gene mutants (GAF, GAF-PHY, GAF-PHY-PAS) were transformed into *E. coli*BL21(DE3) cells by heat shock method. For the transformation, LB agar plate with streptomycin (50µg/ml, stock 50mg/ml) was used. Two starter cultures were prepared by introducing a single colony from transformation plates with grown colonies (3 colonies from 3 plates) into 5ml LB medium containing 5µl streptomycin (1µl/ml, stock 50mg/ml) for each mutant (in total 6 starter culture) and left it to grow in shaker incubator for 18 hours at 37°C, 200 rpm. Eight 2L Erlenmeyers containing 500 ml of fresh LB medium supplemented with 500µl of streptomycin (1µl/ml, stock 50mg/ml) were prepared for protein expression of each mutant. 1ml from the starter culture was inoculated per Erlenmeyers and incubated in a shaking incubator (New Brunswick Excella E25 Incubator Shaker Series, at 37°C, 230 rpm) and monitored the optical density (OD600) after 3 hours. When the OD600 reached 0.6, an SDS-PAGE sample was taken, and then protein expression was

induced by adding 500µl of 1M IPTG (concentration in culture 1mM) to the culture. IPTG induced the expression of the target protein by triggering the production of T7 RNA polymerase that started the transcription of the target gene in the plasmid. The cells were then allowed to grow under this induced condition overnight at 18°C, 230 rpm, facilitating target protein production. To prevent light-induced effects on cells, expressing light-sensitive PCB, all procedures were conducted in darkness and under safe green light conditions (Multamäki *et al.* 2021).

After overnight incubation, cells were harvested via centrifugation to yield culture. Cultures were collected in six 1L centrifugation bottles. The centrifugation of the cultures was done at 6000rpm for 15 minutes at 4°C (Thermo Scientific, Sorval RC6+, and F9S-4x1000Y rotor). After discarding the supernatant, 15 ml of buffer (30mM Tris, 150mM NaCl, pH 8.0) was added to the bacterial pellet to dissolve the pellet, and then the solutions were transferred into 50ml falcon tubes. The suspensions were stored at -80°C for further processing for purification.

2.4 Protein purification

2.4.1 Cell lysis and harvesting

The culture cells were lysed by applying high pressure in the Avestin Emulsiflex-C3 homogenizer system to liberate the RfpA holoprotein. Previously collected cultures were diluted with buffer and introduced into the sample cylinder, followed by thorough rinsing with water. Subsequently, pressure gradually increased to 1000-1500 bars. Under this pressure, bacterial cells were lysed through two cycles of circulating suspension before the lysate was collected. The cell lysate was then centrifuged using an RC6+ centrifuge and F21S-8×50 rotor at 20,000 rpm and 4°C for 30 minutes, yielding a supernatant ready for purification. Prior to purification, a solubility test was conducted via SDS-PAGE analysis to confirm the presence of the target protein in the soluble fraction of the cell lysate. SDS-PAGE samples were collected from supernatant and pellets.

2.4.2 Affinity chromatography

To achieve expressed protein with high purity, two stages of purification have been proceeded. The first step in purifying the protein involved affinity chromatography, utilizing a Bio-rad FPLC system and a Ni-NTA affinity chromatography HisTrap™ column (GE healthcare). This column specifically binds to the His-tagged protein. A six His-tag encoding gene was located in the

vector along with desired protein sequences which facilitate the production of recombinant proteins containing poly His-tag at the C-terminal. Therefore, when passing through the column, the protein binds with the nickel resin and gets separated. During purification, a lysis buffer or buffer A (Table 2), also known as wash buffer, was employed to flush out other proteins from the column, thereby facilitating the binding of the desired protein to nickel. An elution buffer or buffer B (Table 2) containing imidazole was then used to collect the protein as a fraction.

TABLE 2. Buffers and reagents, used in protein purification. Ethanol was used to store the column and system after purification. All the reagents were de-gassed before use in the system.

Buffer A or lysis buffer	30mM Tris, 150mM NaCl, pH 8.0
Buffer B or elution buffer	30mM Tris, 150mM NaCl, 500mM Imidazole, pH 8.0
SEC buffer	30mM Tris, 150mM NaCl, pH 8.0
20% ethanol	
H_2O	

A manual program was executed with a flow rate of 0.5ml/minute and a maximum pressure of 0.7MPa while passing the sample through the column. As the sample passed through the column, the mutant proteins containing His-tag were selectively bound to the column, while non-specific molecules crossed over. Subsequently, the column was washed with buffer A (Table 2) to eliminate any loosely bound or non-specific molecules. Following this, buffer B (Table 2) was introduced to the column containing imidazole, a structural analog of histidine that detached the bound protein from the column and eluted it. The protein fraction was chosen based on the chromatogram. The fractions showing absorption at 710 nm were known to contain phytochrome, which was collected and then subjected to analysis using SDS-PAGE. The fractions that contain the phytochrome were pooled and diluted the pool with 50ml SEC buffer (Table 2). Also, the concentration of the protein was measured by Nanodrop One (Thermo scientific). The diluted pools were concentrated to approximately 3ml, using an Amicon ultra 10 kDa concentration tube at 4800 rpm, 10 minutes, +4°c.

2.4.3 Size exclusion chromatography

For further ultra-purification, size-exclusion chromatography (SEC) was employed. This method effectively separates proteins based on their size and shape, allowing for the removal of any remaining impurities. The column used in size exclusion contains pores within its matrix that selectively trap smaller molecules. Consequently, larger molecules elute first, followed by smaller molecules. SEC buffer (Table 2) was utilized for protein separation, and the

column used was HiLoad™ 26/600 Superdex™ 200 pg (GE Healthcare). The column was equilibrated overnight with SEC buffer. Before injecting into the column, the sample was filtered by using 0.45µm syringe filter. The elution profile obtained from the SEC analysis displayed distinct peaks corresponding to different molecular weights. The effluent was collected based on the chromatograms, and like affinity chromatography, the fraction that showed absorbance at 710nm was collected. Throughout the run, absorbance was measured at wavelengths of 280nm, 405nm, 645nm, and 710nm to monitor protein elution. To confirm the proper separation of proteins, SDS-PAGE analysis of collected fractions was conducted, providing assurance of the purification process, and the protein concentration was also measured by Nanodrop.

Purified protein solutions were concentrated using Amicon ultra 10 kDa concentration tube to the volume of 3ml. Concentrated proteins were aliquoted to 200µl patches into light-safe tubes. These tubes were flash-frozen using liquid nitrogen and stored at -80°C.

Before the SDS-PAGE run, samples were exposed to heat treatment for 10 minutes at +98°C after adding 2x dye, which makes the protein liner and aids in determining the proper band of the protein. Ready-made 12% polyacrylamide SDS (Biorad Mini-Protean TGX Gels) gels were used with 1x SDS loading buffer (Stock solution in 10x – 0.25 M Tris, 1.92 M Glycine, 1% (w/v) SDS, pH 8.3). The loaded gels with samples were run using Bio-Rad cell at 200v for 40 minutes. After removing the gel, it was stained with 0.1% Coomassie blue dye for 15 minutes, destained with destaining solution, and left for 3-4 days to remove the dye. Finally, the gel picture was captured by Bio-Rad ChemiDoc MP Imaging System.

2.5 UV-Vis spectroscopy and spectral analysis

For analyzing the spectral properties of RfpA and its variants, UV-Vis (ultraviolet-visible) absorption spectroscopy was performed within the range of 240–850 nm. For investigating the photocycle of RfpA variants, absorption spectra were measured under various conditions, including dark, far-red light, and red light conditions, by using a Cary 8454 UV/vis spectrometer (Agilent Technologies) in a 10 mm quartz cuvette (110-QS type). Since two variants of RfpA (GAF variants and GAF-PHY variants) were successfully purified, spectral analysis was carried out on these variants. Continuous illumination with 655 nm (red, 7-milliwatt output power) and 785 nm (far-red, 9-milliwatt output power) Lasers were applied, with absorption measurements taken promptly. The entire procedure was conducted under green light to minimize any unintended photochemical reactions. A prototypical phytochrome, DrBphP, from *Deinococcus*

radiodurans served as a control for comparison. For preparing the samples, the absorption ratio of 280 nm to 700 nm of the protein was calculated and diluted the concentrated protein to adjust the absorbance value for 700 nm at approximately 0.2. SEC buffer (30mM Tris, 150mM NaCl, pH 8.0) was used as a solvent for dilution for both DrBphP and constructed mutants. For each sample including control, initial absorption measurements were taken in the dark without light exposure prior to red light or far-red light illumination. Different adapters were used for red light and far-red light conditions. The light illumination procedure was executed as follows.

Regarding DrBphP, reference measurement was done first, which measured the spectra of cuvette containing only buffer (990µl) to autozero the base level. After that, the control sample DrBphP (10µl) was added to the buffer, and spectra were measured. The control sample was in a resting state because there had been no light exposure before. Following this step, it was illuminated by far-red light with a 785 nm wavelength and measured for spectra. Subsequently, after collecting spectra, it was again exposed to red light at 655 nm for 3 minutes, and then the sample was left in the dark for an hour to facilitate dark reversion. After this incubation period, it was again measured to obtain the spectra of the dark state. Spectra were recorded at 1 minute and 3 minutes for both illuminations. Following red light and far-red light illumination, time was allocated for dark reversion of the sample.

In the case of the GAF-PHY mutant, the spectrum of buffer solely was measured before starting with the sample to serve as blank and baseline data. The sample was diluted by adding the sample (25µl) to the buffer (975µl). The protein variant was initially left in the dark for 10 minutes, and then spectrum was taken. After this, the sample was exposed to far-red light at 785 nm for 40 minutes and incubated again in the dark for 20 minutes. Subsequently, the variants were exposed to red light for 6 minutes and kept in darkness for a further 30 minutes for dark reversion. For every exposure, including dark incubation, spectra were recorded at 1 minute, 3 minutes, and then every 5 minutes to get a proper average value for each exposure. The exposure time was decided until there were changes in spectra for each case.

For GAF variants, the same procedure as GAF-PHY was followed, but the exposure time was slightly different. For instance, far-red light was exposed for 32 minutes, and the sample was left in the dark after that for 16 minutes. The red light exposure time was 20 minutes, which was further followed by 10 minutes of dark incubation.

2.6 Data analysis

The analysis of plasmid sequences and sequencing results was conducted using Benchling and Chromas software (Version 2.6.1), respectively. Data obtained from the Nanodrop, and chromatography results were compiled into tables and analyzed using Microsoft Excel and GraphPad software (Version 10).

The images of agarose gel and SDS polyacrylamide gel were analyzed utilizing the Bio-Rad ChemiDoc MP Imaging System. Spectral analysis data was calculated using Microsoft Excel (Microsoft office 11), and graphs were generated using GraphPad (GraphPad Prism 10.2.1.395) and origin (Origin 2024SR1) software. The spectral data were simply normalized before graphing.

3 RESULTS

3.1 Site-directed mutagenesis was performed successfully

After cloning, the transformed plates that were incubated overnight contained a few colonies (approximately ~ 5–10 colonies) in mutant plates, whereas the positive control plate was full of bacterial colonies. The colony size in the mutant plates was almost the same in every plate. Three colonies from each plate were picked and subjected to restriction digestion with the EcoRI restriction enzyme. The results of the GAF-PHY and GAF-PHY-PAS mutant plasmids for all colonies showed expected digestion, while one colony from GAF showed unusual digestion. Therefore, for GAF-PHY and GAF-PHY-PAS, a random colony was taken from those three colonies, and for GAF, one colony was chosen between the two proper colonies that were tested before and performed restriction digestion and gel electrophoresis to check the digestion again and perform further procedure. The EcoRI restriction enzyme cleaves the DNA at the EcoRI site and produces two cleaved portions: the unwanted domains and the rest of the plasmid (our desired one). These cleaved portions show specific bands according to their size when run on 1% agarose gel electrophoresis. The expected sizes of the plasmids of the GAF, GAF-PHY, and GAF-PHY-PAS variants were 5.8 kb, 6.4 kb, and 7.1 kb, respectively. The gel electrophoresis result (Figure 7) shows approximate band sizes (highlighted with red arrow) for these three recombinant plasmids. On the other hand, bands marked with green arrows indicate the unwanted domain of the protein that was cleaved out by the restriction enzyme. In gel electrophoresis, lighter fragments go fast through the gel and show the band. For GAF-PHY-PAS, the excised band is smaller than the

GAF variant, and it is shown clearly in the gel picture, which indicates that the digestion of the plasmids was done appropriately.

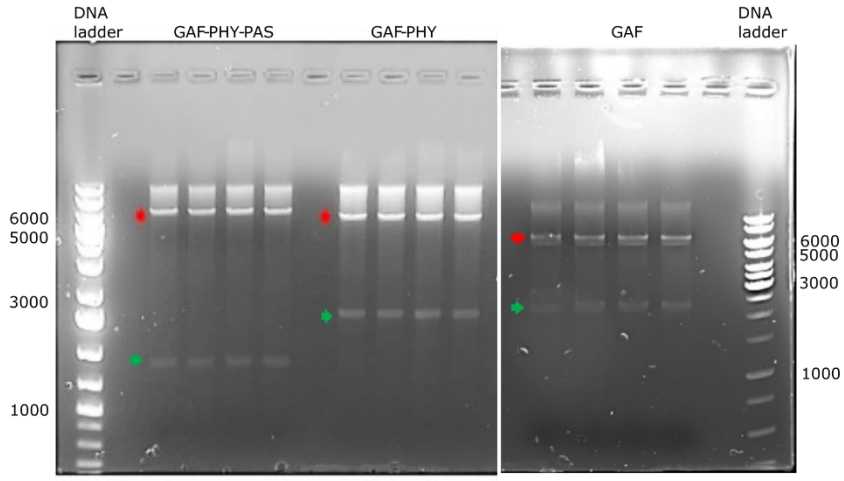

Figure 7. Gel electrophoresis analysis of RfpA variants after restriction digestion. The gel figure illustrates the band size of the DNA fragments. The red-marked bands were the expected recombinant plasmids, and the green-marked bands showed the excise domains. 1kb DNA ladder was used as a marker for assessing the band size. The expected sizes of the GAF-PHY-PAS, GAF-PHY, and GAF were 7100 bp, 6400 bp, and 5800 bp, respectively, and the digested plasmids showed bands near these sizes in the gel picture.

3.2 Expression and purification of recombinant proteins

All the mutants (GAF, GAF-PHY, and GAF-PHY-PAS) and vector expression plasmids were cultured for protein expression. Although the GAF and GAF-PHY mutants were expressed successfully, the full-length protein and the GAF-PHY-PAS mutant proteins were not. After induction of expression with IPTG, all variants showed expression, but during lysis of the cells, full-length protein and the GAF-PHY-PAS version of protein aggregated in the cell pellet. Therefore, GAF and GAF-PHY mutants were considered for further purification. Both steps of purification affinity and size exclusion chromatography for the GAF-PHY mutant proceeded successfully. The expected molecular size was approximately 47 kD for GAF-PHY. This value was retrieved from the mutant sequence and verified by Benchling and the ProtParam tool in the ExPASy online server. In SDS-PAGE analysis, the results showed the accuracy of the expression, purification, and yield of protein. After breaking the cell, the protein was collected from the yield of cell breakage and centrifuged. Although a large proportion of protein remained in the pellet with the degraded bacterial cell (Figure 8A-1), a good amount of protein was collected efficiently from the

supernatant (Figure 8A-2). The gel picture of the elution sample (Figure 8A-3) shows that pure phytochrome was not obtained from affinity chromatography as there were also other proteins that remained with the phytochrome. However, the amount of phytochrome at the approximate size site indicates successful isolation of the phytochrome by affinity chromatography.

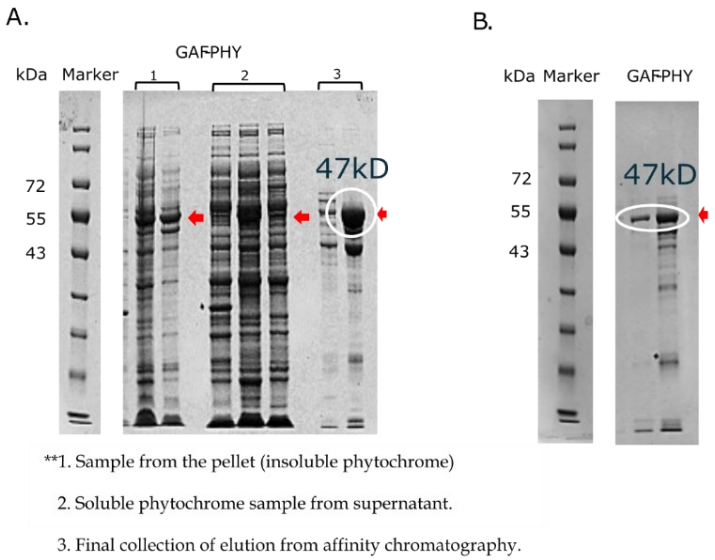

Figure 8. SDS-PAGE analysis of sample from affinity and size exclusion chromatography of GAF-PHY mutants. In figure A SDS gel contains different samples from various stages of expression and purification and figure B illustrates analysis of elution collected after size exclusion chromatography. Sample 1 in figure A was collected from pellet after centrifugation of yield product of cell lysate and sample 2 was from supernatant. Sample 3 was pipetted from pooled fraction after affinity chromatography. The expected size of the GAF-PHY mutant protein was 47kD and in SDS analysis showed a band of nearby size.

Size exclusion chromatography separates the molecules based on their size more precisely than affinity chromatography. The gel picture (Figure 8B) of the elution from SEC illustrates the proper purification of the phytochrome as the band showed molecular size near the expected one. The elution from affinity and SEC chromatography was collected by observing the chromatogram peaks (Figure 9) during the run. Those fractions that showed absorbance at 710nm wavelength were pooled. During affinity chromatography, the protein started to elute in 50 ml. Also, the ratio of 280nm and 710nm absorbance in SEC was close to 1 (Figure 9B). Finally, the fraction pools collected after SEC were concentrated to 20.24mg/ml and made aliquots of 200 µl.

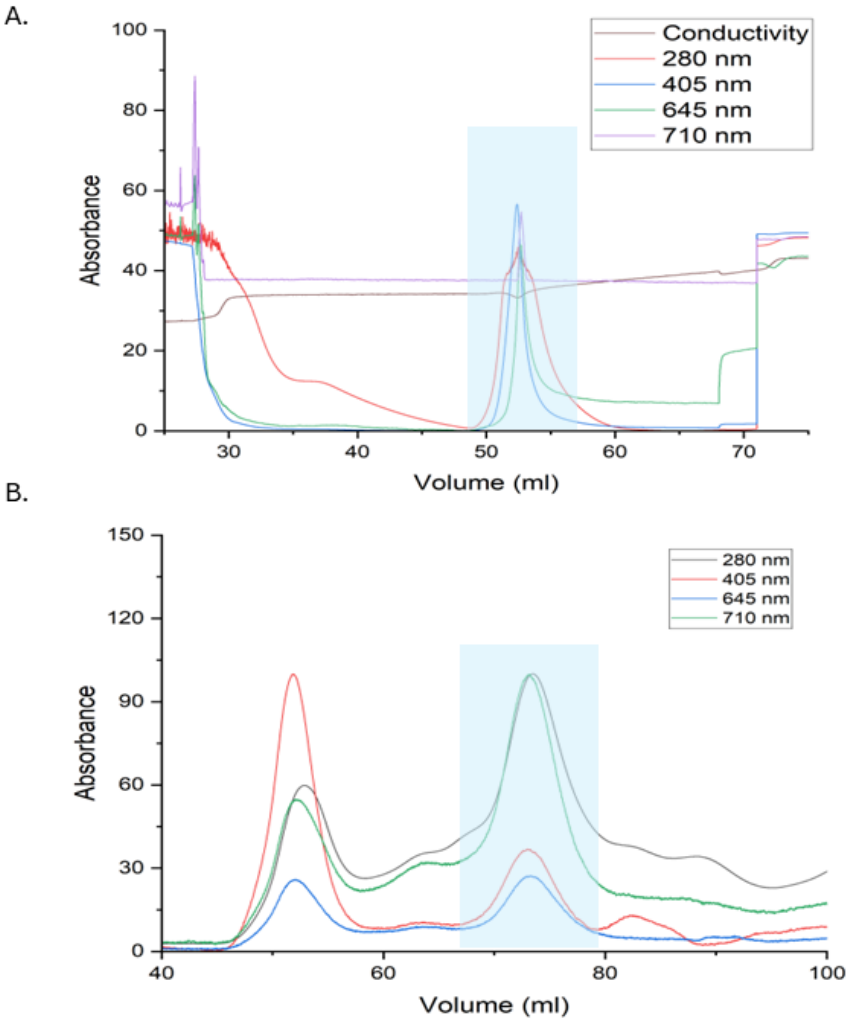

Figure 9. Affinity and Size separation chromatogram of GAF-PHY recombinant. (A) Affinity chromatogram illustrated the collected fractions for size exclusion chromatography (marked with blue bar). (B) Fractions from 710 nm absorbance (blue bar) were collected also for SEC.

With regards to GAF mutant, the expected size was 23 kDa. The value was also retrieved from Benchling and the expacy tool. The SDS-PAGE analysis of the fraction pool collected from SEC purification (Figure 10A) showed the approximate size of GAF variants, which points to the successful purification of the mutants. During the affinity chromatography, the protein started to emerge in 102 ml (Figure 10B). Also, those fractions were pooled which showed absorbance at 710 nm for both affinity and SEC chromatography. In size exclusion chromatography, the chromatogram (Figure 10C) indicates that a good amount of protein was purified, and the quantity was better than the GAF-PHY mutant. After the SEC, the purified protein was stored in aliquots at a concentration of 40.90 mg/mL, 200 µl per aliquot.

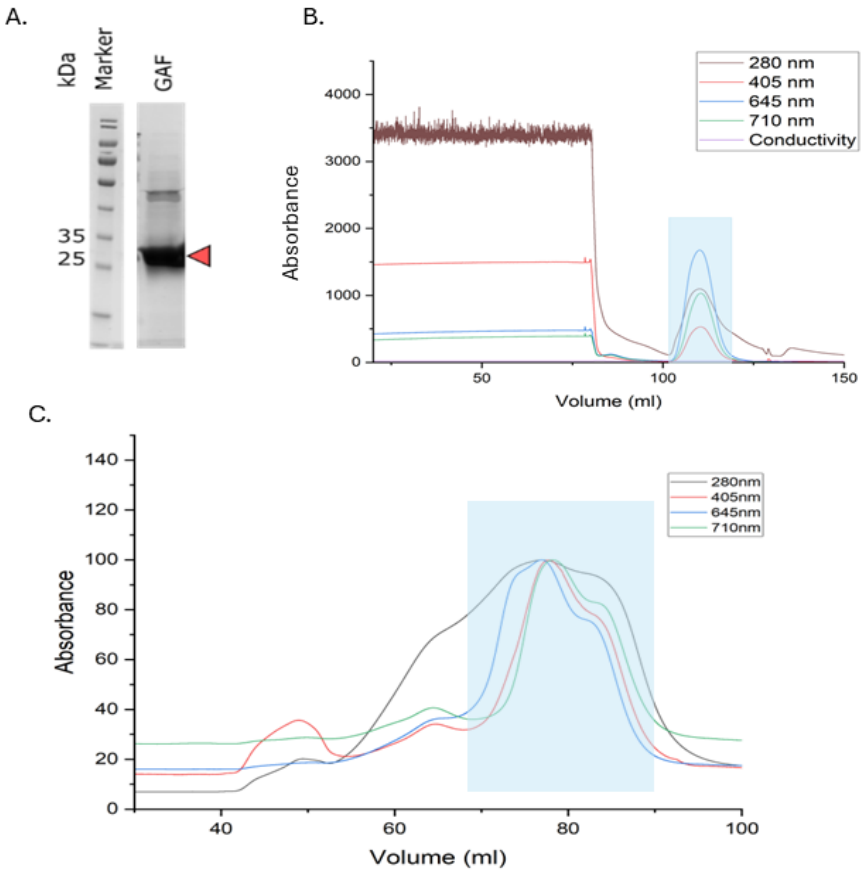

Figure 10. SDS-PAGE analysis and chromatogram of affinity and size exclusion chromatography of the GAF mutant. (A) SDS-PAGE of the elution after size exclusion purification showed a band near the expected size (23 kDa) for GAF, pointing to successful purification. (B) Chromatogram of Affinity Chromatography The collected fractions are highlighted in the blue bar. (C) Chromatogram of SEC. The figure denotes the pooled fractions (blue bar).

3.3 Spectral characterization of RfpA mutants

3.3.1 DrBphP

DrBphP is a prototypical phytochrome; hence, it has a red light absorbing form, or Pr form, as a ground or dark state. Because it has the Pr form in its ground state, it is only photoactive when it is exposed to red light and changes the form to Pfr. The spectral analysis (Figure 11) also shows identical results for the control. Early spectra without light exposure showed absorbance at 700 nm, which is the Pr form for DrBphP. Exposure to the far-red light did not show any

changes in absorbance wavelength, but when it was exposed to red light for 3 minutes, photoconversion occurred, and absorbance showed at 765 nm, which is the Pfr form maxima of DrBphP. Final dark incubation (1 hour 17 minutes) did not change the form of the phytochrome (Figure 11) because dark reversion of DrBphP requires approximately 24 hours.

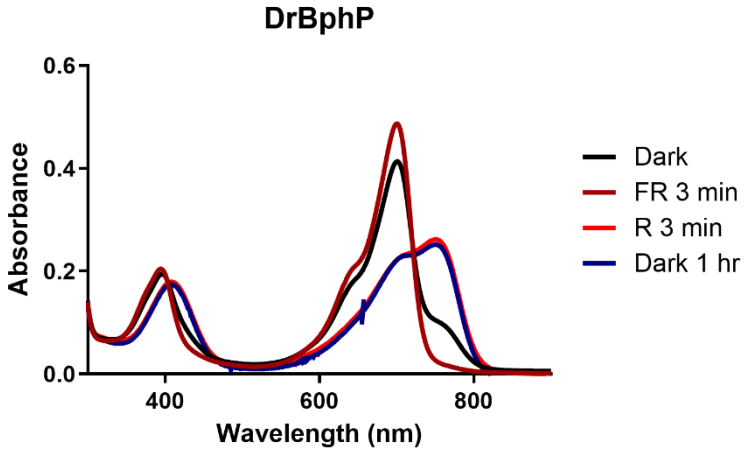

Figure 11. Absorption spectra of the phytochrome of *Deinococcou radiodurans*. Spectra without any light exposure (dark in black) reveal the ground state of the phytochrome. The figure illustrates that the ground state is Pr, as the absorbance occurred around 700 nm. Illumination with 785 nm (FR 3 minutes) did not change the spectra, but 685 nm illumination (R 3 minutes) showed different absorbance (>750 nm), indicating the change of the form into Pfr.

3.3.2 RfpA GAF-PHY

Initially, spectra were measured without light exposure in the dark for 10 minutes (Figure 12A), which showed mixed absorbance at 646 nm and 703 nm. The sample was illuminated with 785 nm (far-red light) and recorded the spectra within every 5 minutes. The spectra showed absorbance maxima at 646 nm, which required 40 minutes of continuous illumination and reported significant photoswitches to Pr form. The sample was left again in the dark for 20 minutes. Later, the variants were exposed to 655 nm (red light) wavelength Laser for 6 minutes, and the spectra changed to 703 nm (Figure 12A) immidiately, which denotes the Pfr form. Following this step, the spectra were taken further in the dark. The dark incubation after both exposures did not demonstrate any change in spectra (Figure 12B, 12C). In the ground state, the absorbance showed the

absorbance maxima at 703 nm, which suggests that this phytochrome has Pfr form as its ground state.

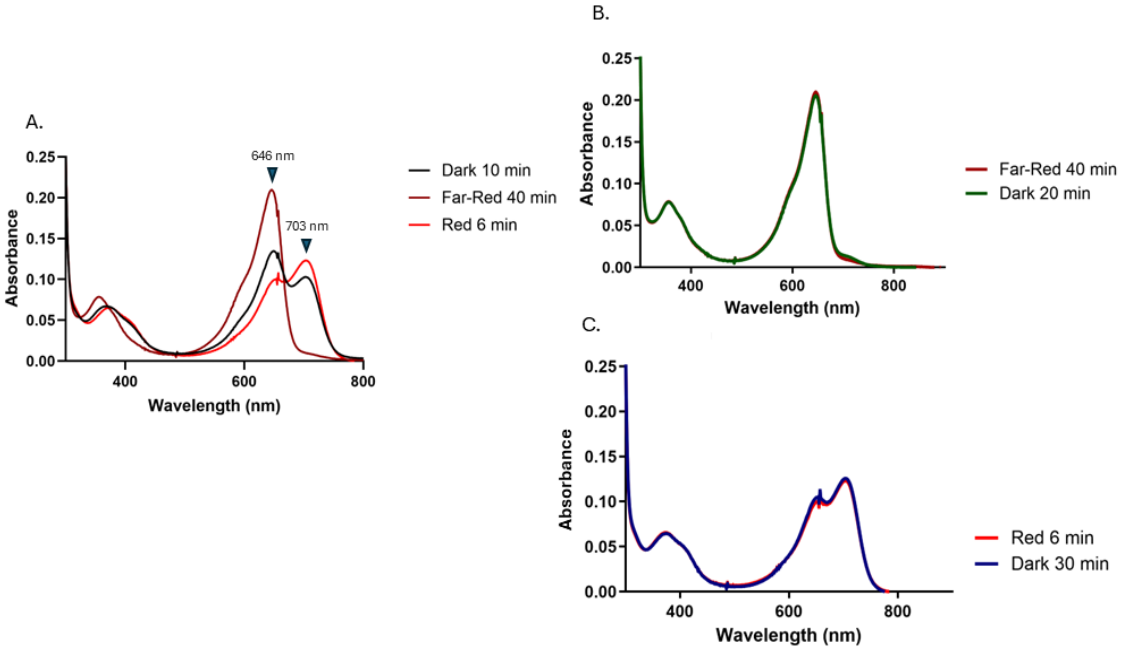

Figure 12. UV-Vis absorption spectra of RfpA GAF-PHY mutants. (A) Spectra without any light exposure (dark, marked in black) showed absorbance at 703nm, suggesting the ground state could be Pfr. Further Illumination with 785 nm (Far-Red) shifted the spectra to 646nm, which specifies the photoconversion from Pfr to Pr. Again, exposure to 685 nm (Red) demonstrated the changes in absorbance at 703 nm, leading to the change of the form from Pr to Pfr. (B. and C.) Dark reversion after 785 nm and 655 nm exposure did not demonstrate any change in spectra, which indicated that dark reversion requires more time.

3.3.3 RfpA GAF

The steps that followed for GAF mutants were similar for GAF-PHY. The spectra were recorded without light exposure in the dark for the first 5 minutes. It was pointed out that the absorbance at 646 nm and 703 nm was the same as for GAF-PHY variants, though the exposure time was little different than GAF-PHY. Exposed to 785 nm and 655 nm proceeded for 32 minutes and 20 minutes, respectively. The first exposure to 785nm wavelength also resulted in a significant shift in the absorbance maxima at 646nm (Figure 13A). This shift of maxima indicated the change of the form Pfr to Pr, suggesting that the sample has Pfr form as the resting state. Also, a further 655 nm illumination paves the conversion from Pr to Pfr. One interesting observation found in the case of GAF variants was that it denotes a different outcome during dark incubation after

Laser exposure. In comparison to the GAF-PHY variants, GAF variants showed instant dark reversion after 785 nm illumination, and during the last dark incubation after 655 nm exposure, the absorbance spectra changed slightly (Figure 13B, 13C).

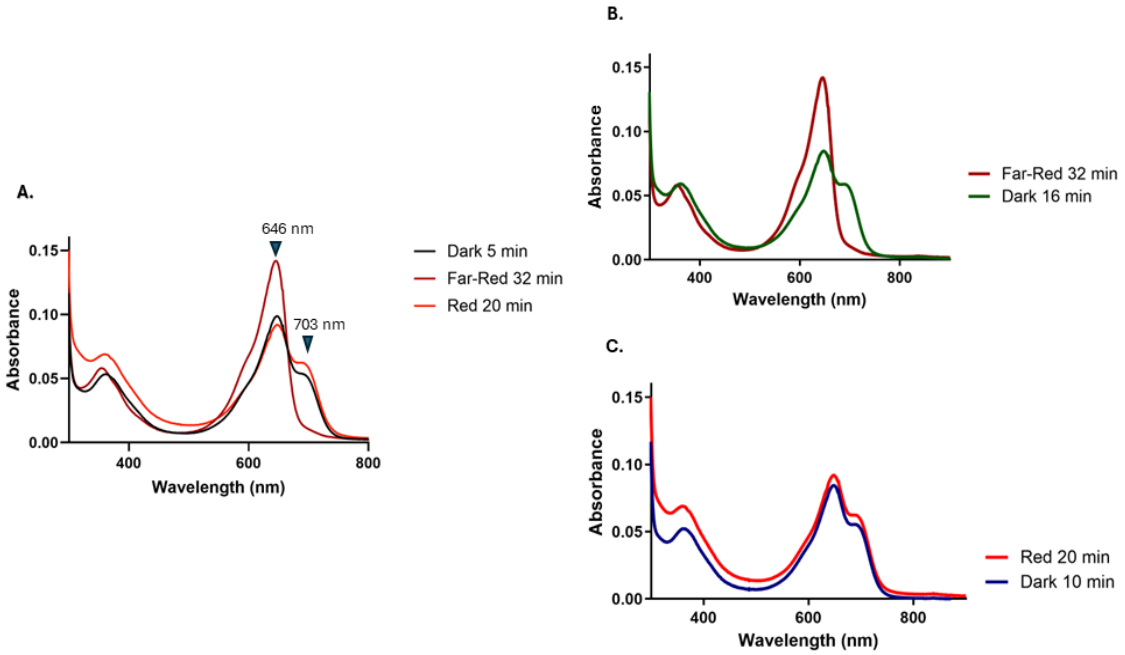

Figure 13. Absorption spectra of RfpA GAF mutants. (A) Spectra taken in the dark showed absorbance at both 646 nm and 703 nm. Excitation at 785 nm delineates the photoconversion from Pfr to Pr as the absorbance changes to 646 nm. Whereas 685 nm (red) illumination generated absorption spectra that were almost similar to the dark state, even more shifted towards 703 nm. (B) After 785 nm exposure, dark incubation notably led to a reversion of the form. (C) Dark reversion after 655 nm exposure insignificantly changed the spectra.

Overall, the spectral analysis showed that the variants of RfpA protein have Pfr form as their resting state and are photoactive when exposed to far-red light illumination. The photoactive form is the Pr form.

4 DISCUSSION

RfpA is an essential protein for FaRLiP cyanobacteria that regulates the synthesis of Chl *d* and Chl *f,* facilitating them to carry out photosynthesis under far-red light. It triggers the transcription of a particular gene apparatus that regulates the FaRLiP acclimation processes. (Gan *et al.* 2014a; Gan and Bryant 2015). However, the knowledge about the interplay between the structure and function of the protein is still constrained, which was the key leader of this study. The essential requirement to gain information about a protein is the purification of the protein, which was one of the core objectives of the study. Our study reveals several insightful findings about the protein. The results suggest that 1) The higher the length of RfpA protein, the less soluble it is. 2) GAF-PHY and GAF variants are readily soluble. 3) GAF-PHY fragments are enough to study the photocycle of the protein. 4) Although GAF demonstrates the photocycle, it is not that prominent. 5) RfpA is a bathy phytochrome.

All three recombinant plasmids were assembled successfully. However, in the first band of the gel picture (Figure 7), two distinct layers of the band were observed. It might be due to incomplete digestion. It could be possible that some cells were not digested by the restriction enzyme, resulting in a higher band than digested one.

Purification of the longest possible variants of the protein was quite challenging throughout the experiment. Although the expression of the protein in culture was successful for all variants, after cell lysis, most of the phytochrome resides in the pellet along with cell debris regarding GAF-PHY-PAS and the full-length version of the protein. For this reason, purification of longer protein (GAF-PHY-PAS, full-length) was not possible during this thesis experiment. Expression of recombinant protein in *E. coli* sometimes creates problems with the solubility of the desired protein due to inefficient folding of secondary structure or production of inclusion bodies (Upadhyay *et al.* 2012). Protein solubilization is also affected by the molecular weight, composition, and sequence of amino acids. Also, the contents of polar and nonpolar groups in amino acids influence their solubility (Nag *et al.* 2022). The literature exemplified that solubility can also be influenced by extrinsic factors, for example, changing the salt concentration, pH, and ionic strength (Taylor *et al.* 2017). We observed that the shorter counterpart of the protein showed higher solubility, but the precipitation problem occurred for all variants, which can relate to this. We used Tris buffer (30mM Tris, 150mM NaCl, pH 8.0) to purify the protein. It is feasible to adjust the buffer salt concentration or change in pH to improve the solubility of full-length

protein. On the other hand, it could also be possible to enhance the binding of the chromophore with the GAF domain by lowering the temperature during expression and eventually by providing additional molecular chaperones could be helpful for stopping aggregation.

The RfpA phytochrome has a different domain structure compared to other phytochrome. They have a simple GAF-PHY module as a photosensory module, whereas the canonical phytochrome has PAS-GAF-PHY (Anders and Essen 2015; Oh and Montgomery 2017). Although the phytochromes are varied in their chemical structure, module organization, or source, all phytochromes exhibit a conventional scheme for the photoconversion mechanism. This conversion initiates by sensing light by the chromophore, which resides in the GAF domain, triggers Z/E isomerization, and ends with the secondary structural changes in the PHY domain (Takala *et al.* 2014). Phytochrome generally share a conserved cysteine residue that abides in GAF domain bilin-binding pocket, which is principally responsible for binding with the bilin chromophore (Cheng *et al.* 2021). The literature reviewed that some phytochrome that is distantly related contain simple photosensors consisting of the GAF domain solely and exhibit photoconversion procedure normally (Fushimi and Narikawa 2019). Also, it has been reported that without the PHY domain, it is feasible to carry out the covalent binding of the chromophore and undergo photoreaction, but the efficiency of the photoconversion is compromised (Fankhauser 2001; Fischer *et al.* 2005). In another study, it has been revealed that some cyanobacteriochrome (CBCR) GAF domains can show unidirectional photoconversion and rapid dark reversion (Ikeuchi and Ishizuka 2008; Fushimi and Narikawa 2019). In figures 12 and 13, the spectra of GAF-PHY and GAF recorded different results during dark reversion after 785 nm excitation. GAF-PHY variants did not reverse to the Pfr form during incubation in the Dark after far-red light exposure, while GAF variants did. The reason might be the lack of PHY domain, as it is required for the retention of photoreaction properly.

In the prototypical phytochrome, photoconversion occurs from Pr to Pfr form, and the most stable thermal state is Pr. On the other hand, some phytochrome shows the inverse mechanism where the most stable form in the dark is Pfr, and they have the Pfr form as their ground state. This specific phytochrome is called bathy phytochrome (Rottwinkel *et al.* 2010; Velázquez Escobar *et al.* 2017; Herder 2022). The spectra analysis clearly showed that in RfpA variants, the stable state in the dark Pfr was dominant, and the photoconversion pathway followed the change of Pfr to Pr, which indicates the possibility of being a bathy phytochrome. Before, the bathy phytochrome was investigated in soil bacteria (Gregor *et al.* 2010). But there is no available information on cyanobacterial bathy phytochrome. Also, during purification,

both mutants showed the holoprotein absorption peak at 710 nm in the chromatogram, which also denotes the possibility of being bathy in nature. However, in spectral analysis, the red light and far-red light illumination was done using 655nm and 785nm wavelength Laser diodes. For both mutants, GAF-PHY (figure 13A) and GAF (Figure 14A), the dark and 685 nm exposure showed mixed absorbance spectra for the Pr and Pfr state. One possible reason behind the mixed spectra that appear in the first dark reading could be that during the handling of the sample, it is exposed to light unintentionally, which could trigger the photoconversion. In the case of red light Laser also, it might be possible that it excited both Pr and Pfr states but with different intensities. On the other hand, the intensity of the 785nm Laser at the position 710nm was not that high. For proper spectral analysis, the spectra should be measured with the nearest wavelength illumination, which will excite our specific region with the possible best intensity without exciting both forms.

Although further research is needed to understand the characteristics of RfpA, this study unveiled some significant features about the protein that can lead to further investigation on this bathy nature. Also, the aim of this experiment was to purify the full-length protein, which was partially accomplished. But the investigation paved the way for the purification of the full-length protein. Future work could also investigate how to solubilize the protein more effectively and to know the functional characteristics of each domain.

5 CONCLUSIONS

Cyanobacteria are prokaryotic microbes that perform oxygenic photosynthesis. Some cyanobacteria have the unique capability to perform photosynthesis under far-red light. This phenomenon is known as FaRLiP, primarily controlled by RfpA phytochrome. However, the mechanisms of RfpA and related signal transduction systems are still unknown. For investigating these signaling mechanisms, purification of the RfpA phytochrome is the initial and crucial step.

In this study, we successfully purified GAF-PHY fragments of RfpA. The UV-Vis spectroscopy result delineates that the GAF-PHY domains are enough to study the photocycle of the protein, which also supports our hypothesis. The spectral characterization of this fragment also reveals that RfpA has characteristics of a bathy phytochrome, as the Pfr form is predominant in the dark state. In addition, this marks the first time a bathy phytochrome is investigated in cyanobacteria. These exceptional features of the RfpA phytochrome make it a more potential candidate for an optogenetic tool and in engineering biotechnological techniques, especially in cyanobacteria, which required further in-depth research to understand the signal mechanisms of RfpA in regulating FaRLiP.

ACKNOWLEDGEMENTS

I would like to express my earnest gratitude to my supervisor Dr. Amit Srivastava, for his instrumental guidance, motivation, and reinforcement throughout the time period of the thesis, from lab work to thesis writing and data analysis. His patience, proficiency, and constructive feedback have been pivotal in shaping the quality and direction of the research.

I am also grateful to Doc. Heikki Takala, for giving me the opportunity to work in his group, and for his insightful explanations and suggestions in thesis reviewing, which have significantly contributed to the improvement and refinement of the thesis. I also want to thank the members of the group for helping with the lab work.

I would like to extend my acknowledgment to my partner and my only daughter for their support, understanding, and sacrifices during this formidable journey.

Jyväskylä May 31, 2024
Syeda Josna Ahmed

REFERENCES

Anders K. & Essen L.-O. 2015. The family of phytochrome-like photoreceptors: diverse, complex and multi-colored, but very useful. *Current Opinion in Structural Biology* 35: 7-16.

Auldridge M.E. & Forest K.T. 2011. Bacterial phytochromes: More than meets the light. *Critical Reviews in Biochemistry and Molecular Biology* 46: 67–88.

Bhoo S.-H., Davis S.J., Walker J., Karniol B. & Vierstra R.D. 2001. Bacteriophytochromes are photochromic histidine kinases using a biliverdin chromophore. *Nature* 414: 776–779.

Burgie E.S. & Vierstra R.D. 2014. Phytochromes: An Atomic Perspective on Photoactivation and Signaling. *The Plant Cell* 26: 4568–4583.

Cheng M.-C., Kathare P.K., Paik I. & Huq E. 2021. Phytochrome Signaling Networks. *Annual Review of Plant Biology* 72: 217–244.

Chia N., Lee S.Y. & Tong Y. 2022. Optogenetic tools for microbial synthetic biology. *Biotechnology Advances* 59: 107953.

Davis S.J., Vener A. V & Vierstra R.D. 1999. Bacteriophytochromes: Phytochrome-Like Photoreceptors from Nonphotosynthetic Eubacteria. *Science* 286: 2517–2520.

Fankhauser C. 2001. The Phytochromes, a Family of Red/Far-red Absorbing Photoreceptors. *Journal of Biological Chemistry* 276: 11453–11456.

Fischer A.J., Rockwell N.C., Jang A.Y., Ernst L.A., Waggoner A.S., Duan Y., Lei H. & Lagarias J.C. 2005. Multiple Roles of a Conserved GAF Domain Tyrosine Residue in Cyanobacterial and Plant Phytochromes. *Biochemistry* 44: 15203–15215.

Fushimi K. & Narikawa R. 2019. Cyanobacteriochromes: photoreceptors covering the entire UV-to-visible spectrum. *Current Opinion in Structural Biology* 57: 39–46.

Gan F. & Bryant D.A. 2015. Adaptive and acclimative responses of cyanobacteria to far-red light. *Environmental Microbiology* 17: 3450–3465.

Gan F., Shen G. & Bryant D.A. 2014a. Occurrence of far-red light photoacclimation (FaRLiP) in diverse cyanobacteria. *Life* 5: 4–24.

Gan F., Zhang S., Rockwell N.C., Martin S.S., Lagarias J.C. & Bryant D.A. 2014b. Extensive remodeling of a cyanobacterial photosynthetic apparatus in far-red light. *Science* 345: 1312–1317.

Ge B., Li Y., Sun H., Zhang S., Hu P., Qin S. & Huang F. 2013. Combinational biosynthesis of phycocyanobilin using genetically-engineered Escherichia coli. *Biotechnology Letters* 35: 689–693.

Golonka D., Fischbach P., Jena S.G., Kleeberg J.R.W.W., Essen L.-O.O., Toettcher J.E., Zurbriggen M.D. & Möglich A. 2019. Deconstructing and repurposing the light-regulated interplay between Arabidopsis phytochromes and interacting factors. *Communications Biology* 2: 448.

Gregor R., Inga O. & Tilman L. 2010. Bathy Phytochromes in Rhizobial Soil Bacteria. *Journal of Bacteriology* 192: 5124–5133.

Herder L.M.S. 2022. Structural investigation of the photocycle of the bacterial bathy phytochrome Agp2 PAiRFP2.

Ho M.-Y., Gan F., Shen G. & Bryant D.A. 2017a. Far-red light photoacclimation (FaRLiP) in Synechococcus sp. PCC 7335. II. Characterization of phycobiliproteins produced during acclimation to far-red light. *Photosynthesis Research* 131: 187–202.

Ho M.Y., Gan F., Shen G., Zhao C. & Bryant D.A. 2017b. Far-red light photoacclimation (FaRLiP) in Synechococcus sp. PCC 7335: I. Regulation of FaRLiP gene expression. *Photosynthesis Research* 131: 173-186

Ho M.-Y., Soulier N.T., Canniffe D.P., Shen G. & Bryant D.A. 2017c. Light regulation of pigment and photosystem biosynthesis in cyanobacteria. *Current Opinion in Plant Biology* 37: 24–33.

Hughes J. 2010. Phytochrome three-dimensional structures and functions. *Biochemical Society Transactions* 38: 710–716.

Ikeuchi M. & Ishizuka T. 2008. Cyanobacteriochromes: a new superfamily of tetrapyrrole-binding photoreceptors in cyanobacteria. *Photochemical & Photobiological Sciences* 7: 1159–1167.

Jeschek M., Gerngross D. & Panke S. 2017. Combinatorial pathway optimization for streamlined metabolic engineering. *Current Opinion in Biotechnology* 47: 142–151.

Knoot C.J., Ungerer J., Wangikar P.P. & Pakrasi H.B. 2018. Cyanobacteria: promising biocatalysts for sustainable chemical production. *Journal of Biological Chemistry* 293: 5044–5052.

Kreslavski V.D., Los D.A., Schmitt F.-J., Zharmukhamedov S.K., Kuznetsov V. V & Allakhverdiev S.I. 2018. The impact of the phytochromes on photosynthetic processes. *Biochimica et Biophysica Acta (BBA) - Bioenergetics* 1859: 400–408.

Lamparter T. 2004. Evolution of cyanobacterial and plant phytochromes. *FEBS Letters* 573: 1–5.

Li Y., Lin Y., Garvey C.J., Birch D., Corkery R.W., Loughlin P.C., Scheer H., Willows R.D. & Chen M. 2016. Characterization of red-shifted phycobilisomes isolated from the chlorophyll f-containing cyanobacterium Halomicronema hongdechloris. *Biochimica et Biophysica Acta - Bioenergetics* 1857: 107-114.

Lin P.-C. & Pakrasi H.B. 2019. Engineering cyanobacteria for production of terpenoids. *Planta* 249: 145–154.

Lindner F. & Diepold A. 2022. Optogenetics in bacteria–applications and opportunities. *FEMS Microbiology Reviews* 46: fuab055.

Möglich A., Ayers R.A. & Moffat K. 2009a. Design and signaling mechanism of light-regulated histidine kinases. *Biophysical Journal* 96: 524a.

Möglich A., Ayers R.A. & Moffat K. 2009b. Structure and signaling mechanism of Per-ARNT-Sim domains. *Structure* 17: 1282–1294.

Möglich A., Yang X., Ayers R.A. & Moffat K. 2010. Structure and function of plant photoreceptors. *Annual Review of Plant Biology* 61: 21–47.

Moon Y.-J., Kim S.Y., Jung K.-H., Choi J.-S., Park Y.M. & Chung Y.-H. 2011. Cyanobacterial phytochrome Cph2 is a negative regulator in phototaxis toward UV-A. *FEBS letters* 585: 335–340.

Multamäki E., Nanekar R., Morozov D., Lievonen T., Golonka D., Wahlgren W.Y., Stucki-Buchli B., Rossi J., Hytönen V.P. & Westenhoff S. 2021. Comparative analysis of two paradigm bacteriophytochromes reveals opposite functionalities in two-component signaling. *Nature Communications* 12: 4394.

Murgida D.H., Stetten D. Von, Hildebrandt P., Schwinté P., Siebert F., Sharda S., Gärtner W. & Mroginski M.A. 2007. The chromophore structures of the Pr states in plant and bacterial phytochromes. *Biophysical Journal* 93: 2410–2417.

Nag N., Khan H. & Tripathi T. 2022. Strategies to improve the expression and solubility of recombinant proteins in E. coli. In: *Advances in Protein Molecular and Structural Biology Methods*, Elsevier: 1–12.

Oh S. & Montgomery B.L. 2017. Phytochromes: where to start? *Cell* 171: 1254–1256.

Pham V.N., Kathare P.K. & Huq E. 2018. Phytochromes and phytochrome interacting factors. *Plant Physiology* 176: 1025–1038.

Pooja P., Lekshmi K.E. & Pradeep N.S. 2023. Biofuels from cyanobacteria-a metabolic engineering approach. *Plant Science Today* 10.

Reshetnikov V. V, Smolskaya S. V, Feoktistova S.G. & Verkhusha V. V. 2022. Optogenetic approaches in biotechnology and biomaterials. *Trends in Biotechnology* 40: 858–874.

Rockwell N.C. & Lagarias J.C. 2017. Phytochrome diversification in cyanobacteria and eukaryotic algae. *Current Opinion in Plant Biology* 37: 87–93.

Rockwell N.C. & Lagarias J.C. 2024. Cyanobacteriochromes from Gloeobacterales Provide New Insight into the Diversification of Cyanobacterial Photoreceptors. *Journal of Molecular Biology* 436: 168313.

Rockwell N.C., Martin S.S. & Lagarias J.C. 2023. Elucidating the origins of phycocyanobilin biosynthesis and phycobiliproteins. *Proceedings of the National Academy of Sciences* 120: e2300770120.

Rockwell N.C., Su Y.-S. & Lagarias J.C. 2006. Phytochrome structure and signaling mechanisms. *Annu. Rev. Plant Biol.* 57: 837–858.

Rottwinkel G., Oberpichler I. & Lamparter T. 2010. Bathy phytochromes in rhizobial soil bacteria. *Journal of Bacteriology* 192: 5124–5133.

Shen G., Canniffe D.P., Ho M.Y., Kurashov V., Est A. van der, Golbeck J.H. & Bryant D.A. 2019. Characterization of chlorophyll f synthase heterologously produced in Synechococcus sp. PCC 7002. *Photosynthesis Research* 140: 77-92.

Sivaprakasam S., Mani V., Balasubramaniyan N. & Abraham D.R. 2021. Cyanobacterial Phytochromes in Optogenetics. *Epigenetics to Optogenetics-A New Paradigm in the Study of Biology. London*.

Smith H. 2000. Phytochromes and light signal perception by plants—an emerging synthesis. *Nature* 407: 585–591.

Srivastava A., Varshney R.K. & Shukla P. 2021. Sigma factor modulation for cyanobacterial metabolic engineering. *Trends in Microbiology* 29: 266–277.

Takala H., Björling A., Berntsson O., Lehtivuori H., Niebling S., Hoernke M., Kosheleva I., Henning R., Menzel A., Ihalainen J.A. & Westenhoff S. 2014. Signal amplification and transduction in phytochrome photosensors. *Nature* 509: 245–248.

Takala H., Edlund P., Ihalainen J.A. & Westenhoff S. 2020. Tips and turns of bacteriophytochrome photoactivation. *Photochemical & Photobiological Sciences* 19: 1488–1510.

Takala H., Lehtivuori H.K., Berntsson O., Hughes A., Nanekar R., Niebling S., Panman M., Henry L., Menzel A., Westenhoff S. & Ihalainen J.A. 2018. On the (un)coupling of the chromophore, tongue interactions, and overall conformation in a bacterial phytochrome. *Journal of Biological Chemistry* 293: 8161–8172.

Taylor T., Denson J.-P. & Esposito D. 2017. Optimizing expression and solubility of proteins in E. coli using modified media and induction parameters. *Heterologous Gene Expression in E. coli: Methods and Protocols*: 65–82.

Trinugroho J.P., Becková M., Shao S.X., Yu J.F., Zhao Z.Y., Murray J.W., Sobotka R., Komenda J. & Nixon P.J. 2020. Chlorophyll *f* synthesis by a super-rogue photosystem II complex. *Nature Plants* 6: 238-244.

Upadhyay A.K., Murmu A., Singh A. & Panda A.K. 2012. Kinetics of Inclusion Body Formation and Its Correlation with the Characteristics of Protein Aggregates in Escherichia coli. *PLoS One* 7: e33951.

Velázquez Escobar F., Buhrke D., Michael N., Sauthof L., Wilkening S., Tavraz N.N., Salewski J., Frankenberg-Dinkel N., Mroginski M.A. & Scheerer P. 2017. Common structural elements in the chromophore binding pocket of the Pfr state of bathy phytochromes. *Photochemistry and Photobiology* 93: 724–732.

Wahlgren W.Y., Claesson E., Tuure I., Trillo-Muyo S., Bódizs S., Ihalainen J.A., Takala H. & Westenhoff S. 2022. Structural mechanism of signal transduction in a phytochrome histidine kinase. *Nature Communications* 13: 7673.

Wiltbank L.B. & Kehoe D.M. 2019. Diverse light responses of cyanobacteria mediated by phytochrome superfamily photoreceptors. *Nature Reviews Microbiology* 17: 37–50.

Wolf B.M. & Blankenship R.E. 2019. Far-red light acclimation in diverse oxygenic photosynthetic organisms. *Photosynthesis Research* 142: 349–359.

Zahra Z., Choo D.H., Lee H. & Parveen A. 2020. Cyanobacteria: Review of current potentials and applications. *Environments* 7: 13.

Zhao C., Gan F., Shen G. & Bryant D.A. 2015. RfpA, RfpB, and RfpC are the master control elements of far-red light photoacclimation (FaRLiP). *Frontiers in Microbiology* 6: 168681.

Zhao C., Gan F., Shen G. & Bryant D.A. 2019. Corrigendum: RfpA, RfpB, and RfpC are the Master Control Elements of Far-Red Light Photoacclimation (FaRLiP). *Frontiers in Microbiology* 10: 447743.

APPENDIX 1.

Table A.1 **Mutagenic primer**

Name	Sequence	Position	GC content	Melting temperature
GAF (Primer 58)	GATTGCGCAGAGCACCCTGCTG GAGCAAACCCGTATCCAG	825-864	60.00%	72.8°C
GAF-PHY (Primer 59)	GGCGATTCAGCAATACGAGCTG TATCTGCAGGTGCATGC	1458-1496	53.85%	69.9°C
GAF-PHY-PAS (Primer 60)	CTTCAAGGACATTACCGAACGTC AGATCGTGGAGCGTATGAAG	2094-2136	48.84%	68°C