



**HAL**  
open science

# Carotenoid translocation and protein evolution in cyanobacterial photoprotection

Fernando Muzzopappa

► **To cite this version:**

Fernando Muzzopappa. Carotenoid translocation and protein evolution in cyanobacterial photoprotection. Bacteriology. Université Paris Saclay (COmUE), 2019. English. NNT : 2019SACLS493 . tel-02682436

**HAL Id: tel-02682436**

**<https://theses.hal.science/tel-02682436>**

Submitted on 1 Jun 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



# Carotenoid translocation and protein evolution in cyanobacterial photoprotection

Thèse de doctorat de l'Université Paris-Saclay

Préparée à l'Université Paris-Sud

École doctorale n°567 - Sciences du végétal : du gène à l'écosystème (SDV)

Spécialité de doctorat : Biologie

Thèse présentée et soutenue à Saclay, le 2/12/2019, par

**Fernando Muzzopappa**

Composition du Jury :

<b>Philippe Minard</b> Professeur, Université Paris Sud et I2BC (UMR 9198 CEA/CNRS/UPS)	Président
<b>Xenie Johnson</b> Directeur de recherche, BIAM (UMR 7265, CEA-Caradache/CNRS/AMU)	Rapporteur
<b>Eric Maréchal</b> Directeur de recherche, LPCV (UMR 5168 CEA-Grenoble/CNRS/UGA)	Rapporteur
<b>Jacques-Philippe Colletier</b> Directeur de recherche, IBS (UMR 5075 CEA-Grenoble/CNRS/UGA)	Examineur
<b>Francesca Zito</b> Directeur de recherche, IBPC (UMR 7099 CNRS/UPD)	Examineur
<b>Diana Kirilovsky</b> Directeur de recherche, I2BC (UMR 9198 CEA/CNRS/UPS)	Directeur de thèse



## *Preface*

---

The work described in this PhD thesis was developed in the framework of the SE2B Marie Curie ITN. This training program consisted in a scientific Network involving 14 laboratories spread all over Europe focus on photosynthesis research. This Network provided an academic environment to 16 students, including myself, to carry out the PhD thesis. In this context, this thesis aims to study the involvement of carotenoid proteins in cyanobacterial photoprotective mechanisms.

In 2001, Diana Kirilovsky proposed the existence of a novel photoprotective mechanism involving the phycobilisome, the cyanobacterial antenna, decreasing the energy arriving at the reaction centers due to an increase of energy thermal dissipation. In 2006, our group confirmed the existence of such mechanism and discovered that a water-soluble carotenoid protein, the Orange Carotenoid Protein (OCP), is essential for it. Several scientific articles and thesis have been devoted to the elucidation of the molecular mechanism behind the role of the OCP. Recently, our group, in collaboration with an American laboratory, showed the first characterization of another family of carotenoid proteins, the Helical Carotenoid Proteins (HCP) which are homologs of one OCP domain.

The starting point of this thesis was the characterization of the homolog of the other domain of the OCP, the CTDH. This characterization and the analysis of the interaction between CTDH and HCP led to discovery of a novel carotenoid translocation mechanism among carotenoid proteins. That mechanism is deeply explored in this thesis, from a molecular point of view.

Finally, the comparison of the OCP system with the CTDH/HCP system raised many questions about the evolution of the OCP. I attempted to shine some light on this problem by studying the role of the linker connecting the OCP domains in the different subclades of modern OCPs. This allowed as to trace the evolutionary history of the OCP.

## *Acknowledgment*

---

During these three years I had the opportunity to learn a lot, about cyanobacteria and photosynthesis as well as how is the life in research. I would like to thank here all the people who were part of this process and also to all those, who contributed to the research described here. However not everything is science in life, moving to a different country leaving friends and family behind is a big challenge. This was only possible thanks to the support of family, old friends and those who I met here and now I am happy to call them *friends*.

First of all, I would like to thank Dr. Diana Kirilovsky for giving me the opportunity of working in this fascinating topic. Thank you for teaching me, not only how to perform an experiment but also how to develop a scientific career. For transmitting your passion for science and for all the advices given. Thank you for taking care of me and all the people of the lab.

Secondly, I want to acknowledge my family. I wouldn't be here now without all you did for me. Ma, Pa and Lea; *gracias por todo*. To my grandparents, especially Roberto, Orlando and Marta who passed away when I was abroad. To Licha, for taking care of me and all the family. To Vir, thanks for everything... *I do not attempt to express the inexpressible*.

I also want to acknowledge Dr. Pierre Sétif and Dr. Dr. Anja Krieger-Liszkay, for their assistance and for being always available to share your knowledge with me.

To all my co-workers, Pablito, Adjélé, Léa, Kathleen, Jiao, Shira, Sandrine, Desire, Hortence, Ginga, Elodie, Marine and Melanie: thank you very much for the good mood, the laughs, the support and the patient. It was great to share the lab with you. I also want to thank all the ESRs, for all the good moments we shared in the SE2B events.

I want to acknowledge also our collaborators, Cédric Montigny, Dvir Harris, Noam Adir, Nikolai Sluchanko and Eugene Maksimov for all the fruitful collaborations we have had. I hope than in the future, the contribution between our labs will keep generating nice results.

Finally, I would like to thank all my friends who supported me. To all my friends in Argentina who showed me that some friendships are endless regardless where we are. To my new friends, the homeless, who showed me different cultures and enriched mine.

## Contents

---

Preface	I	
Acknowledgment	II	
List of abbreviations	V	
<b>Part I: Introduction</b>	<b>1</b>	
<i>The world of Cyanobacteria</i>	<b>2</b>	
<i>Photosynthesis: A general overview</i>	<b>4</b>	
<i>Photosynthetic components: pigments</i>	<b>7</b>	
<i>Photosynthetic components: Antenna complex</i>	<b>9</b>	
<i>Photosynthetic components: Photosystem II</i>	<b>13</b>	
<i>Photosynthetic components: cytochrome b6f</i>	<b>14</b>	
<i>Photosynthetic components: Photosystem I</i>	<b>16</b>	
<i>Photosynthetic electron transport: linear versus cyclic</i>		<b>17</b>
<i>Photoprotection: introduction to oxidative stress</i>		<b>18</b>
<i>ROS scavengers and alternative electron pathways</i>		<b>20</b>
<i>Photoprotection: State transitions</i>	<b>20</b>	
<i>Photoprotection: OCP-related NPQ</i>	<b>22</b>	
<i>The Orange Carotenoid Protein: Structure</i>		<b>26</b>
<i>The Orange Carotenoid Protein: photoactivation</i>		<b>28</b>
<i>The Orange Carotenoid Protein activity</i>		<b>32</b>
<i>The Fluorescence Recovery Protein</i>	<b>35</b>	
<i>The paralogs of the OCP domains</i>	<b>37</b>	
<i>Aim of the thesis</i>	<b>40</b>	
<i>List of publications</i>	<b>41</b>	
<i>Bibliography</i>	<b>43</b>	
<b>Part II: Results</b>	<b>51</b>	
<b>Chapter I: The paralog of the C-terminal domain of the OCP: a novel carotenoid carrier.</b>	<b>51</b>	
<i>Resume of the article</i>	<b>52</b>	
<i>Introduction</i>	<b>55</b>	
<i>Results</i>	<b>59</b>	
<i>Discussion</i>	<b>82</b>	
<i>Materials and Methods</i>	<b>90</b>	
<i>Bibliography</i>	<b>96</b>	
<i>Appendix: supplementary information</i>		<b>100</b>
<b>Chapter II: Analysis of a multidirectional carotenoid transfer mechanism.</b>	<b>109</b>	
<i>Resume of the article</i>	<b>110</b>	
<i>Introduction</i>	<b>113</b>	
<i>Results</i>	<b>116</b>	
<i>Discussion</i>	<b>129</b>	
<i>Appendix: supplementary information</i>		<b>133</b>
<i>Materials and Methods</i>	<b>133</b>	

<i>Bibliography</i>	<b>135</b>	
<b>Chapter III: The CTDH structure: unraveling the role of the CTT in carotenoid translocation.</b>	<b>143</b>	
<i>Resume of the article</i>	<b>144</b>	
<i>Introduction</i>	<b>147</b>	
<i>Results (First part)</i>	<b>149</b>	
<i>Results (Second part)</i>	<b>162</b>	
<i>Discussion (Second part)</i>	<b>170</b>	
<i>Materials and Methods</i>	<b>171</b>	
<i>Bibliography</i>	<b>178</b>	
<i>Appendix: supplementary information</i>		<b>184</b>
<b>Chapter IV: Molecular evolution of the Orange Carotenoid Protein.</b>	<b>193</b>	
<i>Resume of the article</i>	<b>194</b>	
<i>Introduction</i>	<b>197</b>	
<i>Results</i>	<b>199</b>	
<i>Discussion</i>	<b>210</b>	
<i>Materials and Methods</i>	<b>212</b>	
<i>Bibliography</i>	<b>216</b>	
<i>Appendix: supplementary information</i>		<b>223</b>
<b>Part III: Conclusions and Perspectives</b>		<b>233</b>
<i>General discussion</i>	<b>234</b>	
<i>Perspectives</i>	<b>243</b>	
<i>Conclusion</i>	<b>248</b>	
<i>References</i>	<b>248</b>	
<i>Résumé en français (discussion translated)</i>		<b>250</b>

## *List of abbreviations*

**Ana:** (Prefix) protein from *Anabaena* (*Nostoc*) PCC 7120.

**APC:** Allophycocyanin.

**Apo:** (prefix) indicate absence of cofactor in a determined protein.

**ApcC:** Linker protein of phycobilisome core.

**ApcE:** Terminal emitter of phycobilisome and linker of its core.

**ApcD, ApcF:** Terminal emitter of phycobilisome.

**ATP:** Adenosine triphosphate.

**CAN:** Canthaxanthin.

**CrtO:**  $\beta$ -carotene ketolase (required for echinenone and 3'hydroxy-echinenone synthesis).

**CrtR:**  $\beta$ -carotene hydroxylase (required for 3'hydroxy-echinenone and zeaxanthin synthesis).

**CTD:** C-terminal domain of the OCP, isolated from the OCP.

**CTDH:** paralog of the CTD.

**Cyt *b<sub>6</sub>f*:** cytochrome *b<sub>6</sub>f*

**Cys:** cysteine.

**CTT:** C-terminal tail of the OCP-CTD or CTDH

**DNA:** Deoxyribonucleic acid.

**ECN:** Echinenone.

**Fd:** Ferredoxin.

**Flv:** Flavodiiron protein.

**FRP:** Fluorescence Recovery Protein.

**hECN:** 3'hydroxy-echinenone.

**HCP:** Helical Carotenoid Protein, paralog of the NTD.

**Holo:** (prefix or suffix) indicate the presence of cofactor in a determined protein.

**LHCII (LHCI):** Light Harvesting Complex II (I).

**M:** molar

**MD:** molecular dynamics.

**NADP:** Nicotinamide Adenine Dinucleotide Phosphate.

**NTD:** N-terminal domain of the OCP, isolated.

**NTE:** N-terminal extension or arm of the OCP.

**NTF2:** Nuclear Transporter Factor 2.

**OCP:** Orange Carotenoid Protein.

**OCP<sup>R</sup>:** photoactivated OCP (red).

**OCP<sup>O</sup>:** inactive form of the OCP (orange).

**OCP-CTD:** C-terminal domain of the OCP, in the OCP.

**OCP-NTD:** C-terminal domain of the OCP, in the OCP.

**P<sub>680</sub>/P<sub>700</sub>:** Special-pair chlorophyll of photosystem II/special-pair chlorophyll of photosystem I

**PAGE:** polyacrylamide gel electrophoresis.

**PAM:** pulse amplitude modulation.

**PBS:** phycobilisome.

**PC:** phycocyanin.

**PCC:** Pasteur Culture Collection.

**PDB:** Protein Data Bank

**PSI/PSII:** photosystem I / photosystem II.

**PQ/PQH<sub>2</sub>:** plastoquinone / plastoquinol.

**RNA:** Ribonucleic acid.

**SEC:** size exclusion chromatography.

**Te:** (prefix) protein from *Thermosynechococcus elongatus*.

**Tyr:** tyrosine

**Trp:** tryptophan.

**WT:** Wild-type.

“This Abstract, which I now publish,  
must necessarily be imperfect.”

-Charles Darwin.

*On the origin of the Species.* 1859, p.10



# Part I

---

## Introduction

## *The world of Cyanobacteria*

---

Cyanobacteria, formerly known as blue-green algae, are a diverse group of gram-negative prokaryotes. Among the vast group of bacteria, cyanobacteria are the only phylum able to perform oxygenic photosynthesis, a carbon fixation powered by light-driven water reduction. The cyanobacterial phylum is among the most ancient evolutionary lineages, microfossils and phylogenetic analysis set the origin of this taxa at the mid Archean Eon (around 3 billion years ago)<sup>1</sup>. It is commonly accepted that oxygen produced by cyanobacterial photosynthesis had led to the Great Oxygenation Event (GEO), a rapid accumulation of oxygen in the Earth atmosphere 2.3 billion years ago<sup>2,3</sup>.

They are ubiquitous throughout the world; they inhabit a broad range of environments. Cyanobacteria are found in both marine water including coral reef<sup>4</sup> and fresh water (lakes, reservoirs, slow flowing rivers and natural ponds). Cyanobacteria are also present in soils. They can be found in extreme habitats such as Antarctic ice, arid desert<sup>5</sup> and hot springs.

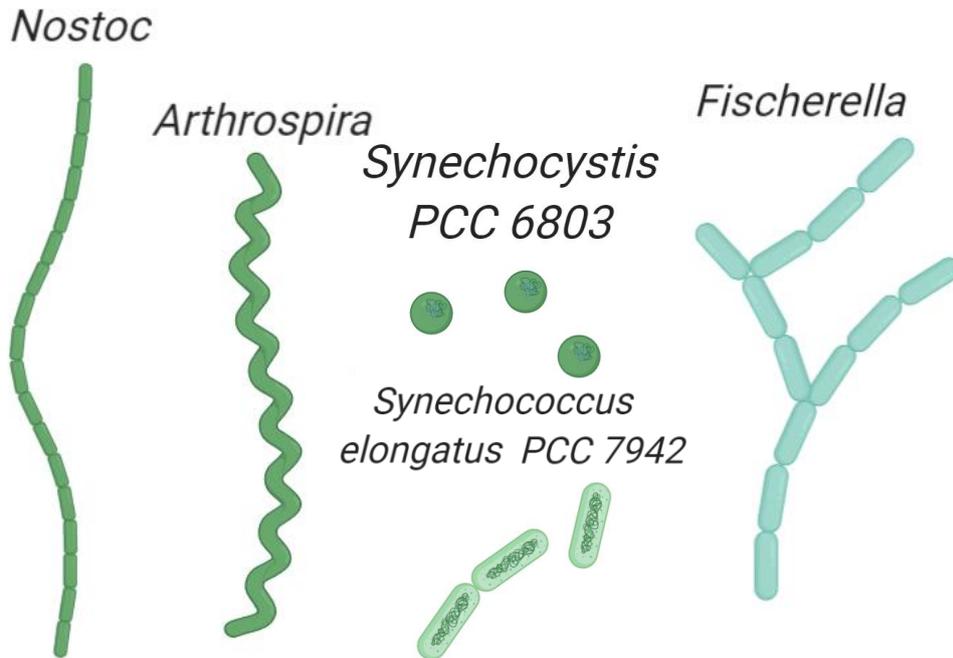
The common cell structure of cyanobacteria is more elaborated than other bacteria. Although they have a broad range of sizes, on average it is higher than most of the prokaryotes<sup>6</sup>. Beside the external peptidoglycan layer of 10 nm and the cytoplasmic membrane<sup>7,8</sup>, cyanobacteria possess, in the cytoplasm, an internal thylakoid membrane system where the photosynthetic apparatus is located (with exception of *Gloeobacter*, a cyanobacterium without thylakoids and a photosynthetic apparatus in the cytoplasmic membrane<sup>9</sup>). Additionally, other compartments or inclusions can be found in cyanobacteria: polyphosphate granules, glycogen granules, cyanophycin (an aspartic acid backbone that works as nitrogen reservoir) and carboxysomes (a protein shell which concentrates the carbon dioxide and contains the ribulose-1,5-bisphosphate carboxylase/oxygenase (RuBisCo) and anhydrases).

Their morphology can vary from unicellular to filamentous depending on the strain. While strains like *Synechocystis* PCC 6803 (hereafter *Synechocystis*) are spherical-shaped unicellular, other strains like *Synechococcus elongatus* PCC 7942 (hereafter *Synechococcus*) and *Thermosynechococcus elongatus* are rod-shaped unicellular and could be found in the environment as isolated, paired, linearly connected (2 to 4 cells), or in small clusters. Finally, other strains such as *Nostoc punctiforme* PCC 7120 (hereafter *Nostoc*) or *Scytonema Hofmannii* PCC 7110 (hereafter *Scytonema*) can form long filamentous structures. The latest allows cell differentiation among different cells belonging to a filamentous, in the case of *Nostoc* four types of cells are possible: vegetative cells (photosynthetic cells, it is the most abundant type under favorable growth conditions), heterocyst (containing-

nitrogenase enzyme cells involved in nitrogen fixation), akinetes (highly resistant spores formed under stress conditions) and hormogonia (gliding filamentous able to disperse in short distances, this cell type is involved in plant symbiosis)<sup>10</sup>. Based on these different properties, cyanobacteria have been classified into five different subsections. Subsections I and II are unicellular coccoids, the first divided by binary fission and the latest can additionally divide by multiple fission. Subsections III-V are filamentous cyanobacteria, subsection III filaments are formed by only vegetative cells, whereas subsection IV and V can undergo differentiation. In addition, subsection V can form branch patterns<sup>11</sup>.

This big diversity of phenotypes also is observed on the impact of cyanobacteria in other organisms, both negative and positive interaction have been reported. A good example of positive interaction is the symbiosis between cyanobacteria, usually belonging to the genus *Nostoc*, and another type of eukaryotic organisms. Generally, cyanobacterial symbionts supply fixed nitrogen to their host. The hosts can vary from fungi or algae to higher plants. On the other hand, cyanobacteria are also capable of form massive, high-dense blooms usually known as "harmful algal blooms" (HAB)<sup>12</sup>. HABs have considerable negative effects on the ecosystem. Cyanobacterial overgrowing depletes oxygen levels and decreases the sunlight in lakes and rivers, killing the other organisms. In addition, many bloom-forming cyanobacteria are toxin producers. Microcystin, produced as secondary metabolite, is one of the most common cyanotoxins. It is a potent inhibitor of protein phosphatases that poisons other organisms including mammals. Most important, the occurrence of HABs is rising due to climate change (temperature increasing at the surface of water reservoirs) and increase of nutrients levels (especially phosphate) in lakes/rivers by human activity.

Symbiosis interaction involving cyanobacteria could have been extremely important in the first stages of eukaryotic evolution. It was proposed in 1905<sup>13</sup> that chloroplast, in eukaryotic cells, were derived from an ancestral cyanobacterium engulfed by another prokaryote. This so-called *Endosymbiotic theory* is supported by many experimental observations such as the similar way to replicate (binary fission) of organelles and prokaryotes and the presence of circular DNA and double membranes in organelles.

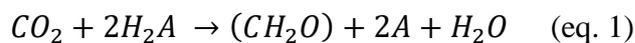


**Figure 1 | Diversity of cell morphology in Cyanobacteria.** Schematic representations of filamentous cyanobacteria *Nostoc* (subsection IV), *Arthrospira* (subsection III) and *Fischerella* (subsection V) are shown. In addition, unicellular cyanobacteria representations of *Synechocystis* and *Synechococcus*, both belonging to the subsection I, are presented.

## ***Photosynthesis: A general overview***

---

One of the most important biological processes on earth is the utilization of sun light to produce biologically compatible energy. This process, called *photosynthesis*, is a light-induced *redox* (reduction-oxidation) chemical reaction that can be formulated in a general way as:



In the equation 1,  $H_2A$  is the electron donor of the redox reaction and  $2A$  is the oxidized sub product. The oxygenic photosynthesis is a special case of photosynthesis where  $H_2A$  is  $H_2O$  and  $2A$  is  $O_2$ . Other types of anoxygenic photosynthesis are possible, for example by using another electron donor such as  $H_2S$ . This *oxidation* reaction is coupled to the *reduction* of  $CO_2$  into organic compounds suitable for biological usage.

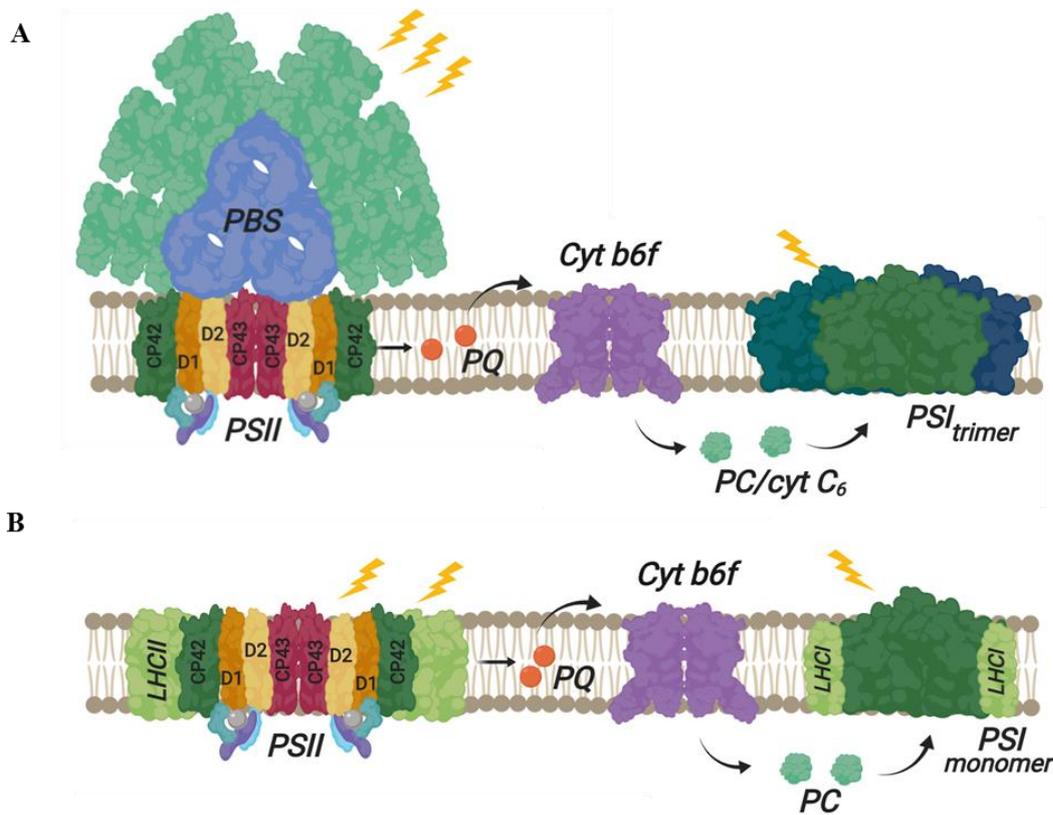
Photosynthetic organisms can be classified according to the bi-product generated from photosynthesis (' $2A$ ' in equation 1). If the bi-product is Oxygen, the photosynthetic organism is considered **oxygenic** and it performs **oxygen-evolving photosynthesis**. If this is not the case, the reaction is considered **anoxygenic**. Six

bacteria phyla are capable of photosynthesis. But only cyanobacteria, together with plants and eukaryotic algae, carry out oxygen-evolving photosynthesis.

Since the first experimental observations<sup>14</sup>, it was suggested that two different processes are involved in oxygenic photosynthesis: a “*light-reaction*” and a “*dark-reaction*”. Now it is clear that the so called *dark* reactions are also light dependent, and many enzymes involved in this phase require light-activation<sup>15</sup>. A better explanation of the photosynthetic mechanisms can be achieved by dividing it into four different phases<sup>16</sup>. The first phase involved the ***light harvesting by the antenna complex*** and the following energy transfer to the photochemical centers. All the photosynthetic organisms contain pigments (see the ‘*Photosynthetic components: pigments*’ section) which absorb light photons that ultimately are converted into chemical energy by photosynthesis. It was proved long time ago that not all the pigments are capable of photochemistry<sup>17</sup>. In fact, a large amount of pigments are part of *antenna systems* that transfer the energy<sup>18</sup> to special pigments to carry out photochemistry. The second phase is the ***charge separation and primary electron transfer in reaction centers***. In this phase, the excitation energy absorbed by the antenna arrives into the reaction centers, a pigment-protein subunit capable of photochemistry. Combined, the antenna system and the reaction center form the *photosynthetic unit (photosystem)*. In this transmembrane multiprotein complex, a special dimer of chlorophylls (green pigment) is promoted to an excited electronic state in which it became a strongly reducer species. This species rapidly transfers an electron to an electron acceptor (EA) generating an ion-pair state ( $P^+EA^-$ ). The electron could return to the chlorophyll-center in a *recombination* reaction, but this is generally avoided by the third phase, ***the stabilization by secondary reactions***. After the charge separation is generated, very fast secondary reactions promote an electron transfer cascade that separate the negative and positive charges in the photosynthetic apparatus, avoiding recombination. This electron transport is coupled to the proton transmembrane movement and generates an electrochemical potential gradient. Afterwards, this gradient is used to generate ATP. The final phase is the ***synthesis of energetic products***. In this phase, by using the ATP and the reduced compound ( $NADPH_2$ ) generated in the first three phases; the  $CO_2$  is fixed into organic molecules. This last part of the photosynthesis is beyond of the scope of this work and it will not be explained in detail.

The first three phases (the light reactions) take place in the photosynthetic apparatus, a combination of pigment-binding protein complexes embedded in the *thylakoid membrane*. In plants, thylakoids membranes are an inner membrane system in the chloroplasts. In cyanobacteria, like in other prokaryotes, organelles (for instance, chloroplasts) are absent. Instead, a membrane system which is different of the cell membrane, harbors both photosynthetic and respiratory electron

transport chains. This is one of the biggest differences between plant and cyanobacterial photosynthesis. Due to the fact that both electron transports are located in the same membrane system, in cyanobacteria some components are shared between photosynthetic and respiratory apparatus<sup>19</sup>. Electron carriers such as plastoquinol and plastocyanin (or cyt *c*<sub>6</sub>), as well as the cytochrome *b*<sub>6</sub>*f*, are involved in electron transport in both apparatuses.



**Figure 2 | Photosynthetic apparatus.** Diagram of the composition of the photosynthetic membrane in cyanobacteria (A) and plants (B). While in Plants the light-harvesting complex II (LHCII) in the membrane is the main antenna of photosystem II (PSII), in Cyanobacteria the soluble phycobilisome (PBS) plays that role. In both cases, the energy is transferred to the PSII. This energy is used by the PSII to reduce plastoquinol (PQ). PQ oxidation by the cytochrome *b*<sub>6</sub>*f* drive reduction of plastocyanin (PC) (or cytochrome *c*<sub>6</sub>, in cyanobacteria).

A schematic representation of the oxygen-evolving photosynthetic apparatus from cyanobacteria and plants are shown in the figure 2. The basic components are the antenna complex (pigment-protein complex that harvest the light), photosystem II (PSII), plastoquinone, cytochrome *b*<sub>6</sub>*f*, plastocyanin and photosystem I (PSI). The self-evident conclusion that can be drawn is that the main structural difference between both apparatuses is the antenna complex. In fact, the antenna complex is

the most diverse component among all photosynthetic organisms. It has been “re-invented” several times during evolution.

### ***Photosynthetic components: pigments***

---

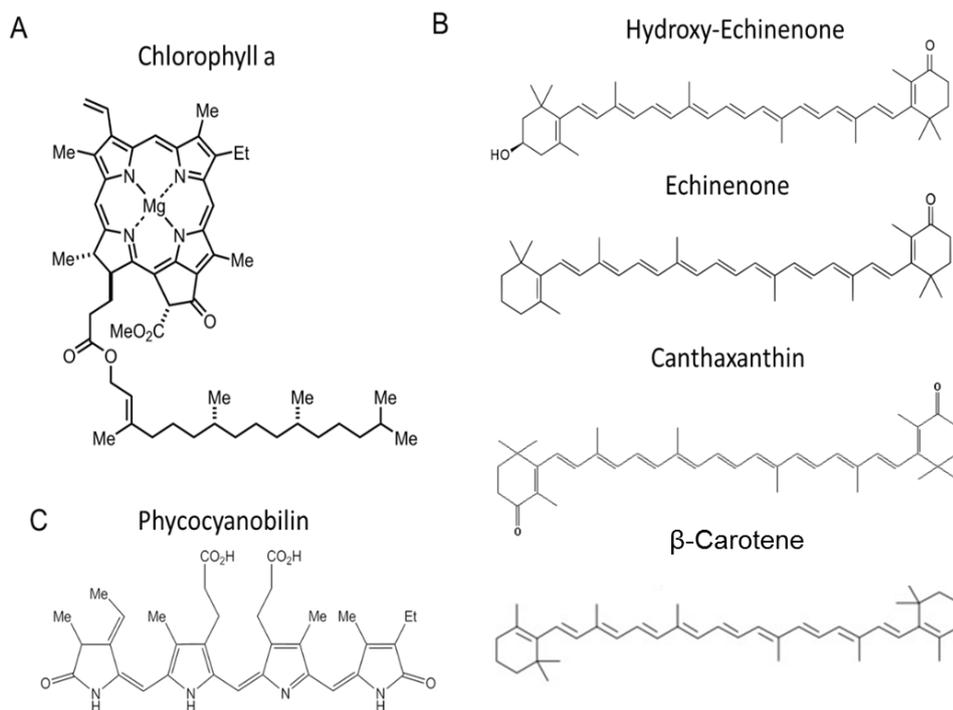
The key players in harvesting the light energy are the pigments. These molecules play a functional and structural role in the photosynthetic apparatus. In this section, the structural and spectral properties of the main pigments involved in photosynthesis are shortly reviewed.

**Chlorophyll.** The main pigment in the reaction centers is the chlorophyll (Chl), a planar organic molecule derived from a pyrrole. Diverse type of chlorophylls can be found in nature as a result of minor enzymatic modifications<sup>20</sup>. The most distributed chlorophyll in oxygenic photosynthetic organisms is **Chl a**, it occurs in reaction centers, in which it plays a major role during charge separation, and in light-harvesting complexes<sup>21</sup>. Oxidation of the methyl group at C-7 in **Chl a** into a formyl group result in a blue-shifted chlorophyll (**Chl b**) synthesis<sup>22</sup>. **Chl b** is found in the accessory light-harvesting complexes in eukaryotic organisms. The only cyanobacteria containing this chlorophyll are the prochlorophytes<sup>23</sup> which lack phycobiliproteins. **Chl c** which lacks the isoprenoid tail is commonly found in brown algae involved in the light-harvesting complex<sup>24</sup>. **Chl d** is found only in cyanobacteria belonging to the cyanobacterial genus *Acaryochloris*, representing the 95-99% of the total chlorophylls in this strain. **Chl d** absorbance is red-shifted compared with **Chl a**, and it is involved in light harvesting and reaction centers in these strains<sup>25</sup>. The recently discovered **Chl f** is the most red-shifted chlorophyll<sup>26</sup>. It can be found as a minor component in some cyanobacteria produced under far-red growth light and it was shown that can occur in reaction centers performing photochemistry<sup>27</sup>. The structure of chlorophyll a, shown in the figure 3A, is composed of five cyclic substructures (rings) with a Magnesium atom coordinated to four Nitrogen atoms which are part of the inner rings. **Chl a** presents two mayor absorption bands, one in the blue region (around 430 nm) called *soret band* and other two in the red region (around 650 nm) called  $Q_x$  and  $Q_y$ . Light absorption in the soret band lead to a second excited state (S2) that rapidly (in femtoseconds) decay to the lowest excited state (S1). This state is involved in all energy transfer process in photosynthesis.

**Carotenoids.** Among the most distributed pigments in the nature, carotenoids are contained in almost all organisms despite the fact that are only produced in plants, algae and some bacteria (for instance, cyanobacteria) and fungi. There are hundreds of different carotenoids that play different roles in nature<sup>28</sup>, from photosynthesis to nutrition and disease prevention. In this work I will focus only on

photosynthetic relevant carotenoids and their function on this process. They are elongated molecules formed by single and double carbon bonds containing one ring at each end. Examples of the carotenoid structure of Hydroxy-Echinenone, Echinenone,  $\beta$ -Carotene and Canthaxanthin are provided in the figure 3B. The conjugated  $\pi$ -electron system exhibits a strong absorption band in the range between 400 nm and 500 nm. This electronic transition is from the ground state ( $S_0$ ) to the second excited state ( $S_2$ )<sup>29</sup>. Transitions to the lowest excited state ( $S_1$ ) are forbidden because of the symmetric of the carotenoid molecule<sup>30</sup>. Excited  $S_2$  relax very fast (femtoseconds) to the  $S_1$  and from there to the ground state in a non-radiative way<sup>31,32</sup>. In addition, ultrafast spectroscopy demonstrated a strong effect of the solvent polarity on the  $S_1$  lifetime<sup>33</sup>. This was attributed to an additional excited state lying below  $S_1$  or coupled to  $S_1$ <sup>34</sup>, in which the electron density is shifted to the terminal ring of the polyene chain, designated as Intramolecular Charge Transfer (ICT)<sup>32,35</sup>. Nevertheless, both excited states ( $S_2/S_1$ ) play an important role in photosynthesis. It is well documented the role of carotenoid in light harvesting. It was shown long time ago, that carotenoid can absorb and transfer energy to chlorophylls<sup>36</sup> in the major antenna complex of plants and green algae (Light Harvesting Complex of Photosystem II, LHCII). In this case, excitation from carotenoid is transfer to  $Q_X$  or  $Q_Y$ . Similar phenomenon had been observed in Photosystem I from cyanobacteria<sup>37</sup>. Maybe the most investigated function of carotenoids in photosynthesis is the quenching of singlet-oxygen species ( $^1O_2$ )<sup>38</sup>. This toxic molecule is generated in stress conditions and carotenoids can both decrease its concentration and prevent its formation (see the '*Photoprotection: introduction to light stress*' section).

**Phycobilins.** In cyanobacteria and red algae, the antenna complex contains different pigments than in plants, open tetrapyrroles that absorb light between 550 nm and 650 nm. The spectral properties of the phycobillins allow these organisms to benefit of a broader light spectrum. There are four kinds of phycobillins present in the nature: phycocyanobilin (PBP), phycoerythrobilin (PEB), phycoerythrocyanin (PEC) and phycoviolobilin (PVB). The structure of phycocyanobilin is shown in the figure 3C.



**Figure 3 | Pigment Structures.** Structure of Chlorophyll a is shown in (A). Representative structures of carotenoid are shown in (B): Hydroxy-Echinenone, Echinenone,  $\beta$ -carotene and Canthaxanthin. Phycocyanobilin structure is shown in (C).

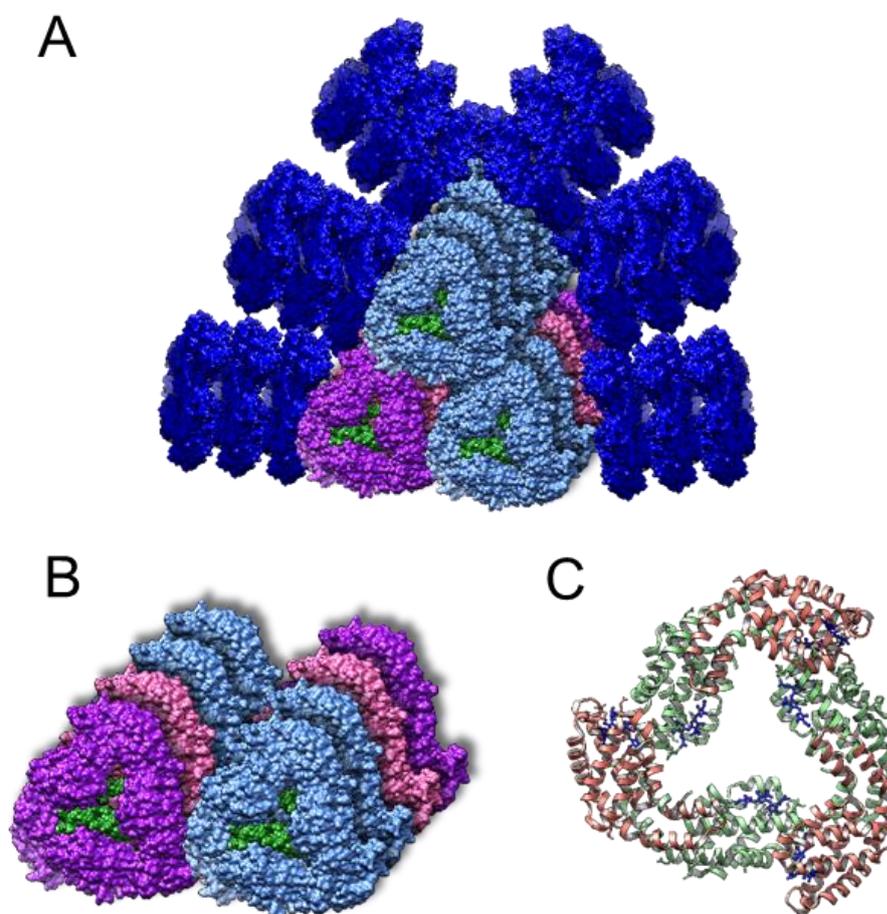
The spectral properties of the pigments described above are finely tuned by the protein environment in which they are embedded. These protein complexes, located in the thylakoid membrane, have evolved to bind and utilize the pigments to harvest and convert the light energy. In the following sections, the main protein complexes involved in oxygen-evolving photosynthesis in cyanobacteria are described and the main differences with plants are pointed out.

### ***Photosynthetic components: Antenna complex***

As it was mentioned before, a large part of the photosynthetic pigments that only collect and transfer the energy are in the antenna. Two big groups can be distinguished: *integral-membrane antennas* and *extrinsic antennas*. Another variant of antenna system are the inner membrane pigment-protein complexes (CP43 and CP47) that are closely associated with the reaction centers and are present in all oxygenic photosynthetic organisms. These antenna systems will be described together with the reactions center in the description of the photosystems.

A well-known example of membrane-bound antenna is the Light Harvesting Complex II (LHCII), present in plant but not in cyanobacteria. It exists as trimer and each monomer consist of a polypeptide of about 25 kilodaltons (kDa) binding several pigments: 8 chlorophylls a, 6 chlorophylls b and 4 carotenoids<sup>39</sup>. The chlorophylls are separated by approximated 10 Å, this close distance enable fast energy transfer between them (exciton equilibration in the LHCII trimer takes place in a few picoseconds) and effective energy transfer to the photosystem II. LHCII is part of a protein family called *lhc*. Other *lhc* genes, *lhc1-4*, are involved in the LCHI, the outer antenna complex associated with photosystem I<sup>40</sup>. The overall structure is well conserved in the *lhc* family as well as the pigment binding sites<sup>41</sup>, but slight differences in the protein environment tune the photochemistry of each antenna component. In addition, between the core of PSII and the trimeric LHCII, there are minor monomeric antennas from the same gene family, namely CP24, CP26 and CP29<sup>42</sup>.

The best example and the most relevant for this work of extrinsic antenna is the ***phycobilisome*** (PBS), the cyanobacterial antenna (Figure 4). The phycobilisome is a large water-soluble pigment-protein complex that carries out light harvesting in cyanobacteria and red algae. The pigments of the phycobilisome are the phycobillins that are covalently attached to proteins, called *phycobiliproteins*. Depending on the number and type of phycobilin that is bound, the phycobiliproteins are divided in 4 kinds<sup>16</sup>: allophycocyanin (APC,  $\lambda_{\max} = 650$  nm), phycocyanin (PC,  $\lambda_{\max} = 620$  nm), phycoerythrin (PE,  $\lambda_{\max} = 560$  nm) and, the less frequent, phycoerithrocyanin (PEC,  $\lambda_{\max} = 575$  nm). Specific lyases catalyze the covalent attachment of the bilin to phycobiliprotein cysteines, for example APC and PC contain phycocyanobilin. The different protein environments of the phycobilin in the different proteins give rise to different spectral properties which allow a very efficient energy transfer within the phycobilisome.



**Figure 4 | Phycobilisome structure.** (A) Side view of a *Synechocystis* phycobilisome. It is composed of 6 rods formed by 3 phycocyanin trimers radiating from the core. The core (zoom view of bottom cylinders in B) is composed of APC trimer (shown in sky blue), APC trimer containing ApcD (shown in violet), the APC trimer containing ApcF and ApcE (shown in pink) and the linker protein, ApcC (shown in green). The rods are formed PC (dark blue). (C) The APC trimer structure is comprised of APC- $\alpha$  (green) and APC- $\beta$  (orange), both bind billins (shown in blue). The APC structure used in the schema and in (C) correspond to PDB: 4PO5. Adapted from Muzzopappa and Kirilovsky 2019.

Despite the fact that different variants can be found in the nature, the most common architecture of the cyanobacterial phycobilisome (called hemidiscoidal PBS) is a fan-like structure with PE/PC rods that radiate from an APC core. All the phycobiliprotein can be found as heterodimer formed by two polypeptides, defined as chain  $\alpha$  or  $\beta$ . Both  $\alpha$  and  $\beta$ , are globular proteins (around 20 kDa) composed of eight helical elements. Evolutionary analysis of these proteins suggested that they evolved from the same ancestor and followed the same evolutionary pathway, giving rise to the same structural patterns (APC, PC, PE and PEC). This equal evolution of the  $\alpha$  and  $\beta$  polypeptides (into APC, PC, PE and PEC) indicate that a phenomenon of co-evolution took place, guided by a strong interaction between  $\alpha$

and  $\beta$ . The basic subunit of the phycobilisome is formed by a heterodimer  $\alpha/\beta$  that are assembled into ring-shaped trimers  $(\alpha/\beta)_3$  and further into hexamers  $(\alpha/\beta)_2$ . Four trimers of APC heterodimers are associated forming cylinders in the phycobilisome core. The number of cylinders in the core depends on the strain: for example, *Synechocystis* have three cylinders arranged in a pyramid-like fashion, *Nostoc* has five and *Synechococcus elongatus* has two. The cylinders are held together due to the binding of specific linker proteins. The basal cylinders which are in contact with the thylakoid membrane have a different composition than the upper cylinder. Two trimers are made up of pure APC  $\alpha/\beta$  subunits. In the third trimer, one  $\beta$  subunit is replaced by ApcF and in the other subunit, an  $\alpha$  subunit is replaced by ApcE. In the last trimer, in one subunit, the  $\alpha$  polypeptide is replaced by ApcD. ApcF, ApcD and ApcE are modified APC proteins, with a red shifted spectrum compared with the bulk of APCs, ApcE also contain a linker domain that maintain the PBS core stability. The linker protein ApcC also stabilize the APC cylinders. In a similar way, the PBS rods are composed of two or three PC hexamers connected by linker proteins. Most of the marine cyanobacteria have PE and/or PEC hexamers at the end of the rods. These phycobiliproteins are blue-shifted compared with PC and expand the usable light range.

The overall arrangement of chromophores in the PBS facilitates the downhill energy transfer from the rods to the core and from there to the photosystems. Due to the intermediary distance between chromophore, around 20-40 Å, the model proposed is the Foster resonance energy transfer that required weak dipole coupling between chromophores and overlapping of emission spectrum of the donor and absorption spectrum of acceptor. Therefore, excitation of PC (Absorption  $\lambda_{\max} = 620$  nm, Emission  $\lambda_{\max} = 645$  nm) is transferred to the APC in the core (Absorption  $\lambda_{\max} = 650$  nm, Emission  $\lambda_{\max} = 660$  nm). Then the energy is transferred towards the terminal emitters (ApcD, ApcF and ApcE, Absorption  $\lambda_{\max} = 670$  nm, Emission  $\lambda_{\max} = 680$  nm) and from there to the chlorophylls in the photosystems.

The PBS interact with the stromal side of the thylakoid membrane possibly through ApcE<sup>43</sup>. PBS diffusion on the membrane surface allow energy supply to both photosystems<sup>43</sup>. Phycobilisome are thought to be the main antenna of PS II, albeit they can also interact and transfer energy to PS I. It seems that the fraction of energy delivery to PSI or PSII depends on the strain, in *Synechocystis*<sup>44</sup> is 50% to each photosystem but in *Arthrospira*<sup>45</sup> 80% is transferred to PSI. Depending on the strain, the different terminal emitters can interact distinctively with the photosystems, creating different energy pathway towards PSII and PSI. In *Synechococcus elongatus* sp. PCC 7002 deletion of ApcD render in a decrease of energy transfer to PSI and deletion of ApcF decreases energy transfer to PSII<sup>46</sup>. On the other hand, in *Synechocystis* ApcF is involved in energy transfer to both

photosystems and deletion of ApcD has a minor role in energy transfer<sup>47</sup>. Recently, it was shown that ApcD was involved in energy transfer to PSI and ApcF had a minor role in energy transfer in *Synechococcus elongatus* PCC 7942<sup>48</sup> (see “appendix 1: Additional publications”). In some strains, these proteins have a role on the regulation of the energy arriving to the photosystems (see below ‘Photoprotection: cyanobacterial state transition’).

## ***Photosynthetic components: Photosystem II***

---

In organism performing oxygen-evolving photosynthesis, such as plants, green algae and cyanobacteria, there are two photosynthetic units: Photosystem I (PSI) and Photosystem II (PSII). PSII is very unique biological complex due to the fact that is capable of carrying out a very thermodynamically demanding reaction, the water oxidation. A great effort was dedicated to isolate and study the photosystem II, leading to the elucidation of the structure of several PSII complexes from different organisms<sup>49–51</sup>, showing high degree of similarity between them.

The photosystem II is a membrane-integral complex composed of two identical monomers, each one of them formed by 19 protein subunits. Each monomer contains 36 chlorophyll a, and 7 carotenoid molecules. A schematic figure<sup>50,52</sup> of cyanobacterial PSII is shown in figure 5A. The core of the reaction center is formed by two homologue proteins, D1 (encoded by the *PsbA* gene) and D2 (encoded by the *PsbD* gene), forming a symmetric heterodimer. Each monomer comprises of five  $\alpha$ -helices exposing the N-termini to the stromal side of the thylakoid membrane. A chlorophyll dimer in the D1/D2, designated as P<sub>D1</sub> and P<sub>D2</sub>, are weakly coupled. The presence of other pigments in each subunit, another chlorophyll a (Chl<sub>D1</sub> or Chl<sub>D2</sub>), one pheophytin (Ph<sub>D1</sub> or Ph<sub>D2</sub>) and one plastoquinol (Q<sub>A</sub> and Q<sub>B</sub>) creates two branches for the electron transfer chain (Figure 5B). Interestingly, only one branch is active. After the excited state of the special pair P<sub>D1</sub>/P<sub>D2</sub> is reached, the energy is enough to promote charge separation and the state P<sub>D1</sub><sup>+</sup>Ph<sub>D1</sub><sup>-</sup> is formed, *via* the intermediary Chl<sub>D1</sub><sup>53</sup>. Due to its spectral properties, P<sub>D1</sub><sup>+</sup> is usually named as P680<sup>+</sup>. The radical pair is quickly stabilized (in about 400 ps) by electron transfer to Q<sub>A</sub>, which is strongly bound to the protein matrix and it is not exchangeable. The Q<sub>A</sub> role is to receive a single electron and mediated the double reduction of Q<sub>B</sub>, which after being doubled reduced (by two successive charge separations) can be replaced by a new plastoquinone from the plastoquinone pool. Attached to the luminal side of D1/D2 heterodimer, there is an inorganic complex, Mn<sub>4</sub>CaO<sub>5</sub>, called Oxygen Evolving Complex (OEC)<sup>51</sup>. This complex catalyzes the oxidation and splitting of water to restore the electron donated by P680<sup>+</sup>. Photo-oxidized P680<sup>+</sup> drags one electron from the Mn atoms of OEC

through the Tyrosine 161 of D1 (Tyr Z). Four sequential oxidations of the OEC allows this complex to oxidize two water molecules and produce O<sub>2</sub> and protons.

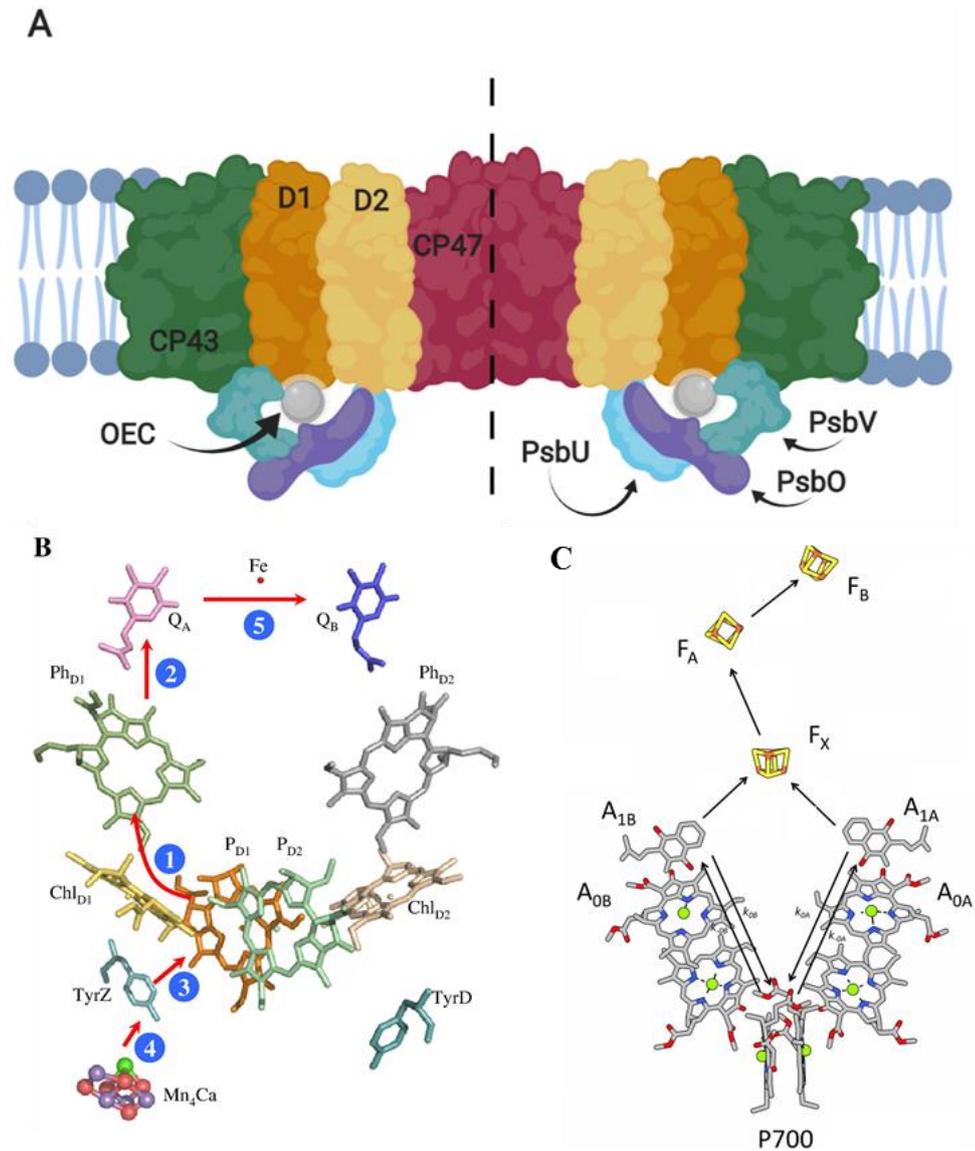
Specific amino acids of D1 (and one from CP43) coordinate the magnesium cluster and others protein (PsbO, PsbU, PsbV and extrinsic loops of D1, D2, CP43 and CP47) are required to stabilize it<sup>54</sup>. In cyanobacteria, there are several isoforms of D1. These isoforms have different tolerance to oxidative damage due to slight difference in the amino acid sequence. Therefore, the isoforms are differentially expressed depending on the light conditions.

Formed by two subunits ( $\alpha$  and  $\beta$ ) and one heme group, the cytochrome *b*<sub>559</sub> (Cyt *b*<sub>559</sub>) is located near the D2. The Cyt *b*<sub>559</sub> is involved in PSII assembly<sup>55</sup> and it is essential for its photochemistry, since the minimum PSII active complex is composed of D1/D2/Cyt *b*<sub>559</sub><sup>56</sup>. The role of Cyt *b*<sub>559</sub> is still enigmatic but several photoprotective mechanisms involving this component have been proposed<sup>57</sup>. To increase the light absorption, two antenna subunits are located surrounding the D1/D2 core. CP43 is composed by five helices and it contain 14 chlorophylls while CP47 is composed by six helices and three  $\beta$ -sheets containing 16 chlorophylls. Two small subunits, PsbI and PsbX stabilize the two peripheral chlorophylls (Chl<sub>ZD1</sub> and Chl<sub>ZD2</sub>, respectively). In the luminal side, PsbV, PsbU and PsbO enclose the OEC.

### ***Photosynthetic components: cytochrome b<sub>6</sub>f***

---

Another essential player of the electron flow in cyanobacterial membranes is the cytochrome *b*<sub>6</sub>*f*. This protein complex is a dimer with eight subunits per monomer. Among these subunits, cytochromes (proteins that covalently bind one or more heme groups) are found<sup>58</sup>. Cytochrome *f* is formed by a single transmembrane helix that anchors the soluble domain, which harbor a heme group, in the luminal side of the membrane. The subunit cytochrome *b*<sub>6</sub> possess four transmembrane helices and coordinates two heme groups (*b*<sub>n</sub> and *b*<sub>p</sub>) and, an additional heme group, *c*<sub>n</sub> that is electronically coupled with *b*<sub>n</sub>. An important component of the cytochrome *b*<sub>6</sub>*f* is the Rieske Fe-S protein. This subunit consists of two domains, a transmembrane N-terminal domain and a soluble domain which coordinate a Fe-S cofactor. Another four small subunits (PetG, PetL, PetM, and PetN), with structural role, and the subunit IV complete the complex monomer. In addition, one carotenoid and one chlorophyll, of unknown function, are embedded on the complex.



**Figure 5 | Photosystem II structure.** (A) Schematic side view of dimeric PSII complex. The schematic representations are based on the PSII structure of *Thermosynechococcus* from Kamiya et al. 2003 and Umena et al. 2011. (B) Electron pathway in the reaction center of PSII, adapted from Cardona et al. 2012. The first charge separation originated the radical pair  $\text{Chl}_{D1}^+ \text{Ph}_{D1}^-$  that rapidly is converted into the more stable  $\text{P}_{680}^+ \text{Ph}_{D1}^-$  (1). Then the  $\text{Ph}^-$  can donate one electron to the plastoquinone in  $\text{Q}_A$  (2).  $\text{P}_{680}^+$  is reduced by the OEC via the TyrZ (3 and 4). Finally, the doubled reduced  $\text{Q}_A$  can reduce the plastoquinone in  $\text{Q}_B$  (5). (C) Electron pathway in the reaction center of PSI, adapted from Makita et al. 2017. Excitation of the reaction center of PSI lead to charge separation of  $\text{P}_{700}^+ \text{A}_0^-$ , that can reduce  $\text{A}_1$  and then Fe-S cluster  $\text{F}_X$ . The electron moves quickly from  $\text{F}_X$  to  $\text{F}_B$  through  $\text{F}_A$ .

The electron transfer within the complex is well known and defined as “Q cycle”<sup>59</sup>. One reduced plastoquinone ( $\text{PQH}_2$ ) binds to the  $\text{Q}_0$  site at cytochrome  $b_6$ , its oxidation lead to one electron transfer to the external plastocyanin, a soluble small copper-binding protein, through the Fe-S cluster and one electron transfer to one oxidized plastoquinone (PQ) bound to the  $\text{Q}_i$  site, through the heme groups  $b_n$

and  $b_p$ . In this first half cycle two protons are released to the luminal side and one proton from the stromal side was consumed during reduction of PQ into PQH in the  $Q_i$  site. During the second half cycle, another PQH<sub>2</sub> is fully oxidized, releasing the 2 protons in the luminal side. One electron is transported to another plastocyanin and the other is used to reduce the PQ<sup>-</sup> in the  $Q_i$  site to PQH<sub>2</sub>, consuming another proton from the stromal side. The full cycle drives the oxidation of one PQH<sub>2</sub>, reduction of two plastocyanin, the transport of two protons and the accumulation of two extra protons in the luminal side.

### ***Photosynthetic components: Photosystem I***

---

The second photosynthetic unit in the oxygen-evolving photosynthetic apparatus is the photosystem I (PSI). PSI, like PSII, is a transmembrane pigment-protein complex. From a structural point of view, PSI from plants and cyanobacteria present some important differences. Maybe, the most remarkable difference is that cyanobacterial PSI, depending on the growth conditions can be found as trimer (*in vivo*<sup>60</sup> and *in vitro*<sup>61</sup>) or monomer<sup>62</sup>. Plant PSI was found to be monomeric. This evolutionary loss of trimeric formation of PSI in plants may be explained by the fact that plant PSI incorporates antenna complex LHCI and the trimeric architecture of PSI does not allow such antenna binding<sup>63</sup>. According to this, LHCI is not found in cyanobacteria<sup>64</sup>. Monomers of cyanobacterial PSI are formed by 12 subunits with a molecular weight around 80 kDa. The core of the reaction center I is a heterodimer PsaA-PsaB which binds a special pair of chlorophylls (defined as P700), four additional chlorophylls (pair A and pair A<sub>0</sub>) and two phylloquinones molecules. An important difference between PSII and PSI is that PSI core also contains one iron-sulfur cluster (Fe-S), F<sub>X</sub>. Another two Fe-S clusters (F<sub>A</sub> and F<sub>B</sub>) are bound to PsaC protein, located next to F<sub>X</sub> in the stromal side. These Fe-S clusters play an important role in the primary electron transfer. Additional pigments surrounding P700 form the intrinsic antenna of PSI, increasing the effective absorption area of PSI. Around 79 chlorophylls are bound to PsaA-PsaB and another 11 chlorophylls and 22 carotenoids are coordinated to the small subunits PsaJ, PsaK, PsaL, PsaM, and PsaX.

When the excitation reaches the PSI core a primary charge separation takes place leading to the oxidation of P700 into P700<sup>+</sup> and the reduction the chlorophyll A<sub>0</sub> (Figure 5C). Then, the electron moves quickly to F<sub>X</sub> through the phylloquinone A<sub>1</sub>. The Fe-S cluster F<sub>X</sub> transfers the electron to F<sub>A</sub>, and from here it is transferred to F<sub>B</sub><sup>65</sup>. Ultimately the electron is used to reduce a soluble protein called Ferredoxin. Due to the symmetry of PsaA-PsaB two branches for the electron transport are present. Interestingly, while in PSII only one branch is active, in PSI both branches

are active. The reduced state of P700<sup>+</sup> is restored by electron transfer from the plastocyanin.

### ***Photosynthetic electron transport: linear versus cyclic***

---

As it was mentioned before, the light harvested by the antenna complexes induces a *redox* reaction in the core of the photosynthetic units PSII and PSI. This triggers an electron transport across the photosynthetic membrane with two important consequences: the generation of reducing power and the generation of a transmembrane potential by coupling the electron transport to proton accumulation in one side of the thylakoid membrane.

The canonical photosynthetic electron transport follows a linear fashion. Upon light absorption by the phycobilisomes (or LHClI), energy is transferred to PSII. This leads to the reduction of the plastoquinol pool consuming 2 H<sup>+</sup> from the stromal side to generate PQH<sub>2</sub>, as it was described before. Water splitting by the OEC produces 4 H<sup>+</sup> in the luminal side. The electrons flow from the plastoquinol pool to the cytochrome *b<sub>6</sub>f* oxidizing PQH<sub>2</sub> into PQ and accumulating four H<sup>+</sup> in the luminal side, as it was detailed before. Plastocyanin transport electrons, one by one, from cytochrome *b<sub>6</sub>f* to PSI. These electrons are used to re-reduce the photo-oxidize P700<sup>+</sup> of PSI that was previously oxidized by the Ferredoxin. Ferredoxin is a soluble protein carrying a Fe-S cluster which is capable of reducing another protein, the Ferredoxin-NADP Reductase (FNR). FNR is a two domains protein. One domain binds Flavin Adenine Dinucleotide (FAD) and its function is to receive one-by-one electrons from Ferredoxin. The other domain binds NADP<sup>+</sup> and catalyzes the two-electron reduction to NADPH<sub>2</sub>. The transmembrane proton gradient generated is ultimately used to catalyze phosphorylation of adenosine diphosphate (ADP) into adenosine triphosphate (ATP) by the ATPase. Both ATP and NADPH are used to fix CO<sub>2</sub> into biologically usable molecules via the Benson-Calvin cycle.

An alternative electron flow can take place if electrons are re-injected in the photosynthetic chain to circulate around PSI<sup>66</sup>. This mechanism leads to the decrease of NADPH<sub>2</sub> formation and an increase of proton accumulation on the luminal side of the membrane, leading to a greater ATP production and less reductive power. In cyanobacteria, two main pathways are involved in cyclic electron transport. The NDH-1 complex, a transmembrane protein complex that resembles the Complex-I of the respiratory chain in mitochondria<sup>67</sup>, is capable to receive electrons from Ferredoxin<sup>68</sup> and transfer them to PSI via the PQ pool. This complex can bind Ferredoxin<sup>69</sup> and NDH-1:PSI supercomplex has been detected<sup>70</sup>. In plants, the second pathway of electron recycling is through a

Ferredoxin/plastoquinone oxidoreductase named PGR5<sup>64</sup>. PGR5 is a transmembrane protein which directly catalyzes the reduction of PQ and the oxidation of Ferredoxin. A second pathway exists in cyanobacteria, PGR5 homologues were suggested for this<sup>71</sup> but the exact identity of the mechanism remains unknown.

### *Photoprotection: introduction to oxidative stress*

---

In photosynthetic organisms, detrimental effects are triggered due to the over-reduction of the photosynthetic electron transport chain. This over-reduction of the chain can be generated by exposure to high lights or other conditions such as nutrient starvation and CO<sub>2</sub> limitation. Under these conditions light-induced reactions produce Reactive Oxygen Species (ROS) in the cell. ROS, that can be hydroxyl radical (<sup>•</sup>OH), hydrogen peroxide (H<sub>2</sub>O<sub>2</sub>), superoxide radical (<sup>•</sup>O<sub>2</sub><sup>-</sup>) and singlet oxygen (<sup>1</sup>O<sub>2</sub>), are highly reactive molecules that can damage DNA/RNA molecules, oxidize lipids, proteins and cofactors. The long-lived excited state of chlorophyll, required for energy transfer during photosynthesis, plays a role in <sup>1</sup>O<sub>2</sub> production. In this long-lived state, the spins of the electrons of excited chlorophyll can rephase and populate the lowest energy-excited state, the chlorophyll triplet state (<sup>3</sup>Chl). <sup>3</sup>Chl can react with <sup>3</sup>O<sub>2</sub> (the most common form of molecular oxygen) producing singlet oxygen. The vast majority of chlorophylls are present in the antenna, in which the presence of nearby carotenoids decreases significantly the <sup>1</sup>O<sub>2</sub> production<sup>73</sup>. The close distance between pigments in these complexes, allow orbital overlap and spin-exchange reaction that quench the <sup>3</sup>Chl by populating the carotenoid triplet state. This state is safely dissipated as heat. Nevertheless, at high light intensity both singlet oxygen production and protein degradation have been detected at the level of the PSII antenna<sup>74</sup>.

Most of the <sup>1</sup>O<sub>2</sub> is generated in PSII (reviewed in<sup>75</sup>). In high light conditions, the plastoquinol pool is completely reduced and the photosynthetic electron transport is blocked. Charge recombination between the primary quinone acceptor (Q<sub>A</sub><sup>-</sup>) and P<sub>680</sub><sup>+</sup> takes place. Two possible pathways are possible, the first one involves the direct recombination into P<sub>680</sub> and Q<sub>A</sub>. The second possible pathway of charge recombination goes through the repopulation of the first radical pair (P<sub>680</sub><sup>+</sup>Phe<sup>-</sup>). This radical pair can also undergo triplet radical formation, <sup>3</sup>[P<sub>680</sub><sup>+</sup>Ph<sup>-</sup>], that ultimately lead to the formation of <sup>3</sup>P<sub>680</sub>, which can react with <sup>3</sup>O<sub>2</sub> and produce singlet oxygen. In the reaction center, the carotenoid molecule is located too far from the chlorophylls to quench the <sup>3</sup>Chl state and singlet oxygen is efficiently

generated. The  $^1\text{O}_2$  can diffuse in the reaction center and damage the D1 subunit leading to its degradation in a process called **photoinhibition**. This process takes place even at low light intensity, but since the repair of PSII is very efficient *in vivo* only at high light intensities a net loss of PSII activity is observed<sup>76</sup>. It is important to note that other physiological conditions such as low temperatures or  $\text{CO}_2$  depletion can favor the limitation of the acceptor side of PSII and increase the probability of charge recombination. Additionally, the midpoint of  $\text{Q}_\text{A}$  can favor the different recombination pathways. In cyanobacteria, an alternative isoform of D1, encoded by *psbA3*<sup>77</sup>, replaces D1 (*psbAI*) at high light conditions. This isoform shifts the redox potential of pheophytin, favoring the non-radiative recombination pathway without  $^1\text{O}_2$  production.

The cytochrome *b<sub>6</sub>f* and PSI also harbor chlorophylls and are theoretically capable of  $^1\text{O}_2$  production. However, it has been proved that the contribution from these complexes is minor compared with PSII. In cytochrome *b<sub>6</sub>f*, the protein environment shorts the lifetime of the excited chlorophyll decreasing the chances of populating the triplet state<sup>78</sup>. Another potential source of  $^1\text{O}_2$  generation is during chlorophyll synthesis, accumulation of precursors can also increase the singlet oxygen formation. In the electron acceptor side of PSI, the excess photo-excitation can lead to reduction of  $\text{O}_2$  into  $\text{O}_2^-$ <sup>79,80</sup> (this reaction is also called Mehler reaction) by the primary electron acceptors, which is further reduced to  $\text{H}_2\text{O}_2$  and  $\text{OH}^-$ <sup>81</sup>. These ROSs can damage and inhibit PSI. Since the PSI turnover time is about days<sup>82</sup>, photosynthetic organisms have mechanisms to keep P700 oxidized and avoid ROS production (see “ROS scavengers and alternative electron pathways.”). In agreement with this, PSI photoinhibition hardly takes place *in vivo*<sup>82</sup>.

In order to cope with the photo-oxidative stress, plants and cyanobacteria have developed many photoprotective mechanisms. Some of these mechanisms are collectively denominated as Non-Photochemical quenching (NPQ) since they decrease the chlorophyll fluorescence by increasing the thermal dissipation and, as consequence, reducing the energy arriving to the photosystems. They can be sorted by the molecular level at which they take place: for instance, preventive mechanisms dissipating the excess of energy at the antenna level or other mechanisms directly deal with already generated ROS. They also can be sorted by the time range when they take place: we can distinguish short-term mechanisms and long-term mechanisms. In the following sections, the principal photoprotective mechanisms of cyanobacteria will be introduced with special focus on the OCP-related Non-Photochemical-Quenching.

## ***ROS scavengers and alternative electron pathways.***

---

Cyanobacteria, as well as many organisms, contain enzymes to detoxify reactive oxygen species. *Synechocystis* uses an Iron Superoxide Dismutase (Fe-SOD) in order to catalyze the dismutation of  $2 \cdot\text{O}_2$  into  $\text{H}_2\text{O}_2$  and  $\text{O}_2$ . However, some marine *Synechococcus* strains use copper/zinc SODs<sup>83</sup>.  $\text{H}_2\text{O}_2$  is detoxified by different enzymes: catalases, peroxidase and peroxiredoxins (reviewed in<sup>84</sup>). Beside the enzymatic reactions that decrease the ROS concentration presented above, cyanobacteria also contain small molecules capable of ROS scavenging. An already described example are the carotenoids. In addition, tocopherol can decrease the  $^1\text{O}_2$  concentration in the reaction center of PSII<sup>85</sup>. Collectively, these enzyme and molecules deal with the ROS already produced in the photosynthetic machinery.

In order to prevent the ROS production, other photoprotective mechanisms constitute an alternative pathway for electrons acting like a valve that releases the pressure from the electron transport chain. An example of this is the Plastid Terminal Oxidase (PTOX), that is present in some cyanobacteria<sup>86,87</sup>, receives electron from the PQ pool and reduce  $\text{O}_2$ . This protect the PQ from over-reduction and as a consequence avoid PSII photoinhibition.

Flavodiiron proteins (Flv) are a family of enzymes composed of two domains: the N-terminal metallo- $\beta$ -lactamase-like domain harboring a non-heme diiron center capable of  $\text{O}_2$  and/or NO reduction and the C-terminal flavodoxin-like domain, containing a flavin mononucleotide (FMN). The Flv found in photosynthetic organism belong to the Flv subgroup C that contains an additional domain that allows NADPH to be used as electron donor (reviewed by Allahverdiyeva *et al.*<sup>88</sup>). Cyanobacterial Flv can be clustered into four different groups, Flv 1 to Flv4. Flv1-Flv3 form a heterodimer that catalyze a “Mehler-like” reaction in the donor side of PSI. In this reaction, Flv1-Flv3 oxidize NAD(P)H and reduce  $\text{O}_2$  to  $\text{H}_2\text{O}$  in a four-electron reaction without ROS production. This photoprotective role is very important under fluctuating light<sup>89</sup>. Some cyanobacteria, for instance *Synechocystis* also contain a second pair of Flavodiiron proteins forming another heterodimer, Flv2-Flv4. Although its mechanism is not clear yet, Flv2-Flv4 are proposed to mediate the electron transfer from PSII (or PQ pool) to an unknown electron sink<sup>88,90</sup> however this model is still under debate.

## ***Photoprotection: State transitions***

---

The light intensity is not the only issue for photosynthetic organisms, the quality the light (the portion of the electromagnetic spectrum irradiated, measured in wavelength) also can affect the photosynthetic capacity. Since in cyanobacteria PSI

contains more chlorophylls than PSII, illumination with blue (absorbed mostly by chlorophylls) or far-red light (absorbed mainly by PSI) will stimulate more PSI activity compared with PSII activity leading to oxidation of the components between the photosystem, such as the plastoquinol (PQ) pool. Since the main antenna of PSII is the PBS, orange illumination (mostly absorbed by PBS) leads to the induction of PSII activity and the reduction of the electron chain (and the PQ pool). This imbalance on the activity of the photosystem can promote the ROS production or/and decrease the photosynthetic capacity.

Plants, algae and cyanobacteria have developed a mechanism to re-balance the energy distribution between the photosystems called “state transition”. The imbalance is sensed by the redox-state of the PQ pool<sup>91</sup>. In darkness, cyanobacterial cells perform respiration (in the same thylakoid membrane) which lead to reduction of PQ pool. Under this condition cells are in *State II*, this state is characterized by a low PSII to PSI fluorescence ratio at 77K and a low PSII chlorophyll fluorescence yield at room temperature. If cells are illuminated with low intensities of blue or far-red light, PSI became more active and the PQ pool more oxidized inducing transition to *State I*. In this state, PSII Chl fluorescence yield increases at room temperature as well as the ratio PSII to PSI fluorescence at 77K. Illumination with orange light will activate PSII and reduce the PQ pool, inducing transition to *State II*.

While it is accepted that in plants the redox state of the PQ pool is sensed by the cytochrome *b<sub>6</sub>f* which activate a specific kinase that phosphorylate LCHII<sup>92</sup> and induce the movement of antenna from PSII to PSI<sup>93</sup>, in cyanobacteria the mechanism is not clear. The involvement of phosphorylation and cytochrome *b<sub>6</sub>f* was controversial, but recently it was proved that none of them are involved<sup>94</sup> in the mechanism of state transitions. Regarding the mechanism, based on diverse experimental observation several models were proposed: 1) Movement of PBS between the phycobilisomes. This model was proposed based on FRAP experiments that showed a high mobility of PBS compared with photosystems<sup>95,96</sup>. According to this model, in state I PBS are bound to PSII and upon transition to state II they move to PSI. In addition, chemicals that alter the PBS-membrane interaction, such as phosphate, sucrose or betaine, inhibit the state transitions<sup>97</sup>. However, it was recently showed that these chemicals also affect processes in the membrane<sup>94</sup>. 2) Spillover between PSII and PSI<sup>98</sup>. This model requires a variable coupling between PSII and PSI. In State I this coupling is weak and the energy harvest by the PBS is transferred to the PSII while in state II strong coupling between PSII and PSI allows further transfer of this energy form PSII to PSI. This photosystem movement on the membrane is in agreement with the strong temperature dependence of the state transitions<sup>99</sup> and electron microscopy

observations<sup>100</sup>. Supporting this hypothesis, Liu and coworkers<sup>101</sup> isolated supercomplexes involving both photosystems and PBS by using crosslinkers. 3) Detachment of PBS from PSII and Quenching in PSII. In this model is proposed that in state I the increase on room temperature fluorescence is due to PBS detachment from PSII. In addition, a recent work in *Synechococcus elongatus* PCC 7942 showed that in state II the PSII is quenched and not involved in spillover<sup>102</sup>.

The PBS are involved in almost all the models, as major antenna complexes, however the specific role of the antenna and the different terminal emitter is not so clear. As it was described before (see “Photosynthetic components: Antenna complex”) depending on the strain, the PBS terminal emitters play different role in the energy transfer to photosystems. In *Synechocystis*<sup>47</sup>, *Synechococcus elongatus* PCC 7002<sup>46</sup> and *Anabaena*<sup>103</sup>, it was shown that deletion of *apcD* impaired state transitions. It was proposed that this was due to a permanent association between PBS and PSII, since *ApcD* was thought to be involved in energy transfer to PSI. On the other hand, deletion of *apcF* impaired state transition in *Synechocystis* but not in *Synechococcus elongatus* PCC 7002. Recently, a work<sup>48</sup> carried out in our group, in which I collaborate by designing and constructing *Synechococcus elongatus* PCC 7942 mutants, showed that the role of these terminal emitters and their involvement on state transitions depend on the strain. In *Synechococcus elongatus* sp. PCC 7942, *ApcD* is involved in energy transfer to PSI but deletion of *ApcF* has no effect on energy transfer from PBS to any photosystem. By contrast, in *Synechocystis* *ApcF* is involved in energy transfer to both photosystems and deletion of *ApcD* does not affect the energy transfer<sup>47</sup>. In addition, while deletion of *apcD* in *Synechocystis* completely inhibit state transition and deletion of *apcF* also showed a strong alteration on this process, neither deletion of *apcD* nor *apcF* affects state transitions in *Synechococcus elongatus* PCC 7942. It is clear that the function of the terminal emitters is not the same in the different strains implying that their sequence or/and the overall architecture of the PBS dictate its role. Their role on state transition is not universal for all the cyanobacteria. Probably state transitions is the result of a series of concerted mechanisms and their relative contribution depends on the specific strain and the light regime that the strain experience.

### ***Photoprotection: OCP-related NPQ***

---

During evolution photosynthetic organisms have found their way to reduce the amount of photo-oxidative damage by just decreasing the energy arriving to PSII, the main source of ROS production. An elegant and efficient approach to achieve this relies on the thermal energy dissipation at the level of the antenna complex. This phenomenon can be followed by measuring how the chlorophyll fluorescence

yield decreases (due to a decrease in the antenna size of PSII) under high light illumination. This component of the (short-term) NPQ, named qE in plants and OCP-related NPQ in cyanobacteria, takes place in the time range of second and minutes. Another important feature of this mechanism is the reversibility of the process under normal light conditions.

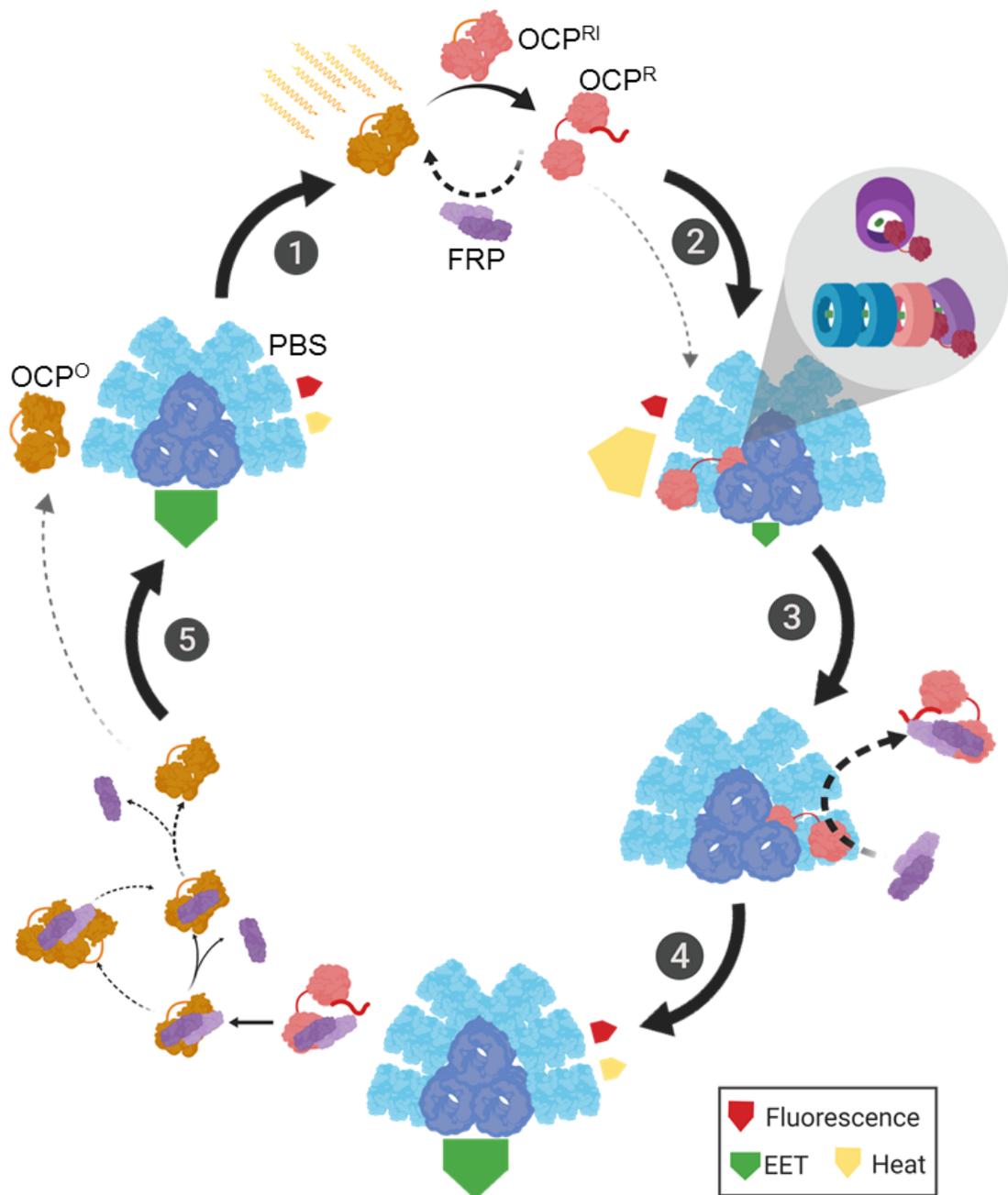
A key factor in these mechanisms is the molecular sensing of the stressful conditions. In plants and green algae this is accomplished by sensing the  $\Delta\text{pH}$  across the thylakoid membrane. Due to the molecular diversity of antenna among the photosynthetic organisms, the short-term NPQ mechanism is different in plants and cyanobacteria. In plants, it is known that protonation of PsbS subunit of LCHII, the xanthophyll cycle activation (conversion of Zeaxanthin to Violaxanthin) and aggregation of the antenna are required for the energy dissipation (reviewed in<sup>104</sup>). Besides that, the exact mechanism is still controversial.

In cyanobacteria, the absence of xanthophyll cycle and the lack of effect of ionic uncouplers on chlorophyll fluorescence (indicating that a  $\Delta\text{pH}$  does not trigger any NPQ mechanism) prompted the researchers to hypothesize<sup>105</sup> that there is not a short-term NPQ mechanism similar to plant qE in this phylum. Instead, the chlorophyll fluorescence changes were attributed to another NPQ mechanisms, state transitions (see above "*photoprotection: state transition*"). This view was challenged by El Bissati and co-workers in 2000 when they provided the first experimental evidence of the existence of this kind of mechanism<sup>99</sup>. They were able to show the existence of a different process which takes place upon strong light illumination and it is independent on the temperature and membrane state. On the other hand, chlorophyll fluorescence changes during state transitions are dependent on the temperature and the membrane physical state. In addition, the authors were able to show that this quenching of the chlorophyll fluorescence was not related to photoinhibition and it can take place at low temperatures which inhibit state transitions. These experimental data pointed towards a novel photoprotective mechanism taking place at the level of the phycobilisome. The next step for describing this new mechanism was provided by Rakhimberdieva and co-workers, by measuring the action spectra of quenched cells (by inducing NPQ at high blue light illumination) they were able to propose that the quenching was induced by a carotenoid molecule<sup>106</sup>.

The decisive breakthrough to decipher the NPQ mechanism on cyanobacteria was achieved by Wilson<sup>107</sup> and co-workers in 2006. They could show that a previously described carotenoid protein, the Orange Carotenoid Protein (hereafter OCP) was directly involved in the cyanobacterial NPQ. The  $\Delta\text{OCP}$  *Synechocystis* mutant showed a lack of blue-light induced NPQ. In addition, 77K fluorescence spectra of quenched and unquenched WT obtained by excitation of Chl or PBS,

demonstrated that the OCP-related NPQ occurs at the level of the phycobilisome. This quenching is relaxed in darkness or dim blue light, and the initial phycobilisome fluorescence is recovered. Later, it was proved that the OCP is not only the quenching inducer but also it is the high light sensor. The OCP can be photoactivated under strong blue light illumination<sup>108</sup>, leading to the formation of the quenching inducer (activated OCP) which is able to bind and induce energy dissipation in the PBS. The final finding to complete the general schema of the OCP-related mechanism was the discovery of an additional protein<sup>109</sup>, the Fluorescence Recovery Protein (hereafter FRP), which interact with the active form of the OCP and accelerate dissipation of the quenching (measured as the recovery of the PBS fluorescence). A general schema of the OCP-related NPQ is presented in the figure 6.

Energy dissipation by OCP-related NPQ mechanism is important for tolerate high light conditions. The  $\Delta$ OCP *Synechocystis* mutant shows a higher sensitivity to high light illumination, evidenced by the decrease of oxygen evolution (PSII activity) faster than in the WT cells<sup>107</sup>. Supporting this result, later it was shown that *Synechococcus* PCC 7942 and *Thermosynechococcus elongatus*, strains which do not contain OCP, were more sensitive to photoinhibition than OCP-containing strains. The latest ones were able to recover faster the photochemical capacity of PSII (measured as the chlorophyll fluorescence under saturating light pulses). In iron depleted conditions, the OCP-related NPQ seems to be important. In this case, *Synechocystis* and *Arthospira* showed an increase in the functional uncoupling of PBS. Under high light, the uncoupled PBS, as well as the coupled PBS, are quenched by the OCP which is up-regulated in iron starved cells<sup>110,111</sup>. Since this work seeks to expand the knowledge about the OCP-related NPQ mechanism, a detailed explanation of its main aspects (OCP structure, OCP photoactivation, FRP interaction, PBS interaction and quenching) will be presented in the following sections.



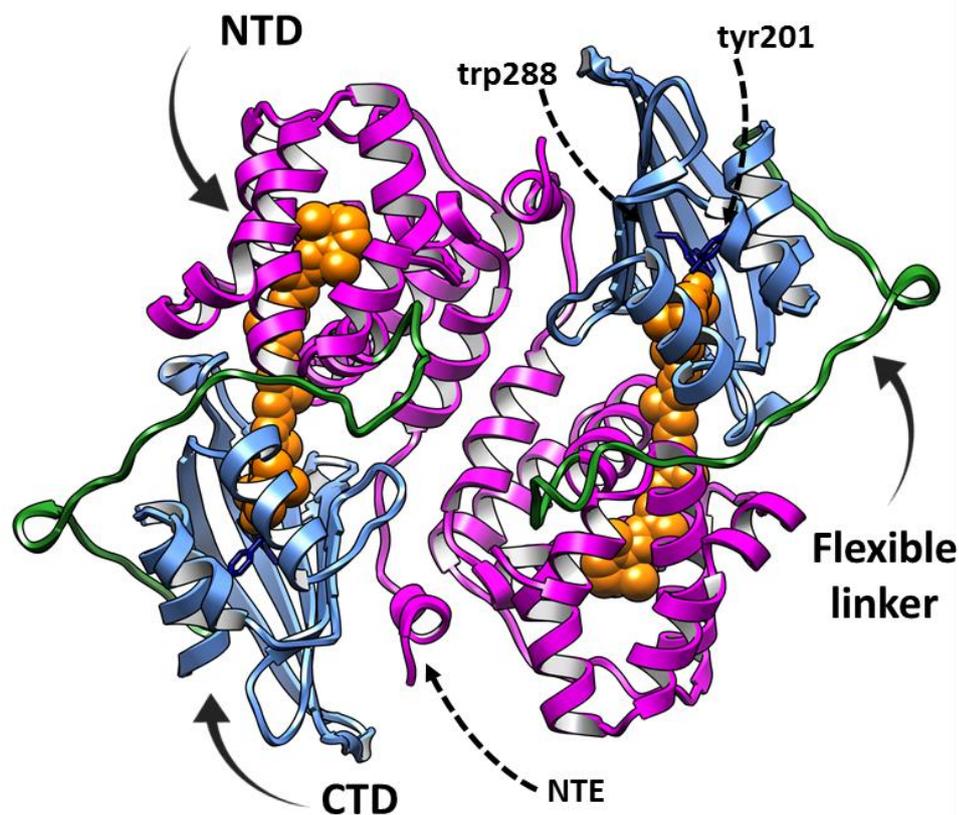
**Figure 6 | OCP-Related NPQ mechanisms.** In darkness and low irradiance, the OCP is in its orange inactive form and non-bound to the PBS. The phycobilisome (PBS) transfers most of the absorbed energy to the photosystems to do photosynthesis. (1) Strong light induces the photoactivation of the OCP<sup>o</sup> which is photoconverted to the active open OCP<sup>R</sup> via intermediary closed red forms OCP<sup>RI</sup>. (2) The OCP<sup>R</sup> binds to the PBS core and induces dissipation of excess absorbed light as heat, PBS fluorescence and EET are minimal. (3) The dimeric FRP helps OCP detachment from PBS under all light conditions. (4) Once detached from the PBS, the OCP<sup>R</sup> is converted to the inactive state OCP<sup>o</sup>. FRP accelerates this reaction. According to the current model, after OCP compaction the FRP can monomerize or another OCP can bind the OCP-FRP dimer complex and via a steric clash splits into two OCP-FRP monomer complexes. Finally, OCP-FRP complex dissociates releasing free OCP<sup>o</sup> to complete the cycle (5). Adapted from Muzzopappa and Kirilovsky 2019.

## *The Orange Carotenoid Protein: Structure*

---

First isolated from *Arthrospira Maxima* by Holt and Krogmann in 1981<sup>112</sup>, the Orange Carotenoid Protein (OCP) is a water-soluble protein that binds non-covalently one molecule of the carotenoid 3'-hydroxyechinenone. Sixteen years later, this protein was isolated from *Synechocystis* and identified as the product of the *Slr1963* gene<sup>113</sup>. In these purification procedures, an additional red fraction was obtained. It was first denominated as Red Carotenoid Protein (RCP) due to its spectral properties, but later it was identified as the N-terminal fragment of the OCP which is cleaved by proteases during the cell extraction. In this work we will refer to this isolated part of the OCP (the N terminal domain) as NTD. The first detailed structural analysis of the OCP<sup>114</sup> was published in 2003 by Kerfeld and co-workers. In that work, the structure of the OCP from *Arthrospira maxima* was solved at a resolution of 2.1 Å. Later, after the photoprotective role of the OCP was elucidated<sup>107</sup>, the structure of the OCP from *Synechocystis sp.* PCC 6803 was solved at a resolution of 1.65 Å. The latter structure is shown in the figure 7. It is important to note that all the OCP structures solved at the moment (*Synechocystis*, *Arthrospira* and *Anabaena*) are of the inactive form of the protein.

The OCP consist of two domains connected by a flexible linker. The N-terminal domain (hereafter OCP-NTD or NTD if it is isolated from the OCP) is formed by two four-helix bundles. The OCP-NTD sequence and protein fold are unique to cyanobacteria. On the other hand, the C-terminal domain (hereafter OCP-CTD or CTD if it is isolated) has an  $\alpha/\beta$  fold (a core of five  $\beta$ -sheets flanked by three  $\alpha$ -helices). It is similar, at the level of structure and sequence, to the NTF2 (nuclear transport factor 2) protein family. Interactions between the domains, such as Asn104-Trp277 and Arg155-Glu244, stabilize the inactive OCP conformation, as it is observed in the crystal structures. Another important interaction that stabilizes the OCP structure is between the N-terminal extension (NTE) and the  $\beta$ -sheet of the OCP-CTD, especially between the Pro 22 and the Phe 277.



**Figure 7 | The crystal structure of OCP from *Synechocystis*.** Homodimer crystal structure of the *Synechocystis* OCP<sup>0</sup>, the OCP-NTD is colored magenta, the OCP-CTD is colored sky-blue, the flexible linker is colored green and the carotenoid is shown in orange. Trp288 and Tyr201 side chains are shown in black. The N-terminal extension (NTE) is pointed with a dotted arrow.

Between the two domains, a hydrophobic cavity contains a keto-carotenoid (Figure 7). The keto terminus of the carotenoid is located at the inner pocket in the OCP-CTD, stabilized by H bonds between its 4-keto-oxygen atom and the side chain of the Trp 288 and Tyr 201<sup>115</sup> (numbering of residue positions in OCP from *Synechocystis*). The other end of the carotenoid is nestled in between the helix bundles of the OCP-NTD and it is stabilized by  $\pi$ - $\pi$  stacking interactions with Tyr 44 and Trp 110<sup>115</sup> (Figure 8B). The carotenoid is in all-*trans* configuration with 16° of planar deviation<sup>114</sup>. A remarkable number (9) of methionines are surrounding the carotenoid, probably tuning the carotenoid properties. OCPs purified from WT *Synechocystis* and *Arthrospira* bind 3'-hydroxyechinenone but it was shown that overexpression of OCP in *Synechocystis* induces echinenone binding and OCP isolated from *Synechocystis*  $\Delta$ CrtO (*knock-out* of the enzyme the  $\beta$ -carotene monoketolase which catalyze the synthesis of echinenone) binds zeaxanthin<sup>116</sup>.

This showed that although 3'-hydroxyechinenone seems to be the native carotenoid of the OCP, it can bind others. Supporting this, Bourcier de Carbon and collaborators were able to set a system of heterologous carotenoid proteins expression in *Escherichia coli* and they were able to selectively produce canthaxanthin binding OCPs<sup>117</sup>.

In all the structures of the OCP published at the moment, the asymmetric unit contains a homodimer. Further biochemical/biophysical data, as size exclusion chromatography<sup>114,118</sup> and mass spectrometry<sup>119,120</sup> had suggested a dimeric form of the OCP at high protein concentrations. Nevertheless, the OCP oligomeric state in the cell is still controversial and the role of dimerization on the activity of the protein is unknown.

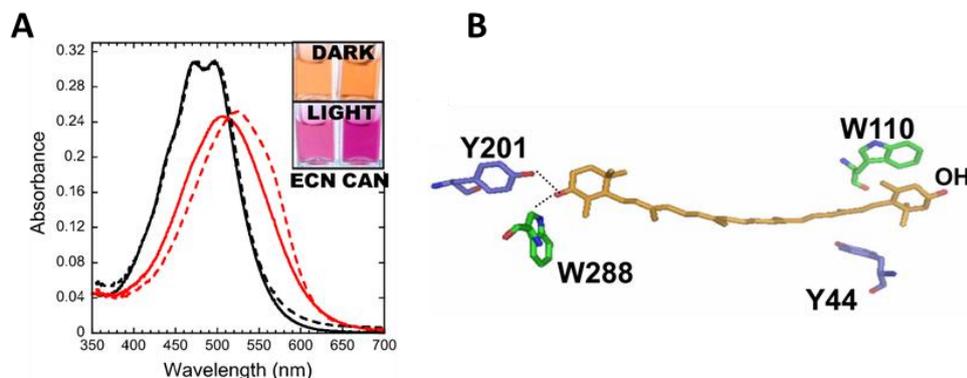
### ***The Orange Carotenoid Protein: photoactivation.***

---

In order to prevent the ROS production by dissipating the energy at the level of phycobilisome, the OCP needs to sense when the environmental conditions are stressful. This is achieved by directly sensing the light through OCP photoactivation under strong blue-green illumination.

In darkness, the absorption spectrum of the OCP (Figure 8A) present the typical shape of carotenoid absorption with two vibrational bands (maximum at 473 nm and 498 nm) well resolved and a shoulder at 450 nm. In solution, the protein present an orange coloration and this inactive form of the OCP will be referred as OCP<sup>O</sup>. After illumination with strong blue-green light, the OCP absorption decreases by a factor of approximated 0.85, the spectrum is red-shifted (maximum at 510 nm for ECN binding proteins and 525 nm for CAN binding proteins) and the vibrational resolution is lost<sup>108</sup>. This change is easily observed as a change in the color of the protein solution to red. Thus, the activated OCP is denominated as OCP<sup>R</sup>. The photoconversion takes place in intact cells and it is correlated with the quenching of the PBS energy<sup>108</sup>. Supporting this, while the OCP<sup>O</sup> cannot induce *in vitro* PBS fluorescence quenching, both OCP<sup>R</sup> and NTD (which have the same spectra) are able to induce it<sup>121,122</sup> (see “The Orange Carotenoid Protein activity”). *In vitro*, dark incubation of OCP<sup>R</sup> (after photoactivation) induces spontaneous back-conversion to the OCP<sup>O</sup>, demonstrating that the OCP<sup>R</sup> is not stable. This reaction is independent on light and strongly dependent on the temperature<sup>108</sup>. On the other hand, the kinetic of OCP<sup>R</sup> accumulation under constant light illumination strongly depends on the light intensity and it is temperature-dependent<sup>123,124</sup>, although to a lesser extent than the back-conversion. The presence of kosmotropes (for example phosphate) inhibit the photoconversion while the presence of chaotropes (for example thiocyanate)

activate the OCP even in the dark<sup>125</sup>. Additionally, photoconversion depends on the pH<sup>126</sup> and presence of Cu<sup>2+</sup><sup>127</sup>.



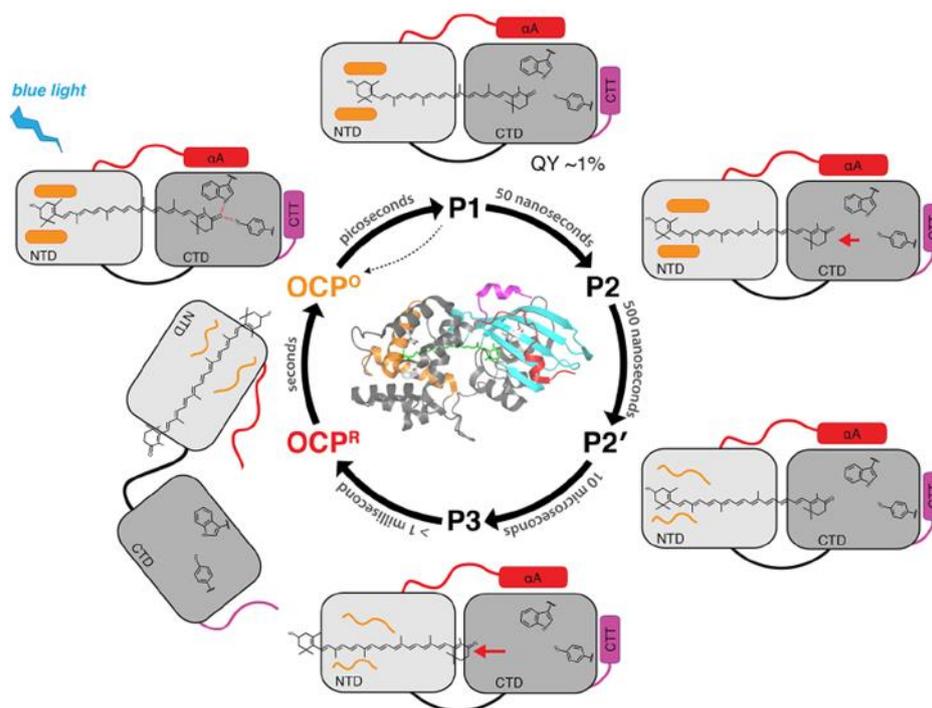
**Figure 8 | OCP photoactivation: steady-state.** (A) OCPO spectrum in black and OCPR spectrum in red, canthaxanthin binding OCP (dotted line) and echinenone binding OCP (solid line) are shown. The inset shows the change in the solution color after OCP photoconversion. Adapted from Bourcier De Carbon et al. 2015. 3'-hydroxyechinenone structure, showing the interaction of  $\beta$ 1-ring with Y201 and W288, and  $\beta$ 2-ring with W110 and Y44 (adapted from Wilson et al. 2008).

The spectral changes upon photoconversion are also associated to deep conformational modifications at the level of both protein and carotenoid. Raman spectroscopy analysis of the OCP<sup>O</sup> confirms the bend conformation of the carotenoid observed in the crystal structure. After illumination, the carotenoid remains as all-*trans*, but its apparent conjugation length increases by about one bond, indicating a more planar structure. Interestingly, various spectroscopically approaches (Raman<sup>128</sup>, steady-state and transient absorption<sup>129</sup>) had indicated heterogeneity in the OCP<sup>O</sup>, which was explained as two different positions in which the carotenoid can be accommodated in the OCP<sup>O</sup>. The major component, which gives rise to the spectral properties of the OCP<sup>O</sup>, and a minor component that is slightly red-shifted, both of them are able to photoconvert. During this photoconversion, the salt-bridge (Glu244-Arg155) that stabilized the OCP<sup>O</sup> conformation is broken allowing the domain separation<sup>130</sup>. After limited-proteolysis of the OCP<sup>R</sup>, isolated NTD containing the carotenoid and Apo-CTD were obtained, suggesting after domain separation the carotenoid remains in the NTD<sup>122</sup>. The crystal structure of canthaxanthin binding NTD<sup>131</sup> showed a different position of the carotenoid compared with OCP<sup>O</sup>, this suggested that upon photoactivation the H-bonds that hold the carotenoid in the OCP-CTD (Tyr201 and Trp288) breaks and the carotenoid is translocated 12 Å into the OCP-NTD. The domain separation leads to an increase in the radius of gyration, observed by SEC and SAXS<sup>132</sup>, and changes in the local structure such as unfolding of the N terminal

helix (NTE)<sup>120</sup> and C terminal helix (CTT). In addition, dimer-to-monomer transition during OCP<sup>O</sup> to OCP<sup>R</sup> were also reported<sup>119</sup>.

The protein-carotenoid interactions as well as the chemical nature of the carotenoid are important for the photoactivity. Zeaxanthin-binding OCP, which lacks the keto group and cannot form H-bonds, is photo-inactive<sup>116</sup>. In a similar way, substitution of the residues involved in the H-bonds, Tyr201 and Trp288, renders the OCP inactive<sup>126</sup>. Furthermore, if the Tyr44 and Trp110 are replaced by a non-aromatic amino acid, the OCP photoactivation is abolished<sup>115</sup> indicating that the  $\pi$ - $\pi$  stacking of the carotenoid  $\beta$ 2-ring is essential for this.

In all the results presented above the photoactivation of the OCP was measured as accumulation of the OCP<sup>R</sup>. This is not the same than the photoactivation process *per se*, in which photon absorption by the OCP leads to the formation of excited state and evolution to the OCP<sup>R</sup>. Despite the fact that OCP<sup>R</sup> accumulation is the direct consequence of the photoactivation, steady-state measurements hardly can shine some light on the fast processes taking place during photocycle. During the last years, a great effort was devoted to their elucidation. Here, the latest advances on the mechanism of photoactivation will be reviewed. The first transient-absorption experiment performed on OCP<sup>108</sup> suggested that after carotenoid S<sup>2</sup> is populated, it decays to a S<sup>1</sup>/ICT state which decay in 1 and 4.5 ps to the ground state. Only 1 % remains as the primary photoproduct, due to the low quantum-yield of the OCP photoconversion (calculated to 0.03%<sup>108,133</sup>). This is consistent with the assumption that the OCP is required only at high light intensity conditions. Therefore, significant OCP photoactivation only takes place at high light, when an excess of photons is available. Nevertheless, this red state is not the final OCP<sup>R</sup> indicating further changes in the millisecond to second timescale. This early intermediate was also observed by dynamic X-ray crystallography<sup>134</sup>, in which the carotenoid became slightly more planar and movement in the protein induce separation of Tyr201/Trp288-Carotenoid and partial domain separation. This intermediary red state, which probably correspond to the P<sub>3</sub> state described below, can relax to the OCP<sup>O</sup> in 300  $\mu$ s<sup>133</sup> if there is no evolution to the OCP<sup>R</sup> final state. The last phases of the photocycle (back conversion) were studied by a combination of techniques that showed that during back-conversion there is an orange intermediary state<sup>133</sup>, in which the carotenoid was already translocated (and probably the H-bonds were restored), that precedes the final OCP<sup>O</sup>, in which the carotenoid adopts the original bent position<sup>133</sup>.



**Figure 9 | OCP photoactivation: fast reactions.** OCP<sup>0</sup> is promoted to the excited state by blue light absorption. Only 1.5-1% can reach the P1 state, the rest relax to the OCP<sup>0</sup> ground-state. In P1, the H-bonds between the carotenoid and the Tyr 201 / Trp 288 are broken. After 50 ns, 40% of P1 evolve to P2 state, in which the carotenoid moves from the original position. In 500 ns, movements of the helices in the NTD lead to formation of P2' and allow the carotenoid to be completely translocated into the NTD in the P3 state in 10  $\mu$ s. In time ranges longer than 0.5 ms the P3 evolve to the final OCP<sup>R</sup>. Figure adapted from Konold et al. 2019.

An extensive spectroscopy analysis (combining UV-visible and IR transient absorption) in the time scale ranging from femtosecond to millisecond, in which I slightly contributed, had unveiled the first photoactivation events<sup>135</sup> (Figure 8C). It was shown that in addition to the S1/ICT states, another state termed as S\* is initially populated in 100 fs. It was proposed that in this state, the structure of the carotenoid is enough distorted to weak the H bonds. Then 1.5 % relaxes to the ground-state of the primary photoproduct (P1,  $\Delta$ absorption maximum at 563 nm) in picoseconds, in which the  $\beta$ 1-ring of the carotenoids remain all-*trans* and it became more planar inducing H-bond breaking. After 50 ns, 40 % of P1 (0.6 % of initial OCP) evolve into the second intermediary (P2,  $\Delta$ absorption maximum at 556 nm), in which probably weakening of  $\pi$ - $\pi$  interaction between the OCP-NTD and the  $\beta$ 2-ring of the carotenoid lead to evolution into a P2' state in 0.5-1  $\mu$ s. P2' is characterized by changes in the helical structure of NTD. These structural changes preceded the evolution to the P3 intermediary state in 10  $\mu$ s, which have a very similar spectrum to the OCP<sup>R</sup> suggesting that the carotenoid is translocated to the OCP-NTD. At room temperature at times longer than 0.5 ms, the NTE and CTE unfold allowing domain separation and the final OCP<sup>R</sup> is reached.

## *The Orange Carotenoid Protein activity*

---

The OCP have two key roles on photoprotection: dissipation of the PBS energy as it was mentioned before and quenching of the singlet oxygen.

As it was described before, *in vivo* measurements were the first hint to determine the quenching induced by the OCP at the phycobilisome levels<sup>107</sup>. The energy transfer within the photosynthetic apparatus can be studied using fluorescence spectroscopy. The OCP-induced quenching in the PBS was demonstrated by three experimental results<sup>107</sup>: **1-** Comparison of 77K the fluorescence spectra of quenched and unquenched cells, in which the decrease of fluorescence correspond with the PBS emission. **2-** The lack of PBS fluorescence quenching in the *Synechocystis* mutant lacking phycobilisomes. **3-** The presence of PBS fluorescence quenching in the *Synechocystis* mutant without active PSII. Furthermore, the room temperature fluorescence spectra of *Synechocystis* upon chlorophyll excitation and addition of DCMU also showed a slight quenching after high blue-light treatment<sup>136</sup>. This was interpreted as an uphill energy transfer from PSII to PBS that lead to a decrease of the fluorescence yield of Chl due to an OCP-induced quenching at the core of PBS (APC). This uphill EET is possible when the reaction centers of PSII are close (not capable of photochemistry), because the Q<sub>A</sub> site is blocked by reduction of the quinone in Q<sub>B</sub> or DCMU binding to Q<sub>B</sub>, and the lifetime of PSII fluorescence are comparable with the rate of EET to the PBS. Other experiments support this interpretation, like the absence of this Chl quenching in PSII-less cell in which the short lifetime of PSI does not allow uphill EET.

An important advance to study the PBS quenching by the OCP was the development of an *in vitro* system to reconstitute this process<sup>121</sup> through the incubation of isolated PBS and OCP in high phosphate concentrations. The high concentrations of phosphate which is required to keep the integrity of the PBS has other effects on OCP photoactivation and PBS-OCP<sup>R</sup> interaction. By using phosphate concentrations in the range 0.5 M to 1.4 M the OCP-induced quenching can be induced at different degrees *in vitro*. Since then, *in vitro* reconstitution of OCP-induced PBS quenching remained as a powerful tool to study the PBS-OCP interaction. Only OCP<sup>R</sup> is able to bind the PBS. This binding is independent of the light<sup>121</sup>. It is important to note that while 6-His C-terminal tagged OCP isolated from *Synechocystis*<sup>121</sup> or *E.coli* (carotenoid producing strain)<sup>117</sup> hardly can detach from PBS in darkness (0.5M or 0.8M phosphate) and almost no fluorescence recovery is observed, 6-His N-terminal tagged OCP isolated from *E. coli* can detach from PBS allowing fluorescence recovery (0.5M phosphate) in about 20 minutes. This suggested that the presence of His-tag in different parts of the protein can modulate the PBS binding. Through mutational analysis of OCP, Wilson and

coworkers demonstrated that the Arg 155 is involved in the interaction with the PBS<sup>130</sup>. This positive charged amino acid is located in the interdomain surface, forming an H-bond with Glu 244, and upon photoactivation is exposed to the surface which interact with the APC trimer. Although the PBS amino acid which interact with OCP R155 was not identified, the negative charged residues around the APC trimer bilins were proposed as binding site. *In vitro* reconstitution of OCP-PBS complexes were used for chemical crosslinking followed by protein digestion and liquid chromatography + mass spectrometry in tandem (LS-MS/MS) in order to determine the exact OCP binding site<sup>119,137</sup>. The first model generated from crosslinking experiments<sup>119</sup> suggested that the OCP-NTD buried between a basal APC660 trimer and a basal APC680 trimer in the PBS core. A second model (Figure 6) generated from a more extensive collection of crosslinked residues<sup>137</sup>, indicated that the OCP-NTD was located between the basal and external trimer containing ApcD and the trimer containing ApcF and ApcE. The OCP-CTD remains accessible to the solvent, in the exterior part of the PBS. In addition, crosslinking between OCP (or ApcF) and ApcC (linker protein) suggested a displacement of the ApcC subunit to a different position upon OCP binding. Kinetic analysis of ApcC-less PBS quenching by OCP indicated a weaker OCP-PBS interaction suggesting that ApcC stabilize this binding. In addition, it was shown that deletion of PB-loop of ApcE decreases the binding of OCP to PBS<sup>138</sup>, indicating that this component may stabilize the OCP-PBS interaction. The quenching amplitude decreases if the OCP interact with the isolated PBS core (hereafter CK)<sup>121</sup> indicating that the presence of rods stabilize the quenched complex.

A central question regarding the quenching mechanism is the site of primary quenching. The *Synechocystis* PBS are composed of 6 phycocyanin rods (Absorption<sub>MAX</sub> = 620 nm, Fluorescence<sub>MAX</sub> = 645 nm) which radiate from the PBS core. The core is comprised of  $\alpha/\beta$  APC subunits (Absorption<sub>MAX</sub> = 650 nm, Fluorescence<sub>MAX</sub> = 660 nm) and terminal emitters, ApcD, ApcF and ApcE (Absorption<sub>MAX</sub> = 670 nm, Fluorescence<sub>MAX</sub> = 680 nm). It was reported that the OCP-related NPQ takes place in *Synechocystis* mutants lacking phycocyanin<sup>107</sup> (it contains only PBS cores) suggesting that the binding and quenching sites were in the PBS core. Analysis of fluorescence spectra at low temperature suggested that probably the quenching site was the APC emitting at 660 nm<sup>139</sup> (hereafter APC660). This was confirmed by time resolved fluorescence experiments<sup>44</sup>. The evolution and decay of the *in vivo*<sup>44</sup> and *in vitro*<sup>140</sup> fluorescence spectrum after a short laser pulse, in the picosecond to nanosecond time range, allows to determine the energy transfer between the photosynthetic components and the rate of such reactions. Tian et al. showed that after 590 nm excitation (absorbed mainly by PBS), the energy equilibration in the rods (lifetime of 6 ps) and the EET to APC660 (77 ps) is not

affected by OCP-related NPQ. On the other hand, the lifetime of the energy transfer from APC660 to APC680 (terminal emitters) was drastically affected by OCP quenching (decrease from 63 ps to 16 ps) confirming APC660 as the quenching site. The molecular quenching rate was calculated as  $250 \text{ fs}^{-1}$  (*in vivo*). The quenching site was further confirmed by Jallet<sup>141</sup> and co-workers by showing that deletions of the terminal emitter ApcD and ApcF or mutation in ApcE, that avoid the covalent bilin binding, have the same OCP-related NPQ quenching than WT cells.

Since the basal cylinders have the same composition, there are two equivalent sites for OCP binding. In fact, single molecule measurements of PBS quenched by the OCP showed two populations of PBS-OCP complexes which were assigned to one or two OCP bound to PBS<sup>142</sup>. On the other hand, OCP quantification of Western Blots using PBS-OCP complexes showed that binding of only one OCP can quench almost all the PBS fluorescence<sup>121</sup>. The single molecule measurements<sup>142</sup> suggested that when excess of OCP are added to the PBS, a first complex with two OCPs per PBS (Q<sub>2</sub>) is more populated than simple 1 PBS : 1 OCP complex (Q<sub>1</sub>). Then, this initially more populated state Q<sub>2</sub> shifts to the state Q<sub>1</sub> over long periods of time. Thus, it is possible that the symmetry of the basal PBS core generate two equivalent binding sites, OCP binding to one of these sites induces structural distortions in the core ( as it was suggested by Harris et al.<sup>137</sup>) and lead to a decrease of the binding affinity of the second site. Since the *in vitro* OCP-PBS complex are prepared (included for single molecule) with an excess of OCP (40 OCP/PBS) PBS-OCP (1:2) complex are likely to be formed but the low affinity of one of the sites leads to unbinding over long time periods. This could explain the Q<sub>2</sub> to Q<sub>1</sub> shift and the 1:1 OCP-PBS ratio quantified from Western-blot. Indeed, another single molecule experiment<sup>143</sup> reported an intermediary state preceding the fully quenched PBS, in which PBS spectral properties suggested a decrease in the connection between rods and the core caused by a transient alteration of PBS structure upon OCP binding. Nevertheless, the number of active binding sites *in vivo* and the exact binding mechanism remain controversial.

Another important question regarding the PBS quenching mechanism is the actual physical mechanisms of quenching. Initially, two possible mechanisms were proposed<sup>107</sup>: conformational changes in PBS induced by OCP binding or direct energy transfer to a quencher specie ( in this case, the OCP). Despite this is still an open question, some advances have been made. On the basis of different spectroscopic characterization of the OCP carotenoid excited state<sup>129,144-146</sup> and the fluorescence lifetime of quenched/unquenched cells<sup>44,140,147</sup>, it was proposed that energy transfer from PBS to OCP was the main dissipative mechanism. The carotenoid in the OCP has different properties than in solution<sup>148</sup>: alterations of the

S1 energy, a short S1 lifetime (only in OCP<sup>O</sup>) and a strongly marked ICT state (compared with carotenoid in solution). Thus, it was proposed that the ICT state was involved in the energy quenching. Nevertheless, NTD-hECN has no ICT state<sup>149</sup> and it is capable of PBS quenching suggesting that ICT state is not directly involved in energy quenching. OCP<sup>O</sup> to OCP<sup>R</sup> transition decreases more the S1 energy<sup>146,150</sup> (from 14700 cm<sup>-1</sup> to 14000 cm<sup>-1</sup>), however both energies are low enough to quench APC660 implying that changes in the S1 energy are more likely to be side effect of photoactivation and the more important change is the exposure of the NTD which can bind the PBS. Due to the extremely fast quenching rate and the lack of a precise structural description of the OCP-PBS complex, the exact quenching mechanism remains unknown. EET<sup>145</sup> and Charge transfer<sup>44,140</sup> mechanisms seems to be the more feasible explanations<sup>44</sup>.

The second photoprotective activity of the OCP is the quenching of singlet oxygen. *In vitro*<sup>114,151</sup> and *in vivo*<sup>151</sup> measurements have shown that the OCP is able to decrease the <sup>1</sup>O<sub>2</sub> concentration by physical and chemical mechanisms. In *Synechocystis* cells, non-photoactivated OCP<sup>O</sup> can also deactivate <sup>1</sup>O<sub>2</sub> protecting PSII<sup>151</sup>. This mechanism could be important in oxidative stress conditions which are not induced by high light, like low CO<sub>2</sub>, in which the OCP is not photoactivated.

### ***The Fluorescence Recovery Protein***

---

The *ocp* gene is present in most of the phycobilisome-containing cyanobacteria. In a high percentage of the freshwater cyanobacteria, the *ocp* gene (*slr1963* in *Synechocystis*) is followed by another conserved gene (*slr1964* in *Synechocystis*). In *Synechocystis*, deletion of this gene does not affect the OCP-related quenching but significantly inhibits the fluorescence recovery after NPQ<sup>109</sup>. This suggested that this gene is not involved in the formation of the quencher, but it is essential to relax the system. Therefore, the protein encoded by this gene was named as Fluorescence Recovery Protein (hereafter FRP). The effect of the ΔFRP mutant can be complemented by overexpressing FRP<sup>109</sup>.

FRP co-immunoprecipitated with OCP *in vivo*, when cell are illuminated with high light<sup>109</sup>, suggesting an interaction between FRP and OCP<sup>R</sup>. *In vitro*<sup>109</sup>, FRP can accelerate the back conversion OCP<sup>R</sup> to OCP<sup>O</sup>. In this way, FRP can modulate the concentration of OCP<sup>R</sup> as well as the rate of NPQ relaxation. Despite the fact that *frp* gene is located next to *ocp* gene, they do not constitute an operon and they are transcribed independently<sup>109</sup>. This is important to balance the FRP:OCP ratio and modulate the rate and amplitude of the NPQ<sup>152</sup>.

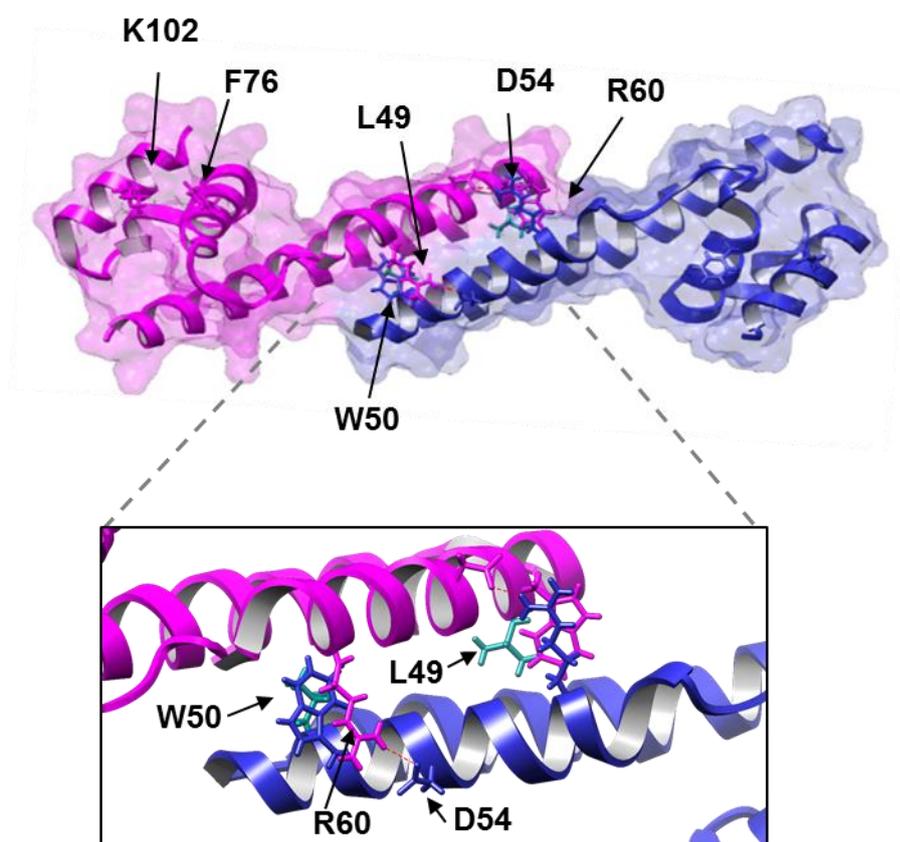
The structure of the FRP was solved at a resolution of 2.5 Å<sup>153</sup> ( Figure 9). Although the asymmetric unit is formed by a mix of one dimer and one tetramer,

only the dimeric form is biologically active<sup>153–155</sup>. This dimeric form is stabilized by hydrophobic interaction including the Leu 49-Leu 49 from each monomer<sup>156</sup> and also H-bonds including the Asp 54 - Arg 60<sup>153</sup>. The position of the Arg 60 is stabilized by  $\pi$ -cation interaction with the Trp 50 (Figure 9). Due to their high conservation, residues in the region 50 to 61 were proposed to be involved in the FRP activity. Indeed, mutations in Trp 50 and Asp 54 decrease FRP activity and mutations in Arg 60 and His 53 completely inhibit the FRP activity on the OCP<sup>R</sup> to OCP<sup>O</sup> acceleration<sup>153</sup>. However, the FRP binding interface with the OCP, is located at the head of the dimeric FRP<sup>157</sup> and the residues Phe 76 and Lys 107 are involved in this interaction. The main site of FRP binding on the OCP is in the  $\beta$ -sheets in the OCP-CTD which are covered by the NTE and exposed to the solvent upon photoactivation<sup>153,158–161</sup>. Supporting this, FRP can bind isolated CTD *in vitro*<sup>153</sup>, and amino acids located in the loops between the  $\beta$ -sheets of the OCP-CTD are important for the OCP back conversion acceleration by FRP, such as Asp 220 and Phe 299<sup>162</sup>. As a consequence of the occlusion of the FRP binding site in the OCP<sup>O</sup>, the binding affinity of FRP to OCP<sup>O</sup> is remarkably low<sup>155</sup> while FRP can easily bind OCP $\Delta$ NTE (in which the NTE was removed)<sup>161</sup> or the more open OCP-W288A<sup>155</sup>. Thermodynamic analysis of OCP back conversion at different temperatures demonstrated that the acceleration by addition of FRP does not change the activation energy of the reaction<sup>155</sup>. Thus, rather than catalyze the OCP<sup>R</sup> to OCP<sup>O</sup> transition, the FRP acts as a scaffold in which the entropy of the system decreases and the OCP compaction is faster<sup>155</sup>.

It was suggested that the FRP undergoes monomerization upon OCP binding<sup>155,157,159,161,163</sup>. However, the mechanisms of FRP binding and acceleration of OCP<sup>R</sup> to OCP<sup>O</sup> back-conversion is unclear. The FRP interaction with OCP is dynamic and transient, it only binds the OCP<sup>R</sup> to accelerate its back conversion and finally it detaches. This interferes with the analysis of the complex formed. Recently, a new work by Sluchanko and coworkers<sup>156</sup> revealed new intermediary complexes during the FRP mechanism, by using monomeric and permanent-dimeric FRP mutants and protein crosslinking. The model proposed on the basis of these results is detailed in the figure 6 (step 5). The FRP bind the OCP only as a dimer and after OCP compaction, it can monomerize. An alternative pathway is that another OCP can bind the second FRP subunit of the dimer and steric clash between the OCPs would lead to splitting of the complex into two OCP:FRP<sup>MONOMER</sup> complexes. Finally, the monomeric FRP is released from the OCP due to its low affinity.

In addition to the acceleration of the OCP back conversion, the FRP has another role in the OCP-related NPQ. It was shown that FRP can accelerate the fluorescence recovery of OCP-quenched PBS in darkness<sup>162</sup>. This was observed also when the

FRP-R60L was used. As mentioned before, the mutation in R60 inhibits the OCP<sup>R</sup> to OCP<sup>O</sup> acceleration. This indicated that faster recovery of the fluorescence by addition of FRP was not a side effect of a decrease of OCP<sup>R</sup> concentration, instead was the result of a second activity of the FRP: promoting the OCP detachment from PBS.

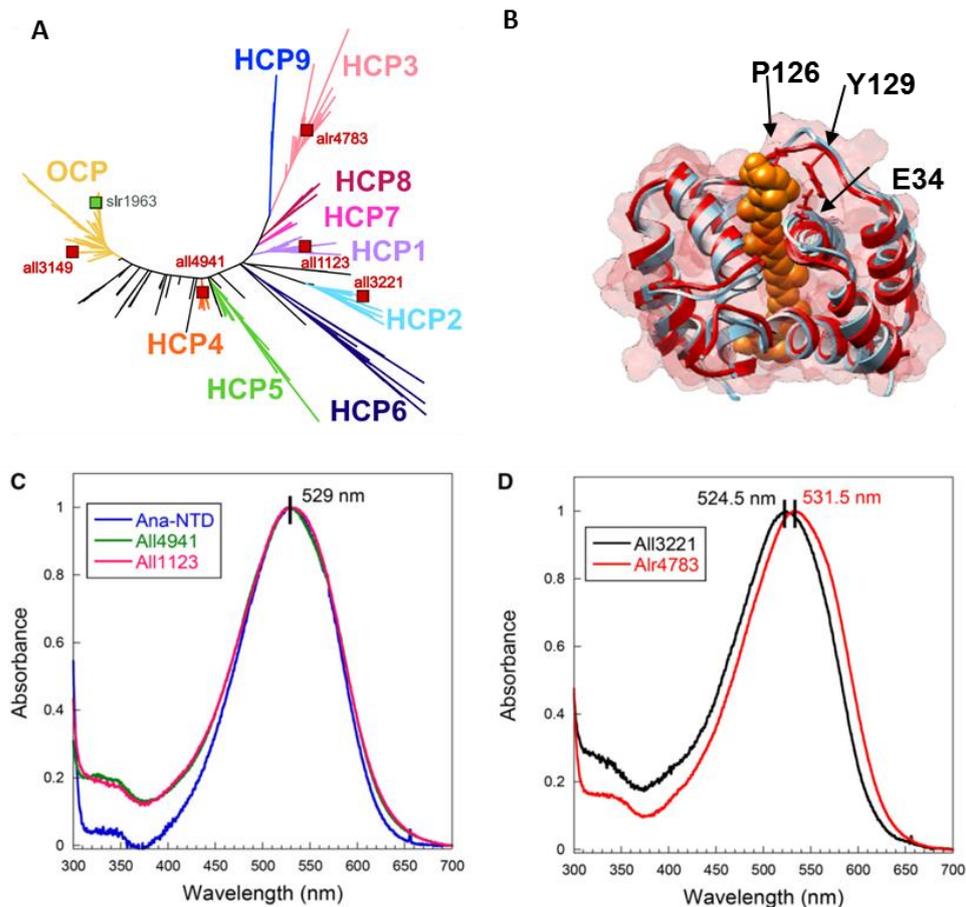


**Figure 9 | Structure of FRP from *Synechocystis*.** The FRP dimer structure (PDB: 4JDX). The amino acids involved in the dimeric interaction and in the FRP binding to OCP are highlighted.

### ***The paralogous of the OCP domains.***

In addition to *ocp* genes, many cyanobacteria also contain genes coding for homologs of the OCP domains<sup>114</sup>. The homologs of the OCP-CTD, are designated as CTDH. The role of the CTDH, as well as its structural features, remained unknown before this thesis work. The homologs of the OCP-NTD are designated as HCP (helical carotenoid proteins) or NTDH. Phylogenetic analysis of the HCP sequence lead to the conclusion that there are at least nine different subclades<sup>164</sup> (named from HCP1 to HCP9). The structure of the HCP1 from *Anabaena* sp. PCC 7120 (*all1123*) was solved<sup>164</sup> and found to be highly similar to the NTD structure (Figure 10). In addition, *Anabaena* contains three more HCPs (HCP2, HCP3 and

HCP4) which have been isolated and characterized. All the *Anabaena* HCPs bind canthaxanthin and show a similar spectrum than the NTD<sup>165</sup>. However, they have marked functional differences<sup>165</sup>. Interestingly, despite having high structural similarities with the NTD, only the HCP4 is able to induce fluorescence quenching of PBS *in vitro*. Only HCP2 and HCP3 are good singlet oxygen quenchers. On the other hand, HCP1 role remains unknown. These different HCP functions and the fact that different HCPs (and OCP) have diverse regulation patterns (at transcriptomic and proteomic level)<sup>164,165</sup> suggested that the HCP subclades have different roles.



**Figure 10 | The HCP protein family.** (A) Phylogenetic tree of HCP and NTD sequences, adapted from Melnicki et al 2016. (B) Comparison of NTD-OCP (PDB: 4XB4, shown in red) and HCP1 (PDB: 5FCX, shown in red), amino acids involved in carotenoid stabilization are marked. (C and D) Spectra of HoloHCP1 (all1123), HoloHCP2 (all3221), HoloHCP3 (alr4783) and HoloHCP4 (all4941), compared with HoloOCP-NTD (from *Anabaena*). Adapted from Lopéz-igual et al. 2016.

The OCP sequences found across all the cyanobacterial world also are quite diverse. Phylogenetic analysis have suggested the existence of at least three subclades, named as OCP1, OCP2 and OCPX. The OCP from *Synechocystis*, *Anabaena* and *Arthospira* belong the subclade OCP1, these canonical OCPs have

been used for all the experiments which elucidate the OCP mechanism. Recently the first OCP2 was described<sup>166</sup>, it was found that OCP2 has a faster accumulation of OCP<sup>R</sup> and a faster back-conversion to OCP<sup>O</sup> than OCP1. In addition, OCP2 cannot interact with FRP and it is monomeric in solution. Based on the biochemical differences between OCP1 and OCP2, it was proposed that OCP2 is more ancestral than OCP1.

Generally, the genomes containing CTDH/HCP also contains OCP, which may indicate that CTDH/HCPs are paralogs of the OCP-CTD/OCP-NTD. The term “paralogs” refers to a special case of homologs (genes or proteins with similar sequences) that have shared ancestry due to an event of gene duplication. Thus, it was proposed that the ancestors of these genes were duplicated and the modern CTDH/HCP and OCP-CTD/OCP-NTD are derived from them. In addition, more than one *hcp* gene is usually present and also *ocp* genes sometime are duplicated in the same genome (see chapter 5 “*Molecular evolution of the OCP: role of the flexible linker and inter-domain interaction*”). Nevertheless, we cannot exclude that this multiplicity of genes was originated by horizontal gene transfer. The presence of genes homologs of the OCP domains and the fact that usually the *ctdh* locus is adjacent to the *hcp4* locus prompted the researchers to suggest a modular origin for the OCP<sup>114,164,167</sup>. According to this hypothesis, after gene duplication the ancestors of the OCP domain underwent gene fusion giving rise to the first primitive OCP. Supporting this hypothesis, two recent works showed that isolated domains of the OCP (CTD and NTD) can interact between them and form a heterodimer complex with highly similar absorption spectrum and photoactivity than the OCP (hereafter OCP-like complex). Two different approaches were used to reconstitute this OCP-like complex, Lechno-Yossef and coworkers<sup>167</sup> overexpress both domains in the same *canthaxanthin-producing E. coli* cells while Moldenhauer and coworkers<sup>160</sup> used *in vitro mixing* of the isolated domains. In the latest approach, the authors found out that only if the carotenoid was incorporated at the CTD at the beginning of the reaction, the OCP-like complex was formed. On the basis of these results, it was claimed that the *ctdh-hcp4* may represent an intermediary step in the evolution of the OCP and they may be able to form a heterodimer with similar characteristic than the OCP.

## *Aim of the thesis*

---

The increasing availability of sequenced cyanobacterial genomes had brought our attention to the high diversity of sequences and combinations of paralogs of the OCP. This motivated our group to investigate the role of the paralogs of the OCP domains (HCP and CTDH) as well as the different OCP subclades. An extensive characterization of the HCPs was carried out in two different works in our laboratory<sup>164, 165</sup>. By contrast, the other paralogs, the CTDH, were not characterized and their biochemical properties and potential roles remained elusive. Was also thought that the study of the CTDH – HCP interaction could provide new insights into the evolutionary process of OCP formation. On the other hand, diversification of OCP into different subclades could have led to the establishment of a diversity of OCP traits that allows a better photoprotective response under different light regimes.

The major goal of this thesis work was to characterize the CTDH family and their interaction with HCPs in the context of a potential new photoprotective mechanism. The second part of this thesis was devoted to the study of the evolutionary origin of the OCP and the modern OCP subclades with a special focus on the role of the flexible linker. The overall structure of this thesis takes the form of six chapters, including this introductory chapter and the “Discussion and Perspective” section at the end. The chapters 2.1 to 2.4 present the experiment realized and the results obtained that are also described in 5 published articles. This thesis aims to address the following questions:

***Are the CTDH carotenoid proteins? Can they interact with HCPs and form a stable and photoactive OCP-like complex? What is their role?*** In the Chapter 2.1, a bioinformatic and biochemical characterization of the CTDHs is presented using CTDH from model organisms *Anabaena* and *Thermosynechococcus*. Then, the interaction between the CTDH and HCP is analyzed and compared with the interaction between CTD and NTD. This study allows the discovery of a carotenoid translocation mechanism between CTDHs and HCPs. (Results described in Muzzopappa et al, Plant Physiol, 2017)

***Is the carotenoid transfer mechanism unidirectional? Are there other interactions between paralogs of OCP-domains that lead to a carotenoid transfer?*** In chapter 2.2, we (in collaboration with Sluchanko and Maksimov groups) have deeply studied the interaction between different paralogs of the OCP-domain and the carotenoid translocation mechanism associated. (Results described in Slonimskiy et al, FEBS journal, 2019).

*What are the structural features of the CTDH? What is the difference with the OCP-CTD? What are the molecular determinants involved in the carotenoid transfer mechanism?* In the chapter 2.3, I present two structures of apo-CTDH from *Anabaena* (one forming dimers and the other forming big oligomers) solved in collaboration with Noam Adir's group. Based on structural and mutational analysis, the CTT was proposed to be involved in the carotenoid translocation mechanism. (Results described in Harris et al, Biol Comm, 2018 and Harris et al 2019 submitted)

*If the OCP-domains can interact and form an OCP-like complex, what is the role of the linker in the OCP? Why a single protein is better than two interacting proteins? Is this role conserved among all the OCP subclades? What are the main differences between the different OCPs?* In the chapter 2.4, we characterize the first OCPX and compared it with the OCP1 and OCP2. Then we studied the role of the linker in these three OCP subclades. The first approach used was to reconstitute the OCP-like complexes in which the linker is absent. The second approach was to selectively mutate the more conserved amino acids in the linker. This analysis together with an extensive phylogenetic analysis allowed as to propose a new evolutionary model for the OCP. (Results described in Muzzopappa et al, 2019 Nature Plants).

### *List of publications*

---

**Muzzopappa F**, Wilson A, Yogarajah V, Cot S, Perreau F, Montigny C, Bourcier de Carbon C, Kirilovsky D. Paralogs of the C-Terminal Domain of the Cyanobacterial Orange Carotenoid Protein Are Carotenoid Donors to Helical Carotenoid Proteins. *Plant Physiol.* 2017 Nov;175(3):1283-1303.

Harris D, Wilson A, **Muzzopappa F**, Sluchanko NN, Friedrich T, Maksimov EG, Kirilovsky D, Adir N. Structural rearrangements in the C-terminal domain homolog of Orange Carotenoid Protein are crucial for carotenoid transfer. *Commun Biol.* 2018 Aug 27;1:125.

Konold PE, van Stokkum IHM, **Muzzopappa F**, Wilson A, Groot ML, Kirilovsky D, Kennis JTM. photoactivation Mechanism, Timing of Protein Secondary Structure Dynamics and Carotenoid Translocation in the Orange Carotenoid Protein. *J Am Chem Soc.* 2019 Jan 9;141(1):520-530.

Slonimskiy YB, **Muzzopappa F**, Maksimov EG, Wilson A, Friedrich T, Kirilovsky D, Sluchanko NN. Light-controlled carotenoid transfer between water-soluble proteins related to cyanobacterial photoprotection. *FEBS J.* 2019 May;286(10):1908-1924.

Calzadilla PI, **Muzzopappa F**, Sétif P, Kirilovsky D. Different roles for ApcD and ApcF in *Synechococcus elongatus* and *Synechocystis* sp. PCC 6803 phycobilisomes. *Biochim Biophys Acta Bioenerg.* 2019 Jun 1;1860(6):488-498.

**Muzzopappa F**, Wilson A and Kirilovsky D. Interdomain interactions reveal the molecular evolution of the orange carotenoid protein. *Nature Plants.* 2019 (in press)

Harris D, **Muzzopappa F**, Glaser F, Wilson A, Kirilovsky D and Adir N. Carotenoid uptake, binding and delivery by water-soluble proteins in cyanobacteria. *PNAS* (submitted).

**Muzzopappa F** and Kirilovsky D. Changing color for photoprotection: The Orange Carotenoid Protein. *Trends in Plant Science* (in revision). **(Review)**

**Muzzopappa F** and Kirilovsky D. Orange Carotenoid Protein and the regulation of energy transfer in cyanobacteria. *Encyclopedia of Biological Chemistry* (Third Edition) (in preparation). **(Review)**

## Bibliography

---

1. Schirrmeister, B. E., Gugger, M. & Donoghue, P. C. J. Cyanobacteria and the Great Oxidation Event: Evidence from genes and fossils. *Palaeontology* **58**, 769–785 (2015).
2. Bekker, A. *et al.* Dating the Rise of Oxygen. *Nature* **427**, 117–120 (2004).
3. Van Kranendonk, M. J. *et al.* A Chronostratigraphic Division of the Precambrian. *Geol. Time Scale* 299–392 (2012). doi:10.1016/B978-0-444-59425-9.00016-0
4. Charpy, L., Casareto, B. E., Langlade, M. J. & Suzuki, Y. Cyanobacteria in Coral Reef Ecosystems: A Review. *J. Mar. Biol.* (2012). doi:10.1155/2012/259571
5. Isichei, A. O. The role of algae and cyanobacteria in arid lands. A review. *Arid Soil Res. Rehabil.* (1990). doi:10.1080/15324989009381227
6. Boone, D. R., Castenholz, R. W. & Garrity, G. M. *Bergey's manual of systematic bacteriology*. (Springer, 2001).
7. Lounatmaa, K., Vaara, T., Österlund, K. & Vaara, M. Ultrastructure of the cell wall of a *Synechocystis* strain. *Can. J. Microbiol.* **26**, 204–208 (1980).
8. Stanier, R. Y., Bazine, G. C. & R. Y. StaIJer and G. Cohen-Bazire. PHOTOTROPHIC PROKARYOTES: THE CYANOBACTERIA. *Ann. Rev. Microbio!* **31**, 50 (1977).
9. Rippka, R., Waterbury, J. & Cohen-Bazire, G. A cyanobacterium which lacks thylakoids. *Arch. Microbiol.* **100**, 419–436 (1974).
10. Meeks, J. C., Campbell, E. L., Summers, M. L. & Wong, F. C. Cellular differentiation in the cyanobacterium *Nostoc punctiforme*. *Arch. Microbiol.* **178**, 395–403 (2002).
11. Stanier, R. Y., Deruelles, J., Rippka, R., Herdman, M. & Waterbury, J. B. Generic Assignments, Strain Histories and Properties of Pure Cultures of Cyanobacteria. *Microbiology* (1979). doi:10.1099/00221287-111-1-1
12. Paerl, H. W. & Otten, T. G. Harmful Cyanobacterial Blooms: Causes, Consequences, and Controls. *Microb. Ecol.* **65**, 995–1010 (2013).
13. Mereschkowsky, C. Über Natur und Ursprung der Chromatophoren im Pflanzenreiche. *Biol. Zent. Bl.* **25**, 593–604 (1905).
14. Emerson, R. & Arnold, W. a Separation of the Reactions in Photosynthesis By Means of Intermittent Light. *J. Gen. Physiol.* **15**, 391–420 (1932).
15. Wirtz, W., Stitt, M. & Heldt, H. W. LIGHT ACTIVATION OF CALVIN CYCLE ENZYMES AS MEASURED IN PEA LEAVES. *FEBS Lett.* **142**, 223–226 (1982).
16. Blankenship, R. E. *Molecular Mechanisms of Photosynthesis*. *Molecular Mechanisms of Photosynthesis* (2008). doi:10.1002/9780470758472
17. Emerson, R. and Arnold, W. the Photochemical Reaction in Photosynthesis. *J. Gen. Physiol.* **16**, 191–205 (2004).
18. Gaffron, H. & Wohl, K. Zur Theorie der Assimilation. *Naturwissenschaften* **24**, 103–107 (1936).
19. Mullineaux, C. W. Co-existence of photosynthetic and respiratory activities in cyanobacterial thylakoid membranes. *Biochim. Biophys. Acta - Bioenerg.* **1837**, 503–511 (2014).
20. Chew, A. G. M. & Bryant, D. A. Chlorophyll Biosynthesis in Bacteria: The Origins of Structural and Functional Diversity. *Annu. Rev. Microbiol.* **61**, 113–129 (2007).
21. Scheer, H. An Overview of Chlorophylls and Bacteriochlorophylls: Biochemistry, Biophysics, Functions and Applications BT - Chlorophylls and Bacteriochlorophylls: Biochemistry, Biophysics, Functions and Applications. in (eds. Grimm, B., Porra, R. J., Rüdiger, W. & Scheer, H.) 1–26 (Springer Netherlands, 2006). doi:10.1007/1-4020-4516-6\_1
22. Tanaka, R. & Tanaka, A. Tetrapyrrole Biosynthesis in Higher Plants. *Annu. Rev. Plant Biol.* **58**, 321–346 (2007).
23. Burger-Wiersma, T., Veenhuis, M., Korthals, H. J., Van de Wiel, C. C. M. & Mur, L. R. A new prokaryote containing chlorophylls a and b. *Nature* **320**, 262–264 (1986).
24. Dougherty, R. C., Strain, H. H., Svec, W. A., Uphaus, R. A. & Katz, J. J. Structure, properties, and distribution of chlorophyll c. *J. Am. Chem. Soc.* **92**, 2826–2833 (1970).
25. Renger, T. & Schlodder, E. The Primary Electron Donor of Photosystem II of the Cyanobacterium *Acaryochloris marina* Is a Chlorophyll d and the Water Oxidation Is Driven by a Chlorophyll a/Chlorophyll d Heterodimer. *J. Phys. Chem. B* **112**, 7351–7354 (2008).
26. Chen, M. *et al.* A red-shifted chlorophyll. *Science* **329**, 1318–9 (2010).
27. Nürnberg, D. J. *et al.* Photochemistry beyond the red limit in chlorophyll f-containing

- photosystems. *Science (80-. )*. **360**, 1210 LP-1213 (2018).
28. Vershinin, A. Biological functions of carotenoids - diversity and evolution. *BioFactors* **10**, 99–104 (1999).
  29. Hashimoto, H., Uragami, C., Yukihiro, N., Gardiner, A. T. & Cogdell, R. J. Understanding/unravelling carotenoid excited singlet states. *J. R. Soc. Interface* **15**, (2018).
  30. Polivka, T., Herek, J. L., Zigmantas, D., Akerlund, H.-E. & Sundstrom, V. Direct observation of the (forbidden) S1 state in carotenoids. *Proc. Natl. Acad. Sci.* **96**, 4914–4917 (1999).
  31. Lustres, J. L. P., Dobryakov, A. L., Holzwarth, A. & Veiga, M. S2→S1 internal conversion in  $\beta$ -carotene: Strong vibronic coupling from amplitude oscillations of transient absorption bands. *Angew. Chemie - Int. Ed.* **46**, 3758–3761 (2007).
  32. Polivka, T. & Sundstrom, V. Ultrafast dynamics of carotenoid excited states - From solution to natural and artificial systems. *Chem Rev* **104**, 2021–2071 (2004).
  33. Bautista, J. A. *et al.* Excited State Properties of Peridinin: Observation of a Solvent Dependence of the Lowest Excited Singlet State Lifetime and Spectral Behavior Unique among Carotenoids. *J. Phys. Chem. B* **103**, 8751–8758 (1999).
  34. Enriquez, M. M. *et al.* The Intramolecular Charge Transfer State in Carbonyl-Containing Polyenes and Carotenoids. *J. Phys. Chem. B* **114**, 12416–12426 (2010).
  35. Vaswani, H. M., Hsu, C.-P., Head-Gordon, M. & Fleming, G. R. Quantum Chemical Evidence for an Intramolecular Charge-Transfer State in the Carotenoid Peridinin of Peridinin–Chlorophyll–Protein. *J. Phys. Chem. B* **107**, 7940–7946 (2003).
  36. Croce, R., Müller, M. G., Bassi, R. & Holzwarth, A. R. Carotenoid-to-Chlorophyll Energy Transfer in Recombinant Major Light-Harvesting Complex (LHCII) of Higher Plants. I. Femtosecond Transient Absorption Measurements. *Biophys. J.* **80**, 901–915 (2001).
  37. Frank L. de Weerd, \*, John T. M. Kennis, Jan P. Dekker, and Grondelle, R. van.  $\beta$ -Carotene to Chlorophyll Singlet Energy Transfer in the Photosystem I Core of *Synechococcus elongatus* proceeds via the  $\beta$ -Carotene S2 and S1 States. (2003). doi:10.1021/JP027758K
  38. Foote, C. S. & Denny, R. W. Chemistry of singlet oxygen. VII. Quenching by  $\beta$ -carotene. *J. Am. Chem. Soc.* **90**, 6233–6235 (1968).
  39. Liu, Z. *et al.* Crystal structure of spinach major light-harvesting complex at 2.72 Å resolution. *Nature* **428**, 287–292 (2004).
  40. Haworth, P., Watson, J. L. & Arntzen, C. J. The detection, isolation and characterization of a light-harvesting complex which is specifically associated with Photosystem I. *Biochim. Biophys. Acta - Bioenerg.* **724**, 151–158 (1983).
  41. Bassi, R., Sandona, D. & Croce, R. Novel aspects of chlorophyll a/b-binding proteins. *Physiol. Plant.* **100**, 769–779 (1997).
  42. Boekema, E. J., van Roon, H., van Breemen, J. F. L. & Dekker, J. P. Supramolecular organization of photosystem II and its light-harvesting antenna in partially solubilized photosystem II membranes. *Eur. J. Biochem.* **266**, 444–452 (1999).
  43. Mullineaux, C. W. Phycobilisome-reaction centre interaction in cyanobacteria. *Photosynth. Res.* **95**, 175–182 (2008).
  44. Tian, L. *et al.* Site, rate, and mechanism of photoprotective quenching in cyanobacteria. *J. Am. Chem. Soc.* **133**, 18304–18311 (2011).
  45. Rakhimberdieva, M. G., Boichenko, V. A., Karapetyan, N. V & Stadnichuk, I. N. Interaction of phycobilisomes with photosystem II dimers and photosystem I monomers and trimers in the cyanobacterium *Spirulina platensis*. *Biochemistry* **40**, 15780–8 (2001).
  46. Dong, C. *et al.* ApcD is necessary for efficient energy transfer from phycobilisomes to photosystem I and helps to prevent photoinhibition in the cyanobacterium *Synechococcus* sp. PCC 7002. *Biochim. Biophys. Acta - Bioenerg.* **1787**, 1122–1128 (2009).
  47. Ashby, M. K. & Mullineaux, C. W. The role of ApcD and ApcF in energy transfer from phycobilisomes to PS I and PS II in a cyanobacterium. *Photosynth. Res.* **61**, 169–179 (1999).
  48. Calzadilla, P. I., Muzzopappa, F., Sétif, P. & Kirilovsky, D. Different roles for ApcD and ApcF in *Synechococcus elongatus* and *Synechocystis* sp. PCC 6803 phycobilisomes. *Biochim. Biophys. Acta - Bioenerg.* (2019). doi:https://doi.org/10.1016/j.bbabi.2019.04.004
  49. Wei, X. *et al.* Structure of spinach photosystem II–LHCII supercomplex at 3.2 Å resolution. *Nature* **534**, 69–74 (2016).
  50. Umena, Y., Kawakami, K., Shen, J.-R. & Kamiya, N. Crystal structure of oxygen-evolving photosystem II at a resolution of 1.9 Å. *Nature* **473**, 55–60 (2011).

51. Ferreira, K. N., Iverson, T. M., Maghlaoui, K., Barber, J. & Iwata, S. Architecture of the Photosynthetic Oxygen-Evolving Centre. *Science (80-. )*. **303**, 1831–1839 (2004).
52. Kamiya, N. & Shen, J.-R. Crystal structure of oxygen-evolving photosystem II from *Thermosynechococcus vulcanus* at 3.7-Å resolution. *Proc. Natl. Acad. Sci.* **100**, 98 LP-103 (2003).
53. Cardona, T., Sedoud, A., Cox, N. & Rutherford, A. W. Charge separation in Photosystem II: A comparative and evolutionary overview. *Biochim. Biophys. Acta - Bioenerg.* **1817**, 26–43 (2012).
54. Shen, J.-R. The Structure of Photosystem II and the Mechanism of Water Oxidation in Photosynthesis. *Annu. Rev. Plant Biol.* **66**, 23–48 (2015).
55. Morais, F., Barber, J. & Nixon, P. J. The chloroplast-encoded alpha subunit of cytochrome b-559 is required for assembly of the photosystem two complex in both the light and the dark in *Chlamydomonas reinhardtii*. *J. Biol. Chem.* **273**, 29315–29320 (1998).
56. Chapman, D. J., Gounaris, K. & Barber, J. Electron-transport properties of the isolated D1-D2-cytochrome b-559 Photosystem II reaction centre. *Biochim. Biophys. Acta - Bioenerg.* **933**, 423–431 (1988).
57. Shinopoulos, K. E. & Brudvig, G. W. Cytochrome b559 and cyclic electron transfer within photosystem II. *Biochim. Biophys. Acta - Bioenerg.* **1817**, 66–75 (2012).
58. Kurisu, G., Zhang, H., Smith, J. L. & Cramer, W. A. Structure of the Cytochrome b6f Complex of Oxygenic Photosynthesis: Tuning the Cavity. *Science (80-. )*. **302**, 1009 LP-1014 (2003).
59. Mitchell, P. The protonmotive Q cycle: a general formulation. *FEBS Lett.* **59**, 137–9 (1975).
60. Hladík, J. & Sofrová, D. Does the trimeric form of the Photosystem I reaction center of cyanobacteria in vivo exist? *Photosynth. Res.* **29**, 171–175 (1991).
61. Jordan, P. *et al.* Three-dimensional structure of cyanobacterial photosystem I at 2.5 Å resolution. *Nature* **411**, 909–917 (2001).
62. Kruip, J., Bald, D., Boekema, E. & Rögner, M. Evidence for the existence of trimeric and monomeric Photosystem I complexes in thylakoid membranes from cyanobacteria. *Photosynth. Res.* **40**, 279–286 (1994).
63. Ben-Shem, A., Frolov, F. & Nelson, N. Evolution of photosystem I – from symmetry through pseudosymmetry to asymmetry. *FEBS Lett.* **564**, 274–280 (2004).
64. Amunts, A., Drory, O. & Nelson, N. The structure of a plant photosystem I supercomplex at 3.4 Å resolution. *Nature* **447**, 58–63 (2007).
65. Golbeck, J. H. Structure and Function of Photosystem I. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **43**, 293–324 (1992).
66. ARNON, D. I., ALLEN, M. B. & WHATLEY, F. R. Photosynthesis by Isolated Chloroplasts. *Nature* **174**, 394–396 (1954).
67. Battchikova, N., Eisenhut, M. & Aro, E.-M. Cyanobacterial NDH-1 complexes: Novel insights and remaining puzzles. *Biochim. Biophys. Acta - Bioenerg.* **1807**, 935–944 (2011).
68. Schuller, J. M. *et al.* Structural adaptations of photosynthetic complex I enable ferredoxin-dependent electron transfer. *Science (80-. )*. **363**, 257 LP-260 (2019).
69. He, Z. *et al.* NDH-1L interacts with ferredoxin via the subunit NdhS in *Thermosynechococcus elongatus*. *Photosynth. Res.* **126**, 341–349 (2015).
70. Gao, F. *et al.* The NDH-1L-PSI Supercomplex Is Important for Efficient Cyclic Electron Transport in Cyanobacteria. *Plant Physiol.* **172**, 1451–1464 (2016).
71. Yeremenko, N. *et al.* Open Reading Frame *ssr2016* is Required for Antimycin A-sensitive Photosystem I-driven Cyclic Electron Flow in the Cyanobacterium *Synechocystis* sp. PCC 6803. *Plant Cell Physiol.* **46**, 1433–1436 (2005).
72. Munekage, Y. *et al.* PGR5 Is Involved in Cyclic Electron Flow around Photosystem I and Is Essential for Photoprotection in Arabidopsis. *Cell* **110**, 361–371 (2002).
73. Edge, R. & Truscott, T. G. Carotenoid Radicals and the Interaction of Carotenoids with Active Oxygen Species BT - The Photochemistry of Carotenoids. in (eds. Frank, H. A., Young, A. J., Britton, G. & Cogdell, R. J.) 223–234 (Springer Netherlands, 1999). doi:10.1007/0-306-48209-6\_12
74. Zolla, L. & Rinalducci, S. Involvement of Active Oxygen Species in Degradation of Light-Harvesting Proteins under Light Stresses. *Biochemistry* **41**, 14391–14402 (2002).
75. Krieger-Liszkay, A. Singlet oxygen production in photosynthesis. *J. Exp. Bot.* **56**, 337–346 (2004).
76. Keren, N., Gong, H. & Ohad, I. Oscillations of reaction center II-D1 protein degradation in vivo induced by repetitive light flashes. Correlation between the level of RCII-QB- and

- protein degradation in low light. *J. Biol. Chem.* **270**, 806–814 (1995).
77. Kós, P. B., Deák, Z., Cheregi, O. & Vass, I. Differential regulation of psbA and psbD gene expression, and the role of the different D1 protein copies in the cyanobacterium *Thermosynechococcus elongatus* BP-1. *Biochim. Biophys. Acta - Bioenerg.* **1777**, 74–83 (2008).
  78. Dashdorj, N. *et al.* The Single Chlorophyll a Molecule in the Cytochrome b6f Complex: Unusual Optical Properties Protect the Complex against Singlet Oxygen. *Biophys. J.* **88**, 4178–4187 (2005).
  79. Mehler, A. H. Studies on reactions of illuminated chloroplasts. *Arch. Biochem. Biophys.* **33**, 65–77 (1951).
  80. Rutherford, A. W., Osyczka, A. & Rappaport, F. Back-reactions, short-circuits, leaks and other energy wasteful reactions in biological electron transfer: Redox tuning to survive life in O<sub>2</sub>. *FEBS Lett.* **586**, 603–616 (2012).
  81. Asada, K. Production and scavenging of reactive oxygen species in chloroplasts and their functions. *Plant Physiol.* **141**, 391–6 (2006).
  82. Shimakawa, G. & Miyake, C. Oxidation of P700 Ensures Robust Photosynthesis. *Front. Plant Sci.* **9**, 1617 (2018).
  83. Ke, W.-T., Dai, G.-Z., Jiang, H.-B., Zhang, R. & Qiu, B.-S. Essential roles of iron superoxide dismutase in photoautotrophic growth of *Synechocystis* sp. PCC 6803 and heterogeneous expression of marine *Synechococcus* sp. CC9311 copper/zinc superoxide dismutase within its sodB knockdown mutant. *Microbiology* **160**, 228–241 (2014).
  84. Latifi, A., Ruiz, M. & Zhang, C.-C. Oxidative stress in cyanobacteria. *FEMS Microbiol. Rev.* **33**, 258–278 (2009).
  85. Krieger-Liszka, A. & Trebst, A. Tocopherol is the scavenger of singlet oxygen produced by the triplet states of chlorophyll in the PSII reaction centre. *J. Exp. Bot.* **57**, 1677–1684 (2006).
  86. McDonald, A. E. *et al.* Flexibility in photosynthetic electron transport: The physiological role of plastoquinol terminal oxidase (PTOX). *Biochim. Biophys. Acta - Bioenerg.* **1807**, 954–967 (2011).
  87. Bailey, S. *et al.* Alternative photosynthetic electron flow to oxygen in marine *Synechococcus*. *Biochim. Biophys. Acta - Bioenerg.* **1777**, 269–276 (2008).
  88. Allahverdiyeva, Y., Isojärvi, J., Zhang, P. & Aro, E.-M. Cyanobacterial Oxygenic Photosynthesis is Protected by Flavodiiron Proteins. *Life* **5**, (2015).
  89. Allahverdiyeva, Y. *et al.* Flavodiiron proteins Flv1 and Flv3 enable cyanobacterial growth and photosynthesis under fluctuating light. *Proc. Natl. Acad. Sci.* **110**, 4111 LP-4116 (2013).
  90. Zhang, P. *et al.* Operon flv4-flv2 Provides Cyanobacterial Photosystem II with Flexibility of Electron Transfer. *Plant Cell* **24**, 1952 LP-1971 (2012).
  91. Mullineaux, C. W. & Allen, J. F. State 1-State 2 transitions in the cyanobacterium *Synechococcus* 6301 are controlled by the redox state of electron carriers between Photosystems I and II. *Photosynth. Res.* **23**, 297–311 (1990).
  92. Wollman, F.-A. & Lemaire, C. Studies on kinase-controlled state transitions in Photosystem II and b6f mutants from *Chlamydomonas reinhardtii* which lack quinone-binding proteins. *Biochim. Biophys. Acta - Bioenerg.* **933**, 85–94 (1988).
  93. Minagawa, J. State transitions—The molecular remodeling of photosynthetic supercomplexes that controls energy flow in the chloroplast. *Biochim. Biophys. Acta - Bioenerg.* **1807**, 897–905 (2011).
  94. Calzadilla, P. I. *et al.* The Cytochrome b6f Complex Is Not Involved in Cyanobacterial State Transitions. *Plant Cell* **31**, 911–931 (2019).
  95. Kaňa, R. Mobility of photosynthetic proteins. *Photosynth. Res.* **116**, 465–479 (2013).
  96. Mullineaux, C. W., Tobin, M. J. & Jones, G. R. Mobility of photosynthetic complexes in thylakoid membranes. *Nature* **390**, 421–424 (1997).
  97. Joshua, S. & Mullineaux, C. W. Phycobilisome diffusion is required for light-state transitions in cyanobacteria. *Plant Physiol.* **135**, 2112–2119 (2004).
  98. Bruce, D., Brimble, S. & Bryant, D. A. State transitions in a phycobilisome-less mutant of the cyanobacterium *Synechococcus* sp. PCC 7002. *BBA - Bioenerg.* **974**, 66–73 (1989).
  99. El Bissati, K., Delphin, E., Murata, N., Etienne, A. L. & Kirilovsky, D. Photosystem II fluorescence quenching in the cyanobacterium *Synechocystis* PCC 6803: Involvement of two different mechanisms. *Biochim. Biophys. Acta - Bioenerg.* **1457**, 229–242 (2000).
  100. Olive, J., M'Bina, I., Vernet, C., Astier, C. & Wollman, F. A. Randomization of the EF particles in thylakoid membranes of *synechocystis* 6714 upon transition from state I to state

- II. *FEBS Lett.* **208**, 308–312 (1986).
101. Liu, H. *et al.* Phycobilisomes supply excitations to both photosystems in a megacomplex in cyanobacteria. *Science* **342**, 1104–1107 (2013).
  102. Ranjbar Choubeh, R., Wientjes, E., Struik, P. C., Kirilovsky, D. & van Amerongen, H. State transitions in the cyanobacterium *Synechococcus elongatus* 7942 involve reversible quenching of the photosystem II core. *Biochim. Biophys. Acta - Bioenerg.* **1859**, 1059–1066 (2018).
  103. Dong, C. & Zhao, J. ApcD is required for state transition but not involved in blue-light induced quenching in the cyanobacterium *Anabaena* sp. PCC7120. *Chinese Sci. Bull.* **53**, 3422–3424 (2008).
  104. Ruban, A. V. Light harvesting control in plants. *FEBS Lett.* **592**, 3030–3039 (2018).
  105. Campbell, D. A., Hurry, V., Clarke, a K., Gustafsson, P. & Oquist, G. Chlorophyll fluorescence analysis of cyanobacterial photosynthesis and acclimation. *Microbiol. Mol. Biol. Rev.* **62**, 667–683 (1998).
  106. Rakhimberdieva, M. G., Stadnichuk, I. N., Elanskaya, I. V. & Karapetyan, N. V. Carotenoid-induced quenching of the phycobilisome fluorescence in photosystem II-deficient mutant of *Synechocystis* sp. *FEBS Lett.* **574**, 85–88 (2004).
  107. Wilson A, Ajlani G, Verbavatz JM, Vass I, Kerfeld CA, K. D. A Soluble Carotenoid Protein Involved in Phycobilisome-Related Energy Dissipation in Cyanobacteria. *Plant Cell Online* **18**, 992–1007 (2006).
  108. Wilson, A. *et al.* A photoactive carotenoid protein acting as light intensity sensor. *Proc. Natl. Acad. Sci.* **105**, 12075–12080 (2008).
  109. Boulay, C., Wilson, A., D'haene, S. & Kirilovsky, D. Identification of a protein required for recovery of full antenna capacity in OCP-related photoprotective mechanism in cyanobacteria. *Proc. Natl. Acad. Sci.* **107**, 11620–11625 (2010).
  110. Boulay, C., Abasova, L., Six, C., Vass, I. & Kirilovsky, D. Occurrence and function of the orange carotenoid protein in photoprotective mechanisms in various cyanobacteria. *Biochim. Biophys. Acta - Bioenerg.* (2008). doi:10.1016/j.bbabi.2008.07.002
  111. Wilson, A., Boulay, C., Wilde, A., Kerfeld, C. A. & Kirilovsky, D. Light-Induced Energy Dissipation in Iron-Starved Cyanobacteria: Roles of OCP and IsiA Proteins. *Plant Cell Online* **19**, 656–672 (2007).
  112. Kay Holt, T. & Krogmann, D. W. A carotenoid-protein from cyanobacteria. *BBA - Bioenerg.* **637**, 408–414 (1981).
  113. Wu, Y. P. & Krogmann, D. W. The orange carotenoid protein of *Synechocystis* PCC 6803. *Biochim. Biophys. Acta - Bioenerg.* **1322**, 1–7 (1997).
  114. Kerfeld, C. A. *et al.* The crystal structure of a cyanobacterial water-soluble carotenoid binding protein. *Structure* **11**, 55–65 (2003).
  115. Wilson, A. *et al.* Structural determinants underlying photoprotection in the photoactive orange carotenoid protein of cyanobacteria. *J. Biol. Chem.* **285**, 18364–18375 (2010).
  116. Punginelli, C., Wilson, A., Routaboul, J. M. & Kirilovsky, D. Influence of zeaxanthin and echinenone binding on the activity of the Orange Carotenoid Protein. *Biochim. Biophys. Acta - Bioenerg.* (2009). doi:10.1016/j.bbabi.2009.01.011
  117. De Carbon, C. B., Thurotte, A., Wilson, A., Perreau, F. & Kirilovsky, D. Biosynthesis of soluble carotenoid holoproteins in *Escherichia coli*. *Sci. Rep.* **5**, 9085 (2015).
  118. Maksimov, E. G. *et al.* A comparative study of three signaling forms of the orange carotenoid protein. *Photosynth. Res.* **130**, 389–401 (2016).
  119. Zhang, H. *et al.* Molecular mechanism of photoactivation and structural location of the cyanobacterial orange carotenoid protein. *Biochemistry* **53**, 13–19 (2014).
  120. Liu, H. *et al.* Mass spectrometry footprinting reveals the structural rearrangements of cyanobacterial orange carotenoid protein upon light activation. *Biochim. Biophys. Acta - Bioenerg.* **1837**, 1955–1963 (2014).
  121. Gwizdala, M., Wilson, A. & Kirilovsky, D. In vitro reconstitution of the cyanobacterial photoprotective mechanism mediated by the Orange Carotenoid Protein in *Synechocystis* PCC 6803. *Plant Cell* **23**, 2631–43 (2011).
  122. Leverenz, R. L. *et al.* Structural and Functional Modularity of the Orange Carotenoid Protein: Distinct Roles for the N- and C-Terminal Domains in Cyanobacterial Photoprotection. *Plant Cell* (2014). doi:10.1105/tpc.113.118588
  123. Jallet, D. *et al.* Specificity of the Cyanobacterial Orange Carotenoid Protein: Influences of Orange Carotenoid Protein and Phycobilisome Structures. *Plant Physiol.* **164**, 790–804 (2014).

124. Maksimov, E. G. *et al.* The Signaling State of Orange Carotenoid Protein. *Biophys. J.* **109**, 595–607 (2015).
125. King, J. D., Liu, H., He, G., Orf, G. S. & Blankenship, R. E. Chemical activation of the cyanobacterial orange carotenoid protein. *FEBS Lett.* **588**, 4561–4565 (2014).
126. Wilson, A., Punginelli, C., Couturier, M., Perreau, F. & Kirilovsky, D. Essential role of two tyrosines and two tryptophans on the photoprotection activity of the Orange Carotenoid Protein. *Biochim. Biophys. Acta - Bioenerg.* **1807**, 293–301 (2011).
127. Liu, H. *et al.* Photoactivation and relaxation studies on the cyanobacterial orange carotenoid protein in the presence of copper ion. *Photosynth. Res.* **135**, 143–147 (2018).
128. Kish, E., Pinto, M. M. M., Kirilovsky, D., Spezia, R. & Robert, B. Echinenone vibrational properties: From solvents to the orange carotenoid protein. *Biochim. Biophys. Acta - Bioenerg.* **1847**, 1044–1054 (2015).
129. Slouf, V. *et al.* Ultrafast spectroscopy tracks carotenoid configurations in the orange and red carotenoid proteins from cyanobacteria. *Photosynth. Res.* 105–117 (2017). doi:10.1007/s11120-016-0302-6
130. Wilson, C. W. A. *et al.* The Essential Role of the N-Terminal Domain of the Orange Carotenoid Protein in Cyanobacterial Photoprotection: Importance of a Positive Charge for Phycobilisome Binding. *Plant Cell* **24**, 1972–1983 (2012).
131. Leverenz, R. L. *et al.* A 12 Å carotenoid translocation in a photoswitch associated with cyanobacterial photoprotection. *Science (80-. )*. **348**, 1463–1466 (2015).
132. Gupta, S. *et al.* Local and global structural drivers for the photoactivation of the orange carotenoid protein. *Proc. Natl. Acad. Sci.* **112**, E5567–E5574 (2015).
133. Maksimov, E. G. *et al.* The photocycle of orange carotenoid protein conceals distinct intermediates and asynchronous changes in the carotenoid and protein components. *Sci. Rep.* **7**, 15548 (2017).
134. Bandara, S. *et al.* Photoactivation mechanism of a carotenoid-based photoreceptor. *Proc. Natl. Acad. Sci.* (2017). doi:10.1073/pnas.1700956114
135. Konold, P. E. *et al.* Photoactivation mechanism , timing of protein secondary structure dynamics and carotenoid translocation in the Orange Carotenoid. *J. Am. Chem. Soc.* **141**, 520–530 (2019).
136. Rakhimberdieva, M. G., Vavilin, D. V., Vermaas, W. F. J., Elanskaya, I. V. & Karapetyan, N. V. Phycobilin/chlorophyll excitation equilibration upon carotenoid-induced non-photochemical fluorescence quenching in phycobilisomes of the cyanobacterium *Synechocystis* sp. PCC 6803. *Biochim. Biophys. Acta - Bioenerg.* **1767**, 757–765 (2007).
137. Harris, D. *et al.* Orange carotenoid protein burrows into the phycobilisome to provide photoprotection. *Proc. Natl. Acad. Sci.* (2016). doi:10.1073/pnas.1523680113
138. Zlenko, D. V. *et al.* Role of the PB-loop in ApcE and phycobilisome core function in cyanobacterium *Synechocystis* sp. PCC 6803. *Biochim. Biophys. Acta - Bioenerg.* **1860**, 155–166 (2019).
139. Rakhimberdieva, M. G., Elanskaya, I. V., Vermaas, W. F. J. & Karapetyan, N. V. Carotenoid-triggered energy dissipation in phycobilisomes of *Synechocystis* sp. PCC 6803 diverts excitation away from reaction centers of both photosystems. *Biochim. Biophys. Acta - Bioenerg.* **1797**, 241–249 (2010).
140. Tian, L. *et al.* Picosecond kinetics of light harvesting and photoprotective quenching in wild-type and mutant phycobilisomes isolated from the cyanobacterium *Synechocystis* PCC 6803. *Biophys. J.* **102**, 1692–1700 (2012).
141. Jallet, D., Gwizdala, M. & Kirilovsky, D. ApcD, ApcF and ApcE are not required for the Orange Carotenoid Protein related phycobilisome fluorescence quenching in the cyanobacterium *Synechocystis* PCC 6803. *Biochim. Biophys. Acta - Bioenerg.* **1817**, 1418–1427 (2012).
142. Squires, A. H. *et al.* Single-molecule trapping and spectroscopy reveals photophysical heterogeneity of phycobilisomes quenched by Orange Carotenoid Protein. *Nat. Commun.* **10**, 1172 (2019).
143. Gwizdala, M. *et al.* Switching an Individual Phycobilisome Off and On. (2018). doi:10.1021/acs.jpcclett.8b00767
144. Berera, R. *et al.* The Photophysics of the Orange Carotenoid Protein, a Light-Powered Molecular Switch. *J. Phys. Chem. B* **116**, 2568–2574 (2012).
145. Berera, R., Gwizdala, M., Van Stokkum, I. H. M., Kirilovsky, D. & Van Grondelle, R. Excited states of the inactive and active forms of the orange carotenoid protein. *J. Phys. Chem. B* **117**, 9121–9128 (2013).

146. Polívka, T., Chábera, P. & Kerfeld, C. A. Carotenoid-protein interaction alters the S1 energy of hydroxyechinenone in the Orange Carotenoid Protein. *Biochim. Biophys. Acta - Bioenerg.* **1827**, 248–254 (2013).
147. Maksimov, E. G. *et al.* The time course of non-photochemical quenching in phycobilisomes of *Synechocystis* sp. PCC6803 as revealed by picosecond time-resolved fluorimetry. *Biochim. Biophys. Acta - Bioenerg.* **1837**, 1540–1547 (2014).
148. Tomas Polivka, Cheryl A. Kerfeld, Torbjorn Pascher, and V. S. Spectroscopic Properties of the Carotenoid 3'-Hydroxyechinenone in the Orange Carotenoid Protein from the Cyanobacterium *Arthrospira maxima* †. *Biochemistry* **44**, 3994–4003 (2005).
149. Chábera, P., Durchan, M., Shih, P. M., Kerfeld, C. A. & Polívka, T. Excited-state properties of the 16kDa red carotenoid protein from *Arthrospira maxima*. *Biochim. Biophys. Acta - Bioenerg.* **1807**, 30–35 (2011).
150. Niedzwiedzki, D. M., Liu, H. & Blankenship, R. E. Excited state properties of 3'-hydroxyechinenone in solvents and in the orange carotenoid protein from *synechocystis* sp. PCC 6803. *J. Phys. Chem. B* **118**, 6141–6149 (2014).
151. Sedoud, A. *et al.* The Cyanobacterial Photoactive Orange Carotenoid Protein Is an Excellent Singlet Oxygen Quencher. *Plant Cell* **26**, 1781–1791 (2014).
152. Gwizdala, M., Wilson, A., Omairi-Nasser, A. & Kirilovsky, D. Characterization of the *Synechocystis* PCC 6803 Fluorescence Recovery Protein involved in photoprotection. *Biochim. Biophys. Acta - Bioenerg.* **1827**, 348–354 (2013).
153. Sutter, M. *et al.* Crystal structure of the FRP and identification of the active site for modulation of OCP-mediated photoprotection in cyanobacteria. *Proc. Natl. Acad. Sci.* **110**, 10022–10027 (2013).
154. Lu, Y. *et al.* Native mass spectrometry analysis of oligomerization states of fluorescence recovery protein and orange carotenoid protein: Two proteins involved in the cyanobacterial photoprotection cycle. *Biochemistry* **56**, 160–166 (2017).
155. Sluchanko, N. N. *et al.* The purple Trp288Ala mutant of *Synechocystis* OCP persistently quenches phycobilisome fluorescence and tightly interacts with FRP. *Biochim. Biophys. Acta - Bioenerg.* **1858**, 1–11 (2017).
156. Sluchanko, N. N. *et al.* OCP–FRP protein complex topologies suggest a mechanism for controlling high light tolerance in cyanobacteria. *Nat. Commun.* **9**, 1–15 (2018).
157. Lu, Y. *et al.* A Molecular Mechanism for Nonphotochemical Quenching in Cyanobacteria. *Biochemistry* **56**, 2812–2823 (2017).
158. Thurotte, A. *et al.* Regulation of Orange Carotenoid Protein Activity in Cyanobacterial Photoprotection. *Plant Physiol.* **169**, 737–747 (2015).
159. Moldenhauer, M. *et al.* Interaction of the signaling state analog and the apoprotein form of the orange carotenoid protein with the fluorescence recovery protein. *Photosynth. Res.* **135**, 125–139 (2018).
160. Moldenhauer, M. *et al.* Assembly of photoactive orange carotenoid protein from its domains unravels a carotenoid shuttle mechanism. *Photosynth. Res.* **133**, 327–341 (2017).
161. Sluchanko, N. N., Slonimskiy, Y. B., Moldenhauer, M., Friedrich, T. & Maksimov, E. G. *Deletion of the short N-terminal extension in OCP reveals the main site for FRP binding.* *FEBS Letters* **591**, 1667–1676 (2017).
162. Thurotte, A. *et al.* The cyanobacterial Fluorescence Recovery Protein has two distinct activities: Orange Carotenoid Protein amino acids involved in FRP interaction. *Biochim. Biophys. Acta - Bioenerg.* **1858**, 308–317 (2017).
163. Slonimskiy, Y. B. *et al.* Functional interaction of low-homology FRPs from different cyanobacteria with *Synechocystis* OCP. doi:10.1101/247882
164. Melnicki, M. R. *et al.* Structure, Diversity, and Evolution of a New Family of Soluble Carotenoid-Binding Proteins in Cyanobacteria. *Mol. Plant* **9**, 1379–1394 (2016).
165. López-Igual, R. *et al.* Different Functions of the Paralogs to the N-Terminal Domain of the Orange Carotenoid Protein in the Cyanobacterium *Anabaena* sp. PCC 7120. *Plant Physiol.* **171**, 1852–1866 (2016).
166. Bao, H. *et al.* Additional families of orange carotenoid proteins in the photoprotective system of cyanobacteria. *Nat. Plants* **3**, 17089 (2017).
167. Lechno-Yossef, S., Melnicki, M. R., Bao, H., Montgomery, B. L. & Kerfeld, C. A. Synthetic OCP heterodimers are photoactive and recapitulate the fusion of two primitive carotenoproteins in the evolution of cyanobacterial photoprotection. *Plant J.* **91**, 646–656 (2017).
168. Berendsen, H. J. C., van der Spoel, D. & van Drunen, R. GROMACS: A message-passing

parallel molecular dynamics implementation. *Comput. Phys. Commun.* **91**, 43–56 (1995).

## **Part II: Results**

---

### **Chapter I**

# **The paralog of the C-terminal domain of the OCP: a novel carotenoid carrier.**

## *Summary of work*

---

This chapter is based on:

Muzzopappa, Fernando, Adjélé Wilson, Vinosa Yogarajah, Sandrine Cot, François Perreau, Cédric Montigny, Céline Bourcier de Carbon, and Diana Kirilovsky. 2017. “The Paralogs to the C-Terminal Domain of the Cyanobacterial OCP Are Carotenoid Donors to HCPs.” *Plant Physiology* 175 (3): 1283–1303.

## *Contribution to this work*

Adjélé Wilson constructed the plasmids for protein expression and isolated some proteins. I isolated some proteins and I performed all the biochemical characterization of the CTDHs and studied the interaction between CTDH and HCP. I also performed the bioinformatic analysis of the CTDH sequences.

## *Resume of the article*

Recent studies explored the interaction between independent NTD and CTD domains from OCP. It was found that as result of this interaction an OCP-like complex is formed. However, if the paralogs of such domains, which are not fully characterized, interact in the same way is unknown. This article attempted to characterize the CTDH family and to elucidate whether the CTDHs and HCPs can interact in the same way than CTD-NTD.

First, a phylogenetic study was done to analyze the CTDH family. We found that CTDH and CTD populate different parts of the phylogenetic tree. Within the CTDH family there are two subfamilies that we called CTDH1 and CTDH2. A particular difference between CTDH1 and CTDH2 is the Cys 103 which is highly conserved in CTDH2 but not present in CTDH1. We studied one CTDH of each subfamily: *T. elongatus* CTDH1 (hereafter TeCTDH), *Anabaena* CTDH2 (hereafter AnaCTDH). We also characterized AnaCTDH with a mutation in Cys 103 in order to study the role of that cysteine (hereafter AnaCTDH-C103F). Finally, the interactions of both CTDH with *T. elongatus* HCP (hereafter TeHCP) and *Anabaena* HCP4 (hereafter AnaHCP4) were studied. CTD and NTD of *Synechocystis* PCC6803 were also analyzed. All proteins were expressed and purified from

*E. coli* as apo-protein and holo-protein (expressed in canthaxanthin producing *E. coli* strain).

We first demonstrated that the CTDH are carotenoid proteins. A first analysis of the oligomeric state of these proteins was done by size exclusion chromatography (SEC) and Native PAGE. We found that TeCTDH(apo) is a monomer, AnaCTDH(apo) is present as dimer and monomer in solution and the mutant AnaCTDH-C103F(apo) is present only as monomer. Holo-CTDHs and holo-CTD are mainly dimers. These results strongly suggested that in presence of carotenoid the CTDH dimer is stabilized.

While holo-CTD and apo-NTD can interact forming an OCP-like complex, which can be irreversibly photoactivated, holo-CTDH and apo-HCP cannot. Absorbance spectra at different times of the interaction suggested that the carotenoid is completely and irreversibly transferred to HCP yielding to holo-HCP formation. This was confirmed by testing PBs fluorescence quenching by holo-HCP. We found that CTD is able to form an OCP-like complex only with NTD. CTD interaction with both HCPs result in carotenoid transfer to HCP (similar to CTDH-HCP interaction). Oxidized holo-AnaCTDH (purified without DTT) was unable to transfer the carotenoid to any HCP or NTD. We hypothesize that Cys103 is forming a disulfide bond avoiding the monomerization needed for carotenoid transfer to HCP. To confirm this, we repeated the experiment using both holo-anaCTDH purified in presence of DTT and holo-AnaCTDH-C103F. We found that in this case anaCTDH is able to transfer the carotenoid to HCPana and HCPte but not to NTD. Holo-teCTDH was able to interact and transfer the carotenoid only to apo-TeHCP.

Finally, we analyzed if this carotenoid-transfer mechanism has a biological meaning. If both HCPs and CTDHs can be synthesized as holo-proteins *in vivo*, the role of carotenoid shutter of CTDHs is meaningless. We found that OCP and CTDH (or CTD) are very efficient in taking carotenoid from membranes, but not the HCPs there is no significant holo-protein formation. Mixing apo-CTDH, apo-HCP and carotenoid containing membranes led to holo-HCP formation, suggesting that the principal role of CTDH *in vivo* can be being the carotenoid donor to HCP.

The most important conclusions of this article are: 1) CTDHs are carotenoid proteins.

- Apo-CTDH are mainly monomeric while the carotenoid binding stabilizes the dimer.

- CTDHs are able to transfer the carotenoid to HCP.
- Only the CTDH can uptake carotenoids from membranes.
- CTDH2 has a conserved Cys which can form a disulfide bond and regulate the carotenoid transfer.

## ***The paralogs to the C-terminal domain of the cyanobacterial Orange Carotenoid Protein are carotenoid donors to Helical Carotenoid Proteins***

---

Fernando Muzzopappa\*<sup>1,2</sup>, Adjélé Wilson\*<sup>1,2</sup>, Vinosha Yogarajah<sup>1,2</sup>, Sandrine Cot<sup>1,2</sup>, François Perreau<sup>3</sup>, Cédric Montigny<sup>1,2</sup>, Céline Bourcier de Carbon<sup>1,2</sup>, and Diana Kirilovsky<sup>1,2a</sup>.

\* These authors contributed equally to this work

### **Affiliations**

<sup>1</sup> Institute for Integrative Biology of the Cell (I2BC), CEA, CNRS, Université Paris-Sud, Université Paris-Saclay, 91198 Gif sur Yvette, France

<sup>2</sup> Institut de Biologie et Technologies de Saclay (iBiTec-S), Commissariat à l'Énergie Atomique (CEA), 91191 Gif-sur-Yvette, France.

<sup>3</sup> INRA, Institut Jean-Pierre Bourgin, UMR 1318, ERL CNRS 3559, Saclay Plant Sciences, F-78026 Versailles, France (F.P.)

### ***Abstract***

The photoactive Orange Carotenoid Protein photoprotects cyanobacteria cells by quenching singlet oxygen and excess excitation energy. Its N-terminal domain (NTD) is the active part of the protein and the C-terminal domain (CTD) regulates the activity. Recently, the characteristics of a family of soluble carotenoid-binding proteins (Helical Carotenoid Proteins or HCPs), paralogs of OCP-NTD, were described. Bioinformatics studies also revealed the existence of genes coding for homologs of OCP-CTD. Here, we show that the latter genes encode carotenoid proteins (CTDHs). This family of proteins contains two subgroups with distinct characteristics. One CTDH of each clade was further characterized and they proved to be very good singlet oxygen quenchers. When synthesized in *E. coli* or *Synechocystis* PCC 6803, CTDHs formed dimers that share a carotenoid molecule and are able to transfer their carotenoid to apo-HCPs and apo-OCP. The CTDHs from clade

2 have a cysteine in position 103. A disulfide bond is easily formed between the monomers of the dimer preventing carotenoid transfer. This suggests that the transfer of the carotenoid could be redox regulated in clade 2 CTDH. We also demonstrate here that apo-OCPs and apo CTDHs are able to take the carotenoid directly from membranes, while HCPs are unable. HCPs need the presence of CTDH to become holo-proteins. We propose that in cyanobacteria the CTDHs are carotenoid donors to HCPs.

## *Introduction*

---

Photosynthetic organisms performing oxygenic photosynthesis use solar energy, water and inorganic carbon to produce all the organic molecules they need. Photosynthesis converts the absorbed energy into chemical energy and reduction power necessary for the assimilation of CO<sub>2</sub> and synthesis of organic carbon molecules. However, photosynthetic organisms cannot control the incoming flux of light and too much light generates secondary reactions creating dangerous species of oxygen leading to cellular damage. Thus, in order to survive, photosynthetic organisms have developed a large variety of photoprotective mechanisms. One of them decreases the energy arriving at the reaction centers by increasing thermal dissipation of excess excitation energy at the level of the antennae (for a review, see (Niyogi and Truong, 2013)). In cyanobacteria, a soluble carotenoid protein, the Orange Carotenoid Protein (OCP) is essential for this mechanism known as the OCP-related Non-Photochemical Quenching (NPQ) mechanism ((Wilson et al., 2006), for a recent review see (Kirilovsky and Kerfeld, 2016)). In addition, OCP has a second photoprotective activity. It is a very good singlet oxygen (<sup>1</sup>O<sub>2</sub>) quencher (Kerfeld et al., 2003; Sedoud et al., 2014).

OCP is a photoactive soluble carotenoid protein (Wilson et al., 2008). It is composed of two globular domains: an  $\alpha$ -helical N-terminal domain (NTD, residues 18-165) that is unique to cyanobacteria and an  $\alpha$  helix/ $\beta$  sheet C-terminal domain (CTD, residues 190-317) which is a member of the nuclear transport factor 2 superfamily (Kerfeld et al., 2003; Wilson et al., 2010). The two domains are connected by a long flexible linker (25 residues). In darkness, strong interactions between these two domains of the orange OCP (OCP<sup>0</sup>) maintain the protein in a closed and compact globular conformation. The interactions occur at the principal interface between the domains, in the middle of the protein (including a salt bridge between residues R155 and

E244), and in the interface between the N-terminal arm (residues 1-20, containing a small  $\alpha$  helix) and the  $\beta$ -sheet of the CTD (Kerfeld et al., 2003; Wilson et al., 2010). In *Synechocystis* and *Arthrospira* cells, the OCP binds a keto-carotenoid, the hydroxyechinenone (hECN), which has a carbonyl (keto) group in one of the rings and a hydroxyl group in the other (Kerfeld et al., 2003; Wilson et al., 2010). The hECN spans both domains and it is in an *all-trans* configuration with its carbonyl group localized in a hydrophobic pocket of the CTD (Kerfeld et al., 2003; Polivka et al., 2005). In the absence of hECN, OCP can bind other keto-carotenoids such as echinenone (ECN) and canthaxanthin (CAN) (Punginelli et al., 2009; Bourcier de Carbon et al., 2015). During photoactivation carotenoid and protein conformational changes are induced (Wilson et al., 2008). Strong blue-green light induces the breaking of the hydrogen bonds between the carotenoid and the protein and those between the CTD and NTD domains, leading to the opening of the protein and the translocation of the carotenoid toward the NTD. The photoactivated OCP is in its red configuration ( $\text{OCP}^{\text{R}}$ ) with the two domains completely separated and the carotenoid buried in the NTD (Wilson et al., 2012; Liu et al., 2014; Gupta et al., 2015; Leverenz et al., 2015; Liu et al., 2016; Maksimov et al., 2017a). In  $\text{OCP}^{\text{R}}$ , NTD interacts with the phycobilisome (PBS) core, the cyanobacterial antenna, inducing an increase of the thermal dissipation of the excess excitation energy absorbed by PBS (Wilson et al., 2012; Leverenz et al., 2014; Harris et al., 2016). For this interaction to occur, NTD-Arg155 located at the interface between the two domains, is essential (Wilson et al., 2012). In  $\text{OCP}^{\text{O}}$ , CTD hinders the binding of NTD to PBS since NTD-Arg155 forms a hydrogen bond with CTD-Glu244 that stabilizes the closed globular form (Kerfeld et al., 2003).

For the recovery process involving the detachment of OCP from PBS and the conversion  $\text{OCP}^{\text{R}}$  to  $\text{OCP}^{\text{O}}$ , another protein is necessary, the Fluorescence Recovery Protein (Boulay et al., 2010; Sutter et al., 2013). FRP interacts with high affinity with the CTD of free and bound  $\text{OCP}^{\text{O}}$  (Boulay et al., 2010; Sutter et al., 2013). In *in vitro* experiments, the presence of FRP accelerates the  $\text{OCP}^{\text{R}}$  to  $\text{OCP}^{\text{O}}$  conversion (Boulay et al., 2010; Sutter et al., 2013; Thurotte et al., 2017) and the recovery of PBS fluorescence (Gwizdala et al., 2011; Thurotte et al., 2015; Thurotte et al., 2017). FRP activity is independent of light: it is active in the dark but also under high light conditions (Sluchanko et al., 2017; Thurotte et al., 2017). FRP affinity to  $\text{OCP}^{\text{R}}$  is strong and its presence during illumination decreases the amplitude of OCP-triggered PBS fluorescence quenching by accelerating the  $\text{OCP}^{\text{R}}$  to  $\text{OCP}^{\text{O}}$  conversion (Gwizdala et al., 2011; Kuzminov et al., 2012; Gwizdala et al., 2013;

Sluchanko et al., 2017; Thurotte et al., 2017) and maybe also by interacting with OCP<sup>O</sup> (with low affinity) and preventing its photoactivation (Sluchanko et al., 2017). FRP is mostly present as a dimer in solution (Sutter et al., 2013; Lu et al., 2017; Sluchanko et al., 2017). It was proposed that it could monomerize upon binding to CTD (Sluchanko et al., 2017). FRP R60 and OCP F299 are essential for the activity of FRP in the conversion OCP<sup>R</sup> to OCP<sup>O</sup> (Thurotte et al., 2017). Less is known about the interaction of FRP with the bound OCP<sup>R</sup>. None of mutations tested until now have prevented the FRP-induced acceleration of OCP detachment from PBS (Thurotte et al., 2017).

Many cyanobacterial genomes contain genes coding for proteins homolog to the N-terminal domain of OCP (Kirilovsky and Kerfeld, 2012, 2013; Melnicki et al., 2016). Last year, a bioinformatics study established that these proteins form a new family of soluble carotenoid proteins (named NTD-like proteins or Helical Carotenoid proteins (HCPs)) containing at least nine different clades (HCP1 to 9, (Melnicki et al., 2016)). The HCP4 clade is the one presenting the highest identity with the NTD of OCP (NTD-OCP). The HCPs from nitrogen-fixing *Anabaena* (*Nostoc*) PCC7120 (hereafter *Anabaena*) were characterized. *Anabaena* possesses 4 HCPs belonging to clades 1, 2, 3 and 4 (Lopez-Igual et al., 2016). All the genes are expressed at similar levels in *Anabaena* filaments grown with or without nitrate (Lopez-Igual et al., 2016). At least two of these proteins, HCP1 (All1123) and HCP4 (All4941) were present in both culture conditions. The HCP1 that was isolated from *Anabaena* cells contained several types of carotenoids with a majority of canthaxanthin, thus confirming that HCPs are carotenoid proteins (Melnicki et al., 2016). When *hcp* genes were expressed in a *E. coli* strain able to synthesize canthaxanthin, all the isolated HCP proteins bound canthaxanthin. The structure of HCP1 was resolved and its tertiary structure was found to be similar to that of the NTD-OCP (Melnicki et al., 2016). The position of the carotenoid is also similar to that in the NTD-OCP. Although all HCPs are able to bind a carotenoid molecule and have similar secondary and tertiary structures, they have distinct activities (Lopez-Igual et al., 2016). Only HCP4 (All4941) is able to bind to PBS and induce energy quenching. HCP2 (All3221) and HCP3 (Alr4873) are good singlet oxygen quenchers. HCP1 has no known activity, it is unable to bind to PBS and it is a bad singlet oxygen quencher (Lopez-Igual et al., 2016).

In addition to HCP proteins, bioinformatics studies also revealed the presence of genes coding homologs to the C-terminal domain of OCP (Melnicki et al., 2016; Bao et al., 2017). The CTD-like homolog (CTDH) of *Anabaena* is expressed in heterocyst as well as filaments grown in the

presence or absence of nitrate (Lopez-Igual et al., 2016). Most of the genes coding for HCP4 proteins are adjacent to a gene coding a CTDH (Melnicki et al., 2016). It was proposed that these proteins might interact, forming an HCP4-CTDH heterodimer and that an ancient fusion event between an HCP4 and CTDH could have given rise to an ancestral OCP (Melnicki et al., 2016). Moreover, it has recently been shown that isolated NTD and CTD can interact in the absence of the linker and of the N-terminal arm suggesting that HCPs can interact with CTDH (Moldenhauer et al., 2017; Lechno-Yossef et al., 2017). In addition, when the *ctd* gene was expressed in carotenoid-containing *E. coli* cells, CTD bound canthaxanthin molecules suggesting that CTDHs could also be carotenoid proteins (Moldenhauer et al., 2017; Lechno-Yossef et al., 2017).

The study described in this article focuses on the CTDH family of proteins. A bioinformatics analysis is first presented indicating that this family is formed by two clades with different characteristics. These proteins conserved some of the amino acids involved in carotenoid interaction in the CTD of OCP. Two CTDHs belonging to different clades, one from *Anabaena* and the other from *Thermosynechococcus elongatus* (thereafter *T. elongatus*) were further studied. In both strains, the gene encoding CTDH is adjacent to the gene encoding an HCP protein able to interact with PBS and quench its fluorescence. The CTDH proteins strongly bound canthaxanthin when synthesized in *E. coli* and *Synechocystis* strains suggesting that they form a new family of carotenoid proteins. The interactions between CTDH and HCP proteins were studied as well as the possible role of CTDH as a carotenoid donor to apo-HCPs and apo-OCPs. Our results support the hypothesis that in cyanobacteria cells, CTDHs are in fact carotenoid donors to HCPs.

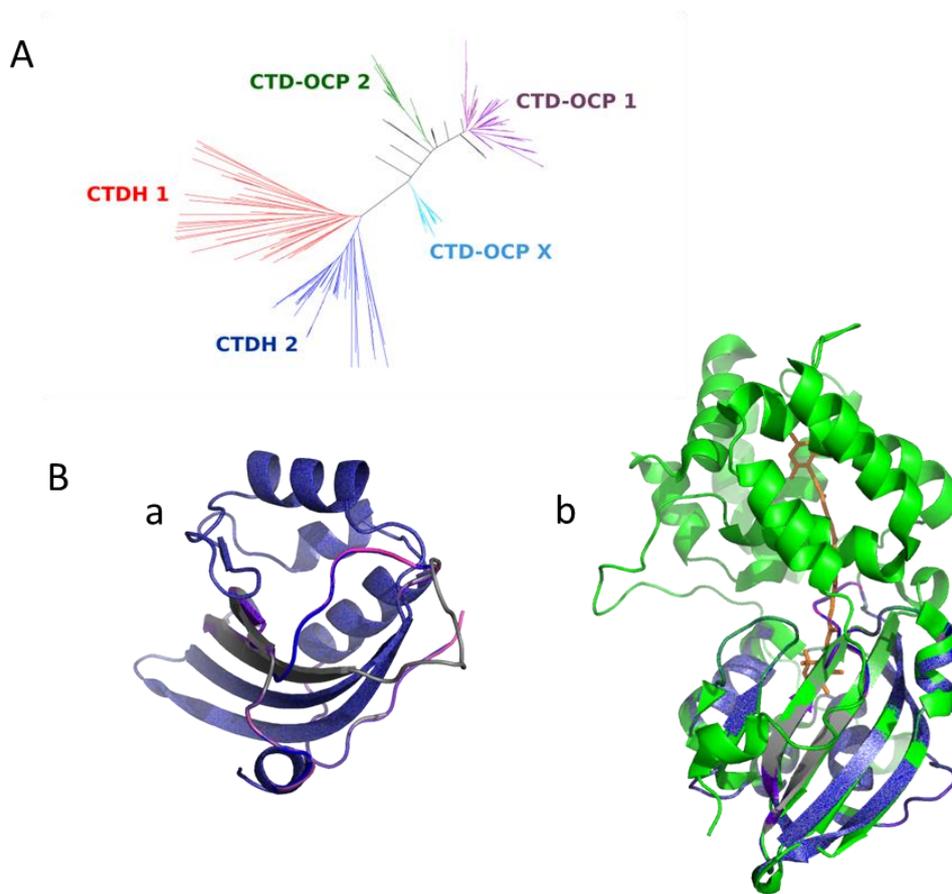
## ***Results***

---

### ***Phylogenetic analysis and overview of CTDH. Structural models***

Homologs of the CTD of OCP (CTDH) were identified using a Blast database search complemented by an Hmm Search, yielding 109 sequences belonging to the cyanobacteria phylum. The sequences of 185 CTDs from full-length cyanobacteria OCPs and the 109 CTDHs were then used to construct a phylogenetic tree (Figure 1A; see Materials and Methods for

details). These sequences belonged to 175 different PBS-containing cyanobacteria strains (see supplemental file 1).

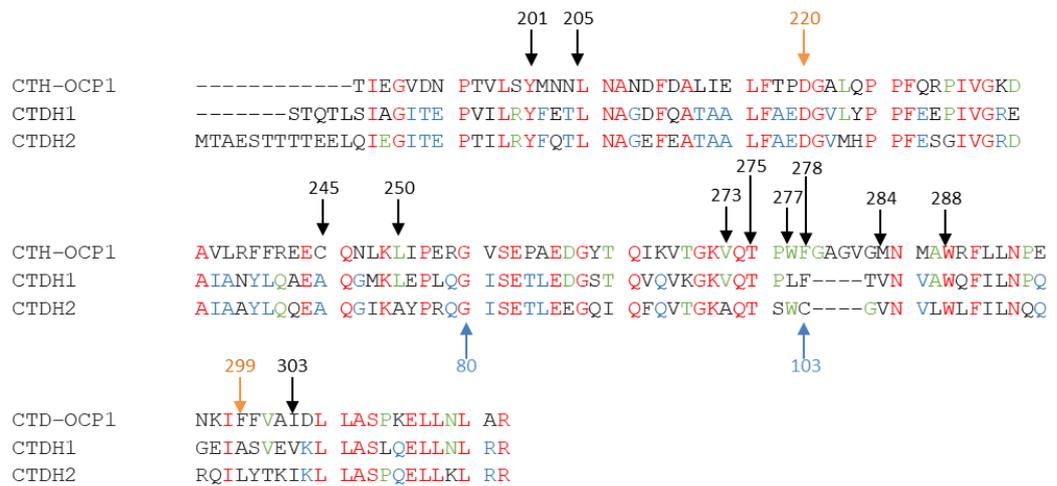


**Figure 1 | The CTDH proteins: Evolution and structure.** A) Evolutionary divergence of CTDH proteins and CTD-OCPs. Unrooted maximum likelihood phylogenetic tree of CTD-OCP and CTDH sequences. The CTD-OCP and CTDHs subtypes were identified as clades designated by numbers and colors. The CTD-OCPs subgroups were named CTD-OCP1, CTD-OCP2 and CTD-OCPX and the clades of CTDH proteins were numerated 1 to 2. B) Comparison of CTD-OCP and CTDH structures. The OCP structure is that of the crystal structure (PDB 1M98). The structures of the CTD-OCP and CTDHs were generated as homology model (see materials and methods). a) Comparison of CTD-OCP (grey) with *Anabaena* CTDH (blue) and *T. elongatus* CTDH (fuchsia). b) Comparison of crystal structure of OCP (green) with CTD-OCP (grey) and CTDH (*Anabaena* (blue) and *T. elongatus* (fuchsia) models. The CTDHs structures of both strains are identical

Our analysis shows that 104 cyanobacteria strains contain one CTDH protein. Two distinct major monophyletic clades of CTDHs were detected: Clade 1 (56 proteins) and Clade 2 (49 proteins) (Figure 11A). The CTDH from *Gloeobacter violaceus* and *Gloeobacter kilaueenses*, which are considered the most ancient cyanobacteria strains, belong to Clade 1. The CTD-OCP sequences are also shown in the phylogenetic tree. Most CTD-OCP sequences are grouped in one large clade (CTD-OCP1). This clade is mapped opposite the CTDH proteins. The two other clades (CTD-OCP2 (26

and CTD-OCPX (26)) contain CTD-OCPs mostly belonging to other families of OCPs (Bao et al., 2017).

In all strains containing CTDH, HCPs are present (supplemental file 1). All the *hcp4* genes are located adjacent to or near a *ctdh* gene while this is the case for only half of the *hcp5* genes (for HCP nomenclature and characteristics see (Melnicki et al., 2016)). In rare cases *ctdh* genes are located adjacent to *hcp* genes belonging to other clades. For example, in *Gloeobacter* strains, the *ctdh* gene is adjacent to a *hcp9* gene.



**Figure 12 | Consensus sequence alignment of different CTDH clades against CTD-OCP.** The amino acids 100% conserved in CTDHs and CTD-OCP are in red; the ones conserved among CTDHs but not present in CTD-OCP are in blue. Amino acids that are conserved in CTD-OCP and at least in one CTDH subgroup are in green. The black arrows indicate the amino acids involved in carotenoid binding in CTD-OCP. The brown arrows mark the two amino acids that are essential for FRP activity. The blue arrows mark the cysteine 103 that is conserved in CTDH2 and position 80.

All CTDH proteins lack 4 amino acids after amino acid 103 (CTDH Clade 2 nomenclature) compared to CTD-OCP. CTDH amino acid 103 is a Phe in Clade 1 and in all CTD-OCPs and is a Cys in Clade 2 (Figure 2). The lack of 4 amino acids seems to have an effect on the tertiary structure of the protein. Figure 1B compares the crystal tridimensional structure of the CTD-OCP with structural homology models of the CTDHs from *Anabaena* (All4940) (Clade 2) and *T. elongatus* (Tll1268) (Clade 1) generated using the Phyre2 web program. Based on these models, the tridimensional structures of both CTDHs are identical. When compared to the CTD-OCP structure, it is observed that the  $\beta$ -sheet region in CTDHs is modified:  $\beta$ -strand 6 is predicted to be shorter and the loop that connects it to  $\beta$ -strand 5 is in a different conformation than in CTD-OCP. This structure is only a model and

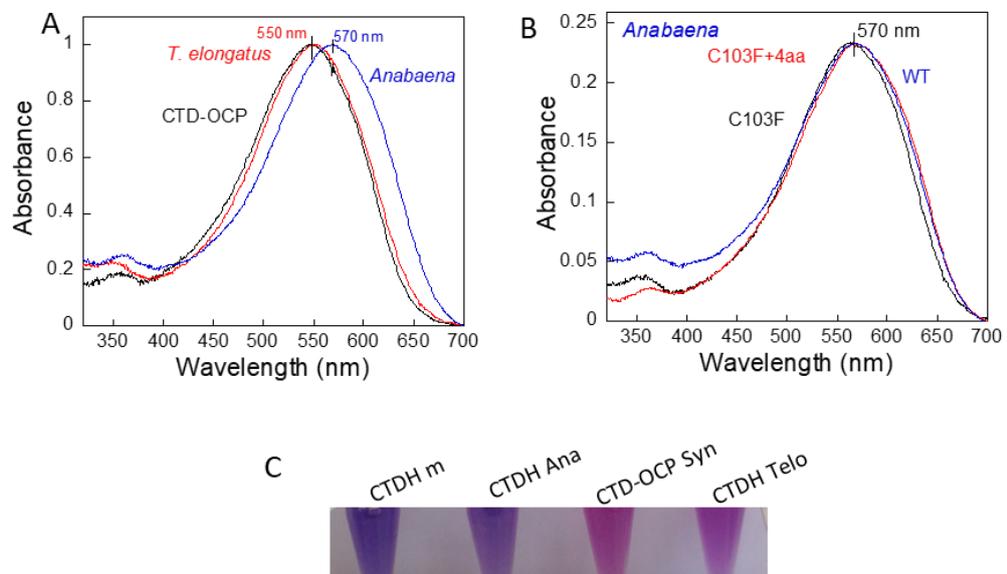
the position of the loop could be different in the protein. However, the models suggest that the absence of these 4 amino acids has an impact on the structure of the protein especially in the region interacting with the NTD in the OCP (Figure 1B). CTDHs of *Gloeobacter* strains not only lack these 4 amino acids, but also other 4 amino acids located before amino acid 80 (Clade 1 nomenclature in Figure 2).

Figure 2 shows the alignment of the consensus sequence of the two major groups of CTDHs compared to CTD-OCP1. The conservation of the amino acids within each clade and subgroup is shown in supplementary Figure S1. The CTDH proteins conserve some of the amino acids known to be proximal to the carotenoid in the orange OCP: Y201, L205, V273, T275, W288 (Figure 2) (Kerfeld et al., 2003; Wilson et al., 2010; Leverenz et al., 2015). This suggests that CTDHs can bind carotenoids like the synthetic CTD-OCP (Moldenhauer et al., 2017, Lechno-Yossef et al., 2017). Moreover, the presence of Y201 and W288 suggests that the carbonyl group of a keto carotenoid could form hydrogen bonds with the protein. Other amino acids are not conserved: C245 (A in CTDHs), I250 (F or Y in CTDHs), M284 (V in CTDHs), W277 (conserved in Clade 1 but not in Clade 2), F278 (conserved in Clade 2 but not in Clade 1). W277 plays also an important role in the interaction between the NTD and CTD domains in orange OCP (Kerfeld et al., 2003; Wilson et al., 2010; Leverenz et al., 2015). F299, which is essential for FRP activity (Thurotte et al., 2017), is absent in the CTDH proteins (Figure 2).

### ***Production of CTDHs by expression in a canthaxanthin producing E. coli strain and characterization***

The CTDHs from *Anabaena* PCC 7120 (Clade 2) and *T. elongatus* (Clade 1) were further characterized and compared to the CTD-OCP from *Synechocystis* PCC 6803. While *Anabaena* cells contain OCP1 and 4 HCPs (HCP1, HCP2, HCP3, HCP4), *T. elongatus* lacks a full-length OCP and contains only one HCP (Lopez-Igual et al., 2016; Melnicki et al., 2016). The HCP from *T. elongatus* maps between the HCP4 and HCP5 clades on the evolutionary tree (Melnicki et al., 2016). It forms a cluster with the HCPs from *Thermosynechococcus* NK55 and *Pseudanabaena* PCC 7387. The *ctdh* gene is adjacent to an *hcp* gene coding for HCP4 in *Anabaena*. In *T. elongatus*, the *ctdh* and *hcp* genes are adjacent.

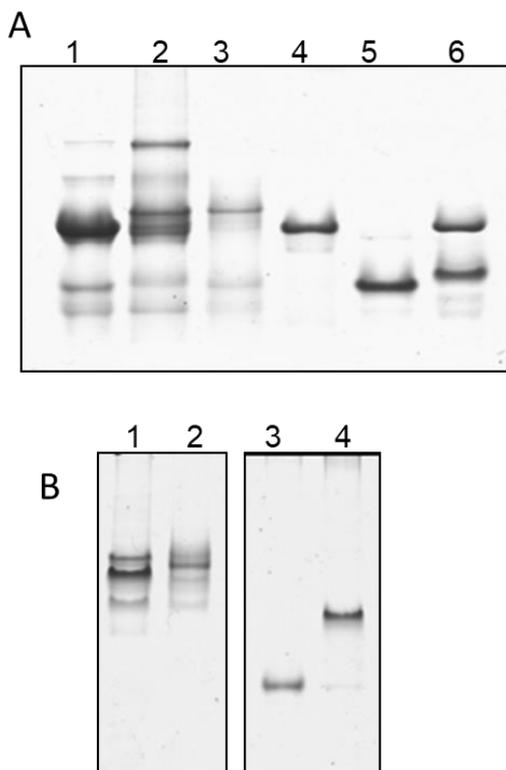
The *Anabaena ctdh* gene (*all4940*), the *T. elongatus ctdh* gene (*tll1268*) and the CTD coding moiety of the *Synechocystis* and *Anabaena ocp* genes (170 to 310) were cloned in a modified pCDFDuet-1 plasmid and expressed in *E. coli* cells in the presence or absence of carotenoids. Two mutants of the *Anabaena* CTDH were also created for this study. In one of the mutants, Cys103 was replaced by a Phe and in the other one, in addition to the Cys substitution, 4 amino acids were added just after the Cys (for details see Materials and Methods). The latter mutant was constructed in view of obtaining a CTDH protein with a structure similar to CTD-OCP (see below). These mutant CTDHs were respectively called C103F CTDH and C103F+4aa CTDH.



**Figure 3 | Absorbance spectra and appearance of different CTDH and CTD-OCP.** (A and B) Absorbance spectra of (A) *Anabaena* CTDH, *Synechocystis* CTD-OCP, and *T. elongatus* CTDH and of (B) *Anabaena* CTDH, CTDH C103F and CTDH C103F+4aa. (C) Photo of isolated CTDHs: CTDH C103F+4aa (CTDHm), *Anabaena* CTDH (CTDH Ana), *Synechocystis* CTD (CTD-OCP Syn) and *T. elongatus* CTDH (CTDH Telo).

All proteins were isolated (supplementary Figure S2) and their absorption spectra were recorded. As expected from previous results (Moldenhauer et al., 2017 and Lechno-Yossef et al., 2017), recombinant *Synechocystis* CTD-OCP (hereafter called CTD-OCP) bound canthaxanthin when expressed in a canthaxanthin producing *E. coli* strain (Figure 3). Analysis of carotenoid content showed that both WT CTDHs and mutated *Anabaena* CTDHs bound 100% canthaxanthin. All the proteins presented broad absorbance spectra without vibronic structure (Figure 3). The CTD-OCP and the *T. elongatus* CTDH are red-violet proteins and their absorbance spectra have a maximum

around 545-550 nm. The *Anabaena* WT and mutated CTDHs are violet with an absorbance maximum red-shifted to 565-570 nm (Figure 3). The differences in the absorbance maxima suggest that the carotenoid-protein interaction and carotenoid conformation might be slightly modified in the different proteins. Interestingly, while the absorbance maximum of *Synechocystis* CTD-OCP is around 550 nm (Moldenhauer et al., 2017 and this work), that of *Fremyella diplosiphon* CTD-OCP is 520 nm (Lechno-Yossef et al., 2017). The comparison of CTDH (CTD) protein concentration (calculated using Bradford) and carotenoid concentration (calculated from carotenoid absorbance) suggested that there was one carotenoid per CTDH dimer: ex. C103F CTDH preparation, 69  $\mu$ M protein and 33.5  $\mu$ M carotenoid (ratio: 2.06).



**Figure 4 | Oligomeric status of apo- and holo-CTDHs** Analysis of CTDH oligomeric state by Coomassie blue stained Native-PAGE. (A) *Anabaena* CTDHs: lanes: 1: *Anabaena* WT apo-CTDH, 2: *Anabaena* WT holo-CTDH, 3: *Anabaena* WT holo-CTDH purified in presence of 1mM DTT, 4: *Anabaena* holo-C103FCTDH, 5: *Anabaena* apo-C103FCTDH, 6: *Anabaena* holo-C103F+4aa CTDH. (B) lane 1: *Synechocystis* apo-CTD-OCP, 2: *Synechocystis* holo-CTD-OCP, 3: *T. elongatus* apo-CTDH, 4: *T. elongatus* holo-CTDH.

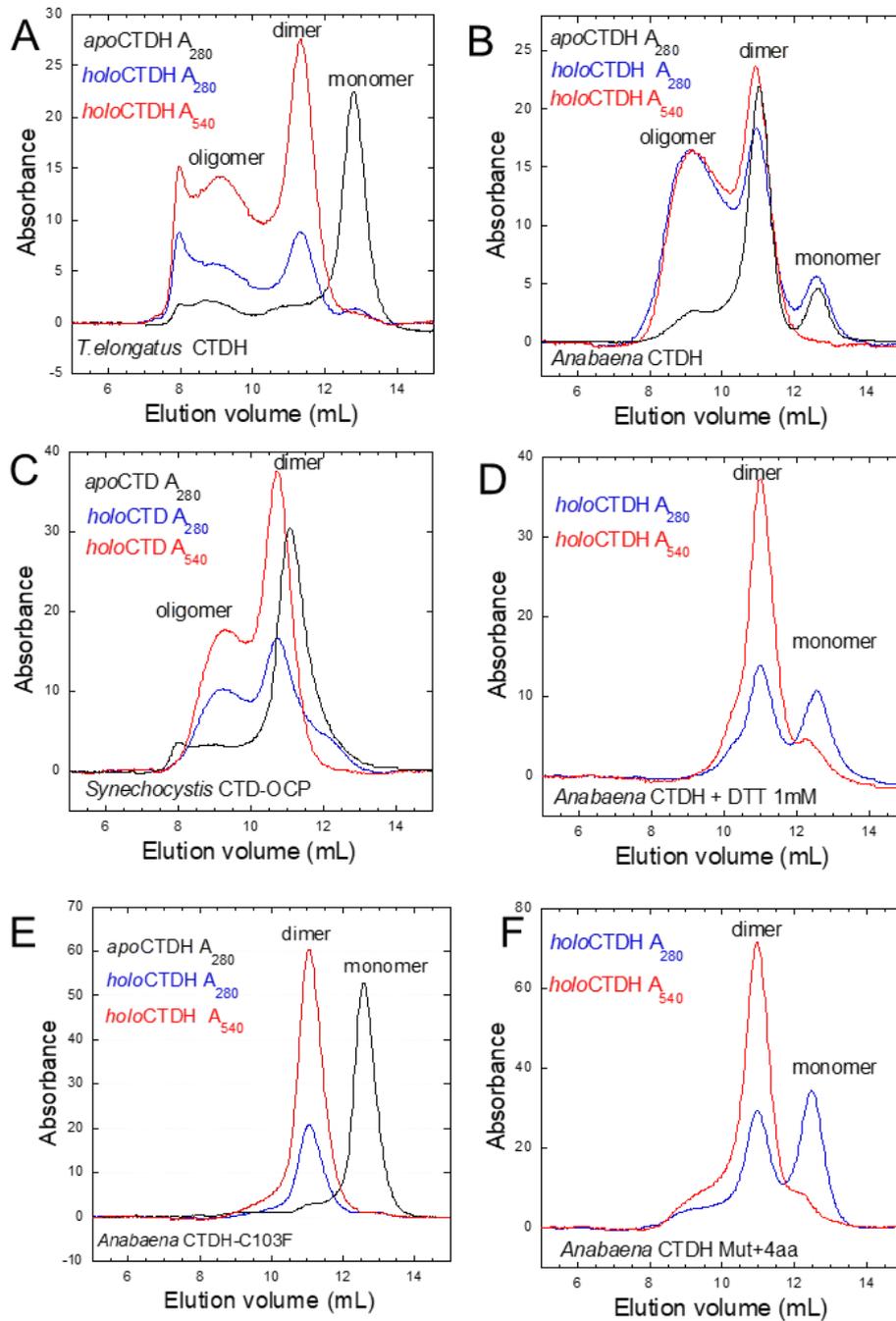
Purified holo- and apo-CTDHs and CTD-OCP were analyzed by Native-PAGE to investigate their oligomeric state (Figure 4). *T. elongatus* apo-CTDH (panel B, lane 3) and *Anabaena* C103F apo-CTDH (panel A, lane 5) migrated mainly as low molecular weight species that may correspond to monomers while *Anabaena* WT apo-CTDH appeared to migrate mainly as a dimer (panel A, lane 1). To confirm these results, purified proteins were also

analyzed by size-exclusion chromatography (SEC; Figure 5). At least at concentrations equal to or lower than 20  $\mu$ M, the *T. elongatus* apo-CTDH (15 kDa, black profile on panel A) and *Anabaena* apo-C103FCTDH (17 kDa, black profile on panel E) behaved mainly as monomers: the elution volumes of the major peaks (13 and 12.6 mL respectively) were close to the elution volume of a standard globular protein of about 17 kDa (Supplementary Figure S3). By contrast, the WT apo-CTDH from *Anabaena* was mostly present as a dimer (34 kDa for a dimer, black profile on panel B), with an elution volume of the major peak (11.1 mL) close to the elution volume of a standard globular protein of about 44 kDa (Supplementary Figure S3). According to the results obtained by Native-PAGE, SEC demonstrated that the C103 residue in *Anabaena* CTDH may favor the dimerization of the protein probably *via* disulfide bonding. When this cysteine is replaced by a phenylalanine (in *T. elongatus* CTDH, or by site-directed mutagenesis in *Anabaena* CTDH) the dimer is destabilized (Figures 4 and 5).

The results obtained with the apo-CTD-OCP were unexpected. SEC analysis of the purified apo-CTD-OCP showed that it behaved mainly as a dimer in solution (34 kDa, black profile on Panel C). This dimer was stable although a Phe residue is present instead of a Cys at position 103. Nevertheless, Moldenhauer et al, (2017) showed that apo-CTD-OCP dimers present a strong concentration-dependent dissociation: the percentage of monomers largely increased at concentrations lower than 20  $\mu$ M. In native gels, the apo-CTD-OCP appeared as multiple bands that migrated slower than the CTDH dimer bands (Figure 4, panel B). Some of these bands can be related to large oligomers and monomers which were also detected in SEC.

In Native-PAGE, the holo-CTDH from *T. elongatus* migrated much slower than the apo-protein, suggesting it was in the form of a dimer (Figure 4, Panel B lane 4). This was confirmed by SEC gel filtration where the major peak corresponded to a dimer (elution peak at 11.3 mL, Figure 5, blue profile on panel A). Recording of absorbance at 540 nm indicates that the carotenoid is mainly bound to the dimer (Figure 5, red profile on panel A). While mainly one principal band corresponding to the dimeric form appeared in the native gel, in SEC the presence of bigger complexes (and aggregates) carrying carotenoid was detected. Monomers carrying a carotenoid molecule were also detected in SEC in one of the preparations (supplementary Figure S3). Interestingly, the maximum of the absorbance spectra of CTDH monomers and dimers was different: 560 nm for dimers and 532 nm for monomers (supplementary Figure S3). The fact that CTDH monomers could bind carotenoids and present a different absorbance spectrum than that of dimers

suggests that there might be an alternative carotenoid position completely inside the CTDH monomer.



**Figure 5 | Oligomeric status of apo- and holo-CTDHs.** Analysis of CTDH oligomeric state by SEC. The holo-proteins were detected at 280 nm (blue) and 540 nm (to detect the presence of carotenoid) (red). The apo-proteins were detected only at 280 nm (black) (A) *T. elongatus* apo- and holo CTDHs; (B) *Anabaena* WT apo- and holo-CTDH ; (C) *Synechocystis* apo- and holo-CTD-OCP; (D) *Anabaena* WT holo-CTDH prepared with 1 mM DTT; (E) *Anabaena* apo- and holo-C103FCTDH; (F) *Anabaena* holo-C103F+4aa CTDH. The elution peaks around 11-11.3 mL contained CTDH dimers and around 12.6-13 mL contained CTDH monomers. Calculated from calibration profile (Materials and Methods and supplementary Figure S3).

The *Anabaena* WT holo-CTDH migrated mostly as a dimer in the native gel (Figure 4). In addition, it had the propensity to form big oligomers. In SEC, two major peaks appeared: one corresponding to the CTDH dimer (elution peak at 11.1 mL) and the other indicating the presence of big oligomers (Figures 5, Panel B, blue profile). Both forms bound canthaxanthin. However, the ratio between  $A_{540\text{ nm}}$  and  $A_{280\text{ nm}}$  (red and blue profile respectively) was lower for *Anabaena* WT holo-CTDH than for other holo-proteins, suggesting that not all dimers and oligomers bound carotenoids. In Native-PAGE, two close bands are related to dimers (Figure 4, Panel A, lane 2). They may correspond to the holo- and apo- forms of the dimer: the slowest band may correspond to the holo-form given the mobility of the apo-protein (lane 1) and of the DTT treated *Anabaena* holo-CTDH (lane 3). When the *Anabaena* CTDH was isolated in the presence of 1mM DTT, the big oligomers disappeared and the dimer was destabilized (Figure 5, Panel D). The protein was present as a mix of dimers and monomers. The dimers largely bound canthaxanthin while the monomers were mostly carotenoid-free. The C103F CTDH, which was monomeric in absence of carotenoid, was mainly dimeric when it bound a carotenoid molecule (Figure 4, panel A, lane 4 and Figure 5, blue and red profile on panel E). This indicated that addition of the carotenoid favors the dimerization as observed for *T. elongatus* holo-CTDH in which a Phe residue is present in position 103. No oligomers or monomers were detected on the SEC profile for this C103F mutant. Surprisingly, the addition of the 4 extra amino acids after F103 (C103F+4aa CTDH mutant) partially prevented (or destabilized) CTDH dimerization: Native SDS-PAGE (Figure 4, panel A, lane 6) and SEC analysis (Figure 5, blue profile on panel F) showed that monomers and dimers coexisted. The dimers largely bound canthaxanthin but the monomers carried only some carotenoid molecules.

Addition of carotenoid to the CTD-OCP (the holo CTD-OCP form) have a limited impact on the electrophoretic mobility (Figure 4, panel B, lane 2). On the SEC profile (Figure 5, blue and red profile on panel C) peaks appeared corresponding to big carotenoid binding oligomers and carotenoid-free monomers. Nevertheless, the highest peak corresponded to the dimer form. Interestingly, a small shift in the elution volume is observed between the apo- and the holo- forms suggesting that the binding of the carotenoid affected the overall structure of the dimer. This resulted in an increase of the Stokes radius of the dimer, probably reflecting some conformational changes. This effect was also observed in the holo *Anabaena* CTDH to a lower extent. The presence of big oligomers was not previously detected in CTD-OCP preparations (Moldenhauer et al., 2017 and Lechno-Yossef et al., 2017). This

suggests that their presence is mainly related to protein expression and isolation conditions than to intrinsic CTD-OCP and CTDH characteristics.

These results showed that the CTDH proteins, like isolated CTD-OCP, form carotenoid dimers. The presence of carotenoid stabilized the CTDH dimers when C103 was replaced by a Phe or under reducing conditions, since in its absence the complex monomerized. Under oxidizing conditions, in the presence of C103, the formation of an S-S bond stabilized the dimer even in the absence of carotenoids.

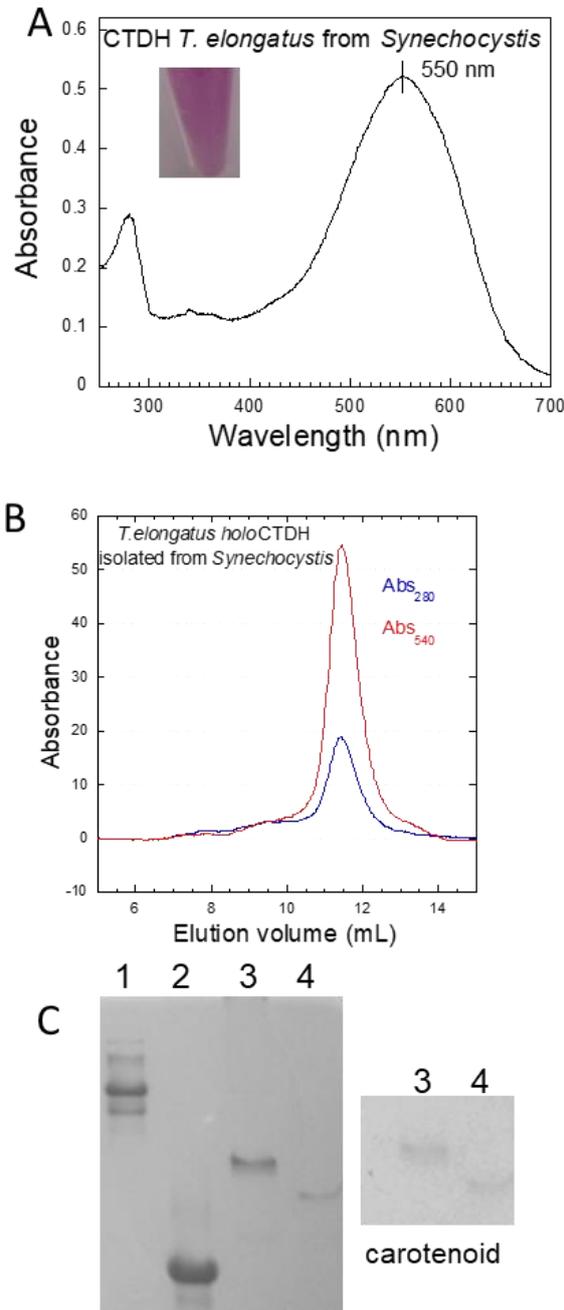
The  $^1\text{O}_2$  quenching activity of CTDH was measured *in vitro* as described by (Sedoud et al., 2014). Electron paramagnetic resonance (EPR) spin trapping was applied for  $^1\text{O}_2$  detection using TEMPD-HCl (2,2,6,6-tetramethyl-4-piperidone). When this nitron reacts with  $^1\text{O}_2$ , it is converted into the stable nitroxide radical, which is paramagnetic and detectable by EPR spectroscopy. The production of  $^1\text{O}_2$  was induced by illumination of methylene blue. The concentration of CTDH dimers that decreased the EPR signal to 50% ( $I_{50}$ ) was 1.3  $\mu\text{M}$  for both WT CTDHs indicating that they are very good  $^1\text{O}_2$  quenchers, at least *in vitro*. The CTDH  $^1\text{O}_2$  quenching activity was similar to that of OCP (Sedoud et al., 2014).

### ***Production of CTDHs by expression in Synechocystis cells and characterization***

The *Anabaena ctdh* gene (*all4940*) and the *T. elongatus ctdh* gene (*tll1268*) were also cloned into the pPSBA2 ampicillin-resistant vector (Lagarde et al., 2000) containing the strong *psbA2* promoter in order to synthesize these CTDHs in *Synechocystis* cells.

When CTD-OCP was expressed in WT *Synechocystis* cells, the isolated protein did not bind any carotenoid (Sutter et al., 2013). This result suggested that CTDHs are not able to bind carotenoids or to stabilize this binding in cyanobacteria cells. Since WT *Synechocystis* cells do not contain canthaxanthin (Punginelli et al., 2009) we decided to overexpress the *ctd-ocp* and *ctdh* genes in a *Synechocystis* mutant lacking the CrtR hydrolase. This mutant lacks zeaxanthin and hydroxy-echinenone but produces echinenone and canthaxanthin (Wilson et al., 2011). The CTD-OCP protein bound only traces of canthaxanthin (Supplementary Figure S4). By contrast, *T. elongatus* CTDH bound to carotenoid molecules and had an absorbance spectrum similar to that produced in *E. coli* cells (Figure 6). The carotenoid content of two independent preparations of *T. elongatus* CTDH was analyzed. It

contained mostly canthaxanthin (73-74%). Unidentified carotenoids (most probably oxidized canthaxanthin derivatives) were also detected.



**Figure 6 | The *T. elongatus* CTDH isolated from *Synechocystis*.** (A) Absorbance spectra of *T. elongatus* CTDH; (B) Analysis of *T. elongatus* CTDH oligomeric state by SEC. The elution peak is at 11.5 mL indicating that the CTDH is a dimer. The holo-proteins were detected at 280 nm (blue) and 540 nm (red). (C) *T. elongatus* CTDH isolated from *E. coli* (lane 3) and *Synechocystis* (lane 4) in Native-PAGE stained by Coomassie blue. The carotenoid is visible on the non-stained gel (panel on the right). Marker 30 kD (lane 1), Marker 14 kD (lane 2)

*T. elongatus* CTDH was stabilized as a dimer based on Native-PAGE gel analysis and SEC chromatography (Figures 6). Nevertheless, the dimer obtained in *Synechocystis* cells had a slightly different conformation from that obtained in *E. coli* cells: it migrated faster in the native gel and eluted in a slightly greater volume in SEC (elution peak at 11.5 mL). No bigger complexes were detected in the *Synechocystis* preparation.

The *Anabaena* CTDH was also able to bind carotenoids (73% canthaxanthin) and to form dimers but the quantity of protein obtained was very small and complete isolation was not possible (supplementary Figure S4).

### *Interaction between holo-CTDH and apo-HCP*

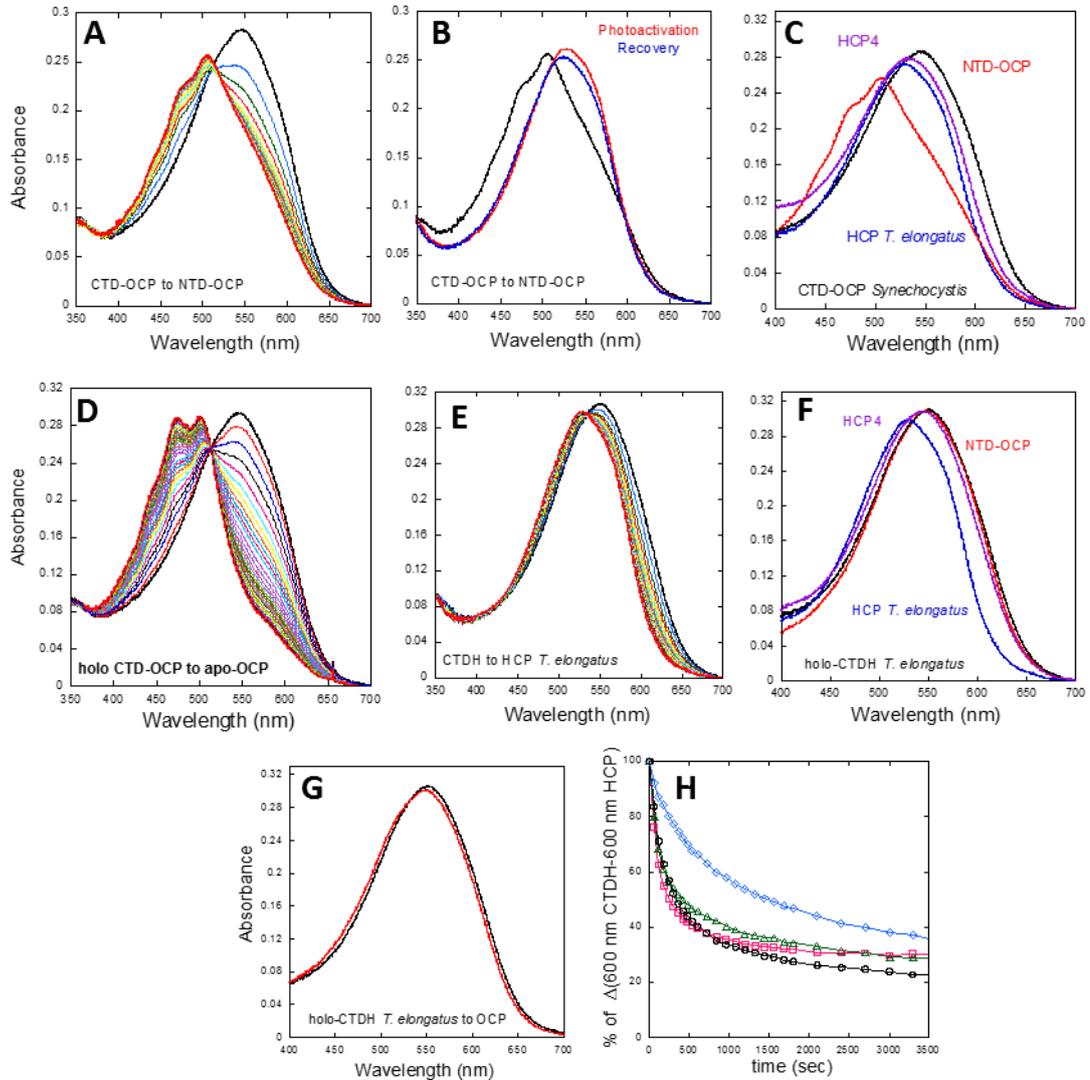
Recently it was shown that when holo CTD-OCP is incubated in the presence of apo-NTD-OCP, CTD can partially give the carotenoid to the NTD (Moldenhauer et al., 2017). Moreover, a partial apparition of OCP<sup>o</sup> was observed. Holo-CTD-OCP was also able to partially give the carotenoid to an apo-OCP (Moldenhauer et al., 2017; Maksimov et al., 2017b). In this work, we further study the interaction between CTD-OCP and OCP and NTD-OCP and tested whether the CTDHs are also capable to give the carotenoid to apo-HCPs and form new OCP<sup>o</sup>.

The genes coding for the HCP4 from *Anabaena*, the HCP from *T. elongatus* and the NTD-OCP from *Synechocystis* were expressed in a canthaxanthin containing *E. coli* strain as described in (Lopez-Igual et al., 2016). The three proteins presented broad absorbance spectra without vibronic structure and a maximum absorbance at 526-527 nm (supplementary Figure S5). These genes were also expressed in a carotenoid-lacking *E. coli* strain to obtain apo-proteins.

We studied the interaction between the holo-CTDHs and the apo-HCPs and *Synechocystis* apo-OCP by following changes in absorbance spectra during incubation of apo-HCPs (NTD-OCP and OCP) and holo-CTDHs (CTD-OCP) dimers. A ratio of 2.5 HCPs to 1 CTDH dimer (mol:mol) was used and the proteins were incubated for 1 hour at 23 °C in darkness. The results are shown in Figure 7, Figure 8 and Table I.

When holo-CTD-OCP interacted with apo-OCP, large changes were observed in the absorbance spectrum during dark incubation (Figure 7D). After 1 hour, the spectrum presented the typical vibronic structure of OCP<sup>o</sup> and the spectral fit indicated that 90% of the carotenoid was transferred from holo-CTD-OCP to apo-OCP (Table I). The new OCP<sup>o</sup> was photoactive and completely converted to OCP<sup>R</sup> upon 5 min illumination with strong light (supplementary Figure S6). The photoactivation was completely reversible in darkness (supplementary Figure S6). The photoactivated OCP was able to significantly quench PBS fluorescence (Figure 9). Interestingly, when apo-OCP and holo-CTD-OCP dimers interacted in a 1:0.5 ratio, almost all the

carotenoids were transferred to apo-OCP however less than half the apo-OCP was converted to OCP<sup>0</sup> confirming that there is one carotenoid per dimer (supplementary Figure S6).



**Figure 7 | Changes in absorbance spectra during carotenoid transfer.** From holo-CTD-OCP (A to D) and *T. elongatus* holo-CTDH (E to G) to apo-HCPs (or apo-NTD) and apo-OCP and (H) kinetics of carotenoid transfer to apo-HCPs. A) holo-CTD to apo-NTD. Black spectrum: Time 0, 100% carotenoid in CTD. Red spectrum: after one hour incubation with apo-NTD. Intermediary spectra are also shown. B) Photoactivation and Recovery of the species formed in A. C) Last spectra after a one-hour incubation of CTD-OCP in the presence of NTD (red), *Anabaena* HCP (violet) and *T. elongatus* HCP (blue). Spectrum of CTDH at time 0 is also shown (black). D) holo-CTD to apo-OCP. Black spectrum: Time 0, 100% carotenoid in CTD. Red spectrum: after one hour incubation with apo-OCP. Intermediary spectra are also shown. E) holo-*T. elongatus* CTDH to *T. elongatus* apo-HCP. Black spectrum: Time 0, 100% carotenoid in CTD. Red spectrum: after one hour incubation with *T. elongatus* apo-HCP. Intermediary spectra are also shown. F) Last spectra after one hour incubation of *T. elongatus* CTDH with NTD (red), *Anabaena* HCP (violet) and *T. elongatus* HCP (blue). Spectrum of CTDH at time 0 is also shown (black). G) holo *T. elongatus* CTDH

to apo-OCP. Time<sub>0</sub> (black) and Time<sub>60 min</sub> (red) are shown. H) Kinetics of carotenoid transfer from CTD-OCP and *T. elongatus* CTDH to different HCPs followed by measuring the decrease of absorbance at 600 nm: CTD-OCP to *Anabaena* HCP4 (blue diamonds), CTD-OCP to NTD-OCP (red squares), CTD-OCP to *T. elongatus* HCP (green triangles) and *T. elongatus* CTDH to *T. elongatus* HCP (black circles). Representative curves are shown. The measurements were repeated at least once.

The interaction between holo-CTD-OCP and apo-NTD-OCP gave a final spectrum with a maximum at 500 nm and a shoulder at 467 nm suggesting that an OCP<sup>O</sup>-like compound was also formed during the incubation (Figure 7A). The spectral fit indicated that the final mix was composed of 67 % OCP<sup>O</sup> and 33% CTD-OCP. When the sample was illuminated, all the OCP<sup>O</sup> was converted to holo-NTD-OCP and/or OCP<sup>R</sup> (Figure 7B). This conversion was irreversible (Figure 7B). The illuminated sample was able to quench PBS fluorescence (Figure 9B), however the quenching was lower than expected suggesting that although the carotenoid is in the NTD, the presence of apo-CTD hinders its interaction with PBS.

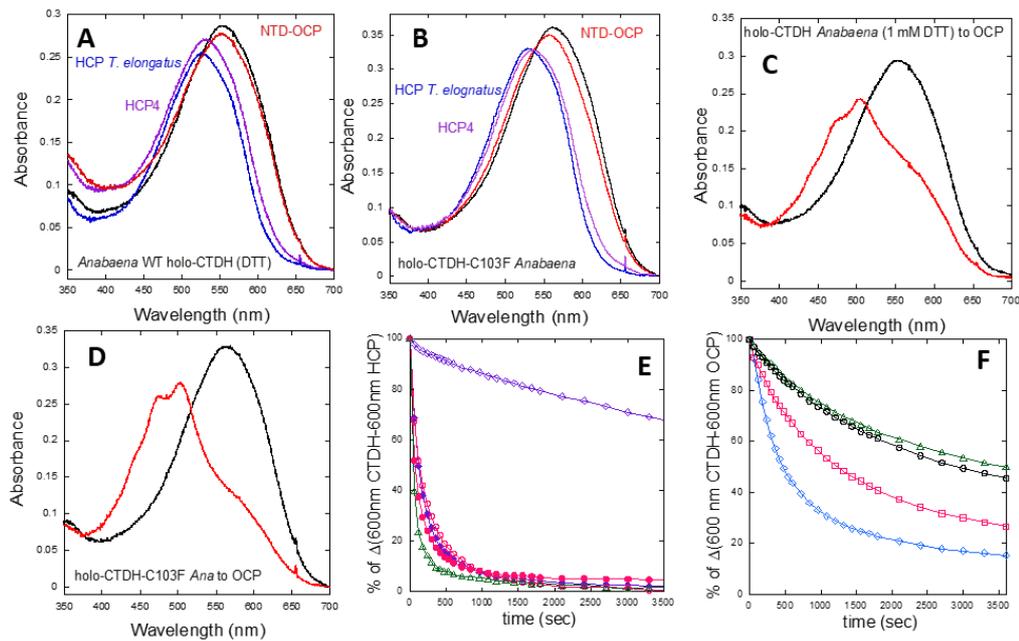
	NTD-OCP	<i>Anabaena</i> HCP4	<i>T.elongatus</i> HCP4/5	<i>Synechocystis</i> OCP
CTD-OCP	67.4 (0.3)	66.5 (3.9)	73.2 (0.4)	89 (3.1)
<i>T. elongatus</i> CTDH	4 (1.0)	22 (7.8)	77 (5.4)	0.03 (0.27)
<i>Anabaena</i> CTDH	0	0	0	13 (0.013)
<i>Anabaena</i> CTDH (DTT)	8 (0.0105)	87.2 (1.7)	91.2 (4.9)	52 (4.6)
<i>Anabaena</i> C103F CTDH	16 (1.9)	96.4 (8.5)	96 (2.3)	72 (2.3)
<i>Anabaena</i> C103F+4aa CTDH	0	29.6 (1.9)	94 (0.4)	48 (0.05)

**Table I: Transfer of carotenoid from holo-CTDH to apo-HCPs and OCP.** The percentage of carotenoid transferred after a one-hour incubation in darkness at 23°C is shown (ST value in brackets). The OCP:CTDH dimer and HCP:CTDH dimer ratios were equal to 2.5. The experiments were repeated at least once.

Addition of *T. elongatus* and *Anabaena* HCPs to holo-CTD-OCP also caused dramatic changes in the absorbance spectra: the final spectra looked more like the HPC spectra than the original CTD-OCP spectrum (Figures 7C). A spectral fit indicated that at the end of the incubation the amount of carotenoid transferred to the *T. elongatus* HCP and *Anabaena* HCP4 was respectively about 71% and 66% (Table I). The holo-HCPs formed induced a

large PBS fluorescence quenching (Figure 9). Thus, holo-CTD-OCP is able to give the carotenoid with a relative high efficiency to all NTD-like proteins. Illumination of the mix containing *T. elongatus* HCP and *Anabaena* HCP4 did not induce any difference in absorbance spectra. No evidence for the formation of an OCP<sup>0</sup>-like protein was observed.

We tested whether holo-CTDH proteins were able to give the carotenoid to the HCPs. The holo-CTD of *T. elongatus* transferred most of its carotenoid (83%) to the *T. elongatus* HCP resulting in holo-HCP formation. However, no traces of OCP<sup>0</sup> were observed (Figure 7E). The holo-HCP that was formed was able to significantly quench PBS fluorescence with the same efficiency as that of the holo-HCP directly isolated from *E. coli* (Figure 9). By contrast, the holo-CTDH of *T. elongatus* was not able to transfer its carotenoid to the *Anabaena* HCP4, nor to NTD-OCP or to apo-OCP (Figure 7 and Table I). Illumination of different samples did not influence the result. Similar results were obtained when the holo-CTDH of *T. elongatus* isolated from *Synechocystis* cells was used (supplementary Figure S7).



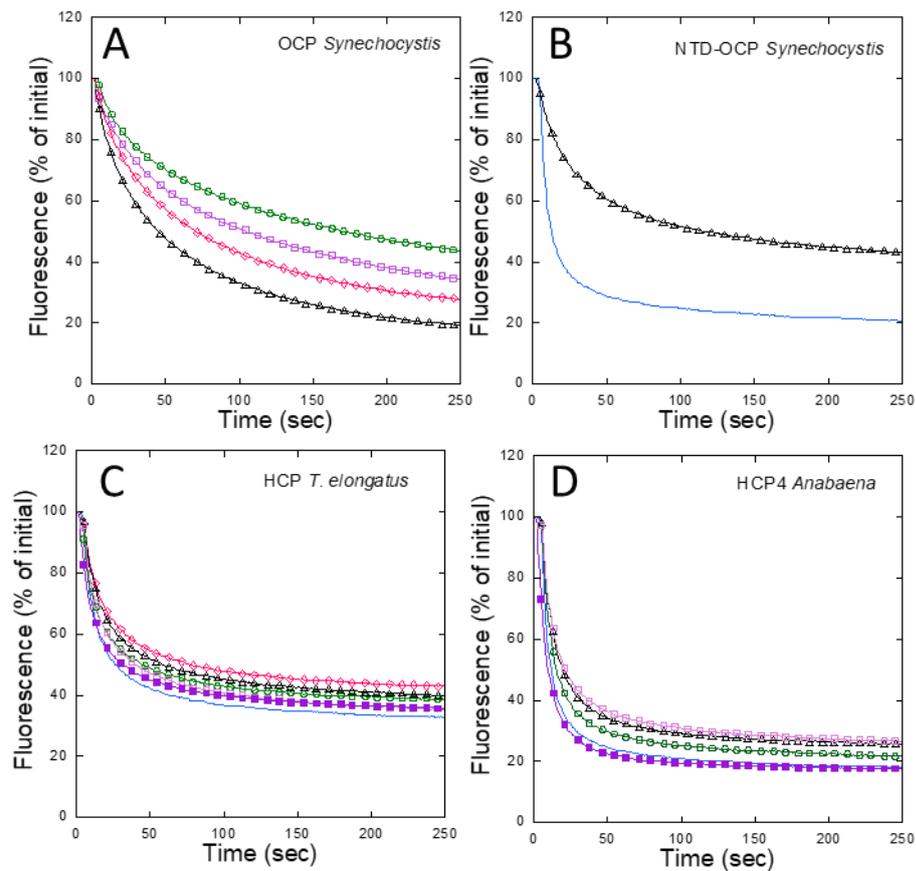
**Figure 8 | Changes in absorbance spectra induced by carotenoid transfer.** From *Anabaena* WT isolated in the presence of DTT (A and C) and *Anabaena* C103F CTDH (B and D) to apo- HCPs (A and B) and apo-OCP (C and D) and kinetics of carotenoid transfer from CTDHs to HCPs (E) and apo-OCP (F). (A and B) Black spectrum: Time 0, all carotenoid in CTDH. Last spectra after one hour incubation of WT CTDH (A) and mutant CTDH (B) with NTD-OCP (red), *Anabaena* HCP (violet) and *T. elongatus* HCP (blue). (C and D) WT CTDH and mutant CTDH to apo-OCP. Black spectrum: Time 0, all carotenoid in CTD. Red spectrum: after one hour incubation with apo-OCP. (E and F) Kinetics of carotenoid transfer from WT and mutant CTDHs to HCPs (E) and OCP (F) were followed by measuring the decrease of absorbance at 600 nm. (E) WT CTDH (DTT) to *Anabaena* HCP4 (green open triangles), C103F CTDH to *Anabaena* HCP4 (fuchsia open circles),

C103F CTDH to *T. elongatus* HCP (fuchsia closed circles), C103F+4aa CTDH to *Anabaena* HCP4 (violet open diamonds), C103F+4aa CTDH to *T. elongatus* HCP (violet closed diamonds), (F) Carotenoid transfer to apo-OCP: CTD-OCP (blue diamonds), *Anabaena* C103FCTDH (red squares), *Anabaena* C103F+4aa CTDH (black circles) and *Anabaena* WT CTDH (1mM DTT) (green triangles).

The isolated *Anabaena* CTDH was unable to give the carotenoid to the HCPs and to apo-OCP (Table I and supplementary Figure S8). In the previous section, we showed that *Anabaena* CTDH is present as a dimer even in absence of carotenoids. All clade 1 CTDHs have a Cys in position 103 that could be involved in an S-S bond between two CTDH monomers. We hypothesized that this bond prevents monomerization of the holo-CTDH dimer and transfer of the carotenoid. To test this hypothesis, the CTDH was incubated with dithiothreitol (DTT) (S-S bond reducer) and the experiment was repeated in its presence. Under these conditions, the *Anabaena* holo-CTDH gave around 50 % of the carotenoid to the *Anabaena* HCP4 suggesting that the presence of an S-S bond inhibited CTDH monomerization and transfer of the carotenoid in the absence of DTT.

To confirm this hypothesis, we also tested the *Anabaena* CTDH isolated under reducing conditions (1mM DTT). This *Anabaena* holo-CTDH preparation gave around 87-91% of its carotenoid to the *Anabaena* HCP4 and to the *T. elongatus* HCP (Figure 8 and Table I). No traces of OCP<sup>0</sup> were observed. The reduced *Anabaena* CTDH was also able to transfer its carotenoid to apo-OCP forming a photoactive OCP (52%) (Table I, Figure 8C and supplementary Figure S6). The holo-HCPs and photoactivated OCPs obtained after carotenoid transfer were able to quench PBS fluorescence with high efficiency (Figure 9). By contrast, *Anabaena* holo-CTDH was unable to transfer the carotenoid to NTD-OCP. The behavior of the *Anabaena* CTDH without a disulfide bond between the monomers of the dimer was further studied using the C103F-CTDH mutant. The C103F-CTDH protein was able to transfer most of its carotenoid to *Anabaena* HCP4 (96.4 %) and to *T. elongatus* HCP (96%) (Figure 8 and Table I). Moreover, it also gave the carotenoid to *Synechocystis* apo-OCP (72%) leading to the formation of photoactive OCP<sup>0</sup> (Figure 8, supplementary Figure S6 and Table I). The kinetics of carotenoid transfer was slower than that observed with CTD-OCP (Figure 8H). All holo-HCPs and photoactivated OCP were able to induce extensive PBS fluorescence quenching (Figure 9). The *Anabaena* C103F CTDH was also able to give its carotenoid to *Anabaena* apo-OCP but with a much lower efficiency (supplementary Figure S8D). *Synechocystis* NTD-OCP took very low quantities of the carotenoid from this CTDH (Table I).

Similar results were obtained using *Anabaena* NTD-OCP (supplementary Figure S8C). Thus, in the absence of a disulfide bond between the monomers, the *Anabaena* CTDH is less selective than the *T. elongatus* CTDH which gave the carotenoid only to its own HCP. Neither were able to give a carotenoid to the *Synechocystis* and *Anabaena* NTD-OCPs. These results suggest that the CTDHs transferred the carotenoid to apo-OCP via CTD. This hypothesis was also suggested by Maksimov et al (2017b) based on their results obtained with CTD-OCP.



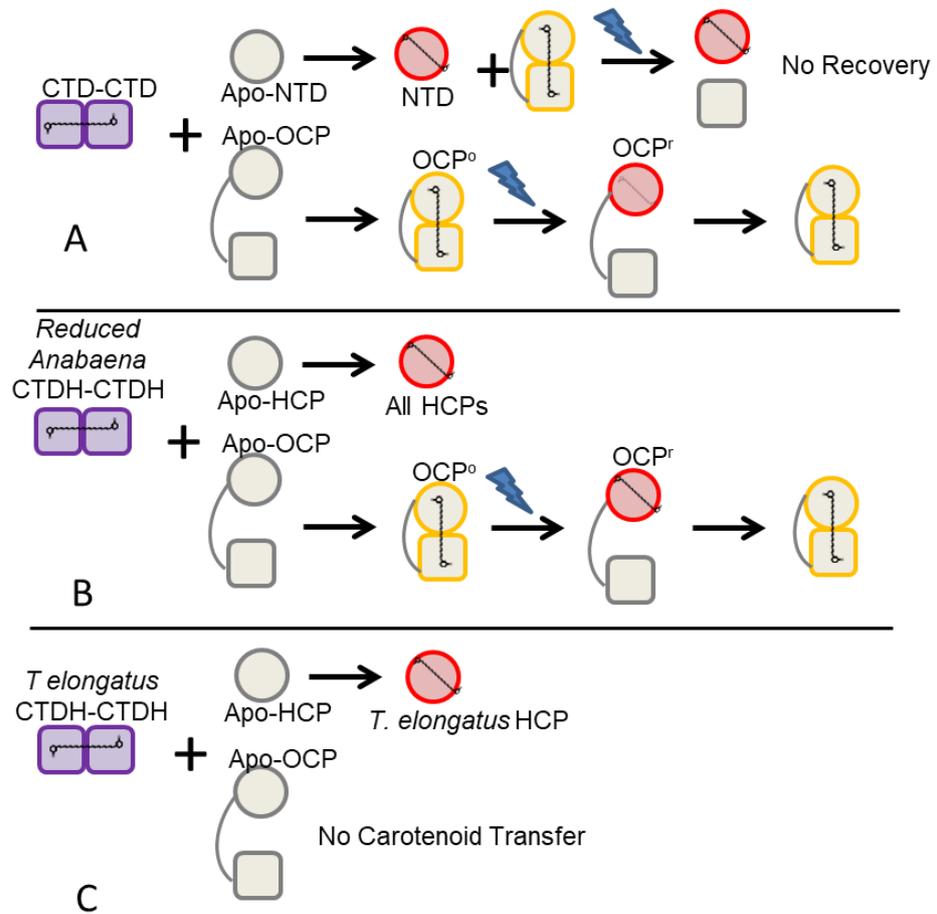
**Figure 9 | Quenching of phycobilisome fluorescence induced by holo-OCP and holo-HCPs formed by carotenoid transfer from the CTDH.** Fluorescence quenching induced by A) photoactivated *Synechocystis* OCPs formed by carotenoid transfer from CTD-OCP (black triangles), *Anabaena* C103F CTDH (fuchsia rhomboids), *Anabaena* C103F+4aa CTDH (violet squares) and *Anabaena* WT CTDH (green circles); B) Isolated NTD obtained in *E. coli* (bleu, without symbols) and NTD formed by carotenoid transfer from CTD *Synechocystis* (black triangles); C) isolated *T. elongatus* HCP obtained in *E. coli* (blue, without symbols), *T. elongatus* HCP formed by carotenoid transfer from *T. elongatus* CTDH (fuchsia diamonds), CTD-OCP (black triangles), WT *Anabaena* CTDH (green circles), C103F CTDH (violet close squares) and C103F+4aa CTDH (rose open squares); D) *Anabaena* HCP4 isolated from *E. coli* cells (blue without symbols), *Anabaena* HCP formed by carotenoid transfer from CTD-OCP (black triangles), WT *Anabaena* CTDH (green circles), C103F CTDH (violet close squares) and C103F+4aa CTDH (rose open squares). The decrease of fluorescence was followed in a PAM fluorimeter using a ratio HCP to PBS equal to 5 at 1.4M phosphate in darkness and a ratio OCP to PBS equal at 10 in 0.8M

phosphate under illumination with blue-green light ( $900 \mu\text{mol photons m}^{-2} \text{s}^{-1}$ ). The OCP were pre-illuminated with strong white light to photoactivate them.

Both WT and mutated CTDHs lack 4 amino acids after F (C)103 (compared to CTD-OCP) leading to a change in the tridimensional structure. We tested whether the addition of these 4 amino acids could enable CTDH interaction with NTD-OCP. We created a mutant of the *Anabaena* CTDH in which Cys 103 was replaced by a Phe and 4 amino acids (those existing in the CTD-OCP) were added after the Phe (C103F+4aa CTDH). A homology model of the C103F+4aa CTDH mutant was generated and it showed a structure very close to that of CTD-OCP, underlying a possible role of the lack of 4 amino acids in the folding of the CTDHs (supplementary Figure S9). The mutated CTDH also was able to bind canthaxanthin resulting in a maximum in its absorption spectrum at 570 nm (Figure 3). It formed dimers that were stabilized by the presence of the carotenoid (Figure 4 and 5). However, during native gel electrophoresis and SEC the carotenoid was easily lost by the dimer which partially monomerized (Figure 4 and 5). This mutated *Anabaena* CTDH transferred its carotenoid to *Anabaena* HCP4 and to *T. elongatus* HCP respectively with 30% and 94% efficiency but was unable to transfer the carotenoid to NTD-OCP (Table I and supplementary Figure S8). Carotenoid transfer to apo-OCP was 48% (Table I). Thus, the addition of amino acids did not improve the interaction with NTD. On the contrary, it decreased carotenoid transfer to HCP4 and apo-OCP compared to the CTDH-C103F mutant (Table I and Figure 8E and F).

Figure 10 summarizes the carotenoid transfer experiments from CTDHs to HCPs and OCP described above. CTDHs can easily monomerize (in the absence of C103 or in reducing conditions) and are able to transfer their carotenoid to various HCPs. The *T. elongatus* CTDH is more specific and can give the carotenoid only to the *T. elongatus* HCP. By contrast, the *Anabaena* CTDH is able to transfer the carotenoid not only to apo-HCPs but also to apo-OCPs. HCP characteristics also seemed to influence carotenoid transfer. Both CTD-OCP and *Anabaena* CTDH transferred their carotenoid to *Anabaena* HCP4 with slower kinetics than to *T. elongatus* HCP (Figures 7 and 8).

It is important to note that no heterodimeric complex was observed by size exclusion chromatography after the carotenoid transfer from *Anabaena* CTDH to *Anabaena* HCP4 (Supplementary figure S12). Instead, a mix between holo-HCP and holo-CTDH was obtained. This indicates that after carotenoid translocation the CTDH and HCP dissociate.

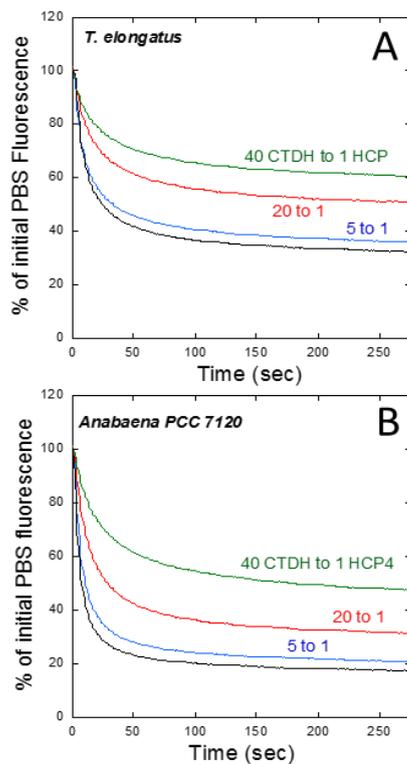


**Figure 10 | Summary of carotenoid transfer from CTD-OCP and CTDHs dimers to apo-NTD-OCP, apo-HCPs and apo-OCP.** (A) The CTD-OCP dimer is able to transfer its carotenoid with high efficiency to the apo-OCP forming a photoactive holo-OCP and to the apo-NTD forming a photoactive holo-OCP-like protein and holo-NTD. The CTD-OCP transfers the carotenoid also to the apo-HCPs, forming holo-HCPs (B) While oxidized *Anabaena* CTDH dimer is unable to transfer its carotenoid, the reduced protein transfers the carotenoid to the OCP and to all HCPs forming photoactive OCP and holo-HCPs respectively. (C) *T. elongatus* CTDH dimer can transfer its carotenoid only to the *T. elongatus* apo-HCP forming holo-HCP.

### *Interaction between holo-HCP and apo-CTDH*

When holo-HCP and NTD-OCP were incubated with the corresponding apo-CTDH and CTD-OCP no changes in the absorption spectra were induced. This indicated that holo-HCPs were not able to give the carotenoid to CTDHs and suggested that the HCPs have a higher affinity for the carotenoid than for CTDH. Interestingly they were also unable to give the carotenoid to apo-OCP. We then tested whether the presence of increasing concentrations of apo-CTDH hinders the interaction between HCPs and PBS.

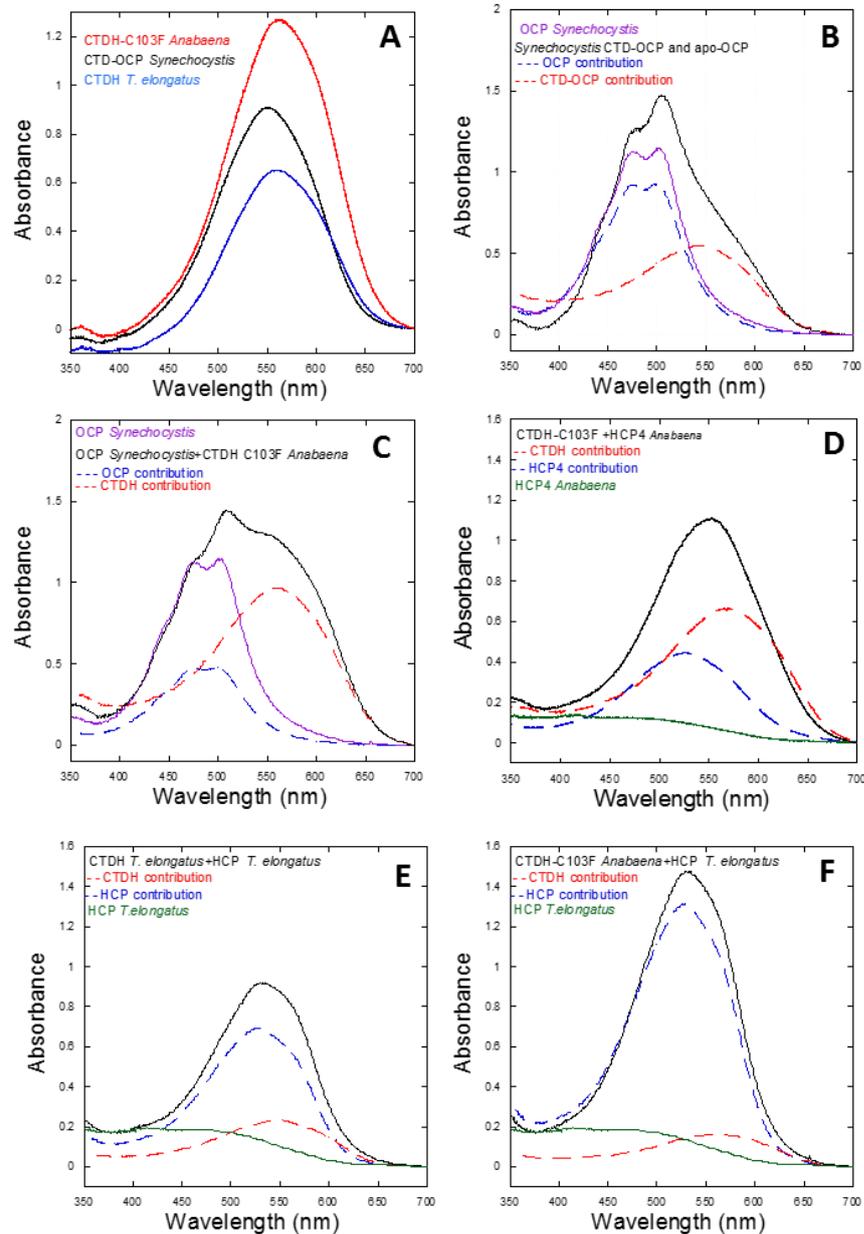
Figure 11 shows that when CTDH was added in small excess (5 CTDHs to 1 HCP) the fluorescence quenching induced was only slightly reduced. By contrast, a large excess of CTDH (20 or 40 CTDHs to 1 HCP) significantly decreased the amplitude of fluorescence quenching (Figure 11). This effect was specific since addition of 40 Bovine Serum Albumin (BSA) per 1 HCP had not influence on the amplitude of fluorescence quenching. Thus, although CTDH and HCP can interact, this interaction is weak and only a large excess of CTDH can inhibit fluorescence quenching.



**Figure 11: Quenching of phycobilisome fluorescence:** induced by *T. elongatus* (A) and *Anabaena* holo-HCPs (B) isolated from *E. coli* in the absence and presence of different concentrations of *T. elongatus* (A) and *Anabaena* (B) CTDHs. Kinetics of phycobilisome fluorescence quenching induced by HCP in the absence (black) and in the presence of CTDH. CTDH to HCP ratios: 5 to 1 (blue), 20 to 1 (red) and 40 to 1 (green). Fluorescence decrease was measured using a PAM fluorimeter and a ratio of HCP to PBS of 5 to 1 in 1.4M phosphate in darkness.

### ***Carotenoid transfer from membranes to CTDHs, HCPs and OCPs: Role of CTDHs***

Our results suggest that the role of CTDH could be to provide carotenoid molecules to OCPs and HCPs as previously proposed by (Moldenhauer et al., 2017). However, approximately half all known OCP-containing cyanobacteria strains lack CTDHs. In addition, when *ocp* genes are expressed in *E. coli* cells, they are able to bind carotenoids in the absence of CTDH or another soluble carotenoid protein (Bourcier de Carbon et al., 2015).



**Figure 12: Carotenoid transfer from membranes to apo-proteins.** Spectra of holo-proteins formed after a one-hour incubation of apo-proteins (12 $\mu$ M) with canthaxanthin containing *E. coli* membranes at 33°C. (A) apo-CTDH incubation with membranes: *Anabaena* C103F CTDH (red), *Synechocystis* CTD-OCP (black) and *T. elongatus* CTDH. (B) *Synechocystis* holo-OCP formation after a one-hour incubation of apo-OCP with membranes (violet spectrum) and of a mix of apo-OCP and apo-CTD-OCP with membranes (black spectrum) (C) *Synechocystis* holo-OCP formation after a one-hour incubation of apo-OCP with membranes (violet spectrum) and of a mix of apo-OCP and *Anabaena* apo-CTDH-C103F (black spectrum). In B and C the deconvolution of the black spectra are shown: OCP contribution (dashed blue spectrum) and CTD contribution (dashed red spectrum). (D-F) holo-HCPs formation after a one-hour incubation of apo-HCPs and membranes (green) and, of a mix of apo-HCPs and apo-CTDHs with membranes (final spectra, black) : (D) *Anabaena* C103FCTDH and *Anabaena* HCP4, (E) *T. elongatus* CTDH and *T. elongatus* HCP, (F) *Anabaena* C103FCTDH and *T. elongatus* HCP. HCP contribution (dashed blue spectrum)

and CTDH contribution (dashed red spectrum) calculated from deconvolution of black spectra are shown.

We decided to test whether OCPs, HCPs and CTDHs are able to directly take the canthaxanthin present in isolated *E. coli* membranes. Using apo-*Synechocystis* OCP, we first tested different temperatures, protein concentrations and carotenoid to protein ratios to fix the experimental conditions (see supplementary Figure S10). The best temperature for carotenoid transfer from membranes to different proteins was found to be 33°C. The results also showed that the carotenoid to protein ratio influences the percentage of carotenoid translocation: the higher the ratio, the larger the carotenoid translocation observed (supplementary Figure 10). All the results shown in Figure 12 and Table II were obtained using a carotenoid to protein ratio of 4:1, the protein concentration was 12 µM and the temperature 33°C. When CTDH was present in addition to HCP or OCP, its concentration was also 12 µM (dimer). For these experiments we used the *T. elongatus* CTDH and the *Anabaena* C103F CTDH mutant which can easily monomerize and transfer the carotenoid to HCPs and OCP.

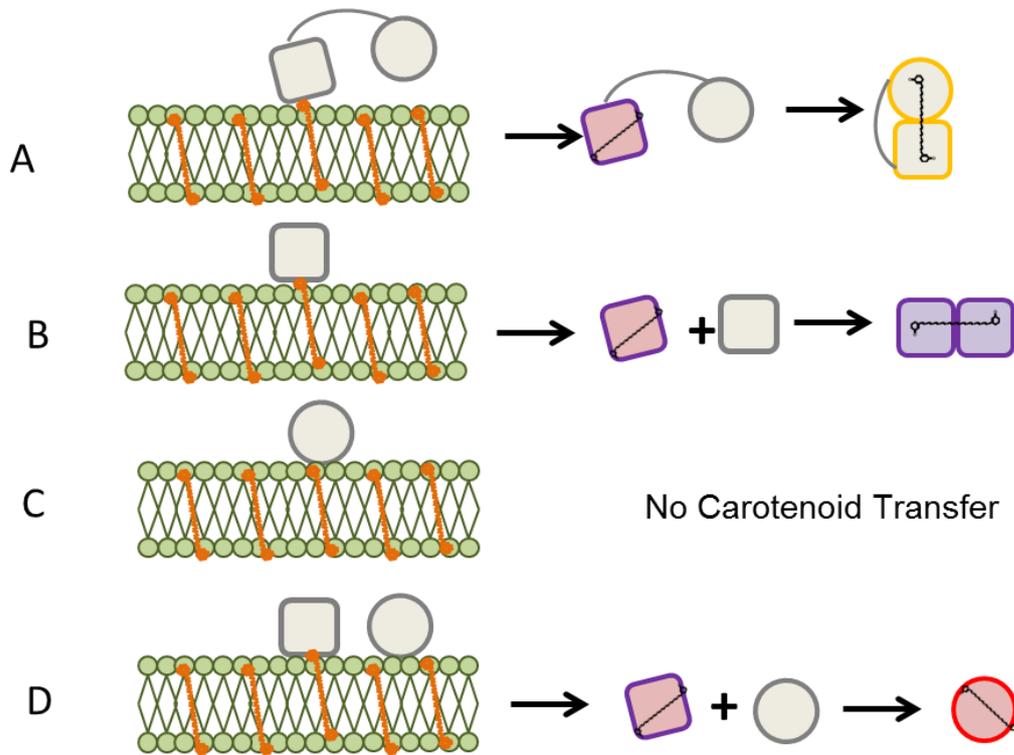
The apo-OCP and the apo-CTDHs “took” the carotenoid directly from the membrane with high efficiency: 77 % of *Synechocystis* apo-OCP, 65% of *Synechocystis* apo-CTD-OCP, 82% of *Anabaena* apo-C103FCTDH and 48% of *T. elongatus* CTDH were converted to holo-proteins after a one-hour incubation (Table II and Figure 12). By contrast, the HCPs were unable to take the carotenoid directly from the membrane (Table II and Figure 12). When apo-CTDHs were present during the incubation in addition to HCPs and membranes, large quantities of apo-HCPs were converted into holo-HCPs. In the presence of *Anabaena* C103F CTDH, 100% of *T. elongatus* apo-HCP was converted to holo-HCP while in the presence of *T. elongatus* CTDH only 55% was converted to holo HCP. One possible explanation could be that *Anabaena* CTDH was more efficient in taking the carotenoid from the membrane than the *T. elongatus* CTDH (Table II). 40% of *Anabaena* apo-HCP4 was converted to holo-HCP in the presence of *Anabaena* C103F CTDH. These experiments clearly show that CTDHs can take the carotenoid from the membrane and then transfer it to HCPs. The presence of CTDHs is essential for carotenoid insertion into HCPs.

Protein	% of apo-protein converted to holo-protein
<i>Synechocystis</i> OCP	77%(6)
<i>Anabaena</i> OCP	26%
<i>Synechocystis</i> CTD-OCP and OCP	65%(0.1) OCP
<i>Synechocystis</i> OCP+ <i>Anabaena</i> CTDH-C103F	39%(3.16) OCP
<i>Anabaena</i> C103FCTDH	82%(9.7)
<i>Anabaena</i> CTDH-C103F+HCP4	40%(1.2) (HCP4)
<i>Anabaena</i> C103FCTDH + <i>T.elongatus</i> HCP	100%(9.7) ( <i>T. elongatus</i> HCP)
CTDH <i>T.elongatus</i>	48%(2.3)
<i>T.elongatus</i> CTDH+ <i>T.elongatus</i> HCP	55%(2.3) ( <i>T.elongatus</i> HCP)
CTD-OCP	60%(3.8)

**Table II: Carotenoid Translocation from membranes to apo-OCPs, CTDHs and HCPs.** The percentage of holo-protein formed after a one-hour incubation of apo-proteins and membranes in darkness at 33°C is shown (ST value in brackets). The ratios OCP to CTDH dimer and HCP to CTDH dimer were equal to 1. The experiments were repeated at least once.

By contrast, the presence of *Synechocystis* CTD-OCP and CTDH had an opposite effect on carotenoid uptake by the apo-OCP (Figure 12 and Table II). Only *Anabaena* C103FCTDH was tested since *T. elongatus* CTDH was not able to give its carotenoid to OCPs. When *Anabaena* apo-OCP was incubated with apo-CTDH and membranes, no carotenoid was integrated into OCP (Supplementary Figure S8F). When *Synechocystis* apo-OCP was used in the experiment, 65% of the apo-OCP was converted to holo-OCP in the presence of CTD-OCP and only 39% in the presence of the *Anabaena* CTDH. This was lower than holo-OCP formation in the absence of CTDH (77%).

Figure 13 shows a model of carotenoid translocation from membranes to apo-OCPs, CTDHs and HCPs.



**Figure 13: Model of carotenoid transfer from membranes to apo-OCP, CTDH and HCP.** Apo-OCPs (A) and CTDHs (B) can take the carotenoid directly from membranes and to form holo-photoactive OCPs and holo-CTDH dimers, respectively. It is the CTD-OCP that takes the carotenoid. First, the carotenoid is in an intermediary position completely buried in the CTD (or the CTDH) and then it is shared with the NTD to form OCP<sup>0</sup> (A) or with a CTDH apo-monomer to form a CTDH dimer (B). In contrast, HCPs are unable to take the carotenoid from membranes (C). The presence of CTDH is essential to transfer the carotenoid from the membrane to the HCPs (D).

## Discussion

---

### *The CTDHs are carotenoid binding homodimers*

Conservation across CTDHs of most of the amino acids involved in the interaction between the carotenoid and OCP suggests that CTDHs could be carotenoid proteins (Bao et al, 2017 and this work). CTDH capacity to bind canthaxanthin was shown to occur not only when CTDHs are synthesized in canthaxanthin-producing *E. coli* cells but also when they are produced in *Synechocystis* cells. Both *T. elongatus* and *Anabaena* CTDHs bound

canthaxanthin and formed a homodimer containing one carotenoid molecule. These two CTDHs belong to different subgroups (or clades) exhibiting different characteristics. Our results clearly show that the absence of carotenoid destabilizes the *T. elongatus* CTDH dimer but not the *Anabaena* CTDH dimer. We observed that the latter formed a disulfide bond between the two monomers of the dimer. Thus, Cys 103 involved in the S-S bond in *Anabaena* CTDH must be located on the interaction face between the monomers. These results confirm the model proposed by Moldenhauer et al., 2017 about CTD-CTD interaction in CTD-OCP dimers. In this model, the carotenoid is shared by the two CTDs forming the dimer and the interaction involves the same CTD face that interacts with NTD in OCP<sup>o</sup>. Cys103 (Phe103) is present on this face.

In the *Anabaena* CTDH dimer, the S-S bond stabilized the dimer even in the absence of the carotenoid and also inhibited the transfer of the carotenoid from CTDH to the HCPs. Reduction of the disulfide bond by DTT allowed monomerization of the dimer and shuttling of the carotenoid to HCPs. Furthermore, replacement of Cys103 by Phe, which did not destabilize the carotenoid *Anabaena* CTDH dimer, accelerated the transfer of the carotenoid to HCPs and OCP. Inside cells reducing conditions are prevalent, thus this disulfide bond may be absent in WT *Anabaena* CTDH allowing the opening of the dimer and carotenoid transfer. This Cys is absent in the CTDH of *Gloeobacter* strains (the most ancient cyanobacteria strains) and in all the CTDHs of Clade 1. The Cys acquired in the course of evolution would have brought to some cyanobacteria cells the possibility of regulating, via redox changes, the opening of the dimer and the transfer of the carotenoid. Under this hypothesis, under normal conditions the disulfide bond is oxidized and only reduced under stress conditions, leading to an increase in HCP production.

According to the structural model, the addition of 4 amino acids after Phe103 in *Anabaena* CTDH changed the secondary structure of the interaction face. This change significantly destabilized the carotenoid dimer: large quantities of carotenoid-free monomers were detected in the native gel and by SEC. By contrast, the carotenoid CTD-OCP dimer, which based on the structural model has a secondary structure of the interaction face similar to that of the C103F+4aa CTDH mutant, was stable under the same experimental conditions. These results suggest that the interactions between the monomers are slightly different in CTDH and CTD-OCP dimers.

The isolated *Synechocystis* CTD-OCP overexpressed in WT *Synechocystis* did not bind any carotenoid (Sutter et al., 2013) while CTD-OCPs synthesized in *E. coli* formed carotenoid homodimers when expressed in the presence of a large canthaxanthin concentration (Moldenhauer et al., 2017, Lechno-Yossef et al., 2017 and this work). In contrast, we have shown here that both *T. elongatus* and *Anabaena* CTDHs are able to bind canthaxanthin when they are expressed in *Synechocystis* cells. Moreover, the *T. elongatus* CTDH isolated from *Synechocystis* and *E. coli* cells presented the same capacity to transfer the carotenoid to the *T. elongatus* HCP. Our results clearly demonstrate that the CTDHs are cyanobacterial carotenoid proteins.

The next step in our research will be to isolate the CTDHs from *Anabaena* and *T. elongatus* to identify the carotenoids that bind to them in the original strains. *Anabaena* cells naturally contain large quantities of canthaxanthin which can stabilize CTDH dimers by binding to them. It has already been shown that *Anabaena* HCP is mostly bound to canthaxanthin (45-63%) and echinenone (20-27%) in *Anabaena* cells. By contrast, *T. elongatus* lacks  $\beta$ -carotene ketolases and does not contain canthaxanthin nor echinenone (Liang et al., 2006; Iwai et al., 2008). In *T. elongatus* the HCP and CTDH proteins bind zeaxanthin, caloxanthin and nostoxanthin, which all contain hydroxyl groups on both rings (Iwai et al., 2008). When *T. elongatus* CTDH was expressed in WT *Synechocystis* (instead of  $\Delta$ CrtR *Synechocystis*), it bound 49% canthaxanthin, 40% hydroxyechinenone and 10% myxoxanthophyll. This suggests that the *T. elongatus* CTDH can bind other carotenoids.

### ***Carotenoid translocation during CTDH interaction with HCPs***

Recently, *Synechocystis* and *Fremyella diplosiphon* apo- and holo-NTD-OCP and CTD-OCP proteins were obtained by expression of synthetic *ntd* and *ctd* genes in *E. coli* cells, in the absence and presence of carotenoids (Lopez-Igual et al., 2016; Moldenhauer et al., 2017, Lechno-Yossef et al., 2017). Both NTD and CTD-OCP were able to bind canthaxanthin in *E. coli* cells (Lopez-Igual et al., 2016; Moldenhauer et al., 2017, Lechno-Yossef et al., 2017). In the case of CTD-OCP, binding of the carotenoid stabilized the homodimer (Moldenhauer et al., 2017). When holo-CTD-OCP was incubated with apo-NTD-OCP, it partially gave its carotenoid to the apo-NTD-OCP forming an OCP<sup>o</sup>-like protein (Moldenhauer et al., 2017). Moreover, an NTD-CTD heterodimer was formed when the domains were expressed as separate polypeptides in *E. coli* (Lechno-Yossef et al., 2017). It has recently been

reported that *Synechocystis* CTDH can also efficiently transfer its carotenoid to apo-OCP, resulting in the formation of photoactive OCP (Maksimov et al., 2017b). In this work, we confirmed that the holo-CTD-OCP dimer is able to give its carotenoid to apo-NTD-OCP and to apo-OCP. Under our experimental conditions, it gave 90% of the carotenoid to OCP and 50% to NTD-OCP respectively resulting in the formation of OCP<sup>o</sup> and OCP<sup>o</sup>-like proteins. Both proteins were photoactive, indicating that the carotenoid was in an identical position to that in WT OCP where it hydrogen bonds with Y201 and W288. The photoactivated proteins were able to interact efficiently with PBS and to induce fluorescence quenching. OCP<sup>r</sup> converted to the inactive orange form in darkness. By contrast, the OCP<sup>r</sup>-like protein remained red even after a long period of darkness indicating that once in NTD, the carotenoid was not able to translocate into CTD to form an OCP<sup>o</sup>-like protein again (Moldenhauer et al., 2017, Lechno-Yossef et al., 2017 and this work). Also, when holo-NTD-OCP was incubated with apo-CTD-OCP (even in large excess) the carotenoid remained in NTD. These results indicate that the presence of the carotenoid in CTDH is essential to induce a “correct” interaction with HCP in order to create the hydrophobic “tunnel” that allows translocation of the carotenoid from one protein to the other. When the carotenoid is in the NTD, the presence of the OCP loop linker seems to be essential for OCP<sup>r</sup> to OCP<sup>o</sup> conversion (Figure 10).

The CTDHs were also able to give their carotenoids to their respective HCPs but OCP<sup>o</sup>-like proteins were not formed. The carotenoid remained completely buried inside the HCPs. Thus, while CTD-OCP is able to share the carotenoid with the NTD-OCP retaining (or recapturing) half of the carotenoid in it, CTDH is not able to do it. Furthermore, the CTDHs presented different characteristics: *T. elongatus* CTDH was selective and gave its carotenoid only to *T. elongatus* HCP, while the *Anabaena* WT and mutated CTDHs transferred their carotenoids to both HCPs. This could be related to the fact that in *Anabaena* cells there are 4 HCPs belonging to different clades: the CTDH must be less selective to give the carotenoid to all of them. It is tempting to propose that the differences between CTDHs and CTD-OCP are at the origin of the different selectivity of CTDH proteins and their inability to form OCP<sup>o</sup>-like proteins. By contrast, the tertiary structure of HCPs and NTD-OCP and the position of the carotenoid in these proteins are similar (Melnicki et al., 2016). Furthermore, the *Anabaena* HCP4 and *T. elongatus* HCP contain an N-terminus prolongation similar to the N-terminal arm of NTD-OCP which could play a role in stabilizing the CTDH-HCP heterodimer. However, some of our results suggest that HCPs also have an

influence on the formation and stabilization of OCP<sup>o</sup>-like proteins. The *Synechocystis* CTD-OCP was able to give the carotenoid to the *T. elongatus* HCP and *Anabaena* HCP4/HCP5. However, in this case no OCP<sup>o</sup>-like protein was formed. In addition, none of the WT and mutated CTDHs tested were able to give their carotenoid to the *Synechocystis* and *Anabaena* apo-NTD-OCPs. *T. elongatus* CTDH gave its carotenoid only to *T. elongatus* HCP. Finally, *T. elongatus* HCP took the carotenoid from CTDHs with a higher efficiency than *Anabaena* HCP4. These results suggested that although the tertiary and secondary structures of HCPs and NTD-OCP are similar (but not identical), small differences in the structure and/or sequences have a considerable influence on their interactions with CTDH.

Although the *Anabaena* CTDH was not able to transfer the carotenoid to NTD-OCP, it readily gave the carotenoid to the *Synechocystis* and *Anabaena* apo-OCP forming photoactivable OCP<sup>o</sup>. This result suggests that the CTDH can form, as an intermediary step, a transient complex with CTD-OCP to give it its carotenoid. Once the carotenoid is in the CTD, the complex is disassembled and OCP orange is formed. Surprisingly, *Synechocystis* apo-OCP was able to take more carotenoid than *Anabaena* apo-OCP indicating that OCP characteristics also play a role in the carotenoid transfer mechanism. *T. elongatus* CTDH does not have the ability to transfer its carotenoid to an OCP. Most probably this ability was lost because OCP is absent in *T. elongatus* cells.

Our results show that CTDHs are able to interact with their respective HCPs and transfer their carotenoids to them. The next important step will be to understand the differences between CTD and NTD-OCP interactions, which form photoactive OCP<sup>o</sup>-like proteins, and CTDH and HCPs interactions, which do not. One area to investigate is the role of the N-terminal arm in the stabilization of OCP<sup>o</sup>. It is also important to understand why once in NTD (or in HCP) the carotenoid cannot transfer back to CTD (or CTDH). Our results suggest that the presence of the OCP linker is a requirement for carotenoid translocation from NTD to CTD during NTD-CTD interaction, but this needs to be further investigated.

### ***CTDHs are essential for carotenoid translocation from the membranes to HCPs***

Moldenhauer and coworkers were the first to show that the CTD-OCP dimer is able to give its carotenoid to NTD-OCP and they proposed that the

role of CTDHs in cyanobacteria cells could be to provide carotenoid molecules to OCPs and HCPs (Moldenhauer et al., 2017). Because approximately half of OCP-containing cyanobacteria strains lack CTDH proteins, and OCPs are able to bind carotenoids in the absence of CTDH or another soluble carotenoid protein when they are synthesized in *E. coli* cells (Bourcier de Carbon et al., 2015), another mechanism must also be involved. This work clearly demonstrated that apo-CTDHs and apo-OCPs are able to take the carotenoid directly and with high efficiency from membranes without the help of CTDHs. Moreover, the presence of CTD-OCP slightly decreased carotenoid translocation from membranes to OCPs. In reducing conditions, *Anabaena* WT CTDH and C103FCTDH were able to partially give the carotenoid to the OCPs in the absence of membranes. However, when carotenoid-containing membranes were incubated with apo-OCPs and apo-CTDH, the presence of CTDH had a negative effect on translocation: *Synechocystis* apo-OCP took less carotenoid and *Anabaena* apo-OCP did not take any in the presence of *Anabaena* CTDH. Apparently, the OCP takes the carotenoid more efficiently from membranes than from CTDH. At least this is true when *E. coli* membranes and large amounts of canthaxanthin are used. Results might be different in cyanobacteria cells with low concentrations of canthaxanthin and Hydroxy-echinenone. Thus, we cannot completely discard that in *Anabaena* cells, under specific conditions requiring large quantities of OCP, CTDH could help accelerate the formation of holo-OCPs.

HCPs did not take carotenoids directly from membranes in *in vitro* experiments. This result could explain why HCP preparations obtained in *E. coli* always contained a high percentage of apo-proteins (Lopez-Igual et al., 2016, this work). By contrast, in the presence of CTDHs, large quantities of holo-HCPs were formed. *T. elongatus* HCP took the carotenoid more efficiently from CTDHs than *Anabaena* HCP4, and this was also observed in the absence of membranes (this work). Although CTDHs are not needed to form holo-OCPs, our results clearly demonstrate that their presence is essential to translocate the carotenoid from the membranes to HCPs.

Until now, nothing was known about how the carotenoid is incorporated in OCP. Previous results obtained in *E. coli* cells suggested that the carotenoid must be added during the translation of the protein: faster translation increased the concentration of apo-protein while slower protein synthesis favored holo-OCP formation (Thurotte et al., 2015). The results of the present work suggest that carotenoid incorporation can also take place after apo-protein synthesis. More importantly, they indicate that it is the C-terminal domain (CTD) of OCP that takes the carotenoid from the membrane. The

CTD then interacts with NTD and transfers the carotenoid to it, prompting the OCP to adopt the “closed” orange structure. The interaction between the CTD and NTD must be very rapid to avoid carotenoid exposure to polar cell constituents. We propose that the carotenoid has a transitory position in which it is completely buried inside the CTD. We have isolated carotenoid-containing *T. elongatus* CTDH monomers presenting an absorption spectrum different to that of the CTDH dimer (supplementary Figure S3) suggesting that this is the case.

### *Can CTDH inhibit the HCP-PBS interaction?*

Here, we have shown that when *T. elongatus* HCP binds canthaxanthin, it is able to interact with PBSs and efficiently quench their fluorescence. We do not know whether this activity is possible in native *T. elongatus* cells, which only contain non-keto carotenoids. In *Anabaena* cells, HCP4 binds ketocarotenoids and is able to permanently bind PBS and quench its fluorescence. It was proposed that, by interacting with HCP4, CTDH can “shut off” or prevent HCP4-induced constitutive PBS fluorescence quenching (Lopez-Igual et al., 2016; Melnicki et al., 2016). Since *Anabaena* cells are not constitutively quenched, the proposition that quenching is prevented by interaction with CTDH is tempting. However, here we have shown that although the presence of CTDH decreased the amplitude of PBS fluorescence quenching, a large excess of CTDH was necessary to obtain significant inhibition. In addition, we observed that once holo-CTDHs gave their carotenoid to apo-HCPs and holo-HCPs were formed, they were able to induce very efficient PBS fluorescence quenching even in the presence of CTDH. Thus, the interaction between CTDH and HCP must be weak since it did not prevent HCP-PBS interaction and generated only unstable heterodimers. In *Anabaena*, both *hcp4* (*all4941*) and *ctdh* (*all4940*) genes are expressed in filaments and in heterocystis (Lopez-Igual et al., 2016). Under the experimental conditions used, their transcription levels were similarly low (Lopez-Igual et al., 2016). Transcriptomics and proteomics studies will be necessary to elucidate the conditions in which a large excess of CTDHs is present in the cells.

### *Origin of the contemporary HCPs and CTDHs*

It was proposed that OCP arose from the fusion of two independent carotenoid domains, homologs of NTD and CTD, which could be the

ancestors of the contemporary HCP4/HCP5s and CTDHs (Melnicki et al, 2016, Lechno-Yossef et al., 2017). This fusion created a photoactive protein with a more complex function than that of the separate domains. The CTDHs and most HCPs are good singlet oxygen quenchers (at least *in vitro*). It is likely that the first role of these proteins was related to protection against oxidative stress. Our results have shown that the presence of CTDH is essential to form holo-HCP4/HCP5. This could be the principal or secondary role of CTDHs before the creation of the OCP. Since HCP4 and HCP5 are able to interact with PBSs and quench their fluorescence, it was also proposed that CTDH regulates this activity by interacting with the HCP and hindering its interaction with PBS (Melnicki et al, 2016, Lopez-Igual et al, 2016). Contemporary CTDHs do not display an efficient interaction with HCP but nothing is known about their ancestors.

The question arises whether contemporary HCPs and CTDHs evolved from the HCP and CTDH ancestors or whether they appeared as the consequence of OCP duplication and subsequent domain separation into HCP and CTDH. In the phylogenetic basal genus *Gloeobacter*, OCP is already present and the OCP-related NPQ mechanism is functional (Bernát et al., 2012). Moreover, *Gloeobacter violaceus* contains two *ocp* genes. Other cyanobacteria strains also contain two or three *ocp* genes indicating a duplication of the *ocp* gene in the past (see supplemental file 1). Since excess of OCP could be detrimental under non-stress conditions, copies of the *ocp* gene decreased (or completely disappeared) with time. A separation into domains may have been the first step in the elimination of extra copies. HCP and CTDH disappeared in strains with smaller genomes and remained in others, especially in heterocyst-forming nitrogen-fixing species having the largest and most expanded genomes. The presence of HCPs and CTDHs which are likely to be good singlet oxygen quenchers could increase photoprotection under high light and oxidative conditions. Nevertheless, it is clear that CTDHs were primarily conserved to provide carotenoids to HCPs. Once the HCP domains separated from the CTDs, they were unable to take the carotenoid from the membranes. All the strains containing HCPs also contain one CTDH protein.

## Conclusions

CTDHs are homodimer proteins in which a carotenoid molecule is shared between the monomers. *In vitro*, they are very good quenchers of singlet oxygen, as good as OCP and some HCPs. They have the surprising ability to

transfer their carotenoid to apo-HCPs with high efficiency. This activity is essential for the synthesis of holo-HCPs. While CTDHs and OCPs share the ability to take carotenoid from membranes, HCPs are unable to do this, at least in *in vitro* experiments. CTDHs are also able to interact with holo-HCPs but this interaction is weak. They must be present in large excess in order to hinder HCP interaction with PBSs suggesting that this activity occurs only in very specific environmental conditions.

## ***Materials and Methods***

---

### ***Strains and culture conditions***

The freshwater cyanobacteria *Synechocystis* sp. PCC 6803 wild type and mutants were grown photoautotrophically in a modified BG11 medium (Herdman et al., 1973), containing double the amount of sodium nitrate. Cells were kept in a rotary shaker (120 rpm) at 30°C, illuminated by fluorescent white lamps giving a total intensity of about 30–40  $\mu\text{mol photons m}^{-2} \text{s}^{-1}$  under a CO<sub>2</sub> atmosphere. The cells were maintained in the logarithmic phase of growth and were collected at optical density at 800 nm = 0.6 to 0.8.

### ***Construction of OCP, NTD and HCP Plasmids for expression in E. coli cells***

The constructs for expression of OCP from *Synechocystis* and *Anabaena* sp. PCC 7120 have been described (Bourcier de Carbon et al., 2015). Construction of plasmids pCDF-RCP-Syn-1-165Ctag, pCDF-NTD-ana-1-165Ctag and pCDF-HCP4-4941Ctag were described previously in (Leverenz et al., 2015; Lopez-Igual et al., 2016).

#### ***pCDF-HCP-Te-1269Ctag***

The *tll1269* gene was amplified by PCR using genomic DNA of *T. elongatus* as template and oligonucleotides F<sub>tll1269</sub>EcoRI and R<sub>tll1269</sub>NotI (supplementary Table I). The PCR fragment was cloned between the EcoRI and NotI sites of pCDFDuet-1. Then the N-terminal extension containing the His-tag initially present in this plasmid was excised and nucleotides coding for 6 His were added at the 3'-end of the *tll1269* gene by site-directed mutagenesis using synthetic primers (supplementary Table I) to create pCDF-HCP-Te-1269Ctag.

## ***Construction of CTD Plasmids for expression in E. coli cells***

### ***pCDF-CTDHAna-4940Ctag and mutants***

The all4940 gene was amplified by PCR using genomic DNA of *Anabaena* sp PCC 7120 as template and Fall4940NcoI and Rall4940NotI as primers. The PCR fragment was cloned between the NcoI and NotI sites of the pCDFDuet-1 plasmid. Nucleotides encoding for 6 His tag were added in the 3'-end of the all4940 gene by site-directed mutagenesis using the synthetic primers to create the pCDF-CTDHAna-4940Ctag plasmid. The point mutation C103F and the replacement of C103 by 5 amino acids (FGAAV) were obtained by mutagenesis to create pCDF-CTDHAna-C103F and pCDF-CTDHAna-C5 plasmids respectively. Primer sequences are described in supplementary Table I.

### ***pCDF-CTDHTe-1268Ctag***

The tll1268 gene was amplified by PCR using genomic DNA of *T. elongatus* as template and Ftl1268NcoI and Rtl1268NotI as primers (supplementary Table II). The PCR fragment was cloned between the NcoI and NotI sites of the pCDFDuet-1 plasmid. Nucleotides encoding for 6 His were added in the 3'-end of the tll1268 gene by site-directed mutagenesis using synthetic primers (supplementary Table I) to create pCDF-CTDHTe-1268Ctag plasmid.

### ***pCDF-CTD6803-Ctag***

To obtain the CTD of *Synechocystis* OCP, the nucleotides encoding the N-terminal domain (8-169) of the ocp gene (slr1963) were deleted by site-directed mutagenesis using plasmid pCDF-OCPsynCtag (Bourcier de Carbon et al., 2015) as template and synthetic primers (supplementary Table I) to create the pCDF-CTD6803-Ctag plasmid.

All plasmids were checked by sequencing.

## ***Holoprotein production in canthaxanthin-producing E. coli***

BL21 (DE3) cells from Agilent Technologies were transformed simultaneously with three plasmids: (1) pACBETA, (2) pBAD-CrtW and (3) pCDF-CTDH (HCP, CTD-OCP, NTD-OCP, OCP) plasmids. The construction of pACBETA and pBAD-CrtW plasmid was described elsewhere (Bourcier de Carbon et al., 2015). The *crtBEIY* operon in pACBETA was constitutively expressed under the control of the *crtE*

promoter, whereas the *crtW* gene was under the control of the arabinose-inducible promoter araBAD and the *ocp*, *ctdh*, *hcp* genes were under the control of a T7 RNA polymerase promoter and their expression was enhanced by addition of IPTG. The expression of different genes and the isolation of apo- and holo CTDHs and HCPs were described in (Bourcier de Carbon et al., 2015; Lopez-Igual et al., 2016).

### ***Construction of Synechocystis Mutant Strains***

The construction of the *Synechocystis* strain overexpressing the His-Tagged C-terminal domain of OCP (CTD-OCP) was described previously (Sutter et al., 2013). To obtain a strain overexpressing *CTDH<sup>Te</sup>* and *CTDH<sup>Ana</sup>*, the *all4940* and *all1268* genes with a sequence encoding for 6-His in 3' (previously described) were amplified by PCR using the restriction sites-creating primers NdeI-For and HpaI-Rev. The digested PCR fragment was cloned into the pPSBA2 ampicillin-resistant vector (Lagarde et al., 2000) between the NdeI and HpaI restriction sites. A 1.3-kb kanamycin resistance gene was then inserted in the unique HpaI restriction site of pPSBA2. The resulting plasmids were used to transform  $\Delta$ CrtR *Synechocystis* PCC 6803 by double recombination. Segregation of the mutant was confirmed by PCR. Primer sequences can be found in supplementary Table I.

### ***Protein purification***

His-tagged CTD overexpressed in *Synechocystis* was isolated as described in (Wilson et al., 2008). OCP, NTD, HCP, CTD holo-proteins and apo-proteins expressed in *E. coli* were isolated as described in (Lopez-Igual et al., 2016). Briefly, cells were resuspended in lysis buffer (40 mM Tris pH 8/10% glycerol/300mM NaCl/1mM EDTA/1mM PMSF, 1mM caproic acid/1mM benzamidic acid/ 50  $\mu$ g mL<sup>-1</sup> DNase) then broken in dim light using a French press. The membranes were pelleted, and the supernatant was loaded on a nickel affinity column (Ni-Probond resin, Invitrogen). Proteins were eluted with 250 mM imidazole and then dialyzed against 40 mM Tris-HCl pH 8.

### ***Extraction and Analysis of Carotenoids***

The carotenoids were extracted with acetone from isolated CTDHs as described in (Sedoud et al., 2014). Liquid chromatography–UV–mass spectrometry analysis was conducted using a Quattro LC instrument

(Micromass), MassLynx software, an Alliance 2695 separation module (Waters), and a Waters 2487 dual UV light detector as described (Punginelli et al., 2009). Carotenoids were identified and quantified as described in (Punginelli et al., 2009)

### *Size Exclusion Chromatography*

Purified proteins were injected into a Superdex 75 HR 10/30 column (Pharmacia) equilibrated with 40mM Tris-HCl pH 8 and 150mM NaCl. The Akta Purifier FLPC system (GE Healthcare) was equipped with a UV detector set up at 280 nm and 540 nm for the present experiments and it was running at a flow rate of 0.4 mL/min. The gel filtration standards (BioRad) included thyroglobulin (670 kDa), bovine  $\gamma$ -globulin (158 kDa), chicken ovalbumin (44 kDa), equine myoglobin (17 kDa), and vitamin B12 (1.35 kDa). The calibration profile is shown in supplementary Figure S3.

### *Absorbance measurements and experiments of carotenoid transfer*

Absorbance spectra, kinetics of photoactivity (illumination with 5000  $\mu\text{mol photons m}^{-2} \text{s}^{-1}$  of white light) and dark recovery of the OCP were measured in a Specord S600 spectrophotometer (Analyticjena) at 23 °C. To study the interaction between the holo-CTDHs and apo-HCPs (or apo-OCPs), holo-CTDHs (2.5  $\mu\text{M}$ ) were incubated with 12.5  $\mu\text{M}$  apo-HCP (or apo-OCP) in 40mM Tris-HCl buffer pH 8 for one hour at 23 °C in darkness. Absorbance spectra were recorded for 1 hour. To study the carotenoid transfer from membranes to HCPs, CTDHs and OCPs, 12  $\mu\text{M}$  apo-protein were incubated with an *E. coli* canthaxanthin-containing membrane suspension (48  $\mu\text{M}$  canthaxanthin (measured by acetone extraction)) for 1 hour at 33°C in darkness. Holo-protein formation was measured by absorbance spectroscopy after precipitation of membranes. The percentage of holo-protein formed was determined by comparing the spectra of 100% holo-proteins (at 12  $\mu\text{M}$ ) to those of the supernatant. When two species were present in the supernatant the absorbance spectra were deconvoluted (see below).

### ***Determination of percentage of carotenoid transfer: Spectra deconvolution***

In order to determine the real contribution of each holo-protein to the final spectra of the protein-protein interaction experiments, a spectral deconvolution was performed using Excel to fit the data to the sum of the reference spectra of the holo proteins involved in the experiment. The spectra of 100% holo-proteins (at the same protein concentration) were used as reference spectra. (See example of deconvolutions in Supplementary Figure S11). The percentage of carotenoid transferred from CTDH to apo-HCPs and apo-OCPs was determined by comparison of carotenoid concentration in CTDHs at the beginning and the end of the experiment. The correlation between the decrease in CTDH carotenoid content and the increase in HCP and OCP carotenoid content was verified.

### ***Isolation of PBS and Fluorescence measurements***

The purification of PBS from *Synechocystis* PCC 6803 was performed as previously described in (Gwizdala et al., 2011). Fluorescence yield quenching was monitored using a pulse amplitude modulated fluorometer (101/102/103-PAM, Walz). Measurements were made in a 1 cm path length stirred cuvette. The kinetics of PBS quenching induced by holo-HCPs were measured in 1.4 M potassium phosphate buffer (pH=7.5) at 23°C in darkness. The PBS concentration used was 0.012  $\mu\text{M}$  and the ratio of carotenoid to PBS was 5:1 in the interaction mix. The PBS quenching of holo-OCP was measured in 0.8M potassium phosphate buffer (pH=7.5) at 23°C in strong blue-green light (900  $\mu\text{mol photons m}^{-2} \text{s}^{-1}$ ). The PBS concentration used was 0.012  $\mu\text{M}$  and the ratio of carotenoid to PBS was 10:1. OCP samples were pre-illuminated with 5000  $\mu\text{mol photons m}^{-2} \text{s}^{-1}$  of white light. Carotenoid concentration was calculated from the carotenoid absorbance spectra of holo-CTDH or holo-HCP at the beginning of the interaction ( $A_{1\%} 1\text{cm carotenoid} = 2158$ )

### ***$^1\text{O}_2$ detection by EPR spin trapping***

Electron Paramagnetic Resonance (EPR) spin trapping was applied for  $^1\text{O}_2$  detection using TEMPD-HCl (2,2,6,6-tetramethyl-4-piperidone) (100 mM). When this nitrene reacts with  $^1\text{O}_2$ , it is converted into the stable nitroxide radical, which is paramagnetic and detectable by EPR spectroscopy. The production of  $^1\text{O}_2$  was induced by illumination of the photosensitizer

methylene blue (10 $\mu$ M). The measurements were done in 40 mM Tris-HCl buffer (pH 8) in the absence or presence of different concentrations of purified CTDHs and *T. elongatus* HCP. The samples were illuminated for 3 min with white light (1000  $\mu$ mol photons m<sup>-2</sup> s<sup>-1</sup>). The EPR settings were as follows: hall center field = 3467.270 G, microwave frequency = 9.74 Ghz, power = 4.450 mV and number of scans = 12.

### ***Protein separation and Native gel-electrophoresis***

Proteins were analyzed by SDS–PAGE on 15% polyacrylamide/2M Urea in TRIS/MES system (Kashino et al., 2001). Non-denaturing gel electrophoresis was performed to determine the oligomeric state of CTDHs. Purified proteins were applied to 15% native polyacrylamide gels (pH 8). Electrophoresis was carried out using 25 mM Tris/192 mM glycine buffer pH 8.  $\alpha$ -Lactalbumin (14 kDa) and Carbonic anhydrase (29 kDa) were used as marker proteins.

### ***Homology Modeling***

Homology models for *T. elongatus* and *Anabaena* CTDHs were built using the default settings in Phyre2 (Kelley et al., 2015), using each CTDH sequences as the seed and 1M98 (*Arthrospira* OCP) as the template.

### ***Phylogenetic analysis***

To identify CTDH homologs, we used Blastp against the RefSeq protein database using as seed the *all4940* protein sequence and an E-value cut-off of 1x10<sup>-4</sup>. Blast hits were filtered by sequence length (<200 pb) in order to discard full OCP sequences. Sequences obtained were then aligned using default ClustalOmega (Sievers et al., 2011) and searched against cyanobacterial proteins in the Uniprot database of reference proteomes using Hmmer (Finn et al., 2011). All sequences obtained were used to generate a database of CTDH proteins. Similar approaches were done in order to build databases of HCP genes and OCP genes. CTD-OCP sequences were extracted from full OCP, removing the amino acids after position 190 in these sequences. The CTDH and CTD-OCP sequences were aligned using default ClustalOmega and columns with more than 70% gaps were removed using Gap Strip/Squeeze v2.1.0. Maximum likelihood analysis was done using

PhyML 3.1 (Guindon et al., 2010) in order to build a phylogenetic tree, with the following parameters: LG substitution model, NNI topology search and aBayes algorithm for branch support. The tree was visualized and analyzed with the online tool iTol (Letunic and Bork, 2016). Each subfamily was aligned with default ClustalOmega, a consensus sequence was calculated with the online AlignmentViewer Toolkit (Alva et al., 2016) and a sequence logos was built using Skylign (Wheeler et al., 2014).

### *Accession Numbers*

Sequence data from this article can be found in the GenBank/EMBL data libraries under the following accession numbers: slr1963, BAA18188; all4941, BAB76640; all4940, BAB76639; all3149, BAB74848; tll1268, BAC08820; tll 1269, BAC08821.

### *References*

---

Alva V, Nam SZ, Soding J, Lupas AN (2016) The MPI bioinformatics Toolkit as an integrative platform for advanced protein sequence and structure analysis. *Nucleic Acids Res* 44: W410-415

Bao H, Melnicki MR, Kerfeld CA (2017) Structure and functions of Orange Carotenoid Protein homologs in cyanobacteria. *Curr Opin Plant Biol* 37: 1-9

Boulay C, Wilson A, D'Haene S, Kirilovsky D (2010) Identification of a protein required for recovery of full antenna capacity in OCP-related photoprotective mechanism in cyanobacteria. *Proc Natl Acad Sci U S A* 107: 11620-11625

Bernát G, Schreiber U, Sendtko E, Stadnichuk IN, Rexroth S, Rögner M, Koenig F (2012) Unique properties vs. Common themes: The atypical cyanobacterium *Gloeobacter violaceus* PCC 7421 is capable of state transitions and blue-light-induced fluorescence quenching. *Plant Cell Physiol* 53:528-542

Bourcier de Carbon C, Thurotte A, Wilson A, Perreau F, Kirilovsky D (2015) Biosynthesis of soluble carotenoid holoproteins in *Escherichia coli*. *Sci Rep* 5: 9085

Cunningham FX, Jr., Pogson B, Sun Z, McDonald KA, DellaPenna D, Gantt E (1996) Functional analysis of the beta and epsilon lycopene cyclase enzymes of *Arabidopsis* reveals a mechanism for control of cyclic carotenoid formation. *Plant Cell* 8: 1613-1626

Finn RD, Clements J, Eddy SR (2011) HMMER web server: interactive sequence similarity searching. *Nucleic Acids Res* 39: W29-37

Fraser PD, Miura Y, Misawa N (1997) In vitro characterization of astaxanthin biosynthetic enzymes. *J Biol Chem* 272: 6128-6135

Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O (2010) New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst Biol* 59: 307-321

Gupta S, Guttman M, Leverenz RL, Zhumadilova K, Pawlowski EG, Petzold CJ, Lee KK, Ralston CY, Kerfeld CA (2015) Local and global structural drivers for the photoactivation of the orange carotenoid protein. *Proc Natl Acad Sci U S A* 112: E5567-5574

Gwizdala M, Wilson A, Kirilovsky D (2011) In vitro reconstitution of the cyanobacterial photoprotective mechanism mediated by the Orange Carotenoid Protein in *Synechocystis* PCC 6803. *Plant Cell* 23: 2631-2643

- Gwizdala M, Wilson A, Omairi-Nasser A, Kirilovsky D (2013) Characterization of the *Synechocystis* PCC 6803 Fluorescence Recovery Protein involved in photoprotection. *Biochim Biophys Acta* 1827: 348-354
- Harris D, Tal O, Jallet D, Wilson A, Kirilovsky D, Adir N (2016) Orange carotenoid protein burrows into the phycobilisome to provide photoprotection. *Proc Natl Acad Sci U S A* 113: E1655-1662
- Herdman M, Delaney SF, Carr NG (1973) A new medium for the isolation and growth of auxotrophic mutants of the blue-green alga *Anacystis nidulans*. *J. Gen. Microbiol.* 79: 233-237
- Iwai M, Maoka T, Ikeuchi M, Takaichi S (2008) 2,2'-beta-hydroxylase (CrtG) is involved in carotenogenesis of both nostoxanthin and 2-hydroxymyxol 2'-fucoside in *Thermosynechococcus elongatus* strain BP-1. *Plant Cell Physiol* 49: 1678-1687
- Jallet D, Thurotte A, Leverenz RL, Perreau F, Kerfeld CA, Kirilovsky D (2014) Specificity of the cyanobacterial orange carotenoid protein: influences of orange carotenoid protein and phycobilisome structures. *Plant Physiol* 164: 790-804
- Kashino Y, Koike H, Satoh K (2001) An improved sodium dodecyl sulfate-polyacrylamide gel electrophoresis system for the analysis of membrane protein complexes. *Electrophoresis* 22: 1004-1007
- Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJ (2015) The Phyre2 web portal for protein modeling, prediction and analysis. *Nat Protoc* 10: 845-858
- Kerfeld CA, Sawaya MR, Brahmamdam V, Cascio D, Ho KK, Trevithick-Sutton CC, Krogmann DW, Yeates TO (2003) The crystal structure of a cyanobacterial water-soluble carotenoid binding protein. *Structure* 11: 55-65
- Kirilovsky D, Kerfeld CA (2012) The orange carotenoid protein in photoprotection of photosystem II in cyanobacteria. *Biochim Biophys Acta* 1817: 158-166
- Kirilovsky D, Kerfeld CA (2013) The Orange Carotenoid Protein: a blue-green light photoactive protein. *Photochem Photobiol Sci* 12: 1135-1143
- Kirilovsky D, Kerfeld CA (2016) Cyanobacterial photoprotection by the orange carotenoid protein. *Nat Plants* 2: 16180
- Kuzminov FI, Karapetyan NV, Rakhimberdieva MG, Elanskaya IV, Gorbunov MY, Fadeev VV (2012) Investigation of OCP-triggered dissipation of excitation energy in PSI/PSII-less *Synechocystis* sp. PCC 6803 mutant using non-linear laser fluorimetry. *Biochim Biophys Acta* 1817: 1012-1021
- Lagarde D, Beuf L, Vermaas W (2000) Increased production of zeaxanthin and other pigments by application of genetic engineering techniques to *Synechocystis* sp. strain PCC 6803. *Appl Environ Microbiol* 66: 64-72
- Lechno-Yossef S, Melnicki M, Bao H, Montgomery B, Kerfeld CA (2017) Synthetic OCP heterodimers are photoactive and recapitulate the fusion of two primitive carotenoproteins in the evolution of cyanobacterial photoprotection. *Plant J* doi: 10.1111/tj.13593
- Letunic I, Bork P (2016) Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Res* 44: W242-245
- Leverenz RL, Jallet D, Li MD, Mathies RA, Kirilovsky D, Kerfeld CA (2014) Structural and Functional Modularity of the Orange Carotenoid Protein: Distinct Roles for the N- and C-Terminal Domains in Cyanobacterial Photoprotection. *Plant Cell* 26: 426-437
- Leverenz RL, Sutter M, Wilson A, Gupta S, Thurotte A, Bourcier de Carbon C, Petzold CJ, Ralston C, Perreau F, Kirilovsky D, Kerfeld CA (2015) PHOTOSYNTHESIS. A 12 A carotenoid translocation in a photoswitch associated with cyanobacterial photoprotection. *Science* 348: 1463-1466
- Liang C, Zhao F, Wei W, Wen Z, Qin S (2006) Carotenoid biosynthesis in cyanobacteria: structural and evolutionary scenarios based on comparative genomics. *Int J Biol Sci* 2: 197-207
- Liu H, Zhang H, King JD, Wolf NR, Prado M, Gross ML, Blankenship RE (2014) Mass spectrometry footprinting reveals the structural rearrangements of cyanobacterial orange carotenoid protein upon light activation. *Biochim Biophys Acta* 1837: 1955-1963
- Liu H, Zhang H, Orf GS, Lu Y, Jiang J, King JD, Wolf NR, Gross ML, Blankenship RE (2016) Dramatic Domain Rearrangements of the Cyanobacterial Orange Carotenoid Protein upon Photoactivation. *Biochemistry* 55: 1003-1009

- Lopez-Igual R, Wilson A, Leverenz RL, Melnicki MR, Bourcier de Carbon C, Sutter M, Turmo A, Perreau F, Kerfeld CA, Kirilovsky D (2016) Different Functions of the Paralogs to the N-Terminal Domain of the Orange Carotenoid Protein in the Cyanobacterium *Anabaena* sp. PCC 7120. *Plant Physiol* 171: 1852-1866
- Lu Y, Liu H, Saer RG, Zhang H, Meyer CM, Li VL, Shi L, King JD, Gross ML, Blankenship RE (2017) Native Mass Spectrometry Analysis of Oligomerization States of Fluorescence Recovery Protein and Orange Carotenoid Protein: Two Proteins Involved in the Cyanobacterial Photoprotection Cycle. *Biochemistry* 56: 160-166
- Maksimov EG, Sluchanko NN, Mironov KS, Shirshin EA, Klementiev KE, Tsoraev GV, Moldenhauer M, Friedrich T, Los DA, Allakhverdiev SI, Paschenko VZ, Rubin AB (2017) Fluorescent Labeling Preserving OCP Photoactivity Reveals Its Reorganization during the Photocycle. *Biophys J* 112: 46-56
- Maksimov EG, Sluchanko NN, Slonimskiy YB, Mironov KS, Klementiev KE, Moldenhauer M, Friedrich T, Los DA, Paschenko VZ, Rubin AB (2017) The unique protein-to-protein carotenoid transfer mechanism. *Biophys J* doi:10.1016/j.bpj.2017.06.002
- Melnicki MR, Leverenz RL, Sutter M, Lopez-Igual R, Wilson A, Pawlowski EG, Perreau F, Kirilovsky D, Kerfeld CA (2016) Structure, Diversity, and Evolution of a New Family of Soluble Carotenoid-Binding Proteins in Cyanobacteria. *Mol Plant* 9: 1379-1394
- Moldenhauer M, Sluchanko NN, Buhrke D, Zlenko DV, Tavraz NN, Schmitt FJ, Hildebrandt P, Maksimov EG, Friedrich T (2017) Assembly of photoactive orange carotenoid protein from its domains unravels a carotenoid shuttle mechanism. *Photosynth Res*
- Moldenhauer M, Sluchanko NN, Tavraz NN, Junghans C, Buhrke D, Willoweit M, Chiappisi L, Schmitt FJ, Vukojevic V, Shirshin EA, Ponomarev VY, Paschenko VZ, Gradzielski M, Maksimov EG, Friedrich T (2017) Interaction of the signaling state analog and the apoprotein form of the orange carotenoid protein with the fluorescence recovery protein. *Photosynth Res*
- Niyogi KK, Truong TB (2013) Evolution of flexible non-photochemical quenching mechanisms that regulate light harvesting in oxygenic photosynthesis. *Curr Opin Plant Biol* 16: 307-314
- Polivka T, Kerfeld CA, Pascher T, Sundström V (2005) Spectroscopic properties of the carotenoid 3'-hydroxyechinenone in the orange carotenoid protein from the cyanobacterium *Arthrospira maxima*. *Biochemistry* 44: 3994-4003
- Punginelli C, Wilson A, Routaboul JM, Kirilovsky D (2009) Influence of zeaxanthin and echinenone binding on the activity of the Orange Carotenoid Protein. *Biochim Biophys Acta* 1787: 280-288
- Sedoud A, Lopez-Igual R, Ur Rehman A, Wilson A, Perreau F, Boulay C, Vass I, Krieger-Liszka A, Kirilovsky D (2014) The Cyanobacterial Photoactive Orange Carotenoid Protein Is an Excellent Singlet Oxygen Quencher. *Plant Cell* 26: 1781-1791
- Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Soding J, Thompson JD, Higgins DG (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol* 7: 539
- Sluchanko NN, Klementiev KE, Shirshin EA, Tsoraev GV, Friedrich T, Maksimov EG (2017) The purple Trp288Ala mutant of *Synechocystis* OCP persistently quenches phycobilisome fluorescence and tightly interacts with FRP. *Biochim Biophys Acta* 1858: 1-11
- Sutter M, Wilson A, Leverenz RL, Lopez-Igual R, Thurotte A, Salmeen AE, Kirilovsky D, Kerfeld CA (2013) Crystal structure of the FRP and identification of the active site for modulation of OCP-mediated photoprotection in cyanobacteria. *Proc Natl Acad Sci U S A* 110: 10022-10027
- Thurotte A, Bourcier de Carbon C, Wilson A, Talbot L, Cot S, Lopez-Igual R, Kirilovsky D (2017) The cyanobacterial Fluorescence Recovery Protein has two distinct activities: Orange Carotenoid Protein amino acids involved in FRP interaction. *Biochim Biophys Acta*
- Thurotte A, Lopez-Igual R, Wilson A, Comolet L, Bourcier de Carbon C, Xiao F, Kirilovsky D (2015) Regulation of Orange Carotenoid Protein Activity in Cyanobacterial Photoprotection. *Plant Physiol* 169: 737-747
- Wheeler TJ, Clements J, Finn RD (2014) Skylign: a tool for creating informative, interactive logos representing sequence alignments and profile hidden Markov models. *BMC Bioinformatics* 15: 7

Wilson A, Ajlani G, Verbavatz JM, Vass I, Kerfeld CA, Kirilovsky D (2006) A soluble carotenoid protein involved in phycobilisome-related energy dissipation in cyanobacteria. *Plant Cell* 18: 992-1007

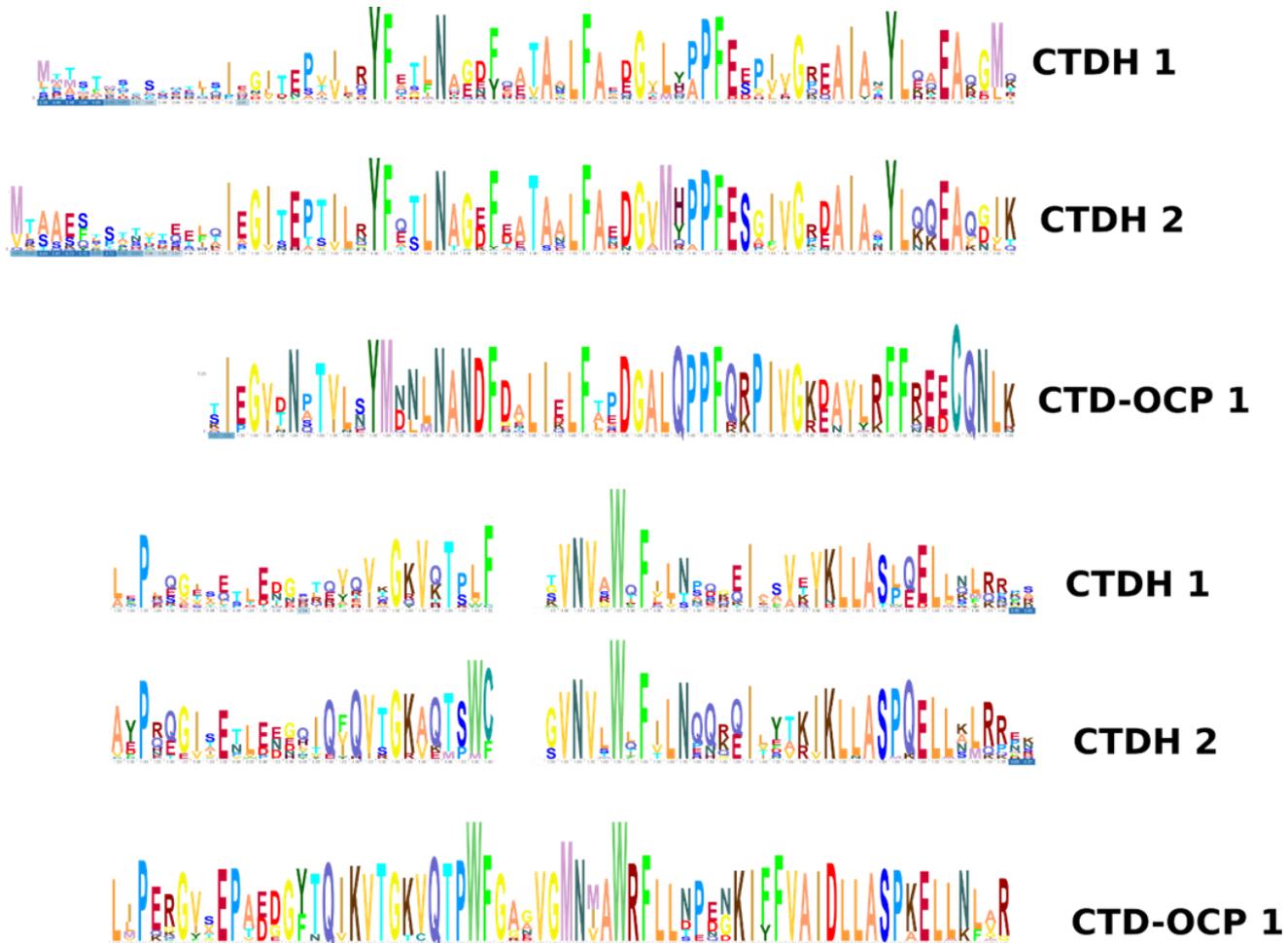
Wilson A, Gwizdala M, Mezzetti A, Alexandre M, Kerfeld CA, Kirilovsky D (2012) The essential role of the N-terminal domain of the orange carotenoid protein in cyanobacterial photoprotection: importance of a positive charge for phycobilisome binding. *Plant Cell* 24: 1972-1983

Wilson A, Kinney JN, Zwart PH, Punginelli C, D'Haene S, Perreau F, Klein MG, Kirilovsky D, Kerfeld CA (2010) Structural determinants underlying photoprotection in the photoactive orange carotenoid protein of cyanobacteria. *J Biol Chem* 285: 18364-18375

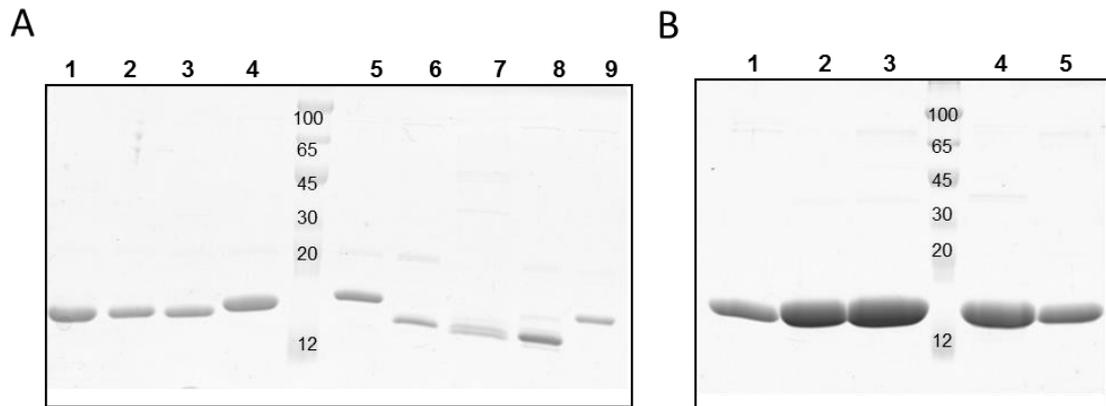
Wilson A, Punginelli C, Couturier M, Perrau F, Kirilovsky D (2011) Essential role of two tyrosines and two tryptophans on photoprotection activity of the Orange Carotenoid Protein. *Biochim Biophys Acta* 1807: 293-301

Wilson A, Punginelli C, Gall A, Bonetti C, Alexandre M, Routaboul JM, Kerfeld CA, van Grondelle R, Robert B, Kennis JT, Kirilovsky D (2008) A photoactive carotenoid protein acting as light intensity sensor. *Proc. Natl. Acad. Sci. U. S. A.* 105: 12075-12080

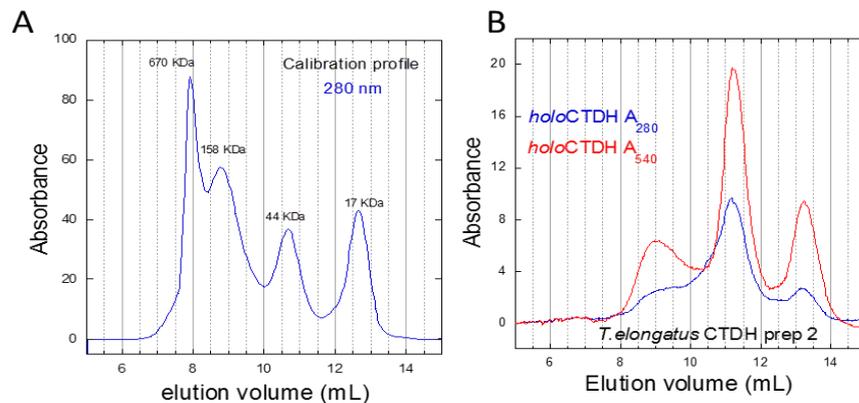
## Appendix: Supplementary Figures



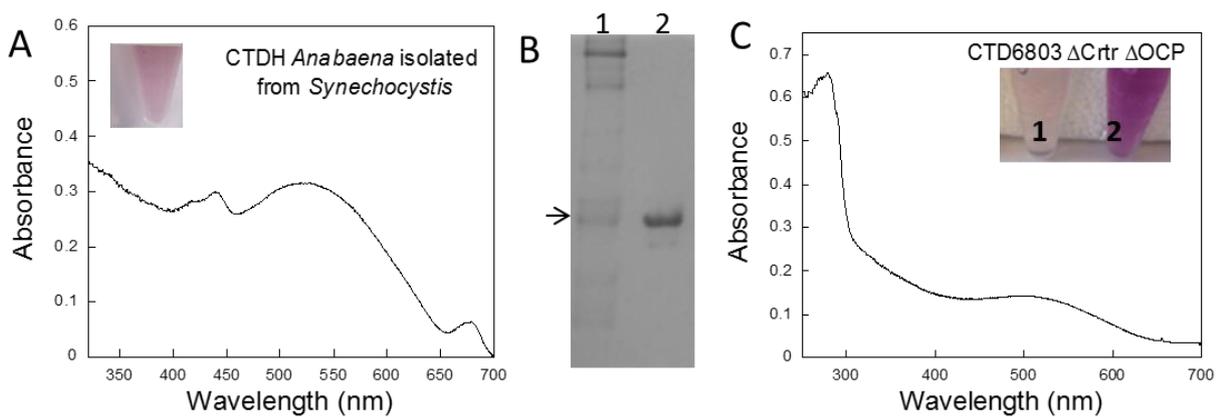
**Supplementary Figure S1. Conservation of amino acids in CTD-OCPs and in the two CTDH clades.** HMM logos were built using the SkyLign web server for aligned sequences from each clade. The height of the stack corresponds to the conservation at that position, and the height of each individual letter represents the frequency of that amino acid within the distribution.



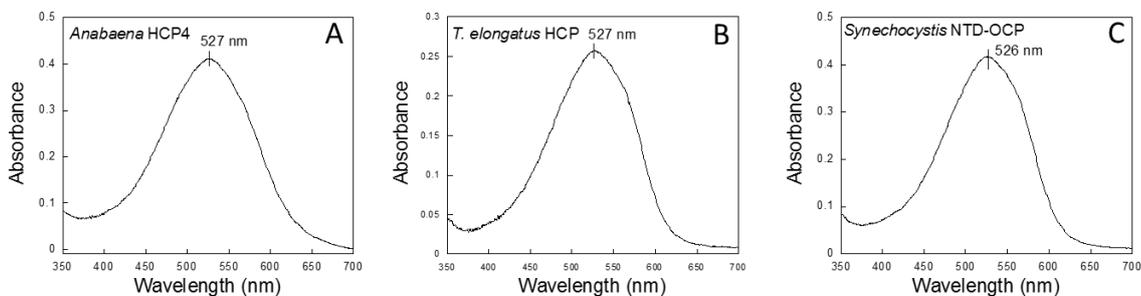
**Supplementary Figure S2 : Analysis of purity of CTDHs and HCPs used in this work.** Coomassie-blue-stained SDS gel electrophoresis. (A) lane1: *Anabaena* WT holo-CTDH ; lane 2: *Anabaena* WT holo-CTDH purified in presence of 1mM DTT; lane3: *Anabaena* holo-C103FCTDH; lane 4: *Anabaena* holo-C103F+4aa CTDH; lane 5: *Synechocystis* holo-CTD-OCP; lane 6: *T. elongatus* holo-CTDH; lane 7: *T. elongatus* holo-CTDH isolated from *Synechocystis* cells; lane 8: *T. elongatus* apo-CTDH and lane 9: *Synechocystis* apo-CTD-OCP. (B) lane 1: *Synechocystis* apo-NTD-OCP; lane 2: *T. elongatus* apo-HCP; lane 3: *Anabaena* apo-HCP4; lane 4: *Anabaena* apo-NTD-OCP; lane 5: *Anabaena* apo-C103FCTDH



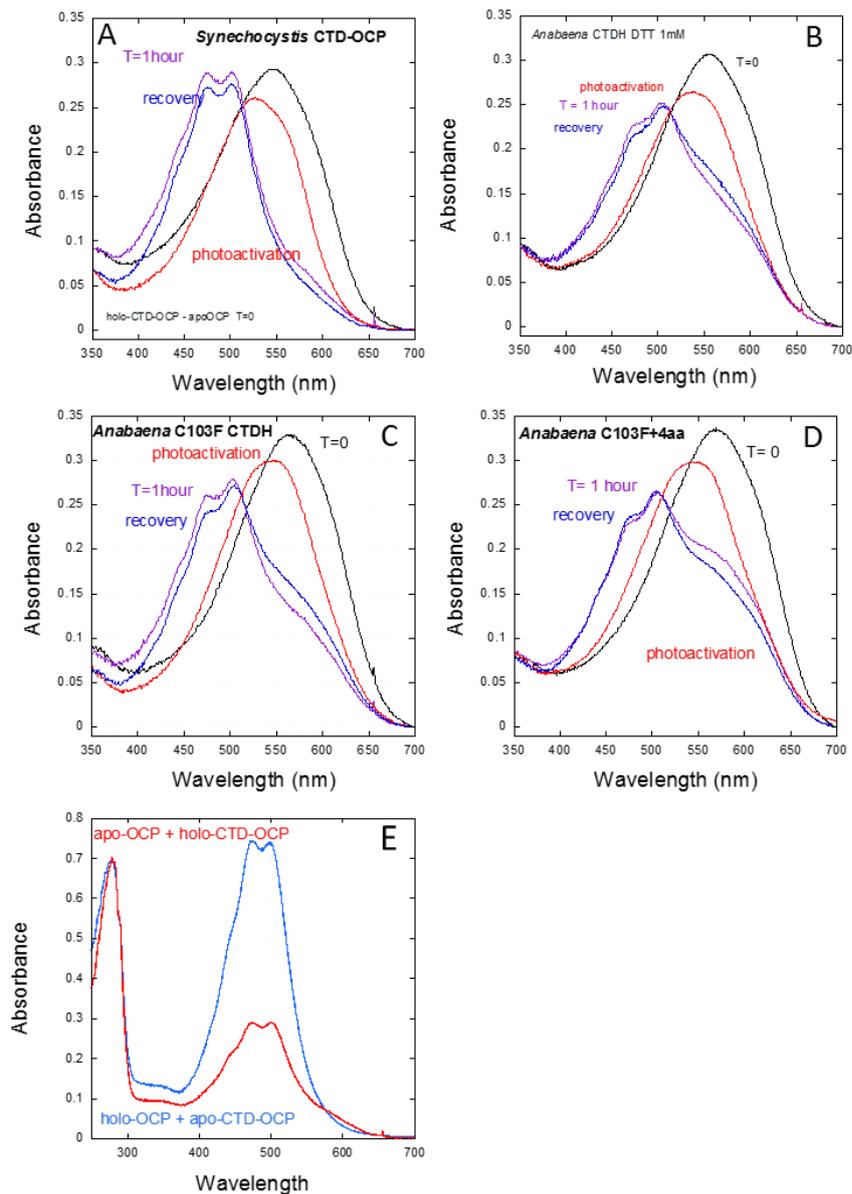
**Supplementary Figure S3: Analytical size-exclusion chromatography (SEC).** (A) Elution of standard proteins with different MW for calibration. Detection at 280 nm. Elution volume: bovine  $\gamma$ -globulin (158 kDa) : 8.7 mL, chicken ovalbumin (44 kDa): 10.6, equine myoglobin (17 kDa): 12.7, (B) Oligomeric status of *T. elongatus* CTDH in preparation 2. Detection at 280 nm (blue) and 540 nm (red). This preparation was especial since contained monomers carrying carotenoid molecules (elution peak at 13.2 mL) in addition to carotenoid binding CTDH dimers (elution peak at 11.2). The samples were injected at a concentration of 20  $\mu$ M. Detection at 280 nm (blue) and 540 nm (red).



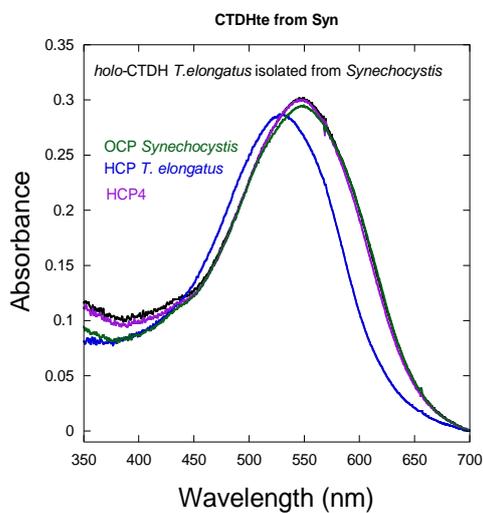
**Supplementary Figure S4: The *Anabaena* CTDH and the *Synechocystis* CTD-OCP expressed and isolated from *Synechocystis*  $\Delta$ CrtR strain.** A) Absorbance spectrum and photograph (insert) of partially isolated *Anabaena* CTDH from *Synechocystis* cells. B) Native-PAGE gel: 1) Partially isolated *Anabaena* CTDH from *Synechocystis*, 2) *Anabaena* CTDH isolated from *E. coli*. The arrow shows the position of the CTDH dimer. C) Absorbance spectrum of completely isolated *Synechocystis* CTD-OCP from the *Synechocystis*  $\Delta$ CrtR strain. Inset in C: photograph of 22 $\mu$ M solutions of *Synechocystis* CTD-OCP isolated from *Synechocystis* (1) and *E. coli* (2) cells.



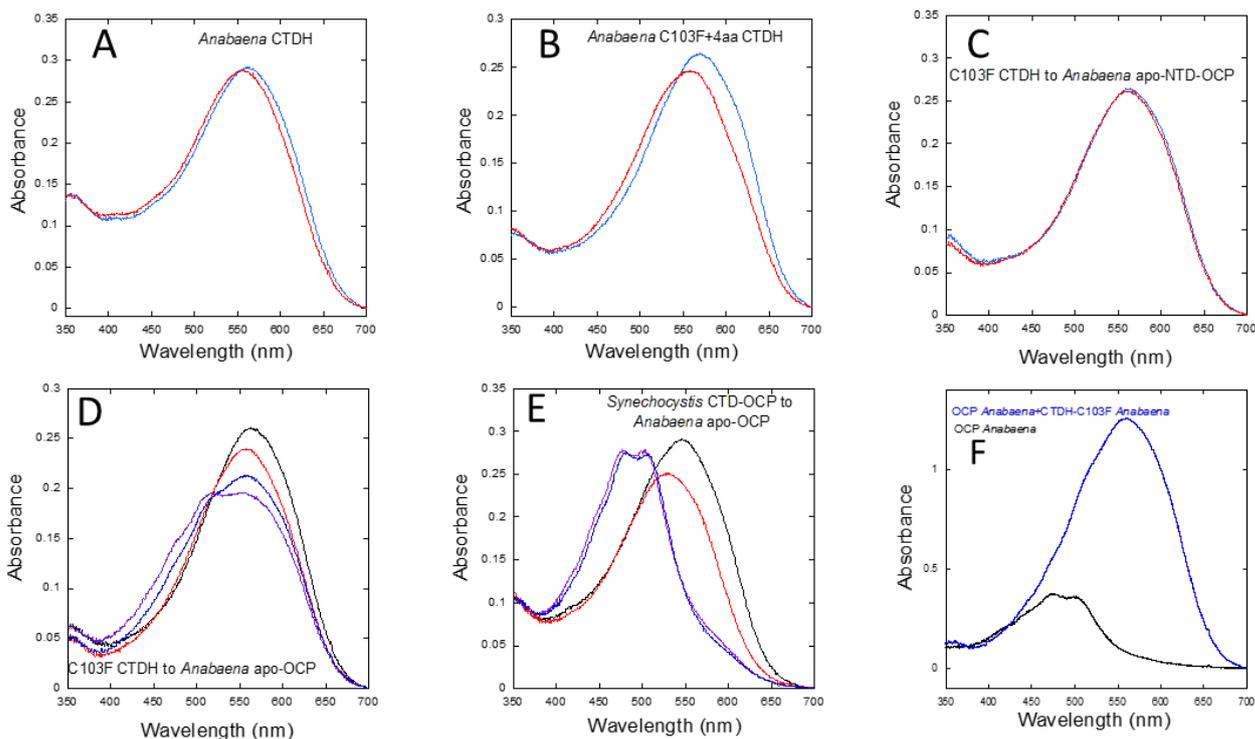
**Supplementary Figure S5: Absorbance spectra of different HCPs and NTD-OCP.** Absorbance spectra of (A) *Anabaena* HCP4, (B) *T. elongatus* HCP (C) *Synechocystis* NTD-OCP.



**Supplementary Figure S6: Changes in absorbance spectra:** induced by carotenoid transfer from holo-CTD-OCP (A) and WT (treated with DTT) (B) and mutated (C and D) *Anabaena* CTDHs to *Synechocystis* apo-OCP, and induced by photoactivity and recovery of formed holo-OCPs. 2.5 apo-OCP per 1 CTDH (dimer) were incubated during 1 hour in darkness (T=0, black; T=1 hour, violet). Then the resulting OCP<sup>o</sup> was illuminated with strong white light during 5 min (photoactivity, red). Finally the mix sample was incubated 20 min in darkness to follow OCP<sup>r</sup> to OCP<sup>o</sup> dark conversion (recovery, blue). (E) Absorbance spectra of a mix holo-OCP and apo-CTD-OCP (blue) (ratio 1 to 1 monomer) with a mix apo-OCP and holo-CTD-OCP (ratio 1 to 1 monomer) after one hour incubation (red). Only 42% of apo-OCP was transformed into holo-OCP while around 85 % of carotenoid was transferred from the CTD-OCP. This result suggests the presence of only one carotenoid per CTD-OCP dimer.

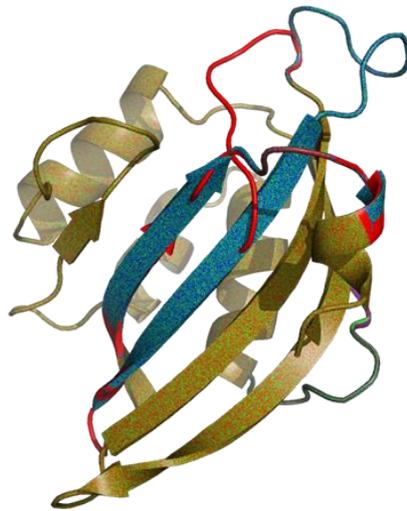
**Supplementary Figure S7:**

Changes in absorbance spectra induced by carotenoid transfer from *T. elongatus* CTDH to *T. elongatus* apo-HCP (blue), *Anabaena* apo-HCP4 (violet) and *Synechocystis* apo-OCP (green). The *T. elongatus* CTDH used in this experiment was isolated from *Synechocystis* cells.

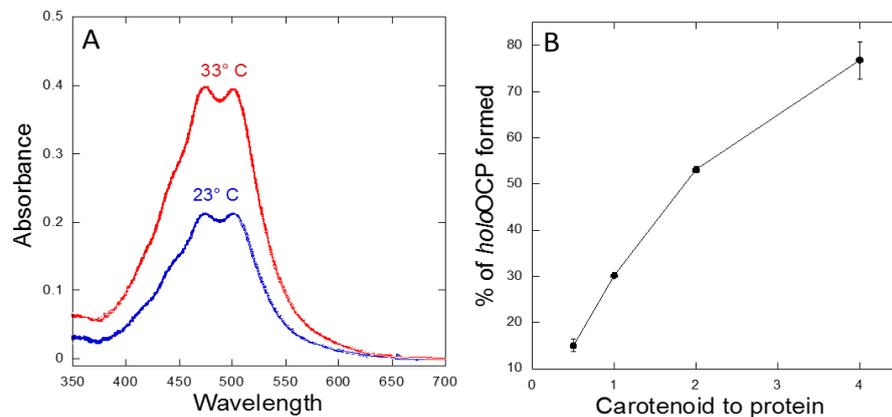
**Supplementary Figure S8: Changes in absorbance spectra induced by**

carotenoid transfer from *Anabaena* WT and mutated *Anabaena* CTDHs to apo-HCPs and *Anabaena* apo-OCP (A) oxidized holo-WT *Anabaena* CTDH to *Anabaena* HCP4. Blue spectrum: Time 0, all carotenoid in CTDH. Red spectrum: after one hour incubation with apo-HCP4. (B) *Anabaena* holo-C103F+4aa CTDH to *Anabaena* HCP4. Blue spectrum: Time 0, all carotenoid in CTDH. Red spectrum: after one hour incubation with apo-HCP4. (C) *Anabaena* holo-C103F CTDH to *Synechocystis* NTD-OCP. Blue

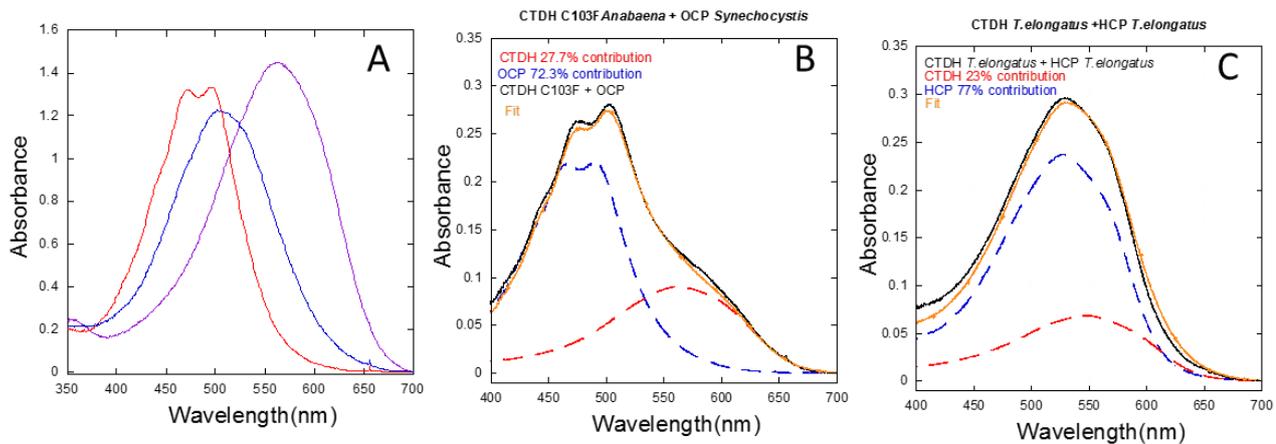
spectrum: Time 0, all carotenoid in CTDH. Red spectrum: after one hour incubation with apo-CTD-OCP. (D and E) *Anabaena* C103F CTDH (D) and *Synechocystis* CTD-OCP (E) to *Anabaena* apo-OCP. Black spectrum: Time 0; Violet: 1 hour incubation with apo-OCP; Red: photoactivation; Blue: recovery. (F) Final absorbance spectra after incubation of *Anabaena* apo-OCP and membranes (black) and a mix of *Anabaena* apo-OCP and *Anabaena* apo-C103F CTDH. All the carotenoid remained in the CTDH. No carotenoid transfer to the apo-OCP was observed.



**Supplementary Figure S9: Comparison of structures** of the CTD-OCP (green) and *Anabaena* WT CTDH (red) and *Anabaena* C103F+4aa CTDH (blue). The structures were generated as homology model (see materials and methods). The CTDH mutant and the CTD-OCP have identical structures based in the homology models.



**Supplementary Figure S10: Carotenoid translocation from membranes to apo-OCP.** (A) Absorbance spectra of *Synechocystis* holo-OCP formed after one hour incubation of canthaxanthin containing *E. coli* membranes with apo-OCP. The ratio carotenoid to protein was 0.5. (B) Percentage of holo-OCP formed after one hour incubation of canthaxanthin containing *E. coli* membranes with apo-OCP. The concentration of the apo-OCP was 12  $\mu$ M. The concentration of membranes varied to obtain carotenoid to protein ratio of 0.5, 1, 2 and 4. The experiment was realized at 33  $^{\circ}$ C.



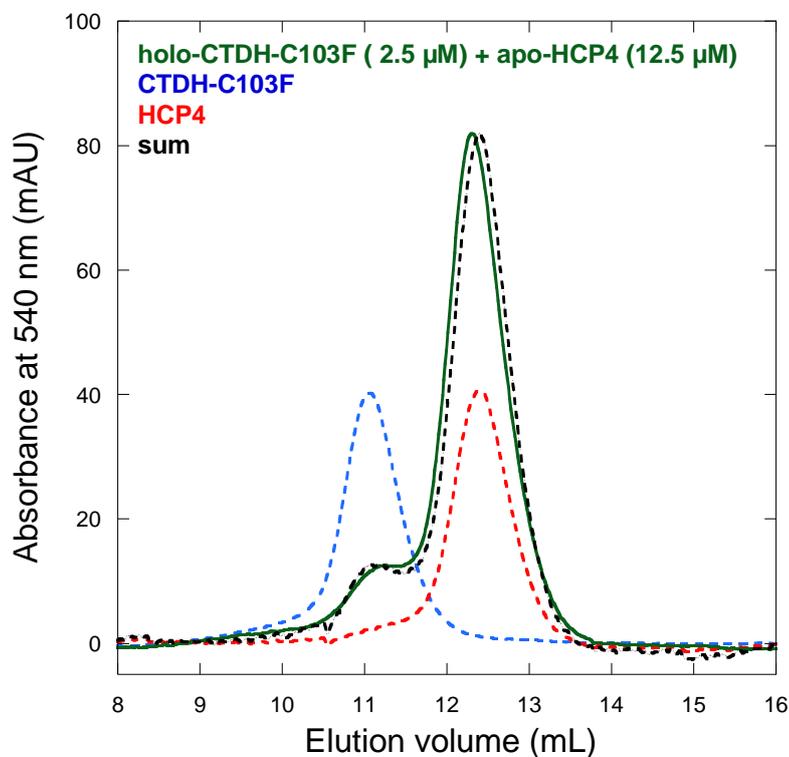
**Supplementary Figure S11: Examples of fit of final absorbance spectra after one hour incubation of CTDH and OCP and HCP.** (A) Absorbance spectra of 12  $\mu$ M 100% holo-CAN-OCP (red), holo-HCP (blue) and holo-CTDH dimer (violet). (B) Final absorbance spectrum of a mixture of *Anabaena* holo-C103FCTDH and apo-OCP (black). (C) Final absorbance spectrum of a mixture of *T. elongatus* holo-CTDH and apo-*T. elongatus* HCP (black). In B and C, the contributions of OCP (B, blue), HCP (C, blue) and CTDH (red) spectra to the black absorbance spectrum is shown as well as the fit (orange).

Primer names	Primer sequences
A. Construction of plasmids for expression in <i>E.coli</i>	
1- Construction of <i>pCDF-HCP-Te-1269Ctag</i> plasmid	
Ftl1269EcoRI	5'-CGATATGAATTCACCTACACCAGCGATACTCG-3'
Rtl1269NotI	5'-CATTATGCGGCCGCGGCAGTTCCTCAGGGGTTTG-3'
Ftl1269NcoI del	5'- TTAATAAGGAGATATACCATGGGCACCTACACCAGCGATACT CGTTTC-3'
Rtl1269NcoI del	5'-CATGGTATATCTCCTTATTAAGTTAAACAAAATTATTCTA- 3'
Ftl1269HIS	5'-CAT CAC CAT CAC CAT CAC TAG TTG AGA TGC CAT GGG TGA CGC-3'
Rtl1269HIS	5'-GTG ATG GTG ATG GTG ATG GACGGCAAGGGG ATCCACACC-3'

<b>2- Construction of pCDF-CTDHTe-1268Ctag plasmid</b>	
FtlI1268NcoI	5'- GATATACCATGGGCAGCAGCACTCATCTTGTTC AAGGGAGC- 3'
RtlI1268NotI	5'-CATTATGCGGCCGCGGGAACAGTTGGCAGCGGTAC-3'
FtlI1268HIS	5'- CATCATCATCATCATCATTAGGGCCCGCCGAAAACACCGCG GAC-3'
RtlI1268HIS	5'-ATGATGATGATGATGATGGCGCTGCTGACGCAGAGGCTG- 3
<b>3- Construction of pCDF-CTDHAna-4940Ctag and mutants plasmids</b>	
Fall4940NcoI	5'-GATATACCATGGGCAAAGCTGCTGAGTCTCTCCC-3'
Rall4940NotI	5'-CAT TAT GCG GCC GC CTT ATT GAT GAC AGC GCA CCC- 3'
FAll4940HIS	5'-CAC CAC CAC CAC CAC CAC TAG TCT TGA TTA ATT ACT TTA ACT TTG-3'
RAI14940HIS	5'-GTG GTG GTG GTG GTG GTG CTG TTC ACG ACG CAA AGC CAG TAA CTC-3'
Fall4940C103F	5'-GCC CAA ACT TCT TGG TTT GGT GTA AAT GTC TTA TGG- 3'
Rall4940C103F	5'-CCA TAA GAC ATT TAC ACC AAA CCA AGA AGT TTG GGC-3'
Fall4940C103FGA AV	5'- TTTGGTGCAGCAGTAGGTGTAAATGTCTTATGGCTATTTACT C-3
Rall4940C103FGA AV	5'- TACTGCTGCACCAAACCAAGAAGTTTGGGCTTTACCAGAAAC -3
<b>4- Construction of pCDF-CTD6803-Ctag plasmid</b>	
Fslr1963NTD del	5'- CATTACCATTGACTCTGCCAAACGCATCGCTGAGCCCGTA- 3'
Rslr1963NTD del	5'- AGAGTCAATGGTGAATGGGCTGCTGCCCATGGTATATCTCCT -3'
<b>1- Construction of psbA2-CTDHAna-4940Ctag plasmid</b>	
Fall4940NdeI	5'-GCA TTC CAT ATG AAA GCT GCT GAG TCT CTC CCT AAT ATC-3'
Rall4940HpaI	5'-GCA TCG GTT AAC AAT GCT CCT GTT TGG CAT TTG TTT GAT C-3'

2- Construction of <i>psbA2-CTDHTe-1268Ctag</i> plasmid	
Fit11268NdeI	5'-GCA TTC <b>CAT ATG</b> ACT CAT CTT GTT CAA GGG AGC CCC -3'
Rit11268HpaI	5'-GCA TCG <b>GTT AAC</b> AGT TGG CAG CGG TAC GCC TGA GTC-3'

**Supplementary Table I: Sequences of primers used in this study.** Listed primers include those used for construction of plasmids for expression in *E. coli* (A) and for expression in *Synechocystis* (B) as described in the Methods section.



**Supplementary Figure S12: Analysis of interaction between CTDH and HCP.** Size exclusion chromatography profiles of holo-CTDH-C103F (dotted blue curve), holo-HCP4 (dotted red curve) and holo-CTDH-C103F (2.5 μM) incubated with apo-HCP4 (12.5 μM) at 23°C for 1 hs (solid green line). The algebraic sum of individual profiles of holo-CTDH and HCP normalized according to their spectral contribution in the mix (Table I) is shown in the dotted black profile. The profiles correspond with the absorbance at 540 nm.

## **Part II: Results**

---

### **Chapter II**

# **Analysis of a multidirectional carotenoid transfer mechanism.**

## *Summary of work*

---

This chapter is based on:

Slonimskiy, Yury B., Fernando Muzzopappa, Eugene G. Maksimov, Adjélé Wilson, Thomas Friedrich, Diana Kirilovsky, and Nikolai N. Sluchanko. 2019. “Light-Controlled Carotenoid Transfer between Water-Soluble Proteins Related to Cyanobacterial Photoprotection.” *FEBS Journal* 286: 1908–24.

Nevertheless, minor modification on the text had been added in order to ensure the compatibility of the terminology used through the different chapters. In addition, results from Harris et al. 2016 (reverse carotenoid transfer from HCP1 to CTDH) have been added in the main text of this chapter for a better presentation of the research project. Consequently, figure 6 and supplementary figure 1 do not belong to the original paper (Slonimskiy et al. 2019) in which this chapter is based.

## *Contribution to this work*

In this work, I collaborate by studying the interaction, and consequent carotenoid transfer, between Holo-CTDH(CAN) and Apo-OCP-CTD proteins. I also performed experiments to assess interaction, and reverse carotenoid transfer, between Holo-HCPs (or NTDs) and Apo-CTDHs. The experiments to characterize both holoforms of AnaCTDH and TeCTDH (binding either CAN or ECN), as well as some experiments of horizontal carotenoid transfer between CTD(H)s and vertical carotenoid transfer from OCP<sup>R</sup>s to CTDH were performed by the Slonimskiy, Maksimov and Sluchanko.

## *Resume of the article*

By using a different expression system for carotenoid proteins, the group of Sluchanko obtained a mixture of CAN- and ECN-binding CTDHs (from *Anabaena*, AnaCTDH, and from *T. elongatus* TeCTDH). Chromatographic isolation of these samples yielded in a CAN enriched CTDH and ECN enriched CTDH. In this article, a biochemical and spectroscopic characterization of the different holoforms is presented. Briefly, it was found that binding of ECN to the CTDH produces a red shift in the absorption spectrum of the carotenoid (CAN:  $\lambda_{\max}$ =534 nm and ECN:  $\lambda_{\max}$ = 560 nm) and

destabilizes the dimeric form. In order to investigate the effect of different carotenoids in the carotenoid transfer mechanisms of CTDH (described in the chapter II) and to explore the directionality of this process, we started a collaboration between the two groups.

In agreement with the previous chapter, carotenoid transfer to SynOCP(apo) is more efficient from OCP-CTD than CTDH. In addition, here we show that ECN transfer from AnaCTDH to SynOCP(apo) is slightly more efficient than CAN transfer. This indicated that the nature of the carotenoid has slight effect on the carotenoid transfer. However, the most exciting finding was that incubation of OCP-CTD(CAN) with AnaCTDH(apo) lead to a spectral shift from 550 nm (maximum of OCP-CTD(CAN)) to 560 nm (maximum of AnaCTDH(CAN)) indicating that carotenoid transfer is also possible between CTD(H)s in a “horizontal” way with heterodimers intermediaries. In addition to this, while interaction between Apo-AnaCTDH and Holo-AnaHCP4 (neither AnaHCP2 nor AnaHCP3) does not induce carotenoid transfer, AnaCTDH can retrieve the carotenoid from AnaHCP1(CAN) (results from Harris et al. 2016). Prompted by this discovery, in Slonimskiy et al. 2019, we explore the multidirectionality of the carotenoid transfer. We found that AnaCTDHs (but not TeCTDH) also can retrieve carotenoid from isolated NTD. This “inverse” carotenoid transfer was also observed when AnaCTDH was incubated with the analog of the OCP<sup>R</sup> (OCP-Y201A/W288A) or the OCP $\Delta$ NTE (in which the interdomain interactions are weaker). The last finding presented in this chapter is that while AnaCTDH cannot retrieve the carotenoid from OCP<sup>O</sup>, upon photoactivation the carotenoid transfer OCP<sup>R</sup> to AnaCTDH is possible.

Overall, this chapter strengthens the idea that the carotenoid transfer among OCP-related proteins (OCP, HCP and CTDH) is a complex process determined by carotenoid affinities as well as protein interactions. As a result of this, the carotenoid transfer can take place between different acceptor/donor pairs at least *in vitro*. These results raise new questions about the potential *in vivo* implication of the carotenoid transfer mechanism.

## ***Light-controlled carotenoid transfer between water-soluble proteins related to cyanobacterial photoprotection***

---

Yury B. Slonimskiy<sup>a,b</sup>, Fernando Muzzopappa<sup>c</sup>, Eugene G. Maksimov<sup>a,d</sup>, Adjélé Wilson<sup>c</sup>, Thomas Friedrich<sup>c</sup>, Diana Kirilovsky<sup>c</sup>, Nikolai N. Sluchanko<sup>a,d,\*</sup>

### **Affiliations**

<sup>a</sup>A.N. Bach Institute of Biochemistry, Federal Research Center of Biotechnology of the Russian Academy of Sciences, 119071 Moscow, Russian Federation

<sup>b</sup>M.V. Lomonosov Moscow State University, Department of Biochemistry, Faculty of Biology, 119991 Moscow, Russian Federation

<sup>c</sup>Institute for Integrative Biology of the Cell (I2BC), CEA, CNRS, Université Paris-Sud, Université Paris-Saclay, 91198 Gif sur Yvette, France

<sup>d</sup>M.V. Lomonosov Moscow State University, Department of Biophysics, Faculty of Biology, 119991 Moscow, Russian Federation

<sup>e</sup>Technical University of Berlin, Institute of Chemistry PC 14, Straße des 17. Juni 135, D-10623 Berlin, Germany

### **Abstract**

Carotenoids are lipophilic pigments with multiple biological functions from coloration to vision and photoprotection. Still, the number of water-soluble carotenoid-binding proteins described to date is limited, and carotenoid transport and carotenoprotein maturation processes are largely underexplored. Recent studies revealed that CTDHs, which are natural homologs of the C-terminal domain (CTD) of the Orange Carotenoid Protein (OCP), a photoswitch involved in cyanobacterial photoprotection, are able to bind carotenoids, with absorption shifted far into the red region of the spectrum. Despite the recent discovery of their participation in carotenoid transfer processes, the functional roles of the diverse family of CTDHs are not well understood. Here we characterized CTDH carotenoproteins from

*Anabaena variabilis* (AnaCTDH) and *Thermosynechococcus elongatus* (TeCTDH) and examined their ability to participate in carotenoid transfer processes with a set of OCP-derived proteins. This revealed that carotenoid transfer occurs in several directions guided by different affinities for carotenoid and by specific protein-protein interactions. We show that CTDHs have higher carotenoid affinity compared to the CTD of OCP from *Synechocystis*, which results in carotenoid translocation from the CTD into CTDH via a stable heterodimer intermediate. Activation of OCP by light, or mutagenesis compromising the OCP structure, provides AnaCTDH with an opportunity to extract carotenoid from the full-length OCP, either from *Synechocystis* or *Anabaena*. These previously unknown reactions between water-soluble carotenoproteins demonstrate multidirectionality of carotenoid transfer, allowing for efficient and reversible control over the carotenoid-mediated protein oligomerization by light, which gives insights into physiological regulation of OCP activity by CTDH and suggests multiple applications.

## ***Introduction***

---

Carotenoids are hydrophobic pigment molecules derived from the isoprenoid metabolism involved in various processes allowing living organisms to interact with light and perceive visual information [1] but also are valuable dietary components [2]. The over 700 carotenoids in nature not only provide for the polychromatic coloration of fruits, plants, algae, and animals [3-5], but more importantly, they are responsible for light harvesting, photosensory and photoprotective reactions as well as defense mechanisms against oxidative stress [6, 7]. Their tremendous role notwithstanding, there is a limited number of water-soluble carotenoid-binding proteins described to date [5, 8], and knowledge about carotenoid transport and carotenoprotein maturation processes is still limited.

The unique carotenoid transfer mechanism between water-soluble proteins related to cyanobacterial photoprotection was discovered in 2017 [9, 10]. Many cyanobacteria, the photosynthetic microorganisms that transformed our planet because of the pioneering oxygen production [11], employ the so-called Orange Carotenoid Protein (OCP) for protecting their sensitive photosystems against the harmful effects of excessively absorbed light [12-14]. OCP is a blue light-triggered photoswitch employing a ketocarotenoid as cofactor, which upon absorption of a blue photon (450-500 nm) undergoes photoconversion from the basal, dark-adapted orange state (OCPO) to the

active red state (OCPR), the latter being able to interact with the phycobilisome antenna complexes to consequently quench their fluorescence and prevent photodamage [15].

OCP is a ~35 kDa water-soluble protein subdivided into two structural domains which encapsulate a single xanthophyll molecule in a common central cavity. The xanthophyll is 3'-hydroxy-echinenone when purified from native cyanobacteria (e.g. *Synechocystis* sp. PCC 6803, further *Synechocystis*, or *Arthrospira maxima* [16, 17]), but photoactivity is also preserved with only 4(4')-ketolated xanthophylls (either mono-ketolated echinenone, ECN, or the diketolated canthaxanthin, CAN) [18-20], whereas insertion of  $\beta$ -carotene or (3,3'-hydroxylated) zeaxanthin results in non-photoconvertible protein [20]. The  $\text{OCP}^{\text{O}} \rightarrow \text{OCP}^{\text{R}}$  photoactivation is triggered by absorption of blue light, making OCP an excellent light intensity sensor over the whole range of ambient light intensities [21, 22]. The  $\text{OCP}^{\text{R}} \rightarrow \text{OCP}^{\text{O}}$  back-conversion occurs spontaneously in the dark, but it can be accelerated by the Fluorescence Recovery Protein (FRP) [23-25]. FRP controls the activity of OCP via direct protein-protein interactions with the photoactivated OCP and protein variants mimicking this state [26-30].

In the basal  $\text{OCP}^{\text{O}}$  state, the protein adopts a compact structure [31] with the all- $\alpha$  N-terminal domain (NTD) and the mixed  $\alpha$ -helical/ $\beta$ -sheet C-terminal domain (CTD) that belongs to the widespread family of nuclear transport factor 2 (NTF-2)-like proteins (Pfam 02136). The  $\text{OCP}^{\text{O}}$  structure is stabilized by multiple protein-protein interactions across the NTD-CTD interface, by the interdomain linker, and by the N-terminal extension (NTE), which directly contacts the CTD [31], as well as protein-carotenoid interactions including two H-bonds between the 4-keto oxygen of the carotenoid and the hydroxyl group of Tyr201 as well as to the imino nitrogen of Trp288 in the CTD (amino acid numbering according to *Synechocystis* OCP) [31, 32]. These amino acids are highly conserved in all known OCP sequences, as their mutation severely interferes with the ability to photoswitch and change spectroscopic and other physicochemical properties of the protein [28, 33, 34]. A recent dynamic crystallography study proposed that the aforementioned H-bonds are the first to break down during OCP photoactivation [35]. Likewise, the Trp288Ala mutation of *Synechocystis* OCP was shown to cause protein transformation into an  $\text{OCP}^{\text{R}}$ -like form with separated protein domains being able to constantly interact with FRP [26, 28]; the double Tyr201Ala/Trp288Ala mutation caused similar effect [27, 34].

Separation of the NTD and CTD accompanying OCP photoactivation is in perfect agreement with the modularity of OCP structure and function [36]. Reassembly of the recombinantly produced individual domains, which stably exist as individual carotenoproteins *in vitro*, allowed researchers to reconstruct metastable photoactive OCP-like species [10, 37], suggesting that a similar process could have arisen in evolution giving birth to the first ancestral OCPs [10, 37], and to reveal the possibility of carotenoid shuttling between the domains [10].

For both domains of OCP, multiple extant homologs exist in nature: up to nine clades of NTD homologs (called Helical Carotenoid Proteins, HCPs) [38, 39] and at least two clades of C-terminal domain homologs (CTDHs) [40, 41], which exist in different cyanobacteria in various combinations regardless whether full-length OCPs are present or not [12, 38, 40]. For example, *Thermosynechococcus elongatus* BP-1 contains only one HCP (clades HCP4/5), one CTDH (clade 1) and no OCP, whereas *Anabaena* sp. PCC 7120 (further *Anabaena*) contains four HCP paralogs (HCP1 to HCP4), one CTDH (clade 2) and one OCP. OCP itself can belong to different clades (e.g., OCP1, OCP2, and OCPx) [42].

Although some HCPs [38, 39] and CTDHs [40, 43] have been structurally and biochemically characterized, their function *in vivo* remains largely speculative. It was demonstrated that the isolated HCPs of *Anabaena* may have different roles in photoprotection: HCP4 interacts with phycobilisomes and induces excitation energy quenching, HCP2 and HCP3 are good quenchers of singlet oxygen and HCP1 might serve as a carotenoid carrier [39]. CTDHs are very good singlet oxygen quenchers, efficiently extract carotenoids from membranes and transfer them to the apoforms of OCP and HCPs [40], as initially demonstrated for the recombinant *Synechocystis* OCP-CTD (also called C-terminal OCP-related Carotenoid Protein, COCP) [9, 10]. The C-terminal tail of CTDH was proposed to be crucial for carotenoid binding, uptake from membranes and transfer to other proteins [43]; however, the molecular details and prerequisites for these processes remain insufficiently understood.

By comprehensive characterization of the apo- and holoforms of CTDHs from *Anabaena* and *T. elongatus* and their interaction with various individual OCP domains and mutant variants, we found that carotenoid transfer is multidirectional, depending on the relative carotenoid affinities of the transfer partners and the efficiency of protein-protein interactions. Of note, this study demonstrates a photoinduced, reversible carotenoid shuttling between OCP

and CTDH, suggesting the hypothesis that CTDH-OCP interactions may contribute to OCP-related photoprotective mechanism in CTDH-containing cyanobacteria independently of the FRP-based process. In addition, these results forward the future design of light-triggered protein-protein interaction cascades and synthetic biology applications, as well as the construction of antioxidant nanocarriers.

## *Results*

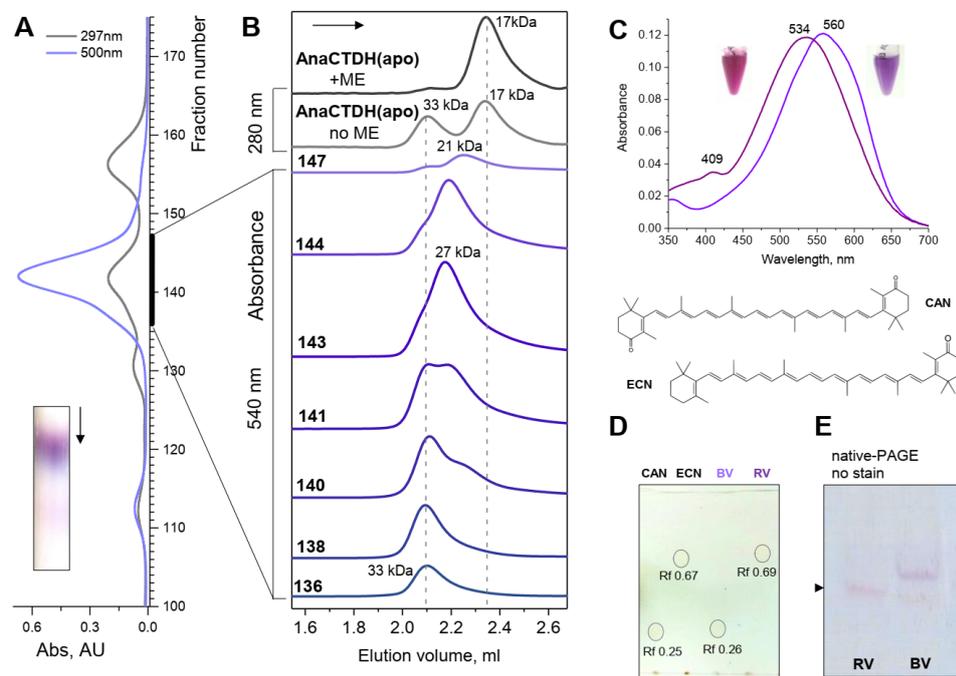
---

### *Characterization of the apo and holoforms of CTDH proteins*

The CTDH from *Anabaena* (AnaCTDH) produced in *Escherichia coli* in the absence of carotenoids (AnaCTDH(apo)) was purified under reducing conditions as a predominantly monomeric species (apparent Mw ~17 kDa; calculated monomer Mw 16.1 kDa) with a propensity to form disulfide-linked dimers (apparent Mw ~33 kDa) upon oxidation (Fig. 1) due to the presence of Cys103, which is typical of clade 2 CTDHs, in agreement with [40]. AnaCTDH produced in an ECN/CAN-producing *E. coli* strain [18] yielded a violet-purple carotenoprotein [40], which according to preparative size-exclusion chromatography (SEC) dissected into two poorly resolved fractions of different color (red-violet, “RV”, or blue-violet, “BV”) in the elution profile (Fig. 1A). Analytical SEC of the fractions obtained during the preparative SEC confirmed the overlap of at least two carotenoid-containing peaks (Fig. 1B), characterized by differences in the visible absorbance spectrum (Fig. 1C).

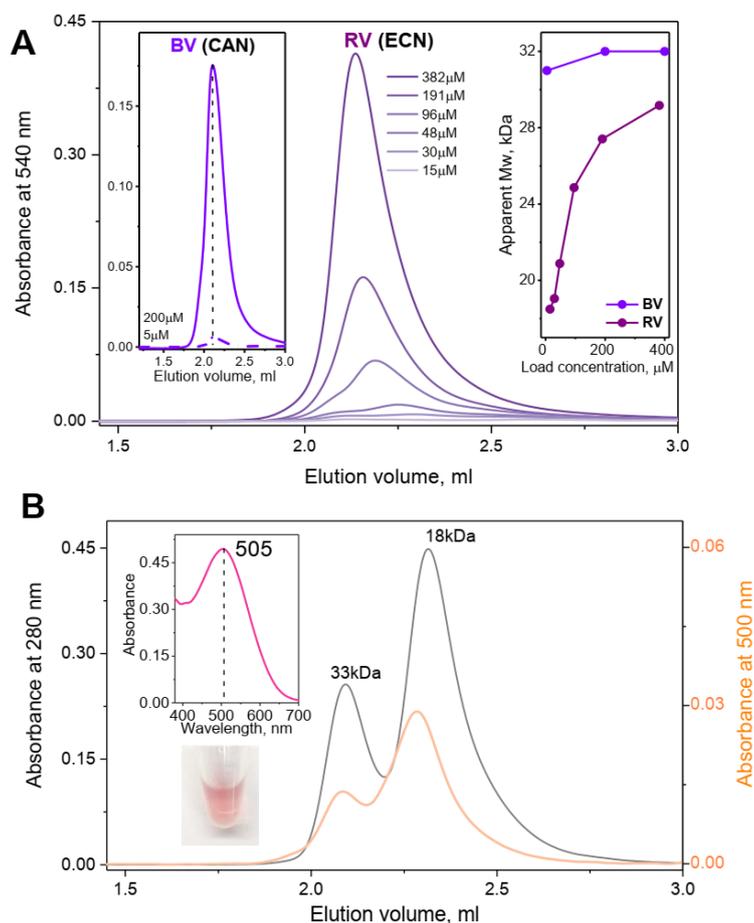
Analysis of the carotenoid content by thin layer chromatography (Fig. 1D) revealed that the BV fraction predominantly contained CAN, whereas the RV fraction contained ECN. CAN bound to the AnaCTDH dimer showed the greatest red shift among all OCP-related species with an absorption maximum above 560 nm [40], which is by ~10 nm more red-shifted compared to OCP-CTD(CAN) [9, 10]. These spectral properties hint at distinct differences in the specific environment formed by the corresponding protein matrix. Given the almost identical energy gap for S0-S2 excitation of ECN and CAN in organic solvents (~460 nm) [44], the about 30 nm bathochromic shift of AnaCTDH(CAN) relative to AnaCTDH(ECN) indicates different carotenoid configuration (conjugation length) and carotenoid-protein interactions. Another spectral signature characteristic of ECN-containing AnaCTDH

species is a small peak at ~409 nm (Fig. 1C), probably corresponding to cis isomerization [45]. In contrast to the symmetric CAN molecule with two ketolated  $\beta$ -ionone rings, ECN cannot simultaneously form hydrogen bonds with Tyr and Trp residues (201 and 288 in *Synechocystis* OCP numbering, respectively) on both ends of the AnaCTDH homodimer. Of note, a similar blue spectral shift was reported due to the Trp288Ala mutation in OCP-CTD [9], supporting the potential role of the H-bonds in spectroscopically relevant carotenoid-protein interactions within CTD (or CTDH). The different dimer stabilities and apparent sizes resulted in different electrophoretic mobility of AnaCTDHs from the RV and BV fractions on native-PAGE (Fig. 1E).



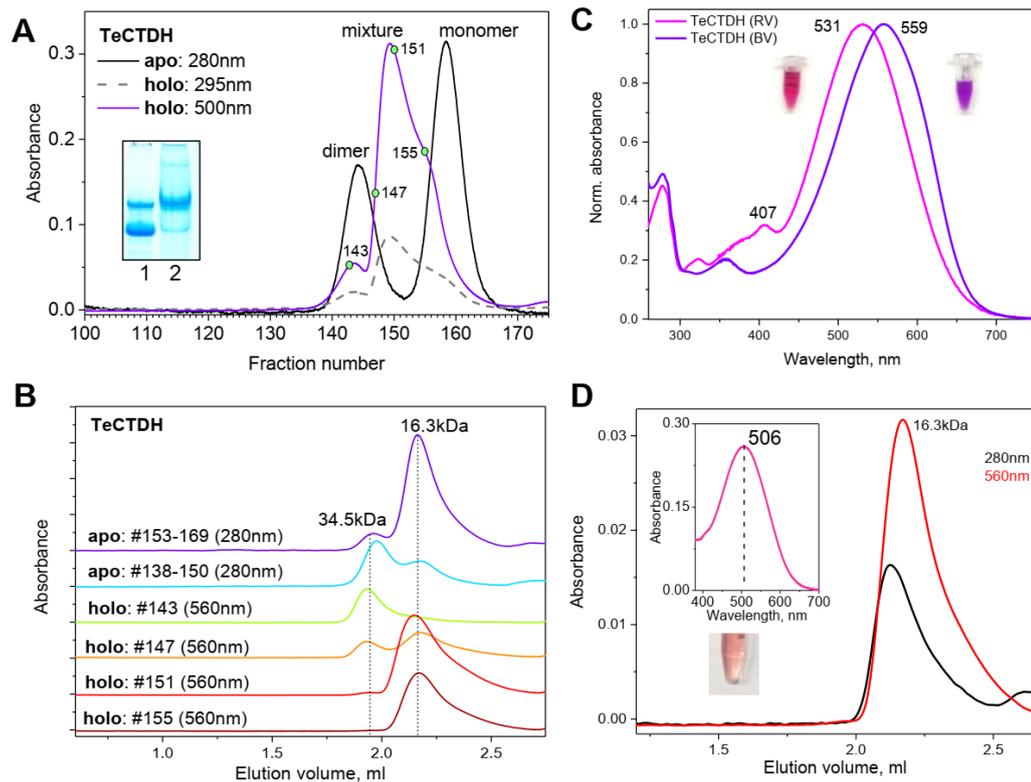
**Figure 1 | Characterization of AnaCTDH preparations obtained from *E. coli* expressing ECN/CAN.** A, B 2D size-exclusion chromatography of the AnaCTDH holoprotein preparation (A – preparative SEC on a Superdex 200 26/600 column at a 1.6 ml/min flow rate monitored using protein (297 nm) and carotenoid-specific (500 nm) absorbance, B – analytical SEC on a Superdex 200 Increase 5/150 column at a 0.45 ml/min flow rate monitored at either 280 or 540 nm). Arrows indicate the direction of the run, apparent Mw values for the main peaks are indicated in kDa based on column calibration. The inset shows a segment of the preparative SEC column during the run. C. Absorbance spectra of the two main fractions of AnaCTDH(holo) obtained, with the natural color of the samples indicated in the insert. Note the presence of the ~409 nm peak characteristic of only the ECN-bound form of AnaCTDH. Chemical formulae of ECN and CAN are depicted below the spectra. D. Thin layer chromatogram of the two fractions, RV and BV, revealing the predominant carotenoid types [46]. E. Native-PAGE analysis of the RV and BV fractions showing different electrophoretic mobility (marked by arrowheads). Note the slight cross-contamination due to incomplete separation of these fractions on preparative SEC.

On analytical SEC, AnaCTDH(CAN) eluted as dimers (Mw of 31-33 kDa) almost independent of protein concentration, whereas AnaCTDH(ECN) demonstrated a pronounced shift upon dilution corresponding to an Mw drop from ~29 to ~18 kDa (Fig. 2A) suggesting dimer dissociation. Most likely, the absence of the second keto-group in ECN hinders coordination of a second AnaCTDH monomer, as in the case of the OCP-CTDW288A mutant, which is lacking one of the ketocarotenoid-coordinating residues [9]. This supports the idea that H-bonds between the keto-groups of CAN and the conserved Tyr/Trp residues in AnaCTDH (as well as in individual OCP-CTD [9]) contribute significantly to dimer stability. Notably, in agreement with [9, 40], we also observed a fraction of AnaCTDH(ECN) monomers (~18 kDa) retaining carotenoid molecules, which exhibited absorbance spectrum with a maximum at ~505 nm, unusual for CTDH (Fig. 2B).



**Figure 2 | Analysis of the carotenoid-coordinating AnaCTDH species by SEC.** A. The concentration-dependent behavior of the BV and RV fractions. While the BV fraction containing mostly CAN is stable to dilution (inset on the left), the RV fraction enriched in ECN shows a remarkable dissociation upon dilution (the comparison is given in the inset on the right). B. The analytical SEC profile of the apparently monomeric fraction obtained during preparative SEC. The color of the preparation and its visible absorbance spectrum are given in the inset.

The behavior of the recombinant CTDH from *T. elongatus* (TeCTDH), which represents the other phylogenetic clade of CTDHs (clade 1) containing a Phe residue in position homologous to Cys103 of clade 2 [40], was similar to that of AnaCTDH. Even in the absence of Cys103, the TeCTDH apoform gave dimeric and monomeric peaks on the SEC profile showing two distinct bands on native-PAGE representing two species in equilibrium with each other (Fig. 3A).



**Figure 3 | Characterization of TeCTDH samples.** A. Preparative SEC on a Superdex 200 26/600 column of either apoform or holoform of TeCTDH produced in *E. coli* in the absence or presence of the carotenoids, respectively. The elution profile was run at a 1.6 ml/min flow rate while recording either protein- or carotenoid-specific absorbance (wavelengths are indicated). The insert shows the electrophoretic mobility of the apoform and the holoform of TeCTDH under non-denaturing conditions. B. Analytical SEC profiles of different fractions collected during the preparative SEC run. Forty  $\mu$ l of each fraction was loaded on a Superdex 200 Increase 5/150 column and run at 0.45 ml/min flow rate while following either protein- or carotenoid-specific absorbance (wavelengths are indicated). The apparent Mw values corresponding to the main peaks were obtained from column calibration (indicated in kDa). C. Absorbance spectra of the two representative fractions of TeCTDH containing predominantly CAN (BV) or ECN (RV). Note the presence of the  $\sim$ 407 nm peak characteristic of only the ECN-bound form of CTDH. D. The analytical SEC profile of the apparently monomeric carotenoid-containing fraction of TeCTDH obtained during preparative SEC. The color of the preparation and its visible absorbance spectrum are given in the inset.

Like AnaCTDH, TeCTDH expressed in *E. coli* in the presence of ECN/CAN yielded on the SEC profile the main overlapping peaks corresponding to CAN- and ECN-containing dimers with the different apparent size (Fig. 3B) and ~30 nm red-shifted visible absorbance spectra (Fig. 3C). The ECN-containing peak represented a dimer-monomer mixture sensitive to dilution, and a stable fraction of carotenoid-carrying TeCTDH monomers with an absorbance maximum at ~506 nm could be isolated (Fig. 3D).

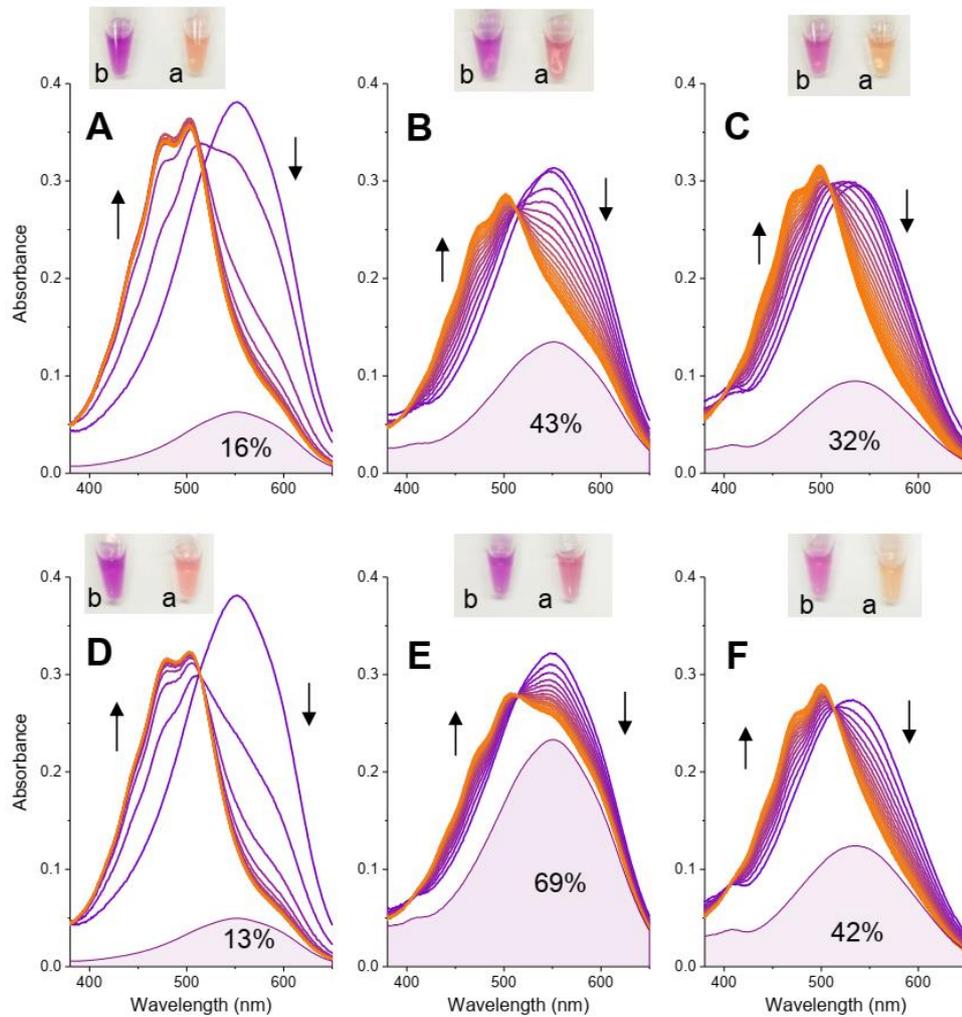
Thus, the two CTDH proteins from different phylogenetic clades (1 and 2) [40] share the ability to coordinate not only di-ketolated CAN but also mono-ketolated ECN, by forming dimeric holofoms with substantially different stability and oligomerization behavior. Both paralogs are able to form monomeric species bearing carotenoids, which are spectrally characterized by the absence of vibronic structure and a significantly blue-shifted maximum compared to other OCP-related proteins described.

### ***Direct carotenoid transfer from CTDH to OCP***

Previous work showed that, when expressed in the *E. coli* strain producing a mixture of ECN and CAN [18], OCP-CTD binds almost exclusively CAN [10] and can transfer it to the OCP apoform [9]. Likewise, AnaCTDH was shown to transfer CAN to the apoforms of OCP, NTD or HCPs, which may be negatively regulated by the formation of the disulfide bridge involving Cys103 [40]. Here, by using fully reduced protein preparations, we found that not only AnaCTDH(CAN) but also AnaCTDH(ECN) was able to transfer the embedded carotenoid to the OCP apoforms of both *Synechocystis* and *Anabaena* (SynOCP and AnaOCP). In both cases, this led to substantial spectral changes and occurrence of pronounced vibronic structure, as typical for OCP<sup>O</sup> (Fig. 4). In contrast to OCP-CTD, which under our conditions transferred ~84% of carotenoid to SynOCP and ~87% to AnaOCP, AnaCTDH was found to be a much less efficient carotenoid donor to both SynOCP [57% of CAN is transferred by AnaCTDH(CAN) and 68% of ECN is transferred by AnaCTDH(ECN)] and AnaOCP [only 31% of CAN is transferred by AnaCTDH(CAN) and 58% of ECN is transferred by AnaCTDH(ECN)] (Fig. 4).

In agreement with the SEC data, the lower stability of AnaCTDH(ECN) dimers may explain the more efficient carotenoid transfer to both OCPs compared to AnaCTDH(CAN) (Fig. 4). The higher transfer efficiency from

OCP-CTD(CAN) (also called COCP [10]) suggests that the specificity of NTD•CTD (or NTD•CTDH) interactions may be important for the carotenoid transfer and that the AnaCTDH(CAN) dimer might have a larger affinity for the carotenoids than OCP-CTD(CAN). This prompted us to test whether AnaCTDH can take up the carotenoid directly from OCP-CTD(CAN).



**Figure 4 | Direct carotenoid transfer:** from OCP-CTD(CAN) (A, D), AnaCTDH(CAN) (B, E), or AnaCTDH(ECN) (C, F) to the apoforms of either SynOCP (A, B, C) or AnaOCP (D, E, F) followed by changes of absorption spectra at 20 °C during 3.5 h recorded with 10 min intervals. The protein concentration of both, carotenoid donors and acceptors, was 20  $\mu$ M. Insets show natural colors of the samples before (b) or after (a) the transfer completion. Note that the process reached equilibrium in each case as no changes occurred after overnight incubation. Decomposition of the resulting spectra using the initial donor spectra and a spectrum of OCPO as a reference allowed us to estimate the fraction of the initial carotenoid-bound species (shown with the filled area under the curve and the percentage of remaining carotenoid donors).

### ***“Horizontal” carotenoid transfer from OCP-CTD to CTDH***

Incubation of *Synechocystis* OCP-CTD(CAN) with AnaCTDH(apo) resulted in an efficient color change from red-violet, typical of OCP-CTD(CAN) [10], to blue-violet, typical of AnaCTDH(CAN) (Fig. 1), indicating at least partial CAN translocation from OCP-CTD to AnaCTDH. This carotenoid transfer between homologous proteins, which we tentatively call “horizontal”, was accompanied by a significant shift of the absorbance maximum from ~552 to ~569 nm (Fig. 5A). Recombinant OCP-CTD from *Anabaena* was also able to transfer the carotenoids to AnaCTDH, producing nearly identical spectral changes, with the final spectrum being equivalent to that of AnaCTDH(CAN) (Fig. 5B).

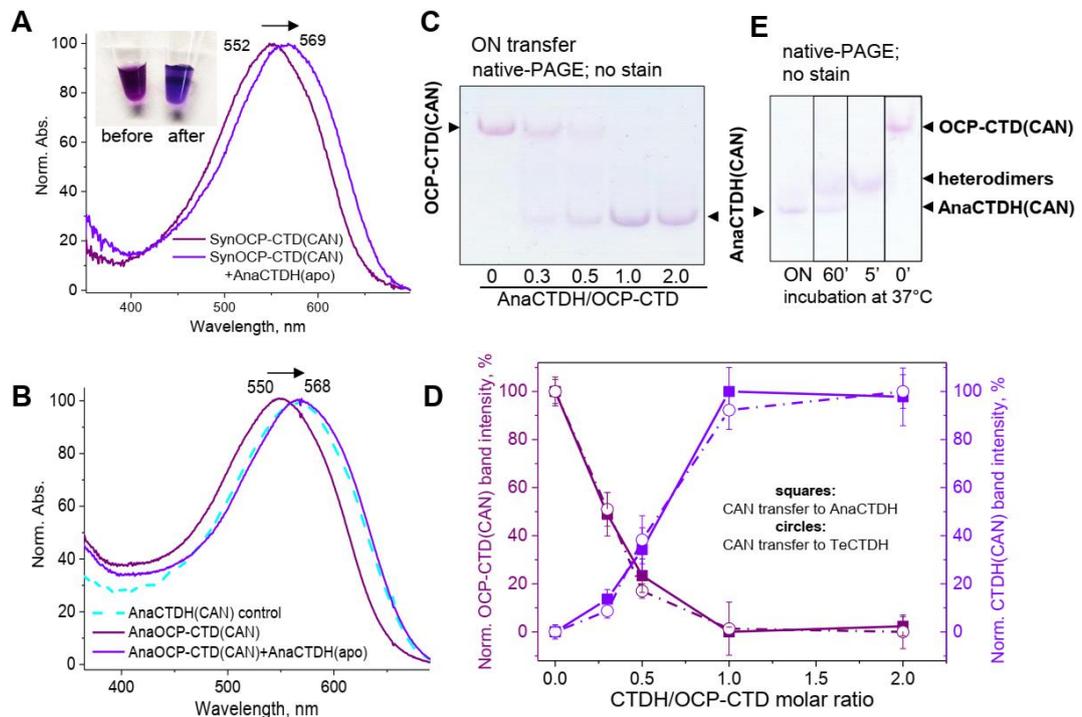
The horizontal carotenoid transfer could be followed by native-PAGE without requiring any staining, owing to the difference in electrophoretic mobility of the carotenoid donor and acceptor (Fig. 5C). Indeed, titration of a fixed OCP-CTD(CAN) concentration (111  $\mu$ M per monomer) by increasing quantities of AnaCTDH(apo) resulted in a disappearance of the colored OCP-CTD(CAN) band with a concomitant appearance of the colored AnaCTDH(CAN) band, which differed by a tint of violet (Fig. 5C). Already at a 1:1 molar ratio and equilibrium conditions (>24 h incubation), we could observe almost quantitative carotenoid transfer from OCP-CTD(CAN) to AnaCTDH and to TeCTDH, demonstrating in both cases nearly identical titration curves (Fig. 5D).

In the course of the transfer, we could observe the appearance of two colored bands of higher mobility on native PAGE gels, suggesting the presence of intermediates, presumably, OCP-CTD(CAN)•AnaCTDH heterodimers, which appeared within minutes after mixing of the samples and disappeared during further incubation time (Fig. 5E). The principal existence of such heterodimers of intermediate mobility sharing a carotenoid molecule implied that the CTD(H)•CTD(H) interactions may mediate other types of carotenoid transfer between related proteins, as an alternative to the NTD•CTD(H) interactions [9, 40].

### ***Inverse transfer of carotenoid to CTDH***

Previously, it was shown that *Synechocystis* OCP-CTD can transfer its CAN to the NTD apoform, resulting in the formation of the metastable photoactive OCPO-like species [10], whereas recombination of the holoforms of NTD (or HCP) with the apoforms of OCP-CTD (or CTDH) did not work

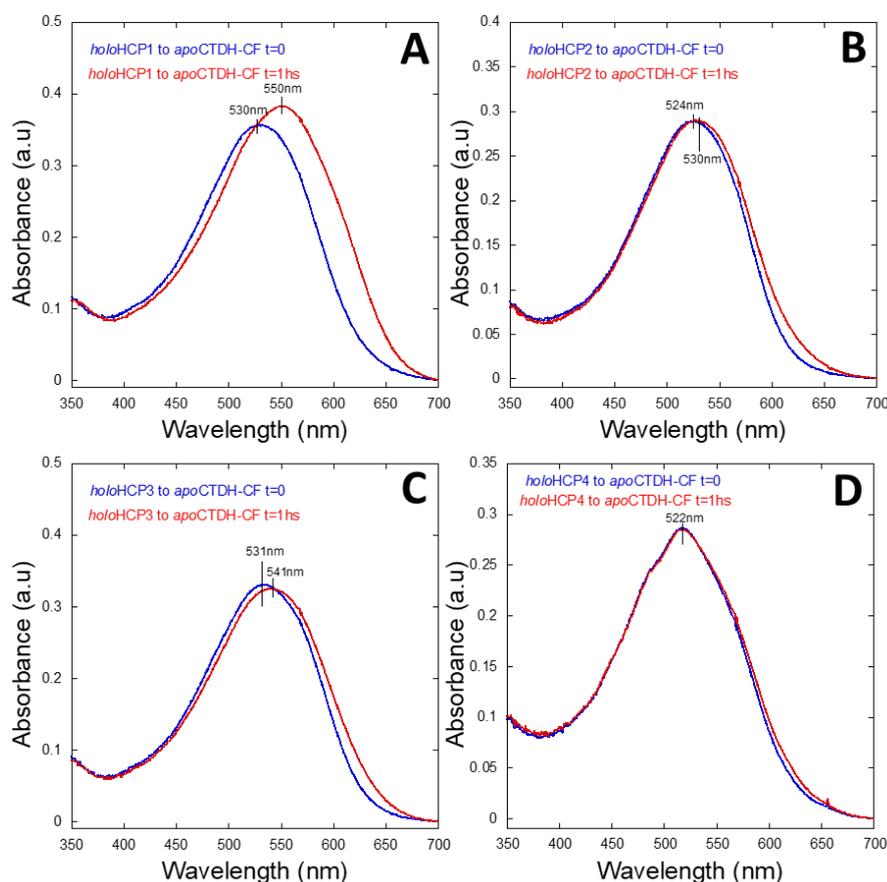
[9, 10, 40], suggesting that carotenoid transfer is a unidirectional process from CTD(H) to NTD (or HCP) and full-length OCP.



**Figure 5 | Horizontal carotenoid transfer from OCP-CTD(CAN) to AnaCTDH(apo).** A. Normalized absorption spectra of the SynOCP-CTD(CAN) sample and the product of carotenoid transfer to AnaCTDH(apo). The inset shows the natural color of the samples. B. Equivalent spectra for AnaOCP-CTD(CAN) used as the carotenoid donor to AnaCTDH(apo). The spectrum of AnaCTDH(CAN) is given as the control. C. Analysis of the horizontal carotenoid transfer at various AnaCTDH/OCP-CTD(CAN) molar ratios (indicated below) by native-PAGE without staining. The gel was run after the completion of the transfer (ON, overnight) and thus represent near equilibrium conditions. D. The same experiment was done for the carotenoid transfer to TeCTDH(apo) and both gels were analyzed by quantitative densitometry using ImageJ 1.48. Error bars represent standard deviation ( $n = 3$ ). E. Analysis of the horizontal carotenoid transfer from SynOCP-CTD(CAN) to AnaCTDH (111  $\mu\text{M}$  per monomer each) at 37 °C at different time points by means of the unstained native-PAGE. The main bands are marked by arrowheads. Note the intermediate band appearing very early upon mixing OCP-CTD(CAN) and AnaCTDH and putatively corresponding to their heterodimers sharing the carotenoid (ON – overnight incubation).

While the physiological relevance of the horizontal carotenoid transfer discovered here is unclear yet, the possibility of carotenoid translocation from OCP-CTD to AnaCTDH strongly suggested high affinity of the latter to carotenoids and supported the idea that carotenoid transfer can occur in many directions, involving both, NTD•CTD(H) and CTD•CTD(H) interactions. In line with the results already obtained with HCP4 (see chapter I), we found that neither AnaHCP2 nor AnaHCP3 are capable of significant carotenoid transfer to CTDH. However, AnaHCP1 (of unknown function at that moment) can transfer the carotenoid to AnaCTDH(apo) (result adapted from

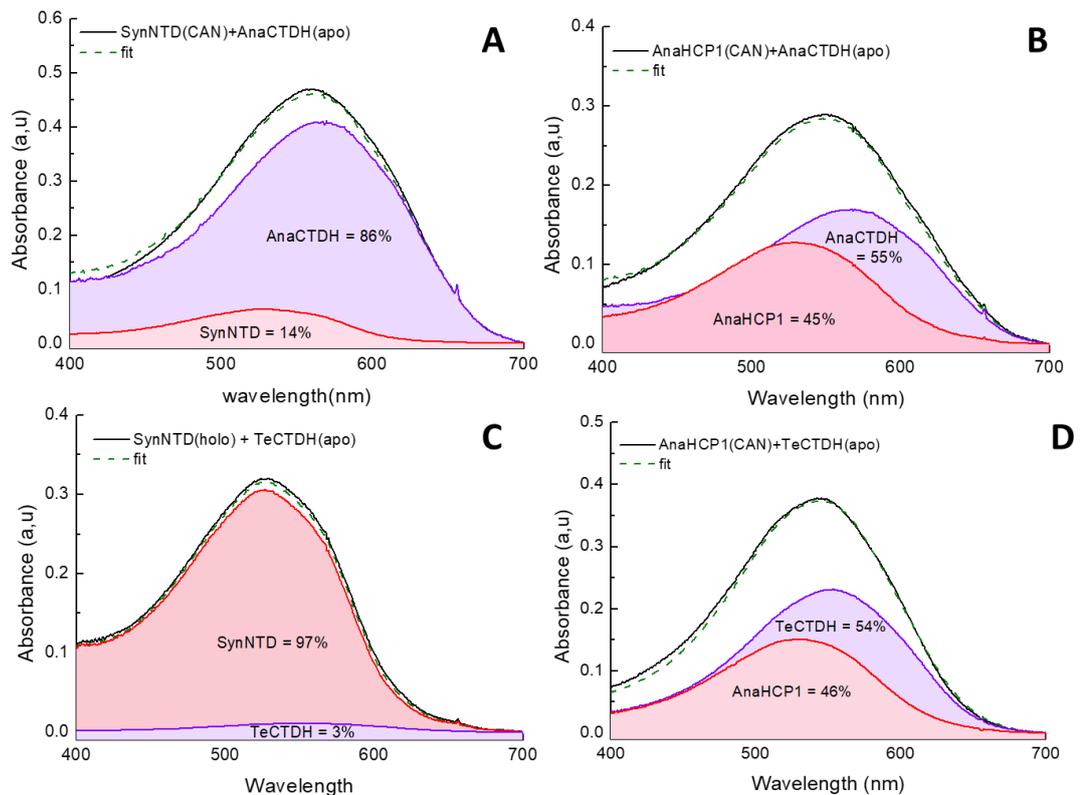
Harris et al 2016 [43], figure 6). Resembling the HCP4, HCP1 was unable to extract the carotenoid from the membrane, and reconstitution of HCP1(CAN) *in vitro* (from apoprotein and carotenoid containing membranes) is only possible in the presence of AnaCTDH(apo) (figure supplementary S1). This suggested that AnaHCP1 may serve as an additional carotenoid carrier. Therefore, we questioned whether the apparently high affinity of AnaCTDH may allow it to retrieve the carotenoid from other OCP-related proteins containing the NTD.



**Figure 6 | Carotenoid transfer from HCPs to Ana-CTDH.** Absorbance spectra at time=0 (blue spectrum) and after a one hour incubation (red spectrum) of AnaCTDH(apo) with (A) AnaHCP1(CAN), (B) AnaHCP2(CAN), (C) AnaHCP3(CAN) or (D) AnaHCP4(CAN). AnaHCPs(CAN) were present at concentration equivalent to OD=0.3 and the AnaCTDH was added in a 2-fold molar excess and incubated 1 hour at 23°C. Adapted from Harris et al. 2016 [43].

First, we examined the ability of CTDHs to retrieve the carotenoid directly from SynOCP-NTD and AnaHCP1 holoproteins. Under the conditions used, AnaCTDH extracted up to 86% of carotenoid from SynOCP-NTD and 55%

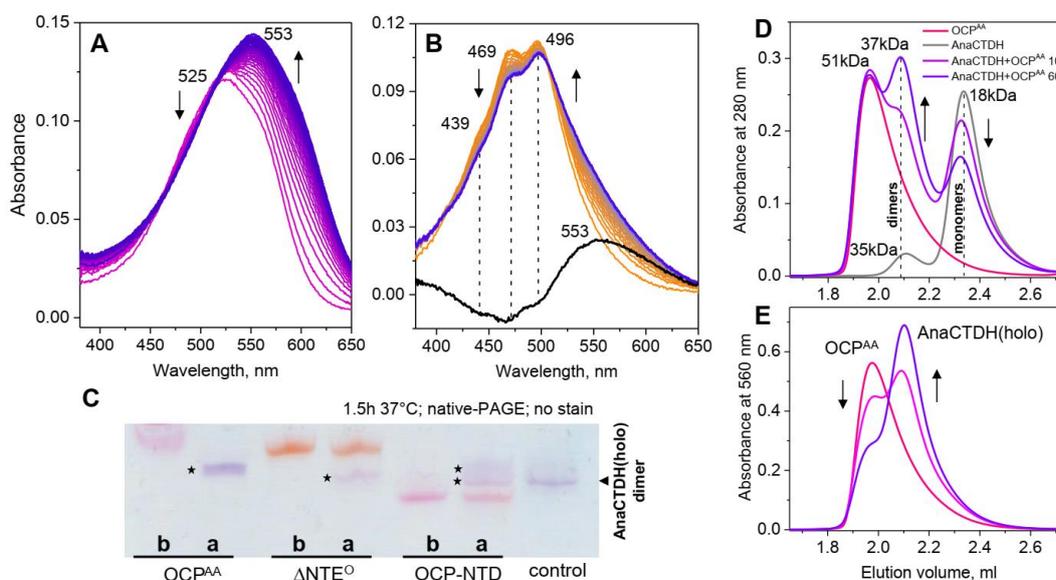
of carotenoid from AnaHCP1 (Fig. 7A, B), as judged from spectral decomposition, confirming the high carotenoid affinity of AnaCTDH and implying favorable NTD•CTDH interactions. Under identical conditions, TeCTDH extracted 55% of carotenoid from AnaHCP1, whereas on contrary, uptake from SynOCP-NTD was barely efficient (~3%) (Fig. 7C, D). These findings emphasize the notion that the transfer process is not only guided by the affinity for carotenoid but also by protein-protein interactions.



**Figure 7 | Inverse carotenoid transfer:** from *Synechocystis* OCP-NTD (A, C) or *Anabaena* HCP1 (B, D) to either *Anabaena* (A, B) or *T. elongatus* CTDH (C, D) followed by absorbance spectroscopy. The carotenoid donors (AnaHCP1, SynOCP-NTD) were present at concentrations equivalent to O.D. = 0.3 at 520 nm and AnaCTDH(apo) was added in a 2-fold molar excess to initiate the transfer. The samples were incubated at 23°C for 40 min. The spectra were deconvoluted in the sum of CTDH and HCP1(or NTD) spectra to quantify the contribution of each component.

Given the modular organization of OCP [36] and the idea of domain separation accompanying OCP photoactivation [18, 33, 47-49], we next questioned whether AnaCTDH can retrieve carotenoids from full-length OCP variants.

Indeed, upon mixing of AnaCTDH(apo) with the analog of photoactivated OCP with separated domains, the OCP<sup>AA</sup> mutant (Y201A/W288A) [27, 34], we observed pronounced changes of the absorption spectrum and a 30-nm shift of its maximum from ~525 nm, characteristic of OCP<sup>AA</sup>, to ~553 nm, characteristic of AnaCTDH(holo) (Fig. 8A). Spectral decomposition using the reference spectra of AnaCTDH(CAN) and OCP<sup>AA</sup> indicated that, under these conditions, up to ~80% of the carotenoid was transferred in the opposite direction compared to the previously characterized process (i.e., CTD/CTDH → HCP/NTD/OCP [9, 10, 40]). This number is close to the efficiency of the transfer from SynOCP-NTD to AnaCTDH(apo) (Fig. 7A), which indirectly indicates that in the OCP<sup>AA</sup> case, the carotenoid is extracted from the NTD, in line with the fact that upon OCP photoactivation, the carotenoid is translocated to the NTD [48-50]. The open conformation of OCP<sup>AA</sup> may favor protein-protein interactions and carotenoid transfer to AnaCTDH, guided by the affinity to a carotenoid.



**Figure 8 | Inverse transfer of carotenoid to AnaCTDH studied by absorption spectrometry (A, B), native-PAGE (C), and analytical SEC (D, E).** Carotenoid transfer from OCP<sup>AA</sup> (A) or  $\Delta$ NTE-OCPO (B) to AnaCTDH(apo) was followed at 30 °C during 4 h by recording absorption spectra each 6 min. AnaCTDH(apo) concentration was 16  $\mu$ M, concentrations of carotenoid donors were adjusted to a similar absorbance at 500-525 nm. Note that the process reached equilibrium in each case as no changes occurred after overnight incubation. Arrows indicate the direction of spectral changes and the black line in panel B represents the difference spectrum (final–start). The vibronic structure characteristic of OCPO in the  $\Delta$ NTE<sup>O</sup> case is marked by dashed lines with the position of the main bands indicated. The color gradient roughly corresponds to the sample colors. C. Carotenoid transfer from various NTD-containing OCP-related proteins to AnaCTDH monitored by native-PAGE without staining. The samples before (b) and after (a) the completion of the transfer (1 h 37 °C) were analyzed. The products/intermediates of the carotenoid transfer are

marked by asterisks. The position of the AnaCTDH(CAN) dimer band is shown by an arrowhead on the right. **D, E.** Inverse carotenoid transfer from OCP<sup>AA</sup> to AnaCTDH followed by subjecting 30  $\mu$ l aliquots of the reaction mixture to analytical SEC monitored by protein-specific (**D**) and carotenoid-specific (**E**) absorbance. The position of AnaCTDH dimers and monomers (dashed lines), apparent  $M_w$  of the peaks (in kDa), and the direction of changes in the course of carotenoid transfer (arrows) are indicated.

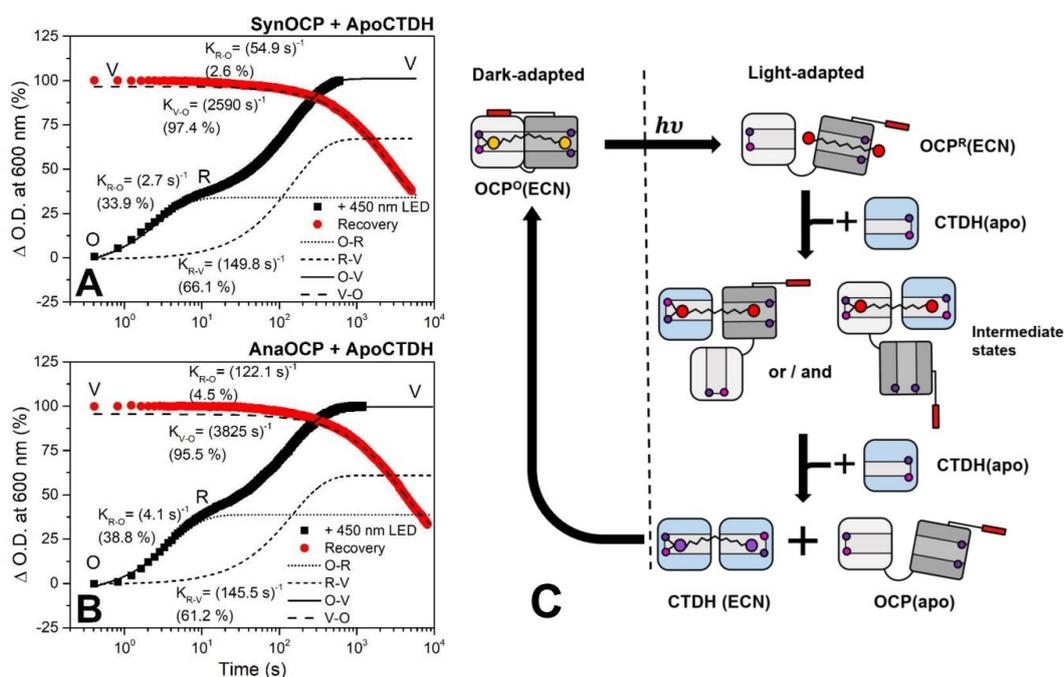
The inverse transfer of carotenoid to AnaCTDH(apo) could be observed not only from purple OCP<sup>AA</sup> but also from the orange compact OCP form lacking the N-terminal extension (NTE), i.e. from the  $\Delta$ NTEO variant, albeit with limited efficiency (15-20% according to spectral decomposition) (Fig. 7B). The product of carotenoid transfer, AnaCTDH(holo), gave a pronounced contribution to the absorbance at  $\sim$ 550 nm (Fig. 7B) and could be easily discriminated on an unstained native-PAGE (Fig. 7C). In Fig. 7C, carotenoid transfer from SynOCP-NTD to AnaCTDH is also shown for comparison. Notably, besides the band matching the AnaCTDH(holo) control, the band with lower electrophoretic mobility could also be detected, especially in the case of the carotenoid transfer from OCP-NTD, most likely representing products and intermediates of the reaction, respectively (marked by two asterisks in Fig. 8C).

The inverse transfer of carotenoid could also be readily monitored by SEC. For example, mixing of OCP<sup>AA</sup> and AnaCTDH(apo) resulted in an increase of the  $\sim$ 37 kDa peak corresponding to the carotenoid-bound AnaCTDH dimers and a concomitant decrease of the  $\sim$ 51 kDa peak of OCP<sup>AA</sup> (followed by 560 nm) and the  $\sim$ 18 kDa peak of the AnaCTDH(apo) (followed at 280 nm) (Fig. 8D, E), indicating physical carotenoid translocation from one protein to the other, in agreement with the data from native-PAGE.

### ***Photoinduced carotenoid shuttle between OCP and CTDH***

Since we found that mixing of AnaCTDH(apo) and the OCP<sup>AA</sup> mutant, which mimics the quenching-competent OCP<sup>R</sup> form [27, 34], resulted in carotenoid translocation into AnaCTDH, as seen from the appearance of violet forms, we questioned if this process could be activated by actinic light, which is known to trigger domain separation in wild-type OCP [18, 33, 47-49]. Mixing of wild-type OCP<sup>O</sup> with ApoCTDH(apo) had absolutely no effect on the absorption spectrum of the samples. This indicates that, as long as the OCP domains are stably closed, AnaCTDH cannot extract carotenoids, and the transfer occurs only upon destabilization of the compact OCP structure (either by the NTE removal or by mutations of the key Tyr/Trp residues in

the C-terminal domain coordinating the carotenoid [28, 34, 51]). We then tested whether carotenoid transfer to AnaCTDH(apo) could be induced by photoactivation of OCP<sup>O</sup> and performed experiments using SynOCP and AnaOCP holoproteins (Fig. 9). In both cases, after the completion of O-R conversion (it took ~10 s to reach the equilibrium), further slow changes of absorption in the red region of the spectrum were observed, associated with the accumulation of the violet forms in the course of the inverse carotenoid transfer process from OCP<sup>R</sup> to CTDH (Fig. 9, “O”, “R”, and “V” states). After ~30 minutes of continuous illumination, the absorption of the samples coincided with that of AnaCTDH(ECN) (see Fig. 1C), without any signatures of the orange or red OCP forms.



**Figure 9 | Photoinduced carotenoid shuttling between holofoms of SynOCP (A), AnaOCP (B) and AnaCTDH.** AnaCTDH(apo) was added into a solution of the dark-adapted OCP<sup>O</sup> holofoms to reach concentration ratio equal to 4 CTDH per single carotenoid molecule. Experiments were conducted at 30 °C, constant stirring and equivalent photon flux densities of actinic blue light (450 nm light-emitting diode). In order to show hierarchy of rate constants and better comparison of A and B, zero on the timescale corresponds to beginning of both - light activated inverse (black squares) and direct carotenoid transfer (red circles), which was observed after the end of illumination by actinic light. Dashed and dotted lines show the main individual components of kinetics, obtained by fitting of the data by decaying exponents. C. Schematic representation of the photoinduced carotenoid shuttle mechanism between OCP and CTDH. OCP<sup>O</sup> expressed in *E. coli* predominantly coordinates ECN [10]. Note that the process occurs via the intermediary formation of heterodimers. NTE is shown by a red rectangle, small violet and magenta circles designate key Trp and Tyr residues, respectively, involved in carotenoid-protein interactions by H-bonding to

carotenoid keto-groups. NTD and CTD are shown in grey and white, respectively, CTDH is shown in light blue.

Thus, *Anabaena* CTDH can gradually extract the carotenoid from the photoactivated forms of OCP as soon as the domains are separated; quantitatively similar results were obtained for *Synechocystis* OCP and *Anabaena* OCP. This directly suggests that potent modulatory effects are possible, when OCP and CTDH co-occur *in vivo*. Importantly, the process was reversible: after the actinic light was turned off, we observed a gradual recovery of the orange OCP forms as the result of the direct carotenoid transfer from AnaCTDH(holo) to Apo-OCPs (Fig. 9A and B, red circles).

## ***Discussion***

---

Lipophilic carotenoids are synthesized at and localized in membrane compartments, whereas OCP and related proteins are water-soluble carotenoproteins, which raises the question how carotenoid delivery to various destinations proceeds through aqueous media. Previous works have shown that the C-terminal OCP domain and its homologs, CTDHs, can efficiently retrieve carotenoids from membranes, and dimerize subsequently while encapsulating the carotenoid molecule. This suggests that these proteins can serve as carotenoid carriers to supply their cargo downstream to the carotenoid acceptors [9, 10, 40]. HCPs were proposed to be an example of such acceptors [40].

Here, we present results that combined with results from Harris et al. 2016 (carotenoid transfer from HCP1 to CTDH, also presented in this chapter) reveal the high complexity of the carotenoid transfer mechanism in cyanobacterial photoprotection. We demonstrated here that the HCP1 proteins are also capable of deliver its carotenoid to other proteins, such as the CTDHs. The carotenoid transfer between AnaCTDH and AnaHCP1 seems to be bidirectional, and the direction is determined by the initial (holo/apo) state of the proteins. However, HCP1 required the presence of AnaCTDH for holoprotein formation (at least *in vitro*). Thus, the carotenoid transfer between OCP-related proteins is a multidirectional process at least *in vitro*. Both, symmetric CAN- and asymmetric ECN-bound CTDH dimers, take part in the initially characterized “direct” carotenoid transfer to OCP from both *Synechocystis* and *Anabaena* similar to the CAN-bound OCP-CTD, leading to the formation of the photoactive OCP species (Fig. 4). However, the properties of the ECN- and CAN-bound CTDH dimers are

remarkably different in terms of spectral characteristics, stability and apparent size (Fig. 1), which can probably explain why ECN is transferred more efficiently (Fig. 4). The substantially lower stability and the concentration-dependent dissociation of the CTDH dimers coordinating asymmetric ECN with only one ketolated  $\beta$ -ionone ring resembled the lower stability of the OCP-CTDW288A mutant lacking one of the H-bond donor residues [9]. This corroborates the fact that the H-bonds to the conserved Tyr/Trp residues play a significant role in carotenoid-protein interactions. Thus, the nature of the carotenoid is an important factor affecting the direction and efficiency of carotenoid transfer as well as the CTDH dimer stability.

While the binding mode and the configuration of the bulky lipophilic carotenoid within the individually stable, but relatively small monomeric CTDH species is completely unclear, the very existence of the latter may imply roles as intermediates of the carotenoid uptake and delivery processes [43].

The occurrence of the so-called horizontal carotenoid transfer between CTD and its homologs, demonstrated here (Fig. 5), indicated that not only NTD-CTD(H) but also CTD(H)-CTD(H) interactions may mediate the translocation of carotenoid between OCP-related proteins. This, in turn, suggests that differences in the affinity of homologous proteins with similar three-dimensional structure [43] and, probably, in the interior of their carotenoid-binding pockets are important prerequisites of carotenoid transfer. The efficiency of carotenoid transfer seems to be also dependent on specific matching between the proteins involved. For example, TeCTDH was capable of carotenoid transfer from AnaHCP1, but not from SynOCP-NTD (Fig. 7), likely associated with the fact that OCP is absent from *T. elongatus* [12] and thus corresponding NTD-CTDH interactions may not have been preserved during the CTDH evolution. Structural elements responsible for such specificity remain to be determined but most likely expand beyond the previously considered difference in the  $\beta 5/\beta 6$  loop length or the presence of the CTDH clade-specific Cys103 residue [40, 43].

According to our observations, the “inverse” carotenoid transfer process, which leads to the formation of the violet-colored dimeric CTDH proteins, can efficiently occur not only from NTD or HCP1 but also from the full-length analog of the photoactivated OCP, OCP<sup>AA</sup> (Fig. 8A). The inability of CTDHs to retrieve the carotenoid from OCP<sup>O</sup> indicated that carotenoid transfer from OCP occurs via domain separation. Indeed, we found that carotenoid transfer from OCP to CTDHs can be reversibly triggered by OCP

photoactivation, which led to the carotenoid-stabilized CTD-CTD(H) dimerization (Fig. 9C). Thus, the affinity for carotenoids in different OCP forms is predefined in such a specific manner that allows carotenoid uptake from membranes by CTD(H), and delivery of the carotenoid into the full-length OCP, the latter retaining the carotenoid in the basal orange form, but readily and reversibly donating it to CTDHs upon photoactivation. Of note, interprotein carotenoid transfer processes described in this study undergo via formation of the intermediates (heterodimers) sharing the carotenoid molecule, before the final product is formed, which makes accurate quantitation a difficult task warranting further detailed thermodynamic analysis.

A reversible, light-controlled carotenoid translocation and carotenoid-mediated protein dimerization allow hypothesizing that a similar process may take place *in vivo* to modulate the efficiency of the OCP-regulated photoprotection by controlling the amount of the quenching OCP form in cyanobacteria expressing CTDH. This may be especially relevant for cyanobacterial strains harboring OCP and CTDH, but lacking FRP genes (See Table 1) [42] and warrants further detailed investigation of the expression levels of CTDH proteins relative to the levels of OCP and HCP. Furthermore, protein engineering can be used to modify the described photo-induced carotenoid transfer process for the future design of light-driven protein-protein interaction cascades and other synthetic biology applications.

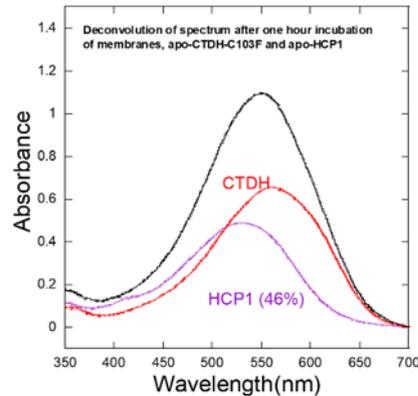
Species	OCP type	HCP type	CTDH	FRP
(i) <i>Gloebacter</i> PCC 7421	X	9	+	-
(i) <i>Gloebacter</i> JS1	X	9	+	-
(i) <i>Chamaesiphon</i> PCC 6605	X	1,2	+	-
(i) <i>Gloeocapsa</i> PCC 7428	2	1,2,3,4,6	+	-
(i) <i>Synechocystis</i> PCC 7509	2, X	1,2,3	+	-
(ii) <i>Pleurocapsa</i> PCC 7327	2, X	4, 5, 7	+	-
(ii) <i>Chroococciopsis</i> PCC 7203	X	1,2,3,4,6	+	-
(iii) <i>Microcoleus</i> FGP-2	X	3,8	+	-
(iii) <i>Oscillatoria</i> PCC 7112	X	3,8	+	-
(iii) <i>Microcoleus</i> PCC 9802	X	3,8	+	-
(iii) <i>Kamptonema</i> PCC 6506	X	3,8	+	-
(iii) <i>Kamptonema</i> PCC 6407	X	3,8	+	-
(iii) <i>Crinalium</i> PCC 9333	2	2,3,8	+	-
(iii) <i>Microcoleus</i> PCC 7113	X	3,8	+	-
(iv) <i>Scytonema</i> PCC 7110	X	1,2,3	+	-
(iv) <i>Rivularia</i> PCC 7116	X	1,4	+	-
(iv) <i>Tolypothrix</i> UTEX 2349	X	1,3,6	+	-
(iv) <i>Nodularia</i> CCY9414	X	4	+	-
(iv) <i>Microchaete</i> PCC 7126	X	1,3	+	-
(iv) <i>Calothrix</i> PCC 7507	X	1,3	+	-
(iv) <i>Nostoc</i> ATCC 29133	X	1,2,3,6	+	-
(iv) <i>Cylindrospermum</i> PCC 7417	X	3,6	+	-
(iv) <i>Anabaena</i> PCC 7122	X	3	+	-
(iv) <i>Anabaena</i> PCC 7108	X	3,4	+	-
(iv) <i>Trichormus</i> 708	X	1,3,4	+	-
(iv) <i>Dolichospermum</i> AWQC310F	X	4	+	-
(iv) <i>Dolichospermum</i> AWQC131C	X	4	+	-
(v) <i>Mastigocladopsis</i> PCC 10914	X	1,2,3,4,6	+	-
(v) <i>Fisherella</i> PCC 7521	2	2,4	+	-
(v) <i>Fisherella</i> JSK-11	2	4	+	-
(v) <i>Fisherella</i> PCC 7414	2	1,2,3,4	+	-

**Table 1 | Examples of cyanobacterial strains containing OCP (OCP2 or OCPX types [42]), HCP and CTDH genes but lacking FRP genes.** Cyanobacterial strains containing OCP1 type were not considered because, utilizing classical FRP regulated system, those strains either contain FRP or do not contain FRP, nor CTDH. Morphological subsections: (i) Chroococcales; (ii) Pleurocapsales; (iii) Oscillatoriales; (iv) Nostocales; (v) Stigonomatales. The data were taken from [42].

We anticipate that the principles of the protein-to-protein carotenoid transfer mechanism reported here may be generally applicable to other related processes involving water-soluble carotenoproteins, especially given the ubiquity of the NTF-2-like family of proteins. In the future, our findings may inspire the development of stable carotenoid nanocarriers for increased carotenoid resistance against photodestruction and for targeted antioxidant delivery, e.g. to retina tissue, where, in the macula lutea, an NTF-2-like STARD3 protein was shown to play a role in photoprotection thanks to its lutein-carrying ability [52].

## Appendix: supplementary figure

---



**Supplementary figure 1 | Carotenoid uptake from membranes.** Absorption spectrum of the supernatant after 1 h incubation of canthaxanthin containing *E. coli* membranes with apo-CTDH-C103F and apo-HCP1 (black). The membranes were removed before spectrum recording. Deconvolution of the graph into HCP1 (46%) and CTDH (54%) components is also shown. The ratio CTDH to HCP1 was 1 during the incubation. The experiments were done at least three times and protein concentrations were similar

## Materials and Methods

---

### Protein expression and purification

AnaCTDH and TeCTDH with the uncleavable His6-tag on the C-terminus were cloned into the pCDFduet-1 vectors [40], TeCTDH with the N-terminal His6-tag cleavable by 3C protease was cloned into the pQE81-M vector, Anabaena OCP with the N-terminal His6-tag cleavable by 3C protease was cloned into the pQE81-M vector. The identity of the constructs and the presence of mutations were verified by DNA sequencing. The obtained plasmids were used to transform chemically competent cells. Protein expression was induced using 0.5 mM isopropyl- $\beta$ -thiogalactoside (final concentration). Apoforms were either purified by isolation from apo-/holo-mixtures due to different hydrodynamic properties or expressed and purified from *E. coli* strains not producing carotenoids. Holoforms were expressed in echinenone (ECN) and canthaxanthin (CAN)-producing *E. coli* cells described in detail earlier [18]. Synthesis of ECN/CAN was achieved by means of the pACCAR25 $\Delta$ crtXZcrtO plasmid carrying the crtO gene encoding a  $\beta$ -carotene ketolase [18], yielding a mixture of both carotenoids. All His6-tagged proteins were purified by immobilized metal-affinity

(IMAC) and size-exclusion chromatography (SEC) to electrophoretic homogeneity and stored at 4 °C in the presence of 3 mM sodium azide [26-28, 51]. Where the construct allowed, the His-tag was removed by human rhinoviral 3C protease and subtractive IMAC [53].

Individual NTD and CTD of *Synechocystis* OCP devoid of the His6-tags as well as His6-tagged HCP1 or OCP-CTD from *Anabaena* were obtained as described in previous works [9, 26, 40], the same applies for the wild-type OCP from *Synechocystis* and its OCPAA and  $\Delta$ NTEO mutants [26, 51]. Each protein was purified in at least three individual batches with qualitatively similar properties.

Protein concentrations were determined at 280 nm using calculated protein-specific molar extinction coefficients. The holoprotein preparations obtained exhibited typical visible-to-UV absorption ratios of 1.6–1.8, indicative of a low apoprotein content. In case of COCP, AnaCTDH(holo) and TeCTDH(holo), this ratio achieved 2.2-2.6. Low apoform content could not be achieved in the case of OCP<sup>AA</sup> and CTDH(holo) monomers and visible-to-UV absorption ratio were ~0.5 and ~0.1, respectively.

### *Thin layer chromatography analysis of the carotenoid content*

Carotenoids were extracted from carotenoproteins by addition of a 2-fold volume excess of pure acetone. Samples clarified by centrifugation were subjected to thin layer chromatography on silicagel plates (Silufol, Kavalier, Czechoslovakia) using a mixture of petroleum ether (85% v/v) and acetone (15% v/v) for 15 min and the results were recorded immediately to prevent oxidation of carotenoids. Previous work reporting R<sub>f</sub> values for different carotenoids was used as a reference [46].

### *Native polyacrylamide gel-electrophoresis (PAGE)*

Protein samples were analyzed by electrophoresis in the glycine-Tris gel system under non-denaturing conditions as described earlier [28, 54]. The gels were run at 350 V and then scanned prior to and after Coomassie brilliant blue staining. Quantitative densitometry was performed using ImageJ 1.48. Each experiment was repeated at least twice, and the most representative results are demonstrated.

### *Analytical SEC*

The oligomeric state of various AnaCTDH and TeCTDH fractions obtained during preparative SEC on a Superdex 200 26/600 column (GE Healthcare, Pittsburg, PA, USA), concentration-dependent dissociation of AnaCTDH(ECN) as well as the results of the inverse carotenoid transfer from OCPAA to AnaCTDH(apo) were analyzed by SEC on a Superdex 200 Increase 5/150 columns (GE Healthcare) as described previously [10, 27, 28, 51]. The column was equilibrated with buffer SEC (20 mM Tris-HCl, pH 7.6, 150 mM NaCl, 0.1 mM ethylenediaminetetraacetate, 3 mM  $\beta$ -mercaptoethanol) and calibrated using bovine serum albumin (BSA) monomer (66 kDa), BSA dimer (132 kDa), BSA trimer (198 kDa), and  $\alpha$ -lactalbumin monomer (15 kDa). The elution profiles were followed by simultaneous UV and visible absorption at indicated wavelengths. Typical results obtained in at least three independent experiments are presented.

### *Absorption spectroscopy*

Steady-state absorption spectra and absorption time-courses at 550 and 600 nm were recorded as described earlier [34, 40, 43]. A blue light-emitting diode (M455L3, Thorlabs, Newton, NJ, USA), with a maximum emission at 455 nm was used for photoconversion of the samples (actinic light for  $\text{OCP}^{\text{O}} \rightarrow \text{OCP}^{\text{R}}$  photoconversion). The temperature of the samples was stabilized by a Peltier-controlled cuvette holder Qpod 2e (Quantum Northwest, USA) with a magnetic stirrer. Amplitudes of photoconversion and the corresponding rates were estimated according to procedures described earlier [33]. To estimate the efficiency of the carotenoid transfer under given conditions, spectral decomposition using reference spectra of carotenoid donors and carotenoid acceptors in the 100% holoform was performed in OriginPro 9.0 (OriginLab, Northampton, MA, USA) by fitting the corresponding contributions of the donor and acceptor spectrum to the spectrum obtained in the end of the mixing experiment. All experiments were repeated at least two times and the most typical results are presented.

### *References*

---

1. Vershinin, A. (1999) Biological functions of carotenoids--diversity and evolution, *Biofactors*. 10, 99-104.

2. Rodriguez-Concepcion, M., Avalos, J., Bonet, M. L., Boronat, A., Gomez-Gomez, L., Hornero-Mendez, D., Limon, M. C., Melendez-Martinez, A. J., Olmedilla-Alonso, B., Palou, A., Ribot, J., Rodrigo, M. J., Zacarias, L. & Zhu, C. (2018) A global perspective on carotenoids: Metabolism, biotechnology, and benefits for nutrition and health, *Progress in lipid research*. 70, 62-93.
3. Khoo, H. E., Prasad, K. N., Kong, K. W., Jiang, Y. & Ismail, A. (2011) Carotenoids and their isomers: color pigments in fruits and vegetables, *Molecules*. 16, 1710-38.
4. LaFountain, A. M., Prum, R. O. & Frank, H. A. (2015) Diversity, physiology, and evolution of avian plumage carotenoids and the role of carotenoid-protein interactions in plumage color appearance, *Arch Biochem Biophys*. 572, 201-212.
5. Gamiz-Hernandez, A. P., Angelova, I. N., Send, R., Sundholm, D. & Kaila, V. R. (2015) Protein-Induced Color Shift of Carotenoids in beta-Crustacyanin, *Angew Chem Int Ed Engl*. 54, 11564-6.
6. Ribeiro, D., Freitas, M., Silva, A. M. S., Carvalho, F. & Fernandes, E. (2018) Antioxidant and pro-oxidant activities of carotenoids and their oxidation products, *Food Chem Toxicol*. 120, 681-699.
7. Hashimoto, H., Uragami, C. & Cogdell, R. J. (2016) Carotenoids and Photosynthesis, *Sub-cellular biochemistry*. 79, 111-39.
8. Bhosale, P. & Bernstein, P. S. (2007) Vertebrate and invertebrate carotenoid-binding proteins, *Arch Biochem Biophys*. 458, 121-7.
9. Maksimov, E. G., Sluchanko, N. N., Slonimskiy, Y. B., Mironov, K. S., Klementiev, K. E., Moldenhauer, M., Friedrich, T., Los, D. A., Paschenko, V. Z. & Rubin, A. B. (2017) The Unique Protein-to-Protein Carotenoid Transfer Mechanism, *Biophys J*. 113, 402-414.
10. Moldenhauer, M., Sluchanko, N. N., Buhrke, D., Zlenko, D. V., Tavraz, N. N., Schmitt, F. J., Hildebrandt, P., Maksimov, E. G. & Friedrich, T. (2017) Assembly of photoactive orange carotenoid protein from its domains unravels a carotenoid shuttle mechanism, *Photosynth Res*. 133, 327-341.
11. Shih, P. M. (2015) Photosynthesis and early Earth, *Curr Biol*. 25, R855-9.

12. Boulay, C., Abasova, L., Six, C., Vass, I. & Kirilovsky, D. (2008) Occurrence and function of the orange carotenoid protein in photoprotective mechanisms in various cyanobacteria, *Biochim Biophys Acta.* 1777, 1344-54.
13. Karapetyan, N. V. (2007) Non-photochemical quenching of fluorescence in cyanobacteria, *Biochemistry (Mosc).* 72, 1127-35.
14. Kirilovsky, D. (2007) Photoprotection in cyanobacteria: the orange carotenoid protein (OCP)-related non-photochemical-quenching mechanism, *Photosynth Res.* 93, 7-16.
15. Kirilovsky, D. & Kerfeld, C. A. (2016) Cyanobacterial photoprotection by the orange carotenoid protein, *Nature plants.* 2, 16180.
16. Holt, T. & Krogmann, D. W. (1981) A carotenoid-protein from cyanobacteria, *Biochim Biophys Acta.* 637, 408-414.
17. Wu, Y. P. & Krogmann, D. W. (1997) The orange carotenoid protein of *Synechocystis* PCC 6803, *Biochim Biophys Acta.* 1322, 1-7.
18. Maksimov, E. G., Moldenhauer, M., Shirshin, E. A., Parshina, E. A., Sluchanko, N. N., Klementiev, K. E., Tsoraev, G. V., Tavraz, N. N., Willoweit, M., Schmitt, F. J., Breitenbach, J., Sandmann, G., Paschenko, V. Z., Friedrich, T. & Rubin, A. B. (2016) A comparative study of three signaling forms of the orange carotenoid protein, *Photosynth Res.* 130, 389-401.
19. de Carbon, C. B., Thurotte, A., Wilson, A., Perreau, F. & Kirilovsky, D. (2015) Biosynthesis of soluble carotenoid holoproteins in *Escherichia coli*, *Sci Rep.* 5, 9085.
20. Punginelli, C., Wilson, A., Routaboul, J. M. & Kirilovsky, D. (2009) Influence of zeaxanthin and echinenone binding on the activity of the orange carotenoid protein, *Biochim Biophys Acta.* 1787, 280-8.
21. Wilson, A., Ajlani, G., Verbavatz, J. M., Vass, I., Kerfeld, C. A. & Kirilovsky, D. (2006) A soluble carotenoid protein involved in phycobilisome-related energy dissipation in cyanobacteria, *Plant Cell.* 18, 992-1007.
22. Wilson, A., Punginelli, C., Gall, A., Bonetti, C., Alexandre, M., Routaboul, J. M., Kerfeld, C. A., van Grondelle, R., Robert, B., Kennis, J. T. & Kirilovsky, D. (2008) A photoactive carotenoid protein acting as light intensity sensor, *Proc Natl Acad Sci U S A.* 105, 12075-80.

23. Boulay, C., Wilson, A., D'Haene, S. & Kirilovsky, D. (2010) Identification of a protein required for recovery of full antenna capacity in OCP-related photoprotective mechanism in cyanobacteria, *Proc Natl Acad Sci U S A.* 107, 11620-5.
24. Gwizdala, M., Wilson, A. & Kirilovsky, D. (2011) In vitro reconstitution of the cyanobacterial photoprotective mechanism mediated by the Orange Carotenoid Protein in *Synechocystis* PCC 6803, *Plant Cell.* 23, 2631-43.
25. Gwizdala, M., Wilson, A., Omairi-Nasser, A. & Kirilovsky, D. (2013) Characterization of the *Synechocystis* PCC 6803 Fluorescence Recovery Protein involved in photoprotection, *Biochim Biophys Acta.* 1827, 348-54.
26. Sluchanko, N. N., Slonimskiy, Y. B., Shirshin, E. A., Moldenhauer, M., Friedrich, T. & Maksimov, E. G. (2018) OCP-FRP protein complex topologies suggest a mechanism for controlling high light tolerance in cyanobacteria, *Nat Commun.* 9, 3869.
27. Slonimskiy, Y. B., Maksimov, E. G., Lukashev, E. P., Moldenhauer, M., Jeffries, C. M., Svergun, D. I., Friedrich, T. & Sluchanko, N. N. (2018) Functional interaction of low-homology FRPs from different cyanobacteria with *Synechocystis* OCP, *Biochim Biophys Acta.* 1859, 382-393.
28. Sluchanko, N. N., Klementiev, K. E., Shirshin, E. A., Tsoraev, G. V., Friedrich, T. & Maksimov, E. G. (2017) The purple Trp288Ala mutant of *Synechocystis* OCP persistently quenches phycobilisome fluorescence and tightly interacts with FRP, *Biochim Biophys Acta.* 1858, 1-11.
29. Sutter, M., Wilson, A., Leverenz, R. L., Lopez-Igual, R., Thurotte, A., Salmeen, A. E., Kirilovsky, D. & Kerfeld, C. A. (2013) Crystal structure of the FRP and identification of the active site for modulation of OCP-mediated photoprotection in cyanobacteria, *Proc Natl Acad Sci U S A.* 110, 10022-7.
30. Thurotte, A., Bourcier de Carbon, C., Wilson, A., Talbot, L., Cot, S., Lopez-Igual, R. & Kirilovsky, D. (2017) The cyanobacterial Fluorescence Recovery Protein has two distinct activities: Orange Carotenoid Protein amino acids involved in FRP interaction, *Biochim Biophys Acta.* 1858, 308-317.
31. Kerfeld, C. A., Sawaya, M. R., Brahmandam, V., Cascio, D., Ho, K. K., Trevithick-Sutton, C. C., Krogmann, D. W. & Yeates, T. O. (2003) The crystal structure of a cyanobacterial water-soluble carotenoid binding protein, *Structure.* 11, 55-65.

32. Wilson, A., Kinney, J. N., Zwart, P. H., Punginelli, C., D'Haene, S., Perreau, F., Klein, M. G., Kirilovsky, D. & Kerfeld, C. A. (2010) Structural determinants underlying photoprotection in the photoactive orange carotenoid protein of cyanobacteria, *J Biol Chem.* 285, 18364-75.
33. Maksimov, E. G., Shirshin, E. A., Sluchanko, N. N., Zlenko, D. V., Parshina, E. Y., Tsoraev, G. V., Klementiev, K. E., Budylin, G. S., Schmitt, F. J., Friedrich, T., Fadeev, V. V., Paschenko, V. Z. & Rubin, A. B. (2015) The Signaling State of Orange Carotenoid Protein, *Biophys J.* 109, 595-607.
34. Maksimov, E. G., Sluchanko, N. N., Slonimskiy, Y. B., Slutskaya, E. A., Stepanov, A. V., Argentova-Stevens, A. M., Shirshin, E. A., Tsoraev, G. V., Klementiev, K. E., Slatinskaya, O. V., Lukashev, E. P., Friedrich, T., Paschenko, V. Z. & Rubin, A. B. (2017) The photocycle of orange carotenoid protein conceals distinct intermediates and asynchronous changes in the carotenoid and protein components, *Sci Rep.* 7, 15548.
35. Bandara, S., Ren, Z., Lu, L., Zeng, X., Shin, H., Zhao, K. H. & Yang, X. (2017) Photoactivation mechanism of a carotenoid-based photoreceptor, *Proc Natl Acad Sci U S A.* 114, 6286-6291.
36. Leverenz, R. L., Jallet, D., Li, M. D., Mathies, R. A., Kirilovsky, D. & Kerfeld, C. A. (2014) Structural and functional modularity of the orange carotenoid protein: distinct roles for the N- and C-terminal domains in cyanobacterial photoprotection, *Plant Cell.* 26, 426-37.
37. Lechno-Yossef, S., Melnicki, M. R., Bao, H., Montgomery, B. L. & Kerfeld, C. A. (2017) Synthetic OCP heterodimers are photoactive and recapitulate the fusion of two primitive carotenoproteins in the evolution of cyanobacterial photoprotection, *Plant J.* 91, 646-656.
38. Melnicki, M. R., Leverenz, R. L., Sutter, M., Lopez-Igual, R., Wilson, A., Pawlowski, E. G., Perreau, F., Kirilovsky, D. & Kerfeld, C. A. (2016) Structure, Diversity, and Evolution of a New Family of Soluble Carotenoid-Binding Proteins in Cyanobacteria, *Molecular plant.* 9, 1379-1394.
39. Lopez-Igual, R., Wilson, A., Leverenz, R. L., Melnicki, M. R., Bourcier de Carbon, C., Sutter, M., Turmo, A., Perreau, F., Kerfeld, C. A. & Kirilovsky, D. (2016) Different Functions of the Paralogs to the N-Terminal Domain of the Orange Carotenoid Protein in the Cyanobacterium *Anabaena* sp. PCC 7120, *Plant Physiol.* 171, 1852-66.
40. Muzzopappa, F., Wilson, A., Yogarajah, V., Cot, S., Perreau, F., Montigny, C., Bourcier de Carbon, C. & Kirilovsky, D. (2017) Paralogs of

the C-Terminal Domain of the Cyanobacterial Orange Carotenoid Protein Are Carotenoid Donors to Helical Carotenoid Proteins, *Plant Physiol.* 175, 1283-1303.

41. Sluchanko, N. N., Slonimskiy, Y. B. & Maksimov, E. G. (2017) Features of Protein-Protein Interactions in the Cyanobacterial Photoprotection Mechanism, *Biochemistry (Mosc).* 82, 1592-1614.

42. Kerfeld, C. A., Melnicki, M. R., Sutter, M. & Dominguez-Martin, M. A. (2017) Structure, function and evolution of the cyanobacterial orange carotenoid protein and its homologs, *The New phytologist.* 215, 937-951.

43. Harris, D., Wilson, A., Muzzopappa, F., Sluchanko, N. N., Friedrich, T., Maksimov, E. G., Kirilovsky, D. & Adir, N. (2018) Structural rearrangements in the C-terminal domain homolog of Orange Carotenoid Protein are crucial for carotenoid transfer, *Communications biology.* 1, 125.

44. Chabera, P., Fuciman, M., Hribek, P. & Polivka, T. (2009) Effect of carotenoid structure on excited-state dynamics of carbonyl carotenoids, *Phys Chem Chem Phys.* 11, 8795-803.

45. Britton, G., Liaaen-Jensen, S. & Pfander, H. (1995) *Carotenoids*, Basel, Boston, Berlin.

46. Grung, M., D'Souza, F. M. L., Borowitzka, M. & Liaaen-Jensen, S. (1992) Algal Carotenoids 51. Secondary carotenoids 2. *Haematococcus pluvialis* aplanospores as a source of (3S, 3'S)-astaxanthin esters, *Journal of Applied Phycology.* 4, 165-171.

47. Gupta, S., Guttman, M., Leverenz, R. L., Zhumadilova, K., Pawlowski, E. G., Petzold, C. J., Lee, K. K., Ralston, C. Y. & Kerfeld, C. A. (2015) Local and global structural drivers for the photoactivation of the orange carotenoid protein, *Proc Natl Acad Sci U S A.* 112, E5567-74.

48. Leverenz, R. L., Sutter, M., Wilson, A., Gupta, S., Thurotte, A., de Carbon, C. B., Petzold, C. J., Ralston, C., Perreau, F., Kirilovsky, D. & Kerfeld, C. A. (2015) A 12 angstrom carotenoid translocation in a photoswitch associated with cyanobacterial photoprotection, *Science.* 348, 1463-1466.

49. Maksimov, E. G., Sluchanko, N. N., Mironov, K. S., Shirshin, E. A., Klementiev, K. E., Tsoraev, G. V., Moldenhauer, M., Friedrich, T., Los, D. A., Allakhverdiev, S. I., Paschenko, V. Z. & Rubin, A. B. (2017) Fluorescent

Labeling Preserving OCP Photoactivity Reveals Its Reorganization during the Photocycle, *Biophys J.* 112, 46-56.

50. Konold, P. E., van Stokkum, I. H. M., Muzzopappa, F., Wilson, A., Groot, M. L., Kirilovsky, D. & Kennis, J. T. M. (2018) Photoactivation mechanism, timing of protein secondary structure dynamics and carotenoid translocation in the Orange Carotenoid Protein, *J Am Chem Soc.*

51. Sluchanko, N. N., Slonimskiy, Y. B., Moldenhauer, M., Friedrich, T. & Maksimov, E. G. (2017) Deletion of the short N-terminal extension in OCP reveals the main site for FRP binding, *FEBS Lett.* 591, 1667-1676.

52. Li, B., Vachali, P., Frederick, J. M. & Bernstein, P. S. (2011) Identification of StARD3 as a lutein-binding protein in the macula of the primate retina, *Biochemistry.* 50, 2541-9.

53. Alexandrov, A., Dutta, K. & Pascal, S. M. (2001) MBP fusion protein with a viral protease cleavage site: one-step cleavage/purification of insoluble proteins, *Biotechniques.* 30, 1194-8.

54. Schaub, M. C. & Perry, S. V. (1969) The relaxing protein system of striated muscle. Resolution of the troponin complex into inhibitory and calcium ion-sensitizing factors and their relationship to tropomyosin, *Biochem J.* 115, 993-1004.



## **Part II: Results**

---

### **Chapter III**

# **The CTDH structure: unraveling the role of the CTT in carotenoid translocation.**

## ***Summary of work***

---

This chapter is based on two articles:

*Harris Dvir, Adjele Wilson, Fernando Muzzopappa, Nikolai N Sluchanko, Eugene G Maksimov, Diana Kirilovsky, and Noam Adir. 2018. "Structural Rearrangements in the C-Terminal Domain Homolog of Orange Carotenoid Protein Are Crucial for Carotenoid Transfer." Communications Biology 1 (1): 125.*

*Harris Dvir, Muzzopappa Fernando, Glaser Fabian, Kirilovsky Diana and Noam Adir. 2019. "Carotenoid uptake, binding and delivery by water-soluble proteins in cyanobacteria". Submitted*

This chapter contains two articles, the first article presented (Harris et al 2018) is comprised of introduction, results (figures 1-5). The discussion of the first article are detailed together with the results. The second article (Harris et al. 2019) continues and complements the first part of this chapter and therefore include only results (figures 6-10) and discussion. Material and methods and references are merged. It also important to note that the carotenoid transfer from HCP1 to CTDH described in the previous chapter was first described in the first article (Harris et al. 2018) of this chapter.

## ***Contribution to this work***

In this work including to articles, I collaborate by isolating the AnaCTDH (C103F) and AnaCTDH $\Delta$ CTT mutants (the molecular cloning was performed by Adjelé Wilson) and performing their biochemical characterization as well as analyzing the interaction and the carotenoid transfer with/to the HCPs. In addition, I also prepared the plasmids to overexpress CTDH-L137D and CTDH-L128D, then I purified these mutants CTDHs and finally I realized their biochemical characterization.

## ***Resume of the article***

In this chapter, the structure of the apoform of CTDH from *Anabaena* had been solved. The sampled used for the crystallization contained dimers and big oligomers as it was shown in the first chapter for the AnaCTDH. To solve the structure our collaborators obtained crystals of the big oligomeric fractions. Then by adding 2 M urea and isolating different fractions through SEC, they obtained crystals of the dimer. In the first article, the structure of

the dimeric (2 M urea) was solved. In this structure, the CTDH is involved in two different types of dimer. A back-to-back dimer (A type), stabilized by interaction between the  $\beta$ -sheets, and a head-to-head dimer (F type). The F type is proposed as the main dimer formed in the holoprotein, since the carotenoid channels of both monomers are connected and the carotenoid tunnel formed can contain the cantaxanthin molecule (and the carotenoid is protected from the solvent). SAXS experiments indicated that the OCP-CTD(apo) forms A-type dimers in solution. The main difference between the CTDH structure and that of the CTD in the OCP<sup>o</sup> is the position of the C-terminal tail (CTT). In the CTD, the CTT adopts an 'open position' in which interact with the  $\beta$ -sheets. On the other hand, in the CTDH the CTT adopts an internal position which partially covers the carotenoid cavity.

In the second article, the dimeric CTDH structure was used as the molecular replacement (MR) model to calculate the electron density map of the big oligomer of CTDH and solve its structure. In this new structure, 12 CTDH subunits form the asymmetric unit. In this dodecamer, the CTT is found populating two different positions: one 'closed' similar to the present in the dimeric CTDH structure and another position, which is intermediary between the open and the closed conformation found in the OCP. Computational simulations suggested that the CTT is highly mobile. Based on the structure and the MD calculations, we proposed that the movement of the CTT between the open (or intermediary) form to the closed position is involved in the switching between the different dimeric forms in carotenoid translocation. To confirm these hypotheses, we constructed and isolated a CTDH without the CTT and two point mutants in the CTT ( L137D and L128D). CTT deletion inhibits carotenoid uptake from membranes (or HCP1) and also decreases carotenoid transfer from CTDH(CAN) to HCP4/1(apo). The CTT has a hydrophobic patch that could interact with the carotenoid. Mutation to negative residues in this patch (L137D) strongly affects the carotenoid uptake/transfer of the CTDH. Mutation in the L128, predicted by computational simulations to stabilize the back-to-back interaction and therefore the A-type dimer, also strongly decreases the carotenoid uptake.

Based on these results, we propose a working model for carotenoid uptake from the membranes and carotenoid transfer from CTDH to other proteins. During carotenoid uptake, the CTT is involved in carotenoid interaction facilitating the carotenoid extraction from membranes and it is also involved in the switching from A-type dimer to F-type dimer. In the carotenoid transfer from CTDH to other proteins, the CTT also plays a role by facilitating the translocation of the carotenoid to the HCP.

## ***Structural rearrangements in the C-terminal domain homolog of Orange Carotenoid Protein are crucial for carotenoid transfer***

---

Dvir Harris<sup>1,2</sup>, Adjele Wilson<sup>3</sup>, Fernando Muzzopappa<sup>3</sup>, Nikolai N. Sluchanko<sup>4,5</sup>, Thomas Friedrich<sup>6</sup>, Eugene G. Maksimov<sup>5</sup>, Diana Kirilovsky<sup>3</sup> & Noam Adir<sup>1,2</sup>.

### **Affiliations**

<sup>1</sup>Schulich Faculty of Chemistry, Technion, 3200003 Haifa, Israel.

<sup>2</sup>Grand Technion Energy Program (GTEP), Technion, 3200003 Haifa, Israel.

<sup>3</sup>Institute for Integrative Biology of the Cell (I2BC), CEA, CNRS, Université Paris-Sud, Université Paris-Saclay, 91198 Gif sur Yvette, France.

<sup>4</sup>A.N. Bach Institute of Biochemistry, Federal Research Center, “Fundamentals of Biotechnology” of the Russian Academy of Sciences, Moscow 119071, Russia.

<sup>5</sup>Department of Biophysics, Faculty of Biology, M.V. Lomonosov Moscow State University, Moscow 119992, Russia.

<sup>6</sup>Technical University of Berlin, Institute of Chemistry PC 14, Straße des 17. Juni 135, 10623 Berlin, Germany.

### **Abstract**

A recently reported family of soluble cyanobacterial carotenoproteins, homologs of the C-terminal domain (CTDH) of the photoprotective Orange Carotenoid Protein, is suggested to mediate carotenoid transfer from the thylakoid membrane to the Helical Carotenoid Proteins, which are paralogs of the N-terminal domain of the OCP. Here we present the three-dimensional structure of a carotenoid-free CTDH variant from *Anabaena* (Nostoc) PCC 7120. This CTDH contains a cysteine residue at position 103. Two dimer-forming interfaces were identified, one stabilized by a disulfide bond between monomers and the second between each monomer's  $\beta$ -sheets, both compatible with small-angle X-ray scattering data and likely representing intermediates of carotenoid transfer processes. The crystal structure revealed a major positional change of the C-terminal tail. Further mutational analysis revealed the importance of the C-terminal tail in both carotenoid uptake and

delivery. These results have allowed us to suggest a detailed model for carotenoid transfer via these soluble proteins.

## *Introduction*

---

Photosynthetic organisms, including cyanobacteria, utilize the energy of the sun to sustain themselves in a wide range of static and dynamically changing conditions. The photosynthetic process is initiated by light absorption by antenna complexes, which are the phycobilisomes<sup>1,2</sup> in cyanobacteria. Excess absorbed energy can become deleterious without proper regulation as overexcitation may cause damage to the photochemical reaction centers, leading to cell stress and eventually cell death<sup>3</sup>. Cyanobacteria have evolved several protection mechanisms including non-photochemical quenching, induced by the Orange Carotenoid Protein (OCP)<sup>4,5</sup>. OCP is a water-soluble, carotenoid-binding, 35 kDa protein found to induce dissipation of phycobilisome excitation energy minimizing energy flux toward the reaction centers<sup>5,7</sup>. OCP can also directly protect against the presence of singlet oxygen<sup>8,9</sup>. OCP is comprised of two domains: the N-terminal domain (NTD) and the C-terminal domain (CTD) connected through an inter-domain flexible linker loop<sup>9</sup>. In darkness and low light, OCP is in its stable inactive orange state (OCP<sup>o</sup>). Upon strong illumination, changes in the carotenoid and the protein lead to the formation of the metastable, active red state (OCP<sup>R</sup>) in which CTD and NTD are completely separated<sup>6</sup>. Only the OCP<sup>R</sup> state interacts with the phycobilisome to induce fluorescence quenching through its NTD<sup>10,11</sup>. The CTD also interacts with the fluorescence recovery protein, which supports the reorganization of OCP<sup>R</sup> to OCP<sup>o</sup><sup>12-14</sup>. The structure of OCP<sup>o</sup> (PDB code: 5UI2) demonstrated that the carotenoid is situated in a hydrophobic cavity formed by both the CTD and the NTD<sup>9,15-18</sup>.

Although the structure of the OCP<sup>R</sup> state remains elusive to date, the structure of the isolated NTD with bound carotenoid (holo-NTD) was determined (PDB code: 4XB4)<sup>16</sup>. This holo-NTD structure revealed that, upon photoactivation, the carotenoid becomes more planar and is translocated by 12 Å into the NTD cavity<sup>16</sup>. The crystal structure of a carotenoid-free CTD (apo-CTD) has not yet been determined. Phylogenetic studies discovered that cyanobacteria contain genes that encode for homologs of both the NTD and CTD<sup>19-21</sup>. NTD homologs have been termed Helical Carotenoid Proteins

(HCPs) for their all  $\alpha$ -helical structure<sup>19,22</sup>. *Anabaena* (*Nostoc*) PCC 7120 (hereafter *Anabaena*) possesses four variants of HCPs (out of nine total clades identified in cyanobacteria<sup>19</sup>). The holo-HCP1 was isolated from *Anabaena* and its structure was solved, proving HCPs to be carotenoid-binding proteins<sup>19</sup>. Its structure exhibited high similarity to the NTD of OCP, which can individually exist as a proteolytic cleavage product of OCP<sup>10</sup>. Previous studies have shown that each of the HCP clades demonstrates different features<sup>22</sup>. Unlike the wide variety of HCPs, the CTD homologs (CTDH) fall only into two clades, of which only one is present in *Anabaena* (a clade 2 CTDH). The CTDH gene in *Anabaena* is located adjacent to the HCP4 gene, therefore it was suggested that HCP4 and CTDH are the progenitors of OCP<sup>19,23,24</sup>.

Recently, it was revealed that isolated CTD of *Synechocystis* OCP and CTDHs from *Anabaena* and *Thermosynechococcus elongatus* (a clade 1 CTDH) are also capable of carotenoid binding<sup>25,26</sup>. Both CTD and CTDH were shown to mediate carotenoid transfer to apo-OCP, apo-NTD, or apo-HCP (from clade 4)<sup>25,26</sup>. OCP and CTDH are also capable of taking up a carotenoid molecule from membranes while HCP4 and isolated NTD are unable to do so<sup>25</sup>. Thus, one likely role of CTDHs is to serve as intermediate transfer molecules required for delivery of carotenoids from membranes to the HCPs. *Anabaena* CTDH was shown to form a disulfide-linked dimer in the presence or absence of carotenoid under oxidizing conditions. By contrast, the CTDH from *T. elongatus* (hereafter TeCTDH) that lacks Cys103 forms a dimer only in the presence of a carotenoid that is shared by the two monomers<sup>25</sup>. When stabilized by the presence of the disulfide bond, the CTDH dimer could not donate its carotenoid to an acceptor molecule; the reduced form of holo-AnaCTDH was able to transfer its carotenoid to HCP. This suggests that the disulfide bond containing clade 2 CTDH dimer has additional regulation and may be controlled by changes in the redox state in the cell, functioning upon stress induction.

We present here three important findings on the carotenoid transfer mechanism by CTDH. First, there is a major structural shift of the C-terminal tail of the CTDH (with respect to its position in the holo-form of CTD-OCP), bringing it into close proximity to the carotenoid-binding pocket. Second, we show that the C-terminal tail has a strong positive impact on both the carotenoid uptake as well as the carotenoid delivery by the CTDH; when absent, the rates of these processes are considerably reduced. Finally, the role of CTDH as carotenoid donor to HCPs was confirmed. Surprisingly, apo-AnaCTDH was shown to be also capable of receiving a carotenoid molecule

from HCPI1, suggesting an increased complexity for carotenoid transfer between different partners.

## ***Results***

---

### ***The AnaCTDH can form a large homogeneous oligomer***

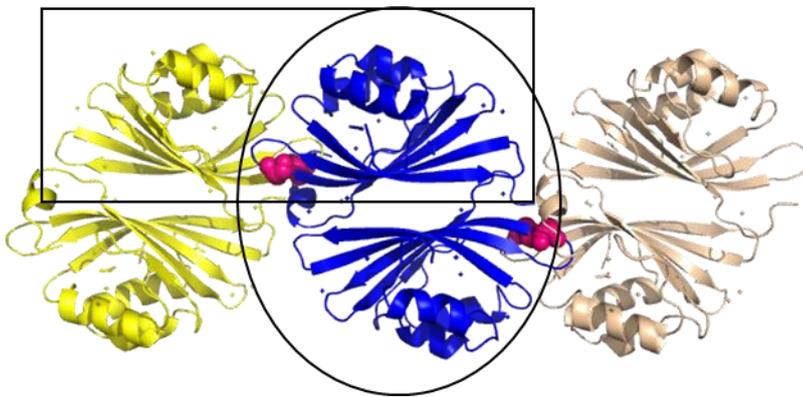
*Anabaena* apo-CTDH was prepared for crystallization as described in Muzzopappa et al. (2017)<sup>25</sup>. Native-polyacrylamide gel electrophoresis (native-PAGE) indicated that the major assembly is a dimeric form, suggesting the presence of a disulfide bond between the CTDH monomers (Supplementary Figure 1A). Under reducing conditions, the CTDH appeared as a monomer (Supplementary Figure 1B). Size exclusion chromatography revealed that the CTDH protein in solution was mostly in a higher oligomeric state (Supplementary Figure 2, dashed bold line) with additional, smaller oligomeric states in solution. This discrepancy concerning AnaCTDH oligomeric state between native-PAGE and size exclusion chromatography has previously been published<sup>25</sup>.

Crystallization trials were attempted on this apparent mixture of forms and large crystals were obtained. Using a laboratory X-ray source, the crystals diffracted to 2.9 Å. Indexing revealed a rather large unit cell (Table S1) with respect to the molecular weight (MW) of the AnaCTDH dimer (34 kDa). The calculated Matthews coefficient indicated that the asymmetric unit could contain 10–24 AnaCTDH monomers.

Whether this large oligomerization state has any biological significance as a quaternary assembly has to be further investigated. The fact that it was successfully crystallized and diffracted well suggests that this state is not a random aggregation and that different faces of interaction between the monomers (or between dimers) must exist. The presence of high MW oligomers was also detected in carotenoid-containing versions of AnaCTDH, TeCTDH, and *Synechocystis* CTD-OCP preparations obtained in *Escherichia coli* cells<sup>25</sup>, suggesting that oligomerization is a more general characteristic of CTD-like proteins. Nevertheless, these large oligomers were not detected in size exclusion chromatography when the holo-proteins were isolated from *Synechocystis* cells or when AnaCTDH was isolated under reducing conditions or in a AnaCTDH Cys103Phe mutant<sup>25</sup>.

### *Crystallization of apo-AnaCTDH dimer*

In order to obtain a more homogeneous preparation of apo-AnaCTDH dimers, we added 2 M urea to the solution<sup>27,28</sup>, which caused a large proportion of the high oligomeric states to disassemble, resulting in predominantly dimeric CTDH (Supplementary Figure 2, solid line). Crystals were obtained in the presence of urea (see Methods). The AnaCTDH(apo) dimer crystals were found to diffract to 2.43 Å with synchrotron radiation (Table S1).

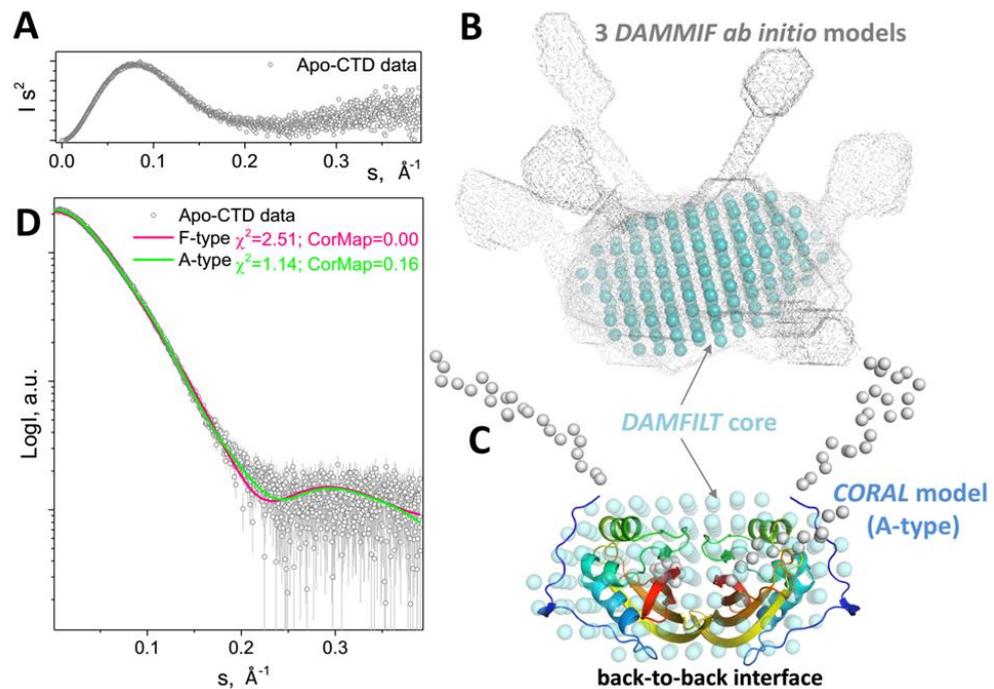


**Figure 1 | Two dimeric interfaces are revealed in the CTDH structure.** Three adjacent asymmetric units are depicted in yellow, blue and wheat. The C103 residues forming the disulfide bonds are presented as pink spheres. The asymmetric unit contains a “back-to-back” dimer interface (called A-type; black oval). The disulfide bonds link the “head-to-head” functional dimer (F-type; black rectangle).

### *Determination of the urea-treated dimeric apo-AnaCTDH structure*

The X-ray crystallographic structure of AnaCTDH(apo) was determined by molecular replacement (MR) to 2.75 Å resolution using the OCP<sup>O</sup>-CTD as the search model (from the 5UI2 structure), with two monomers in the asymmetric unit. Somewhat surprisingly, the disulfide bond between Cys103 residues of two monomers mentioned above was found to be between two different asymmetric units and not between the two monomers of one asymmetric unit. The asymmetric unit dimer (hence called A-type or back-to-back) was formed via an interaction plane between the β-sheets of two monomers, quite distant from Cys103 (Fig. 1). The presence of the oxidized disulfide bond was confirmed by the presence of continuous electron density between the two Cys103 residues in the Fo-Fc omit map (Supplementary Figure 3).

Analysis using the PISA macromolecular interaction tool<sup>29</sup> showed that both the A-type dimer and the functional disulfide-containing dimer (hence called an F-type or head-to-head) exhibited negative calculated  $\Delta G$  values, predicted to be stable. PISA suggested that only the F-type dimer has biological significance, ( $\Delta^i G_{P \text{ value}}$  of 0.067 for the F-type dimer interface). This finding agrees with the suggestion proposed by Moldenhauer et al. (2017)<sup>26</sup> and Muzzopappa et al. (2017)<sup>25</sup> that the interaction plane between the CTD(H) dimer should be at the same site as the CTD and NTD in OCP<sup>o</sup>. In contrast, a  $\Delta^i G_{P \text{ value}}$  of 0.518 in the case of the A-type dimer interface implies possibly crystal-stabilized interface (as a  $\Delta^i G_{P \text{ value}} < 0.5$  is considered necessary for biological significance<sup>29</sup>). However, this interface might still be important for the assembly into large oligomers in solution, which was shown to be condition-dependent<sup>25</sup> as well as a possible different type of dimeric interaction under some specific conditions.



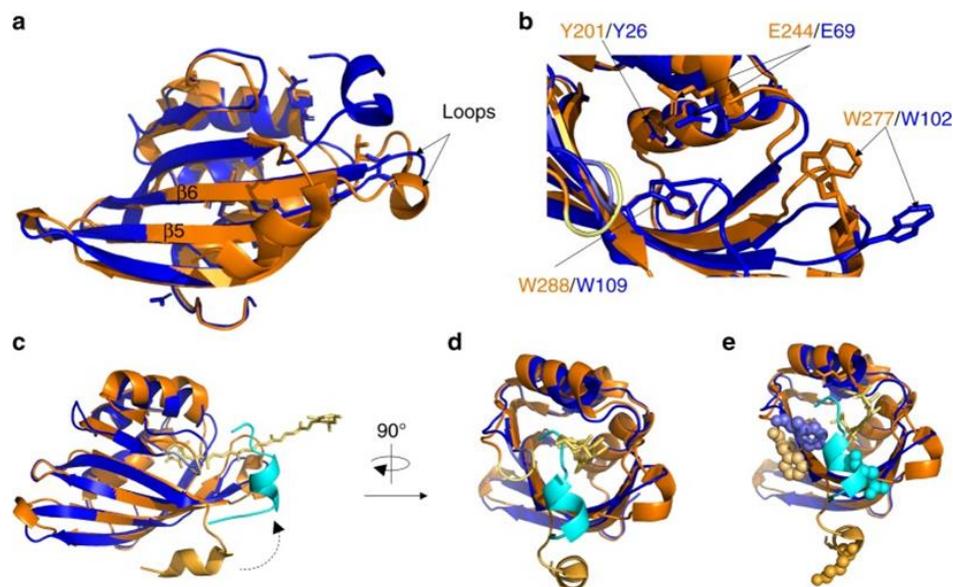
**Figure 2 | Analysis of the solution conformation of *Synechocystis* Apo-CTD dimers using SAXS.** (A) Kratky plot showing that Apo-CTD is represented by particles with a globular core and flexible termini. (B) Ab initio shape reconstruction using DAMMIF. Three best fitting DAMMIF models are shown superimposed to reveal the common core (DAMFILT core, cyan spheres). (C) CORAL-derived structural model of the A-type Apo-CTD dimer obtained upon modeling of the flexible N- (23 residues) and C-terminal (12 residues) tails in order to minimize the discrepancy between the experimental and theoretical SAXS curves calculated from the model. The CORAL model overlaid with the DAMFILT core (cyan spheres) from panel B is shown by ribbons colored by gradient from blue (N-terminus) to red (C-terminus), with the flexible termini represented by Ca atoms (grey spheres). (D) Fitting of the experimental SAXS profile for Apo-CTD by the best-fitting CORAL-derived models corresponding to either A or F-type of dimers.

### *Analysis of the dimeric interfaces in solution*

The presence of the disulfide bond forces the formation of the F-type dimer of AnaCTDH(apo). However, upon the reduction of cysteines or in their absence (in clade 1 CTDHs or in CTD-OCP), the A-type dimerization can become possible. The heterogeneous nature of the apo-AnaCTDH solution and this protein's propensity to monomerize in the absence of carotenoid<sup>25</sup> impedes structural analysis of its non-covalently bound dimers. In contrast, individual apo-CTDs from *Synechocystis* OCP (carrying a Phe instead of Cys at residue position 103) efficiently dimerize without forming higher-order oligomers<sup>26</sup>, making apo-CTD dimers suitable for structural characterization in solution by using small-angle X-ray scattering (SAXS).

At 270  $\mu\text{M}$ , the 18.7 kDa apo-CTD construct gave particles with a radius of gyration ( $R_g$ ) of 2.7 nm and the maximum particle dimension ( $D_{\text{max}}$ ) of 9.5 nm. The bell-shaped Kratky plot with the gradual rise of the curve at high angles (Fig. 2A) suggested particles with a compact globular core and moderate flexibility. The Porod volume of 56 nm<sup>3</sup> suggests that the corresponding MW estimate (35 kDa) is close to that expected for apo-CTD dimers (37 kDa), in line with their known dimerization at concentrations above 150  $\mu\text{M}$ <sup>26,30</sup>. Ab initio molecular shape reconstruction using *DAMMIF*<sup>31</sup> resulted in globular models with pairs of protrusions at variable positions (Fig. 2B), consistent with the Kratky plot. Averaging of the *DAMMIF* models with *DAMAVER*<sup>32</sup> revealed a core, common for the generated models (*DAMFILT* core; shown by cyan spheres in Fig. 2b). Given the similarity of the primary structure and fold of apo-CTD from *Synechocystis* and apo-AnaCTDH ( $C\alpha$  r.m.s.d.  $<1$  Å), the two types of dimeric cores found in the crystal lattice (Fig. 1) were built by spatial overlay, yielding either the F-type or the A-type dimer. To account for local differences between the apo-CTD and apo-AnaCTDH structures, the A-type dimer of apo-CTD was further refined by local protein–protein docking approach using *RosettaDock* server<sup>33</sup>, which resulted in a more connected, realistic structural model free from steric clashes (Supplementary Figure 4A). This model, matching the *DAMFILT*-derived globular core (Fig. 2C), was used as a template to model the unstructured terminal regions (including the C-terminal tails) using *CORAL* (Fig. 2C). Likewise, the F-type of the apo-CTD dimer was slightly corrected to eliminate apparent clashes and then used to build a *CORAL* model with flexible termini. *CORAL*-derived structural models for both types of dimers described the experimental SAXS data reasonably well; however, statistical analysis including  $\chi^2$  and the correlation

map  $P$  value (CorMap)<sup>34</sup> showed much better fits in the case of A-type dimer (Fig. 2D). This favored fit for the A-type apo-CTD dimer may indicate that, in the absence of the Cys–Cys bond, the back-to-back interface might be more favorable in solution. It is likely that the compromised ability of apo-AnaCTDH to form non-covalent dimers, which presumably occurs only at high protein concentrations, is associated with the absence of the functionally important phenylalanine residues in the external side of the  $\beta$ -sheet (in positions 299 and 300 of *Synechocystis* OCP), which in AnaCTDH are occupied by Ile-His residues (see Supplementary Figures 4B and 4C). This hypothesis is in line with the reported inability of apo-TeCTDH lacking a Phe in position homologous to 299 in the CTD of *Synechocystis* OCP to form stable dimers<sup>25</sup>. Attempts to fit the SAXS data with either a monomer or a trimer of apo-CTD gave significantly poorer fits ( $\chi^2 \sim 19$  for a monomer,  $\chi^2 \sim 5$  for a trimer).



**Figure 3 | Crystal structure comparison between apo-AnaCTDH and CTD OCPO (PDB code: 5UI2).** a CTD-OCPO (PDB code: 5UI2, orange) and apo-AnaCTDH (blue) structures are superimposed. Both elongation of  $\beta 6$  strand and shortening of the loop connecting  $\beta 5$  and  $\beta 6$  strands are observed for apo-AnaCTDH. b These differences lead to spatial changes in both E244/E69 and W277/W102. In contrast, Y201/Y26 and W288/W109 overlap well in both structures. c Cartoon representation of CTDH (blue shades) super-positioned with CTD of OCPO (orange shades). The C-terminal tail changes its orientation (indicated by dashed arrow) substantially from an external position (tail in orange) to an internal position for apo-AnaCTDH (tail in light blue). The position of the carotenoid molecule (yellow stick representation) is preserved from the 5UI2 structure. d Same as c but rotated clockwise by  $90^\circ$ . e Same as b but with amino acids known to be substantially perturbed in their solvent accessibility upon holo-to-apo transition in CTD-OCPO

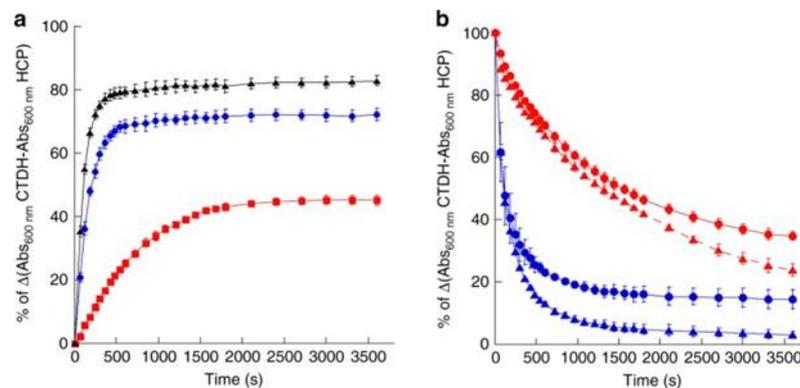
These observations indicate that transitions between the two types of dimers may occur, which may be relevant for the carotenoid transfer mechanism. Owing to the steric constraints, the F-type crystallographic apo-AnaCTDH dimer is not capable of performing carotenoid uptake from the membranes, as the carotenoid channel is partially blocked by the C-terminal tail and filled with the Cys103-containing loop (hereafter  $\beta 5/\beta 6$  loop), while in the A-type dimer, the carotenoid cavity is open. Notably, the carotenoid transfer is suppressed when the dimeric holoform is stabilized by the disulfide bond<sup>25</sup>. Furthermore, the A-type conformation may represent one of the intermediates during carotenoid transfer into the HCPs. One may speculate that, during the transfer, the F-type dimer disassembles, exposing one side of the carotenoid, and yields an intermediate A-type dimer with the carotenoid accessible in one holo-monomer, which then forms a new F-type-like interface with an apo-HCP counterpart.

### *Structural changes related to the shortening of CTDH relative to CTD*

To date, the only high-resolution structural information available on the CTD is when it is in association with the NTD in the OCP<sup>o</sup> state of the full-length OCP protein (closed carotenoid-associated conformation—PDB codes: 5UI2, 4XB5, 5TUX, 3MG1, etc.). As opposed to the CTDs of OCPs, all CTDHs lack the four amino acids that immediately follow residue 103 (residue 278 in *Synechocystis* OCP) in the CTD (Supplementary Figure 4C). Therefore, tertiary structure differences were expected to be present between the proteins<sup>25</sup>. Homology modeling suggested that, in the CTDH, the gap would result in a shortening of the  $\beta 6$  strand, while the loop that connects this strand with  $\beta 5$  would become longer forcing it into a different position<sup>25</sup>. The resolution of the three-dimensional structure of apo-AnaCTDH presented in this study shows that the lack of these four amino acids causes structural changes, different than those predicted by modeling. The  $\beta 6$  strand is longer and the loop between strands  $\beta 6$  and  $\beta 5$  is shorter (Fig. 3A). These structural changes lead to a change in the position of Glu244 and Trp277 (Fig. 3B), which play an essential role in the interactions between the CTD and NTD in the OCP<sup>o</sup> state<sup>15</sup>. These changes in positions can affect assembly processes, which in turn result in diverse activities by different protein pairs. More specifically, when holo-CTD and apo-NTD interact, an OCP-like protein is formed, in which the carotenoid is shared by the CTD and NTD (similarly to OCP<sup>o</sup>); however, when holo-CTDH and apo-HCP interact, the carotenoid is

not shared by both domains since full carotenoid transfer occurs (from CTDH to HCP)<sup>25</sup>. In contrast to the residues noted above, Tyr201 and Trp288 in *Synechocystis* CTD (Tyr26 and Trp109 in *Ana*CTDH), which are found in interaction with the carotenoid in OCP<sup>o</sup>, are in almost identical positions in OCP and in the apo-*Ana*CTDH structure presented here, suggesting that the carotenoid can interact with these amino acids in CTDH dimers.

In an engineered *Ana*CTDH mutant in which the four amino acids missing in the sequence (Supplementary Figure 4C) were reintroduced and the Cys103 was changed to Phe, the holo-dimer was less stable and easily lost the carotenoid<sup>25</sup>. This suggests that the structural change induced by the lack of the four amino acids is important for carotenoid stabilization in the CTDH dimer. Here we show that the addition of the 4 amino acids also hinders carotenoid uptake.



**Figure 4 | Carotenoid transfer between HCPs and CTDHs.** (A) Kinetics of carotenoid transfer from holo-HCP1 to apo-CTDH-C103F-Cter-Histag (blue), apo-CTDH-C103F-Nter-Histag (black), and apo-CTDH-C103F+4aa (red) followed by an increase of absorbance at 600 nm. In all the cases the ratio of holoHCP to apoCTDH is 1:2.5. (B) Kinetics of carotenoid transfer from holo-CTDH-C103F (blue) and from holo-CTDH-C103F- $\Delta$ C-terminal tail ( $\Delta$ C-terminal tail abbreviated to CTT on spectra, red) to apo-HCP1 (triangles) and apo-HCP4 (circles) followed by the decrease of absorbance at 600 nm. The CTDHs used in these experiments have the His-tag in the C-terminus.

HCP1 largely transferred its carotenoid to apo-*Ana*CTDH-C103F (Fig. 4A, blue and black curves). The carotenoid transfer was slower from HCP1 to apo-*Ana*CTDH-C103F+4aa (Fig. 4A, red curve), indicating that this CTDH mutant is hindered in the uptake and stabilization of the carotenoid. In conclusion, in CTDHs, the structural changes generated by the absence of the four amino acids following Cys103 are critical for carotenoid uptake, binding, and stabilization of the dimer.

### *The position of the C-terminal tail in the CTDH structure*

The most striking difference between the apo-AnaCTDH structure presented here and the CTD-OCP<sup>o</sup> structure is a major shift in the position of the C-terminal tail (Fig. 3C). In the CTD-OCP<sup>o</sup> structure, the C-terminal tail is in an external position interacting with the  $\beta$ -sheets and possibly with the N-terminal extension (based on their proximity, as demonstrated in 5UI2 structure). In contrast, in the apo-AnaCTDH structure, the C-terminal tail adopts an internal position, coming closer to the F-type interface between the CTDH monomers (CTD-NTD interface in OCP), partially covering the opening of the designated carotenoid-accommodating cavity in the apo-AnaCTDH structure (Fig. 3D).

This dramatic shift in the position of the C-terminal tail (CTT) is in good agreement with the literature. The structural changes taking place upon OCP<sup>o</sup> to OCP<sup>r</sup> photoconversion were studied previously<sup>17</sup> by examining amino acid solvent accessibility. It was demonstrated that residues Pro225-Pro226-Phe227 (Phe227 is shown as shaded wheat-orange spheres in Fig. 3E) exhibit the second highest decrease in accessibility in the CTD upon phototransformation. It was postulated that this PPF patch serves as a cap for the carotenoid cavity in the CTD upon translocation of the carotenoid in the OCP<sup>r</sup> state. However, Lys310 (bottom-most light orange spheres in Fig. 3E) exhibited the highest decrease in accessibility, yet no explanation was suggested for this observation. The apo-AnaCTDH structure provides a more comprehensive interpretation of these results. Upon loss of the carotenoid to the NTD, the CTT, which points away from the OCP<sup>o</sup> globular structure and was highly accessible, now rotates by nearly 180° and turns into the carotenoid cavity, capping it. This capping function covers the PPF patch completely while internalizing Gln132, the structurally homologous CTDH residue of Lys310 in the CTD (Fig. 3E). In a similar way, we hypothesize that incorporation of the carotenoid into the CTDH releases the CTT from this internal “close position” to adopt its open position (like in the OCP<sup>o</sup>). The CTT, which is comprised of a hydrophobic path followed by hydrophilic path, can be involved in carotenoid translocation.

In order to assess the role of the C-terminal tail, two AnaCTDH-C103F mutants lacking the C-terminal tail were constructed. Both mutated proteins lacked 11 C-terminal amino acids from Leu126 to Leu136. One of the proteins had the His<sub>6</sub>-tag at the C-terminus just after Lys125 while the second mutant had the His<sub>6</sub>-tag at the N-terminus (see Methods). Regardless of the position of the His-tag, the mutated CTDH-C103F- $\Delta$ C-terminal tail failed to

receive a carotenoid from holo-HCP1 (Supplementary figures 5). The carotenoid uptake from the membranes was also largely reduced when the C-terminal tail was absent (only 15% of holo-protein was produced (Figure 9D).

The importance of the C-terminal tail was further examined for the carotenoid delivery by the holo-CTDH. For both apo-HCP1 and apo-HCP4 as recipients, holo-CTDH was able to transfer its carotenoid with or without the C-terminal tail (Figures 4E), yet the absence of the C-terminal tail from the holo-CTDH decreased the rate of carotenoid transfer.

These results demonstrate that the internal, cavity-blocking position of the C-terminal tail in the apo-form does not hinder carotenoid uptake. In fact, it facilitates its uptake from both the membrane and, surprisingly, from holo-HCP1. Moreover, the C-terminal tail facilitates transfer the carotenoid to HCPs. Therefore, the C-terminal tail, which starts with a relatively hydrophobic patch and ends in a relatively hydrophilic patch may serve as a facilitator for carotenoid uptake and delivery. For uptake, the hydrophilic end may bind to the hydrophilic lipid head groups followed by forming an interaction between the hydrophobic patch and the membrane-embedded carotenoid molecule. For delivery, the hydrophilic patch would bind to the appropriate site on the HCP, followed by transfer of the carotenoid to the HCP-binding cavity, facilitated by the hydrophobic patch.

### ***Possible structural changes anticipated in CTDH following carotenoid uptake***

The apo-CTDH structure is expected to change upon carotenoid binding, so we attempted to understand the structural effects of canthaxanthin binding on apo-AnaCTDH. Computational docking of a carotenoid molecule to the apo-AnaCTDH structure was executed using the Swissdock server<sup>35</sup>. When the F-type dimer was used as the host (with or without the C-terminal tails), not a single 1 of the 34 docking clusters predicted a carotenoid positioned inside the cavity formed by the dimer (Supplementary Figures 6A and 6B). Manual positioning of the carotenoid molecule was also unsuccessful. Thus, the F-type dimer cannot accommodate the carotenoid unless structural changes occur. Indeed, size exclusion chromatography and native-PAGE showed different volumes for holo- and apo-AnaCTDH25. The same situation seems to be valid for CTD-OCP, where the holo-dimer was reported to have a larger apparent size than the apo-dimer<sup>26</sup>.

By contrast, when an apo-AnaCTDH monomer was used as the host (with or without the C-terminal tail), 31 out of the top 34 clusters suggested that the carotenoid would be positioned in the pocket that has been shown to bind the carotenoid in OCPO (Supplementary Figures 6C and 6D). In both docking trials, the outward-facing region of the C-terminal tail was found to be a good docking site for the carotenoid molecule (red circles, Supplementary Figures 6B and 6D), suggesting that there could be an interaction between the two (as proposed above).

Two possible structural obstacles for carotenoid accommodation in the F-type CTDH dimer based on the apo-AnaCTDH structure can be suggested. First, the  $\beta 5/\beta 6$  loops (Gln100-Gly106, see Supplementary Figure 7), from both monomers, fill the opposing cavities (Supplementary Figure 7), which together are suspected to host the carotenoid molecule in the holo-AnaCTDH dimer form as inferred from the OCPO structure. When using an apo-AnaCTDH monomer as a docking target, the  $\beta 5/\beta 6$  loop from the other monomer is absent in the cavity, thus possibly enabling the carotenoid molecule to be docked in the monomeric target. Based on the scenario described above, we suggest that, upon carotenoid association, the position of these loops is altered, facilitating the binding of the carotenoid.

In addition, as already noted, the volume of holo-AnaCTDH is larger than that of apo-AnaCTDH. This could be related to a slight separation of monomers to yield a sufficiently sized cavity, long enough to accommodate the carotenoid. The distances between Trp110–Trp110' and Tyr27–Tyr27' in the F-type dimer ( $\sim 25$  and  $\sim 30$  Å, respectively) are too small for the accommodation of a canthaxanthin molecule ( $\sim 30$  Å), especially if these amino acids should be allowed to form H-bonds with the carotenoid carbonyls. The Trp288Ala mutant of CTD-OCP was shown to transfer the carotenoid more easily than the wild type (WT)<sup>30</sup>. This implied that the Trp stabilizes the binding of the carotenoid in CTD(H)s as in full-size OCPs.

It is important to note that carotenoid-containing monomeric forms of CTD(H)s were detected previously<sup>25,30</sup>. We thus hypothesized that a monomer can temporarily accommodate the carotenoid taken up from a membrane. Carotenoid binding could occur during CTDH synthesis, although uptake into apo-AnaCTDH has been shown experimentally<sup>25</sup>. The question arises about how this is possible with the position of the C-terminal tail in apo-AnaCTDH acting as an intermediate cap of the carotenoid-binding site. As discussed, we observed that not only apo-AnaCTDH can take the carotenoid from the membrane but also from HCP1. These results suggest

that the C-terminal tail moves rather easily. Moreover, our results indicate that the C-terminal tail facilitates carotenoid uptake and delivery. It was suggested that the position of the carotenoid in the monomer (completely buried) must be different to that in the dimer in which it is shared by two monomers<sup>25,30</sup>. We propose that, first, the C-terminal tail supports the uptake of the carotenoid and that structural changes (i.e.,  $\beta 5/\beta 6$  loop movement) assist to better host the carotenoid molecule within the first monomer as a moiety of an A-type CTDH dimer. Second, after a local energetic minimum is reached, reorganization of the monomers yields the F-type dimer while undergoing another structural change in the second monomer cavity to facilitate carotenoid binding.

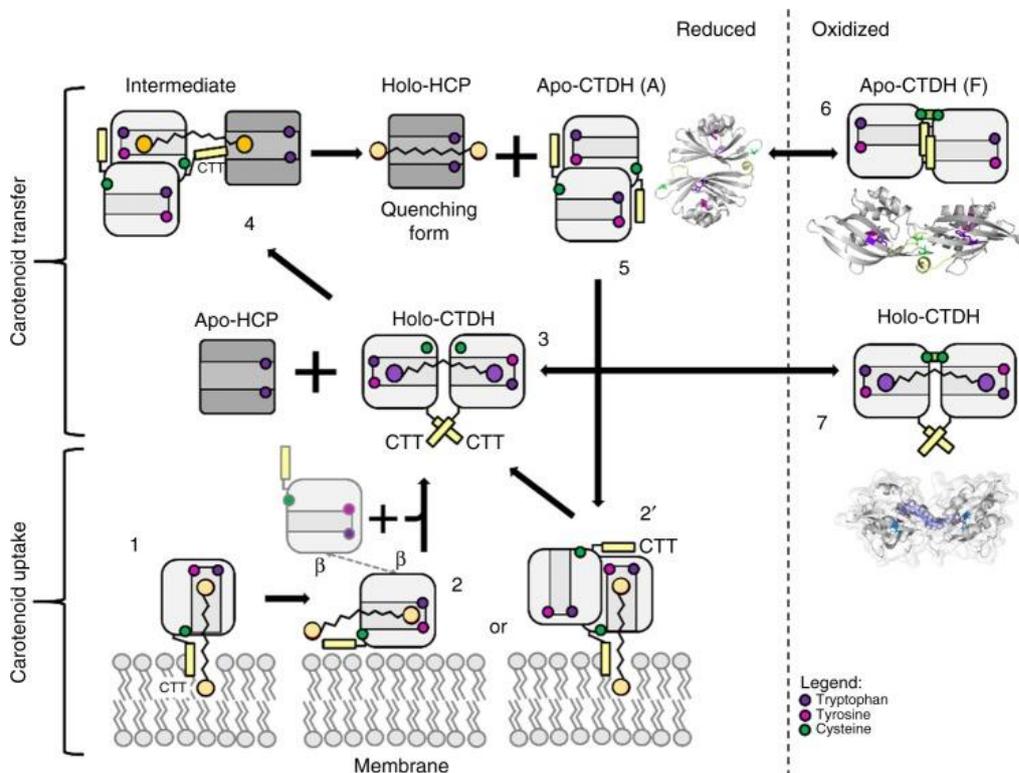
### *Structural position of holo-AnaCTDH subunits yields absorption differences compared to OCPO*

When comparing the OCPO structure and the F-type apo-AnaCTDH dimer, the NTD-equivalent, second CTDH monomer is positioned differently in space with respect to the first (CTD super-positioned) CTDH monomer, as can be seen in Supplementary Figure 8. While the carotenoid cavity in OCP<sup>O</sup> is substantially tilted between the two domains, it is relatively straight in the apo-AnaCTDH dimer. Indeed, the carotenoid in OCP<sup>O</sup> adopts a bent conformation, while it has a planar conformation in the holo-NTD variant<sup>16</sup>. The increased chromophore planarity and rotation of the terminal rings are known to induce the red shift of absorption peak due to elongation of the effective conjugation length<sup>36,37</sup>. While some structural alterations are anticipated upon carotenoid uptake by the apo-AnaCTDH, the overall axial positions of the two monomers are not expected to be dramatically changed. Thus, the carotenoid cavity in the CTDH dimer will allow a planar conformation of the carotenoid. This should be clarified by structural studies using holo-CTD(H) dimer.

### *CTDH carotenoid uptake and delivery cycle*

In chloroplasts and cyanobacteria, various redox-activation processes of enzymes in the stroma exist. These activations occur via ferredoxin, ferredoxin-thioredoxin reductase, and thioredoxin (Trx)<sup>38,39</sup>. Trx can also be reduced by NADPH-Trx reductases. Trx also acts as an electron donor for antioxidant defense systems (e.g., peroxiredoxins, catalases)<sup>40</sup>. From the known functions of Trx, it can be clearly inferred that proteins with oxidized

disulfide bonds exist in cyanobacterial cytoplasm under various conditions. The CTDH dimer therefore could also be a Trx-regulated protein (and/or another regulatory agent).



**Figure 5 |CTDH carotenoid uptake and delivery cycle.** apo-AnaCTDH monomer (1) or a dimer in its A-type state (2') may extract a carotenoid molecule from the membrane with the assistance of the C-terminal tail (labeled as CTT in the figure). Then a holo-dimer is formed (3) in an F-type state (with association of another monomer; 2) Then the holo-CTDH can either be redox regulated in its F-type state between the reduced (3) and oxidized states (7) or interact with apo-HCP to allow carotenoid transfer through the A-type state with the assistance of the C-terminal tail (4). This yields the holo-HCP and the apo-CTDH dimer that can undergo A-to-F type transition and either be in its reduced state (5) and thus can take another carotenoid molecule or be modified to its oxidized states (6). Different carotenoid molecule colors represent different binding modes and thus different spectroscopic characteristics

Based on these data and previous work of our laboratories, we can suggest a mechanism (Fig. 5) describing carotenoid delivery and uptake by a clade 2 CTDH. *In vivo*, in low light conditions or darkness (where no photoprotection is needed), Trx is more oxidized owing to low photosynthetic activity, and thus holo- and apo-CTDH dimers are stabilized by the disulfide bond between monomers that inhibits carotenoid uptake and delivery (Fig. 5, steps 6 and 7). Under conditions where photoprotection against oxidative stress is needed

(reducing conditions, high light, low CO<sub>2</sub>), the Trx pool is largely reduced which in turn reduces both the apo- and holo-CTDH dimer to facilitate carotenoid transfer (Fig. 5, step 3) and uptake (Fig. 5, step 5). When reduced, and with the aid of HCP, the holo F-type dimer is converted to an A-type dimer, leaving the carotenoid bound to one monomer, bringing it one step closer to be transferred to the HCP. Then, by the assistance of the C-terminal tail, the carotenoid can be transferred directly to HCPs (Fig. 5, step 4).

After carotenoid transfer, the C-terminal tail remains near the carotenoid cavity and the CTDH can dimerize to yield the F-type apo-CTDH shown in this work. This apo-dimer is stable and will be able to receive a new carotenoid molecule only upon reduction of the disulfide bond. The uptake of the carotenoid from the membrane (and from HCP1) can be done by either a monomer (Fig. 5, step 1) or a reduced dimer (Fig. 5, step 2) in its A-type state; however, in either case the extension of the C-terminal tail toward the carotenoid site facilitates its translocation toward the designated cavity in the CTDH. A second monomer (Fig. 5, step 2') is then assembled in order to co-encapsulate the carotenoid and form the F-type holo-CTDH (Fig. 5, step 3), which under regular conditions undergoes an oxidation event to yield the carotenoid-locked holo-CTDH (Fig. 5, step 7). We believe that F-type apo-dimer can be a result of two possible scenarios, which are not necessarily mutually exclusive. First, it is possible that carotenoid uptake is already completed during or immediately after CTDH synthesis before the formation of the dimer and disulfide bond. In this case, the F-type apo-dimer will be an intermediate stage along the CTDH carotenoid cycle. Nevertheless, it is also possible that the protein dimer scaffold can be formed prior to carotenoid uptake, and thus interchange between redox states can occur. In this case, F-type apo-dimer would be a preliminary stage of the cycle. These alternatives require further investigation.

## ***Carotenoid uptake, binding and delivery by water-soluble proteins in cyanobacteria***

---

Dvir Harris<sup>1,2</sup>, Fernando Muzzopappa<sup>3</sup>, Fabian Glaser<sup>5</sup>, Adjele Wilson<sup>3</sup>, Diana Kirilovsky<sup>3</sup> & Noam Adir<sup>1,2</sup>.

### **Affiliations**

<sup>1</sup>Schulich Faculty of Chemistry, Technion, Haifa 3200003, Israel.

<sup>2</sup>Grand Technion Energy Program (GTEP), Technion, Haifa 3200003, Israel.

<sup>3</sup>Institute for Integrative Biology of the Cell (I2BC), CEA, CNRS, Université Paris-Sud, Université Paris-Saclay, 91198 Gif sur Yvette, France.

<sup>4</sup>Institut de Biologie et Technologies de Saclay (iBiTec-S), Commissariat à l'Énergie Atomique (CEA), 91191 Gif-sur-Yvette, France.

<sup>5</sup>Lorry I. Lokey Interdisciplinary Center for Life Sciences and Engineering, Technion-Israel Institute of Technology, Haifa, Israel.

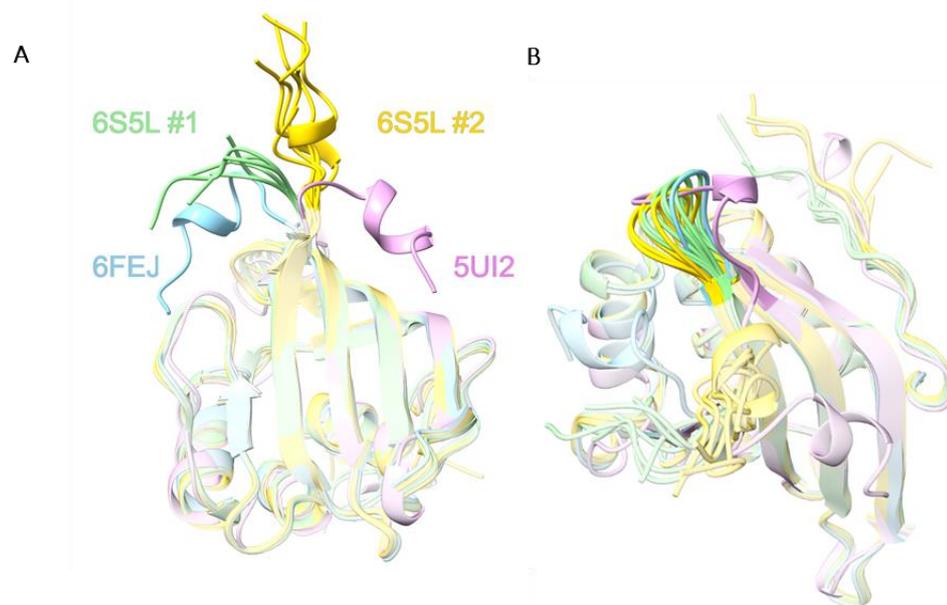
## ***Results***

---

### ***The ApoCTDH structure in absence of urea***

The dimeric CTDH structure was used as the MR search model for the data obtained in the absence of urea. The results of structure determination (PDB code 6S5L) is shown in supplementary figure 9 and the crystallographic data collection and refinement parameters are given in table S1. In order to analyze the structural differences between the two structures obtained, a superposition of the 12 individual apo-protomers was done together with a 6FEJ apo-monomer (in 2M urea). In addition, the 5UI2 holo-monomer (CTD of OCP<sup>0</sup>) was added to this superposition, as can be seen in figure 6. The new high-oligomeric state structure (PDB code 6S5L) revealed that the 12 protomers present in the asymmetric unit can be divided into two populations (comprising 6 protomers each, #1 colored green and #2 colored yellow in Fig. 6), according to their relative CTT position. The CTT can assume two additional positions, intermediate to the two positions already identified (purple and orange positions, respectively). Site #1 is similar to the site identified in the 6FEJ structure (apo-CTDH with 2M urea). Site #2 is

intermediate between the site #1 position and the position of the CTT in the OCP<sup>O</sup> structure (5UI2). The crystal packing interactions vary between the 12 monomers, however CTTs in both site #1 and site #2 are found in these positions without stabilization of nearby monomers. We thus conclude that all three positions seen in Fig. 6 are independently stable. In addition, the  $\beta 5/\beta 6$  loop also displays more subtle dynamics, when compared to the CTT (Fig. 6B), adopting more internal positions (relatively to the protein cavity) as compared to the 6FEJ/5UI2/3MG1 structures. This new oligomeric structure of the CTDH seems to show a dynamic motion of the CTT, stabilized at specific terminal and intermediate states. We thus decided to utilize molecular dynamics (MD), to shed new light upon this process of CTT translocation and its implications.



**Figure 6 | Superposition of all known CTD(H) structures to date.** (A) Superposition of HoloCTD of OCP<sup>O</sup> (violet, PDB code 5UI2, carotenoid is omitted for clarity), 2M urea apoCTDH structure (cyan, PDB code 6FEJ) and the newly found 12 apo-protomers of the CTDH structure in the absence of urea (green and yellow, PDB code 6S5L) demonstrates the different positions CTT adopts. (B) Same as A, but the different  $\beta 5/\beta 6$  loop positions are demonstrated.

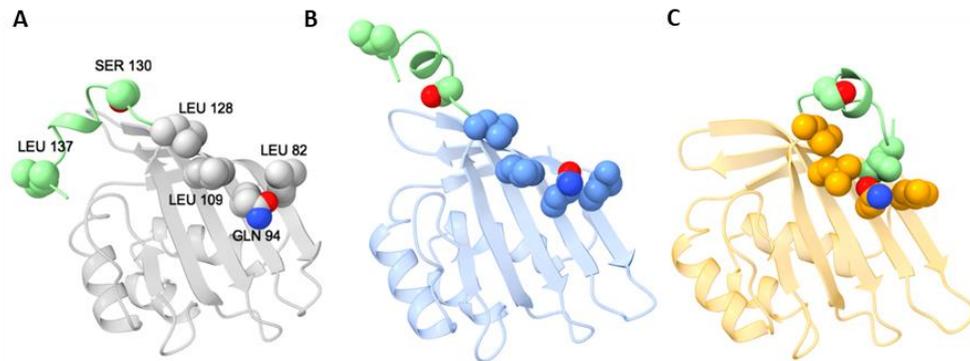
### *CTDH molecular dynamics*

To simulate the effect of carotenoid binding on the dynamicity of the CTT, canthaxanthin (CAN) ligand (extracted from RCP structure, 4XB4) was docked into the 6FEJ structure by Autodock<sup>55</sup>. Interestingly, although

partially blocked by the CTT, CAN could be docked into the pocket region without any constraints, since the pocket was able to accommodate the ligand in its original PDB conformation. The CTT was also monitored along the MD simulation in order to explore how CAN incorporation into the carotenoid-binding cavity induces dynamicity. To do so, we carried on MD runs for apoCTDH and CAN-docked holoCTDH and followed CTT dynamics as a function of simulation duration. In both cases, when CTT arrived to its fully-open position, this position was retained throughout the remainder of the simulation with little-to-none CTT fluctuations, thus implying the existence of a strong binding mode of the CTT by the outward-facing  $\beta$ -sheet. Another very prominent feature emerged from the simulations is that the perpendicular CTT orientation (gold in Fig. Supplementary figure 10) has a non-negligible stability, as the CTT dwells in this position before collapsing to the final position (either open or closed). The perpendicular orientation is retained for much longer period of time in the apo-form versus the holo-form, as derived from the mid RMSD range (10-15Å, Supplementary figure 11A, 11B). This position is almost identical to the #2 population of the 6S5L structure (yellow in Fig. 6A), thus hints towards relevancy of this state. We postulate that this CTT orientation is the one which extracts/delivers the carotenoid, as it has the longest reach towards the respective donating/receiving partner. The hinge point that allows this rotation of the CTT helix between the closed and open conformation is between A129 and S130. The phi and psi angles between those two residues change dramatically from -150 to 50 and from -50 to -150 degrees, respectively, to allow the hinge movement (Supplementary figure 11C).

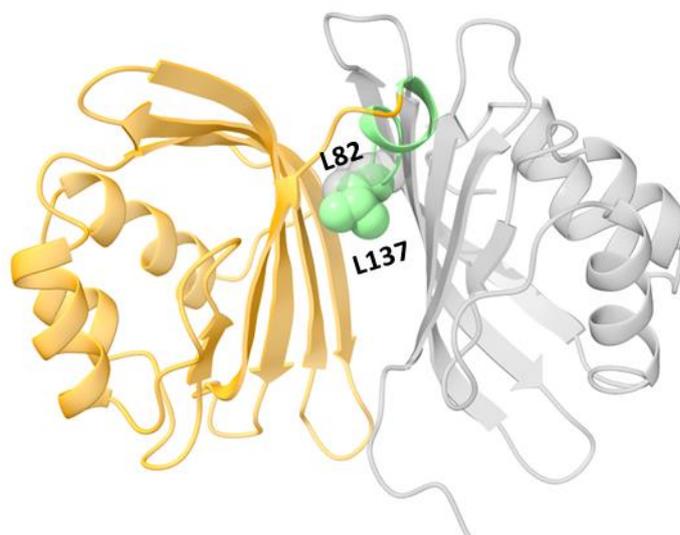
The movement of the CTT into the fully-open conformation of CTT is only possible by disrupting the suggested back-to-back interface, and so we then further examined the putative mechanisms for back-to-back to head-to-head CTDH reorganization. Molecular Mechanics Poisson-Boltzmann Surface Area (MM-PBSA) calculation is a post-processing end-state simulation that allows us to calculate free energies of molecules in solution and the binding free energy in a non-covalently bound, receptor-ligand complex. Furthermore, it is possible also to decompose calculated free energies of interaction and identify specific amino acids (AAs) that stabilize protein-protein interaction. This analysis (Supplementary figure 12) reveals that several hot spot residues that contribute more than -2 Kcal to the change in free energy upon complex formation of the back-to-back dimer. These include L82, Q94, S96 and L109, which are found in the  $\beta$ -sheet forming the back-to-back interface (6FEJ monomer is shown in Fig. 7A). The back-to-

back interface is predicted to be very stable, as most AAs contribute significant negative  $\Delta G$  values, resulting in a total of -41 Kcal. (-20.5 Kcal/monomer).



**Figure 7 | L137 helps to stabilize CTT flip by forming a hydrophobic cluster when CTT adopts “open” position.** (A) Initial MD apoCTDH monomer PDB 6FEJ, shows that while the CTT adopts the closed conformation, the stabilizing AAs (L128, L82, L109 and N94, spheres representation) are available to interact with the other monomer. (B) CTT intermediate conformation. (C) CTT open conformation. A129-S130 phi and psi are the dihedrals that are change dramatically to allow this hinge movement.

After the CTT dwells and passes the interim perpendicular state (Fig. 7B), it adopts the open conformation, in which L137 interacts with the exact same four AAs that stabilize the back-to-back dimerization the most (Fig. 8C). L137 forms a hydrophobic binding cluster with these AAs which remains stable throughout the remainder of the simulation. Interestingly, the polar residues S96 and Q94, which initially pointed towards each other (Oxygens distance  $\sim 4.5$  Å), now point away from each other, with an O-O distance of  $\sim 9$  Å, making room for L137 and allowing the formation of this hydrophobic interaction. Then, we superimposed the CTDH monomer with its opened CTT conformation onto the back-to-back dimer, as depicted in figure 8.



**Figure 8 | The CTT in the open conformation destabilizes the back-to-back dimer.** Upon translocation to the open conformation (orange cartoon), the CTT (green cartoon) is predicted to overlap with the second monomer (gray cartoon) forming the back-to-back interface. L137 (green spheres) of the first monomer overlaps with residue L82 (gray spheres) of the second monomer.

As expected, once the CTT adopts the open conformation, the  $\alpha$ -helix (positions 129 to 138) overlaps with the position of the second monomer. Thus, when in this state the back-to-back dimer will either not form or may be destabilized. Moreover, in the CTT open conformation, residue L137 overlaps with the interfacial residue L82 of the second monomer (Fig. 8) “replacing” its stabilizing role between monomers, creating a transient internal hydrophobic patch between CTT at the monomer-monomer interface patch in a similar way. Therefore, we believe the CTT flip may have two inter-connected roles, pigment ligand binding and complex formation which also affects carotenoid uptake and delivery.

### *The role of the CTT and specific amino acids.*

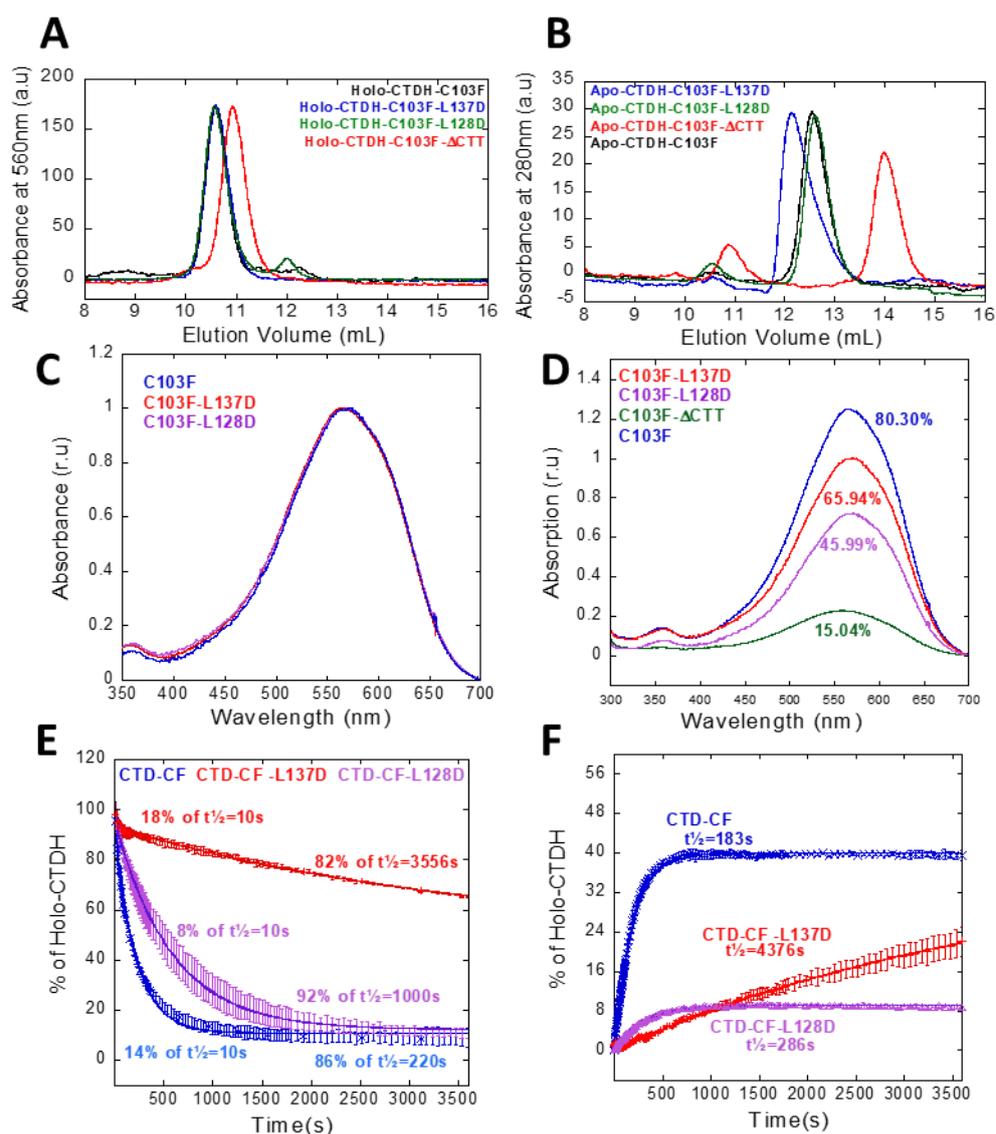
As it was mentioned before, the AnaCTDH- $\Delta$ CTT mutants fail to retrieve carotenoid from the holo-AnaHCP1 (Supplementary figure 5). In addition, the carotenoid uptake from membranes also strongly decreased (Fig. 9D). The carotenoid transfer from AnaCTDH-C103F- $\Delta$ CTT to apo-AnaHCP4(1) also was affected (Fig. 4B). In both cases, the rate of carotenoid delivery decreased.

From the oligomeric structure and the MD simulation it was proposed that the CTT can populate different conformations suggesting that the CTT has a flexible character. This is confirmed by the reduction of the hydrodynamic radius of the protein compared with the WT (C103F) in both apo- and holo-forms, determined by the highest elution volume in SEC experiments (figure 9A and 9B). Although, it is important to note that there is a relatively small population of CTDH-C103F- $\Delta$ CTT(apo) that is present as dimer. It is tempting to hypothesize that the lack of CTT allow more back-to-back dimerization in the apo-form.

Based on the results obtained via the CTDH structures and MD, we rationally introduced point mutations to L128 and L137. As shown above, L137 stands out structurally as the anchoring hook for the CTT movement to its final open position by forming a hydrophobic cluster. Therefore, we replaced the Leucine with Aspartic acid at this site, to hamper the hydrophobic nesting feature. We also hypothesized that Leu 128, might have a prominent stabilizing effect on the back-to-back interface, based on the MMPBSA analysis and its position at the end of the  $\beta$ -strand from which stems the CTT. SEC experiment indicated that the oligomeric state of the holoforms of the mutants is not altered (Fig. 8A and 8B). Nevertheless, while the L128D mutation does not alter the oligomer/radius of the protein, the introduction of the L137D mutation yield in a slightly bigger volume of the CTDH(apo) monomer. This could be due to the destabilization of the “open position” of the CTT and an increase of its flexibility by populating intermediary positions. This finding strongly corroborates our MD simulation-derived hypothesis, which suggests that the CTT is retained in the open position by holding its hydrophobic L137 using the counter hydrophobic cluster.

As the CTDH protein serves as a relay for carotenoid uptake and delivery, we were interested to examine how these point-mutations affect the ability of the apo-mutants to extract a carotenoid from membranes or HCP1 and the ability of the holo-mutants to deliver a carotenoid to apoHCP4. We first examined whether introducing the L137D or L128D mutations affect protein-carotenoid interactions. The absorption spectra of holoCTDH for both mutations are essentially identical (Fig. 8C), suggesting that the protein-chromophore interaction remain similar in all forms of holoCTDH. However, these point mutations have a strong impact on all aspects of CTDH carotenoid-shuttling capabilities. The ability to uptake a carotenoid from membranes (Fig. 8D) is reduced in the case of L137D (66% versus 80% in the case of C103F), and even more so in L128D (46%). A complete truncation

of the CTT results in an apoCTDH with poor ability to uptake a carotenoid from membranes (only 15%), confirming the importance of the CTT in the carotenoid translocation. While L137D can be reasonably explained by decreasing CTT ability to incorporate into the membrane due to the addition of a negative charge, L128D decreased functionality is harder to interpret.



**Figure 9 | Biochemical and Functional characterizations of CTDH point mutations.**

Holo- (A, absorption measured at 560 nm) and apo- (B, absorption measured at 280 nm) CTDH SEC profiles (legend on graph). (C) Absorption spectra of holo point-mutations. (D) Carotenoid uptake from carotenoid-containing membranes by apoCTDH-C103F (80.30% ± 4.00, blue curve), apoCTDH-C103F-L137D (65.94% ± 2.00, red curve), apoCTDH-C103F-L128D (45.99% ± 2.40, violet curve) and apoCTDH-C103F (15.04% ± 0.01, green curve). (E) Carotenoid delivery from holoCTDH to apoHCP4, ratio 1:2.5 (legend on figures). (F) Carotenoid acquisition by apoCTDH from holoHCP1, ratio 1:1 (legend on graph). The differences in amplitude with the figure 4A are due to the differences in apo to holo ratio; in these experiments CTDH and HCP were in equal amounts while in figure 4 there is a 2 fold

excess of apo-CTDH. In all cases, C103F mutant was introduced in the background to enable carotenoid and protein dynamics.

Examining the time-resolved results present in fig. 8E (carotenoid acquisition by apoCTDH from holoHCP1) and fig. 8F (carotenoid delivery from holoCTDH to apoHCP4) enables better understanding of the different effects these point mutations have.

Similarly, to the uptake case, the propensity of L128D to receive a carotenoid from holoHCP1 is hampered in amplitude, yet having similar kinetics, with respect to the control. In contrast, L137D does not hit an amplitude plateau of carotenoid acquisition from holoHCP1 yet having very slow kinetics (Fig. 8E). Analysis of carotenoid transfer from AnaCTDH-C103F and the mutants to ApoHCP4(apo) indicated that the kinetics can be fitted to biexponential functions (Figure 9E). While the fast component ( $T^{1/2} = 10$  s) was almost not affected (amplitude and rate) by the point mutations, the slow component was largely slowed down in the mutants ((AnaCTDH-C103F  $T^{1/2} = 220$  s; L128D  $T^{1/2} = 1000$  s; L137D  $T^{1/2} = 3556$  s). There are two possible causes for this biexponential kinetic. In the first one we consider that the minor fast component corresponds to a small population of CTDH which are able to transfer the carotenoid faster. The nature of this small population is unclear, but since the mutations do not affect this component, this population could be attributed to CTDH(CAN) dimers, in which the CTT is not locked in the open position. The other possible explanation is that the carotenoid transfer is a two-step process. However, the spectral change associated with the two phases is homogeneous implying that if a first specie is formed it should have the same absorption spectrum than the final specie. A possible first step is the monomerization of CTDH(CAN). However, we expect the mutations L137D and L128D to affect the transition between oligomeric forms and therefore the fast component. The discernment between the two possible origins of this biphasic behavior is out of the scope of this thesis, and a deeper kinetic analysis will be required to understand it. We suggest that L128D mainly compromises CTT flexibility, which is more crucial for receiving a carotenoid from a delivering entity (membrane/holoHCP1) than delivering a carotenoid to a receiving entity (apoHCP4). As for L137D, introducing this hydrophobic/hydrophilic switch makes it challenging and kinetically slower for the hydrophobic carotenoid to slide past this point (either on the way into CTDH carotenoid-binding cavity or from it), yet thermodynamically favorable (with respect to L128D).

## Discussion

---

In this work, we had clarified the mechanisms of the CTDH in carotenoid translocation, which involved the C-terminus tail and the switch between the back-to-back and head to head dimeric forms. This was made possible by following the observations that emerged from the determination of the higher oligomeric state apoCTDH structure, which could only be resolved thanks to the revelation of the urea-induced dimeric apoCTDH structure. This high oligomeric apoCTDH structure had the CTTs (of the 12 monomers constituting the oligomer) situated in two different, intermediate positions with respect to holoCTD of OCPO (completely open) and the 2M urea conditioned apoCTDH (almost completely closed). This diverse population of CTT conformations implied that it is highly mobile (supported by the decrease of volume in the  $\Delta$ CTT mutant). MD simulation showed that once the CTT arrives to the open position, it is stabilized by interaction between residues of the CTT (like the L137) and the  $\beta$ -sheet hydrophobic cluster. The same hydrophobic cluster is involved in the stabilization of the back-to-back dimer. Thus, a transition from back-to-back to head-to-head dimer was proposed during carotenoid uptake. In this case, the CTT displace the second monomer and the L137 replace the hydrophobic interaction with the  $\beta$ -sheet. Point mutagenesis has corroborated the hypothesis that L137 serves as the anchoring hook of the CTT in the open position, as well as having a pivotal role in carotenoid delivery, and in carotenoid uptake. Inversely, L128, which was calculated to have a stabilizing effect on back-to-back interaction, was shown to be specifically critical for carotenoid uptake, but less so for carotenoid delivery.

The mechanism for carotenoid uptake, binding and delivery presented in this work can be compared to other types of hydrophobic ligand-binding proteins of other systems. The retinal pigment epithelium 65 (RPE65) enzyme is known to function as retinoid isomerase in the vertebrate visual system, binding an all-trans retinol as its substrate. Interestingly and similar to CTDH, RPE65 is also predominantly dimeric and was shown to adopt a back-to-back orientation in the asymmetric unit<sup>42</sup>. However, in contrast to the dimeric CTDH (which has the cavities of its two dimer-forming monomers facing the opposite directions), the cavities of RPE65 dimer both face the same direction. In addition, while the canthaxanthin cargo of CTDH is roughly twice the length of the carotenoid-binding cavity of a monomer, the retinol-binding cavity of a monomer is more than twice the length of the retinol substrate. These differences may provide the rationale for the different

functions of these proteins, from carotenoid dynamics perspective – while RPE65 is an enzyme that catalyzes stereospecific retinoid isomerization (followed by transfer of the product to other enzymes), CTDH is a temporal carotenoid carrier. Thus, while in RPE65 the desired feature is increased cargo loading (as facilitated by the two cavities both facing the retinol-donating membrane), in CTDH, it is rather the delicate control over carotenoid uptake and delivery. In the CTDH dimer, the cavities facing opposite direction favor the carotenoid uptake by only one monomer. This allow the consequent carotenoid encapsulation and stabilization in the F-type dimer. Moreover, in all RPE65 structures available (including those with bound retinyl analogs), the 110-126 hydrophobic segment is missing due to unclear electron density, derived from high flexibility of this region<sup>43,44</sup>. This segment is adjacent to the cavity entrance and is highly hydrophobic. We propose that this 17 amino-acid missing segment is the highly-dynamic equivalent of the CTT in the RPE65 case, serving as the carotenoid-extracting rod of the protein.

In conclusion, a novel quaternary-structure reorganization process of carotenoid-binding protein was discovered, emphasizing the importance of dynamics on protein-ligand regulation, as well as unraveling critical roles of specific amino acids in this carotenoid uptake and delivery mechanism.

## *Materials and Methods*

---

### *Construction of plasmids containing *Synechocystis ctd-ocp* and *Anabaena ctdh* and *hcp* genes for expression in *E. coli**

Construction of the plasmid pCDF-CTDHana with His-tag in the C-terminus, carrying the CTDH from *Anabaena* (gene all4940) used to overexpress the Apo-CTDH for crystallization was previously reported<sup>25</sup>. The construct for the overexpression of HCP1 (pCDF-HCP1), HCP2 (pCDF-HCP2), HCP3 (pCDF-HCP3), HCP4 (pCDF-HCP4)<sup>22</sup> and CTDH-C103F<sup>25</sup> have been described. The C-terminal tail was deleted by mutagenesis using the plasmid pCDF-CTDHana-C103F and the primers F-4940-ΔC-terminal tail- (5'-GATCATCCACACCCAAATTAACACCACCACCACCACCACTAGTCTTG-3') and R-4940-ΔC-terminal tail (5'-TTTAATTTGGGTGTGGATGATCTGTTTTTCTTGTTAAGAG-3') to

create the pCDF-CTDHAna-C103F- $\Delta$ C-terminal tail. In all these CTDH plasmids, the sequence coding for the His-tag were in the 3' terminus of the gene (C-terminus of the protein).

To obtain N-terminal His-tagged CTDHs, the primers F-4940-BamHI (5'-CACACGGATCCGAAAGCTGCTGAGTCTCTCCC-3') and R-4940-NotI (5'-CATTATGCGGCCGCCTTATTGATGACAGCGCCCC-3') were used to amplify the *ctdh* gene from genomic DNA of *Anabaena* PCC 7120 and cloned in pCDFDuet-1 digested by BamHI and NotI. The point mutation C103F (see Muzzopappa et al. (2017)<sup>25</sup> for primers) and the deletion of the C-terminal tail (using F-4940- $\Delta$ C-terminal tail (5'-GATCATCCACACCCAAATTAATAGTCTTGATTAATTACTTTAAC-3') and R-4940- $\Delta$ C-terminal tail) were introduced by mutagenesis. Then a part of the N-terminal prolongation present in these plasmids was excised by mutagenesis using the primers F-4940- $\Delta$ BamHI (5'-CATCACCATCATCACCACAAAGCTGCTGAGTCTCTCCC-3') and R- $\Delta$ Duet (5'-GTGGTGATGATGGTGATGCATGGTATATCTCCT-3') to create the plasmids pCDF-NHISCTDHAna-C103F and pCDF-NHISCTDHAna-C103F- $\Delta$ CTT. The proteins obtained using these plasmids contained only six His after the first Met and then the CTDH primary sequence.

The point mutations L128D and L137D were introduced by directed mutagenesis, using the pCDF-NHISCTDHAna-C103 plasmid as template and the following mismatching primers:

L128D\_fw:

CCCAAATTAAGCTCGACGCTTCTCCCAAGAGTTACTGGC(for L128D).

L128D\_rv:

GCCAGTAACTCTTGGGGAGAAGCGTCGAGTTTAATTTGGG(for L128D).

L137D\_fw:

CCCAAGAGTTACTGGCTGACCGTCGTGAACAG (for L137D).

L137D\_rv:

CTGTTACGACGGTCAGCCAGTAACTCTTGGGG (for L137D).

### ***HCPs and apo-CTDH production, isolation, and purification***

The production and isolation of apo and holo-AnaCTDH (and mutants) was performed as previously described<sup>25</sup>. Briefly, BL21 (DE3) cells from Agilent Technologies were transformed with the pCDF-CTDH plasmid for the apoprotein and simultaneously with three plasmids: pACBETA, pBAD-CrtW, and pCDF-CTDH plasmids for canthaxanthin-binding protein

production. IPTG (0.2 mM) was added to *E. coli* cultures at OD<sub>600</sub>=1-1.2 to induce protein expression and incubated at 24°C over-night. In order to overexpress holoproteins, a previous step of addition of arabinose (0.03%) and overnight incubation at 37°C was performed in order to induce the expression of enzymes required for carotenoid synthesis. Then, the cells were resuspended in lysis buffer (40mM Tris, pH 8, 10% glycerol, 300mM NaCl, 1 mM EDTA, 1 mM phenylmethylsulfonyl fluoride, 1 mM caproic acid, 1mM benzamidic acid, and 50 mgmL<sup>-1</sup> DNase) and then broken in dim light using a French press. The membranes were pelleted, and the supernatant was loaded on a nickel affinity column (Ni-Probond resin; Invitrogen). Proteins was eluted with 250mM imidazole and then dialyzed against 40mM Tris-HCl, pH8.

### *Size Exclusion Chromatography (SEC)*

Purified proteins (10mM) were injected into a Superdex Increase 75 10/300 column (GE Healthcare) equilibrated with 40mM Tris-HCl pH 8, 150mM NaCl. The Akta FLPC system was equipped with a UV detector at 280 nm, 510nm and 560 nm and was set at a flow rate of 0.8 mL/min. The Gel Filtration Calibration Kit LMW (GE Healthcare) was used to calibrate the column.

### *Starting coordinates*

Initial atomic coordinates of CTDH were downloaded from the RCSB site (PDB 6FEJ). The carotenoid ligand coordinates were extracted from the crystal structure of RCP (PDB 4XB4).

### *Molecular Docking*

The SDF format of the ligand was downloaded, charged and minimized for docking by Maestro Schrodinger<sup>51</sup> V11.9 software, using OPLS2005 force field. Protein target processed by AutoDockTools<sup>55</sup> (ADT) and gasteiger charges were assigned to the target<sup>56</sup>. Docking was then performed between the carotenoid ligand from 4XB4 and the target CTDH structure (6FEJ) – by the AutoDock docking software, using the Lamarckian genetic algorithm. We ran 200 docking simulations with 25,000,000 energy evaluations per run and a grid spacing of 0.375 Ang. The lowest energy pose was then chosen for starting molecular dynamics simulation.

### *Molecular dynamics setup*

The lowest energy docking pose carotenoid conformation was then pre-processed with Antechamber<sup>57</sup> to calculate charges and atom types with GAFF forcefield<sup>58</sup>. All atom types were recognized by the forcefield. Then, Tleap program (version 18 Amber16 suite), processed the protein and ligand, with the FF14SB force field for the protein, GAFF for the ligand and TIP3P for water and added the complex into a truncated solvated octahedron box with a 10 Å nonbonded cutoff, generating coordinates and parameter input files for MD simulation<sup>59</sup>. This process was applied for several different starting conformations and configurations (e.i. holo and apo CTDH) (see below) to obtain all MD data.

### *Molecular Dynamics protocol*

The equilibration protocol consisted of an initial minimization and several steps of heating, and a gradual reduction of initial positional restraints<sup>60</sup>. Equilibration of box volume was carried on at constant pressure while production is run at constant volume. Equilibration consisted of a total of 8 ns of MD with a time step of 1fs (stages 1-8) and 2fs (stage9). First a minimization of water and added H atoms (stage 1); 1ns of MD heating, using restraints on every atom at constant volume (NVT ensemble) (stage2); Then 1ns of MD at constant pressure (NPT ensemble) and constant temperature to adjust the density of water with full restraint on protein (stage 3); Then 1ns of MD with lower restraints at NPT (stage 4); Then a second minimization of the side chains (stage 5); Then three stages of 1ns MD at constant pressure with decreasing restraints on backbone (stages 6-8); And finally a 2ns unrestrained run at NTP (stage 9).

The production MD was then performed at NVT for 2000 ns, which was found to be sufficient for reaching convergent results (based on rmsd). Hydrogen-mass repartitioning was used to increased time step to 4fs (dt=0.004). Periodic boundary conditions and Ewald sums (grid spacing of 1 Å) were used to treat long range electrostatic interactions. During production runs, netCDF (Network Common Data Form) trajectory files for each ns of MD were created (writing frequency of 500 and 500.000 steps in total) were performed.

All analyses were performed using CPPTRAJ module<sup>61</sup>, while structural snapshots were extracted from each simulation using the Visual Molecular Dynamics software<sup>62</sup>.

### *MM-GBSA and MM-PBSA*

Molecular docking results from AutoDock could be used to compute binding free energy. However, the lack of sampling results in an inaccurately binding free energy estimate. It is therefore common to evaluate the binding free energy with more accurate method including the MM/PBSA method<sup>63,64</sup>. Thus we used MM-GBSA and MM-PBSA free energy calculations to compute the  $\Delta G$  for the CTDH monomer-monomer interface using Amber built in MMPBSA.py script<sup>65</sup>. This script computes the end-state free energy of the monomer-monomer binding from an ensemble of representative structures obtained from MD by considering the solvation energies of the interacting molecules in addition to molecular mechanics (MM) energies. The contribution of polar solvation energy is calculated with the implicit solvent model (GB or PB) whereas the nonpolar part of the solvation energy is computed from the solvent accessible surface area (SASA). The free energies of binding ( $\Delta G$  binding) can be estimated from the free energies of the reactants. Here, we used 500 snapshots collected from the last 10 ns (spaced by 20 ps) of the production MD simulation to compute MM-PB(GB)SA and free energy decomposition energies.

### *MD analysis tools*

We used CPPTRAJ<sup>61</sup> to analyze the trajectory data and calculation of root mean squared fluctuation (RMSF) of specified atoms/residues and root mean square deviation (RMSD), and monitoring key distances, hydrogen bonds etc. during the simulation. Hydrogen bond criteria applied a donor–acceptor distance cutoff of 3.5 Å and a 120° cutoff for the donor-hydrogen-acceptor angle.

For trajectory visualization we used tools in Visual Molecular Dynamics (VMD) software<sup>62</sup> for structures that were extracted every 10 ps for each 1000 ns trajectory.

Figures from dynamics simulation were created by ChimeraX (Molecular graphics and analyses performed with UCSF ChimeraX, developed by the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco, with support from NIH R01-GM129325 and P41-GM103311)

### *Absorbance measurements and experiments of carotenoid transfer*

Absorbance spectra, kinetics of carotenoid transfer were measured in a Specord S600 spectrophotometer (Analytic Jena) at 23 °C. To analysis the carotenoid transfer as result of the CTDH-HCP interaction holo-CTDH and apo-HCP4 were incubated in a ratio = 1 CTDH to 2.5 HCP4, or holo-HCP1 and apo-CTDH (ratio = 1 HCP1 to 1 monomeric CTDH), and the spectral change was recorded for 1 hour in darkness. The kinetics were obtained from the changes in absorbance at 600 nm. Spectral deconvolution was performed to determine the percentage of carotenoid transfer, using Excel to fit the data to the sum of reference spectra (Holo-CTDH and Holo-HCP) as it was previously described<sup>25</sup>.

The Holoprotein reconstitution upon carotenoid retrieving from membranes was study by incubating 12  $\mu$ M Apo-CTDH-dimers with an *E. coli* canthaxanthin-containing membrane suspension (48  $\mu$ M canthaxanthin (measured by acetone extraction)) for 1 hour at 33°C in darkness. The percentage of holo-protein formed was determined by comparing the spectra of 100% holo-proteins (at 12  $\mu$ M) to those of the supernatant.

### *Crystallization and data collection*

AnaCTDH oligomer crystals diffracting to 2.9 Å were obtained at 20 °C by hanging drop diffusion method when grown in 1.26 M NaH<sub>2</sub>PO<sub>4</sub> and 0.14 M K<sub>2</sub>HPO<sub>4</sub>, pH 5 (being an optimized condition following INDEX HT screen, Hampton research). Large football-shaped crystals were obtained after several days and data sets were collected at Technion Center for Structural Biology (TCSB). AnaCTDH dimers in 2 M urea were crystallized at 15 °C by hanging drop vapor diffusion method (final urea concentration was 1 M). AnaCTDH crystals diffracting to a resolution of 2.43 Å were obtained by growth in 0.1 M citric acid, 25% w/v PEG 3350, pH 4.2 (optimized following INDEX HT screen, Hampton research). Smaller crystals were obtained after several days, and data sets were collected at the European Synchrotron Radiation Facility (ESRF) on beamline ID 30-A1 using MXPressE automatic data collection service. The data sets for both oligomeric states were scaled and merged using MOSFLM<sup>45</sup> and SCALA<sup>46</sup>, respectively. Molecular replacement runs were carried out using Phaser<sup>47</sup>. While the AnaCTDH oligomer structure was failed to be solved using molecular replacement, the structure of the dimeric AnaCTDH was solved and then refined using both

NCSref and Phenix.refine<sup>48</sup>. Structural solution inspections and manual modifications were made using Coot<sup>49</sup>. PDBREDO server<sup>50</sup> was utilized to minimize errors prior to deposition to the PDB. The structure was then examined and compared to other structures using PyMoL<sup>51</sup>. A section of the composite Fo-Fc omit map calculated using the Phenix protocol and visualized in Pymol is shown in Supplementary Figure 6. The overall B-factors of the CTDH structure are higher in the peripheral surfaces of the protein, perhaps due to the presence of 1 M urea in the crystallization liquor (Table 1). Three urea molecules were modeled into densities too large to be modeled as solvent and not corresponding to bound ionic species. The B-factors of the urea molecules are higher than the average protein B-factors, further indicating that the urea molecules required to avoid oligomerization are weakly bound to the protein elements. In the case of the big oligomer, the data set was scaled and merged using MOSFLM<sup>45</sup> and molecular replacement runs were carried out using Phaser<sup>47</sup>, using the structure of one apoCTDH monomer obtained in the presence of 2M urea (PDB code 6FEJ), with truncated CTT as it was suspected a priori for dynamicity. The structure was then refined using Phenix.refine<sup>48</sup>. Structural solution inspections and manual modifications were made using Coot<sup>49</sup>. PDBREDO server<sup>50</sup> was utilized to minimize errors prior to deposition to the PDB. The structure was then carefully examined and compared to other structures using PyMoL<sup>61</sup>.

### ***SAXS of apo-OCP-CTD***

Apo-CTD from *Synechocystis* (residues 165–317), carrying an uncleavable N-terminal hexahistidine tag<sup>26</sup>, was analyzed in 20 mM Tris-HCl buffer (pH 7.6) containing 150 mM NaCl, 0.1 mM EDTA, 2 mM dithiothreitol, and 3% glycerol by synchrotron SAXS at the EMBL P12 beam line (PETRA III, DESY Hamburg, Germany)<sup>53</sup>. SAXS curves collected in a batch mode (1 s exposure time, collected as 20 × 50 ms frames) at different protein concentrations showed substantial concentration dependence, in line with the previous observations using size exclusion chromatography<sup>26</sup>. Since no extrapolation was possible, to ensure the predominance of the dimeric species, the SAXS data at the highest protein concentration (270 μM per monomer) were analyzed and used for modeling. The averaged SAXS curve for the sample was buffer subtracted in PRIMUS<sup>54</sup>. No radiation damage was detected by inspection of the time course of the scattering for protein frames. The Guinier region was linear and was used to determine experimental radius of gyration, R<sub>g</sub>. Pairwise distance distribution, P(r), was calculated by

GNOM52 at  $s \leq 0.28 \text{ \AA}^{-1}$  to determine the maximum particle dimension,  $D_{\text{max}}$ , and the Porod volume. Ab initio shape reconstruction was performed using DAMMIF<sup>31</sup> and three best-fitting models were averaged using DAMAVER<sup>32</sup> to reveal the average core common among the models (DAMFILT). Theoretical SAXS curves and fitting to the experimental data were calculated using CRY SOL<sup>56</sup>. The core models of the apo-CTD dimers were built by superposition of the Synechocystis OCP-CTDs (PDB code: 4XB5; residues 173–305 out of 317) onto the crystallographic Anabaena apo-CTDH dimers in order to preserve the subunit interfaces. Few clashes in the amino acid side chains and flexible loops were relieved manually in Coot<sup>49</sup> in the case of the F-type apo-CTD dimer preserving the head-to-head orientation. In the case of A-type apo-CTD dimer, the local protein–protein docking of the two subunits was required to account for local differences between apo-CTD and apo-CTDH sequences and to relieve clashes. The RosettaDock server<sup>33</sup> was used with the default set of parameters, which resulted in a more connected and realistic top-scoring model (Supplementary Figure 4), devoid of steric clashes but preserving the back-to-back subunit orientation. On the basis of the fixed core dimers (either A or F) thus obtained, the unstructured parts of the protein (23 residues in the N-terminal tail and 12 residues in the C-terminal tail) were modeled using CORAL<sup>57</sup> to minimize the difference between the model-derived and experimental SAXS curves. The modeling procedure was repeated ten times for each scenario to verify that a stable solution is found and the data corresponding to the best-fitting solution are presented along with the statistical analysis using  $\chi^2$  and correlation map P value<sup>34</sup>. To assess the robustness of the solution, the hypothetical modeling by either apo-CTD monomer or trimer (supplemented with the flexible termini) was also performed but yielded much worse fits to the SAXS data.

## References

---

1. Blankenship RE. Origin and early evolution of photosynthesis. *Photosynth. Res.* 1992;33:91–111. doi: 10.1007/BF00039173.
2. Adir N. Elucidation of the molecular structures of components of the phycobilisome: reconstructing a giant. *Photosynth. Res.* 2005;85:15–32. doi: 10.1007/s11120-004-2143-y.
3. Krieger-Liszky A, Fufezan C, Trebst A. Singlet oxygen production in photosystem II and related protection mechanism. *Photosynth. Res.* 2008;98:551–564. doi: 10.1007/s11120-008-9349-3.

4. Niyogi KK, Truong TB. Evolution of flexible non-photochemical quenching mechanisms that regulate light harvesting in oxygenic photosynthesis. *Curr. Opin. Plant Biol.* 2013;16:307–314. doi: 10.1016/j.pbi.2013.03.011.
5. Wilson A, et al. A soluble carotenoid protein involved in phycobilisome-related energy dissipation in cyanobacteria. *Plant Cell.* 2006;18:992–1007. doi: 10.1105/tpc.105.040121.
6. Wilson A, et al. A photoactive carotenoid protein acting as light intensity sensor. *Proc. Natl. Acad. Sci. USA.* 2008;105:12075–12080. doi: 10.1073/pnas.0804636105.
7. Gwizdala M, Wilson A, Omairi-Nasser A, Kirilovsky D. Characterization of the synechocystis PCC 6803 fluorescence recovery protein involved in photoprotection. *Biochim. Biophys. Acta.* 2013; 1827:348–354. doi: 10.1016/j.bbabi.2012.11.001.
8. Sedoud A, et al. The cyanobacterial photoactive orange carotenoid protein is an excellent singlet oxygen quencher. *Plant Cell.* 2014;26:1781–1791. doi: 10.1105/tpc.114.123802.
9. Kerfeld CA, et al. The crystal structure of a cyanobacterial water-soluble carotenoid binding protein. *Structure.* 2003;11:55–65. doi: 10.1016/S0969-2126(02)00936-X.
10. Leverenz RL, et al. Structural and functional modularity of the orange carotenoid protein: distinct roles for the N- and C-terminal domains in cyanobacterial photoprotection. *Plant Cell.* 2014;26:426–437. doi: 10.1105/tpc.113.118588.
11. Gwizdala M, Wilson A, Kirilovsky D. In vitro reconstitution of the cyanobacterial photoprotective mechanism mediated by the Orange Carotenoid Protein in *Synechocystis* PCC 6803. *Plant Cell.* 2011;23:2631–2643. doi: 10.1105/tpc.111.086884.
12. Sutter M, et al. Crystal structure of the FRP and identification of the active site for modulation of OCP-mediated photoprotection in cyanobacteria. *Proc. Natl. Acad. Sci. USA.* 2013;110:10022–10027. doi: 10.1073/pnas.1303673110.
13. Boulay C, Wilson A, D’Haene S, Kirilovsky D. Identification of a protein required for recovery of full antenna capacity in OCP-related photoprotective mechanism in cyanobacteria. *Proc. Natl. Acad. Sci.* 2010;107:11620–11625. doi: 10.1073/pnas.1002912107.
14. Sluchanko NN, et al. The purple Trp288Ala mutant of *synechocystis* OCP persistently quenches phycobilisome fluorescence and tightly interacts with FRP. *Biochim. Biophys. Acta.* 2017;1858:1–11. doi: 10.1016/j.bbabi.2016.10.005.
15. Wilson A, et al. Structural determinants underlying photoprotection in the photoactive orange carotenoid protein of cyanobacteria. *J. Biol. Chem.* 2010;285:18364–18375. doi: 10.1074/jbc.M110.115709.
16. Leverenz RLL, et al. A 12 A carotenoid translocation in a photoswitch associated with cyanobacterial photoprotection. *Science.* 2015;348:1463–1466. doi: 10.1126/science.aaa7234.
17. Gupta S, et al. Local and global structural drivers for the photoactivation of the orange carotenoid protein. *Proc. Natl. Acad. Sci.* 2015;112:E5567–E5574. doi: 10.1073/pnas.1512240112.

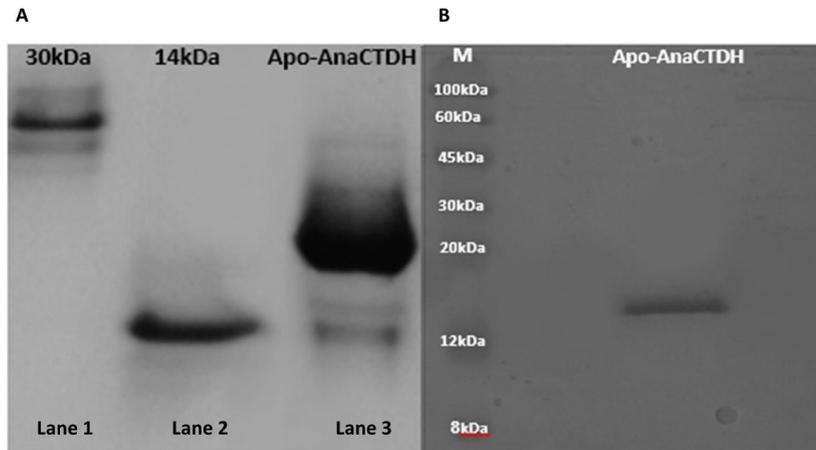
18. Kirilovsky D, Kerfeld CA. Cyanobacterial photoprotection by the orange carotenoid protein. *Nat. Plants*. 2016;2:16180. doi: 10.1038/nplants.2016.180.
19. Melnicki MR, et al. Structure, diversity, and evolution of a new family of soluble carotenoid-binding proteins in cyanobacteria. *Mol. Plant*. 2016;9:1379–1394. doi: 10.1016/j.molp.2016.06.009.
20. Kerfeld CA. Water-soluble carotenoid proteins of cyanobacteria. *Arch. Biochem. Biophys*. 2004;430:2–9. doi: 10.1016/j.abb.2004.03.018.
21. Kirilovsky D, Kerfeld CA. The orange carotenoid protein in photoprotection of photosystem II in cyanobacteria. *Biochim. Biophys. Acta*. 2012;1817:158–166. doi: 10.1016/j.bbabi.2011.04.013.
22. Lopez-Igual R, et al. Different functions of the paralogs to the N-terminal domain of the Orange Carotenoid Protein in the cyanobacterium *Anabaena* sp. PCC 7120. *Plant Physiol*. 2016;171:1852–1866. doi: 10.1104/pp.16.00502.
23. Bao H, Melnicki MR, Kerfeld CA. Structure and functions of Orange Carotenoid Protein homologs in cyanobacteria. *Curr. Opin. Plant Biol*. 2017;37:1–9. doi: 10.1016/j.pbi.2017.03.010.
24. Lechno-Yossef S, Melnicki MR, Bao H, Montgomery BL, Kerfeld CA. Synthetic OCP heterodimers are photoactive and recapitulate the fusion of two primitive carotenoproteins in the evolution of cyanobacterial photoprotection. *Plant J*. 2017;91:646–656. doi: 10.1111/tbj.13593.
25. Muzzopappa F, et al. The paralogs to the C-terminal domain of the cyanobacterial OCP are carotenoid donors to HCPs. *Plant Physiol*. 2017;175:1283–1303. doi: 10.1104/pp.17.01040.
26. Moldenhauer M, et al. Assembly of photoactive orange carotenoid protein from its domains unravels a carotenoid shuttle mechanism. *Photosynth. Res*. 2017;133:327–341. doi: 10.1007/s11120-017-0353-3.
27. Dines M, Sendersky E, David L, Schwarz R, Adir N. Structural, functional, and mutational analysis of the NblA protein provides insight into possible modes of interaction with the phycobilisome. *J. Biol. Chem*. 2008;283:30330–30340. doi: 10.1074/jbc.M804241200.
28. Dines M, Sendersky E, Schwarz R, Adir N. Crystallization of sparingly soluble stress-related proteins from cyanobacteria by controlled urea solubilization. *J. Struct. Biol*. 2007;158:116–121. doi: 10.1016/j.jsb.2006.10.021.
29. Krissinel E, Henrick K. Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol*. 2007;372:774–797. doi: 10.1016/j.jmb.2007.05.022.
30. Maksimov EG, et al. The unique protein-to-protein carotenoid transfer mechanism. *Biophys. J*. 2017;113:402–414. doi: 10.1016/j.bpj.2017.06.002.
31. Franke D, Svergun DI. DAMMIF, a program for rapid ab-initio shape determination in small-angle scattering. *J. Appl. Crystallogr*. 2009;42:342–346. doi: 10.1107/S0021889809000338.
32. Volkov VV, Svergun DI. Uniqueness of ab initio shape determination in small-angle scattering. *J. Appl. Crystallogr*. 2003;36:860–864. doi: 10.1107/S0021889803000268.

33. Lyskov S, Gray JJ. The RosettaDock server for local protein–protein docking. *Nucleic Acids Res.* 2008;36:W233–W238. doi: 10.1093/nar/gkn216.
34. Franke D, Jeffries CM, Svergun DI. Correlation map, a goodness-of-fit test for one-dimensional X-ray scattering spectra. *Nat. Methods.* 2015;12:419–422. doi: 10.1038/nmeth.3358.
35. Grosdidier A, Zoete V, Michielin O. SwissDock, a protein-small molecule docking web service based on EADock DSS. *Nucleic Acids Res.* 2011;39:W270–W277. doi: 10.1093/nar/gkr366.
36. Cianci M, et al. The molecular basis of the coloration mechanism in lobster shell:  $\beta$ -crustacyanin at 3.2-Å resolution. *Proc. Natl. Acad. Sci.* 2002;99:9795–9800. doi: 10.1073/pnas.152088999.
37. Peng PP, et al. The structure of allophycocyanin B from *Synechocystis* PCC 6803 reveals the structural basis for the extreme redshift of the terminal emitter in phycobilisomes. *Acta Crystallogr. D. Biol. Crystallogr.* 2014;70:2558–2569. doi: 10.1107/S1399004714015776.
38. Lindahl M, Kieselbach T. Disulphide proteomes and interactions with thioredoxin on the track towards understanding redox regulation in chloroplasts and cyanobacteria. *J. Proteomics.* 2009;72:416–438. doi: 10.1016/j.jprot.2009.01.003.
39. Florencio FJ, Pérez-Pérez ME, López-Maury L, Mata-Cabana A, Lindahl M. The diversity and complexity of the cyanobacterial thioredoxin systems. *Photosynth. Res.* 2006;89:157–171. doi: 10.1007/s11120-006-9093-5.
40. Lindahl M, Florencio FJ. Thioredoxin-linked processes in cyanobacteria are as numerous as in chloroplasts, but targets are different. *Proc. Natl. Acad. Sci.* 2003;100:16107–16112. doi: 10.1073/pnas.2534397100.
41. Kashino Y, Koike H, Satoh K. An improved sodium dodecyl sulfate-polyacrylamide gel electrophoresis system for the analysis of membrane protein complexes. *Electrophoresis.* 2001;22:1004–1007. doi: 10.1002/1522-2683(200106)22:6<1004::AID-ELPS1004>3.0.CO;2-Y.
42. Kiser, P. D., Golczak, M., Lodowski, D. T., Chance, M. R. & Palczewski, K. Crystal structure of native RPE65, the retinoid isomerase of the visual cycle. *Proc. Natl. Acad. Sci.* 106, 17325–17330 (2009).
43. Kiser, P. D. et al. Structure of RPE65 isomerase in a lipidic matrix reveals roles for phospholipids and iron in catalysis. *Proc. Natl. Acad. Sci.* 109, E2747–E2756 (2012).
44. Kiser, P. D. et al. Catalytic mechanism of a retinoid isomerase essential for vertebrate vision. *Nat. Chem. Biol.* 11, 409 (2015).
45. Battye, T. G. G., Kontogiannis, L., Johnson, O., Powell, H. R. & Leslie, A. G. W. *iMOSFLM*: a new graphical interface for diffraction-image processing with *MOSFLM*. *Acta Cryst.* 67, 271–281 (2011).
46. Evans P. Scaling and assessment of data quality. *Acta Crystallogr. Sect. D. Biol. Crystallogr.* 2006;62:72–82. doi: 10.1107/S0907444905036693.
47. McCoy AJ, et al. Phaser crystallographic software. *J. Appl. Crystallogr.* 2007;40:658–674. doi: 10.1107/S0021889807021206.

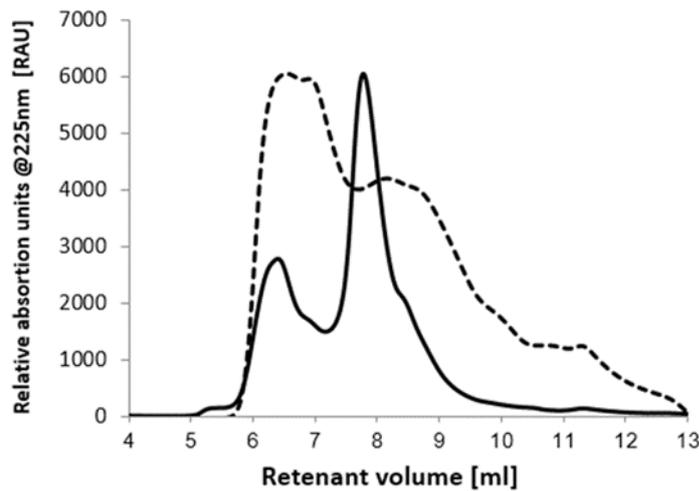
48. Afonine PV, et al. Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr. Sect. D. Biol. Crystallogr.* 2012;68:352–367. doi: 10.1107/S0907444912001308.
49. Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D. Biol. Crystallogr.* 2004;60:2126–2132. doi: 10.1107/S0907444904019158.
50. Joosten RP, Long F, Murshudov GN, Perrakis A. The PDB\_REDO server for macromolecular structure model optimization. *IUCrJ.* 2014;1:213–220. doi: 10.1107/S2052252514009324.
51. DeLano, W. L. The PyMOL Molecular Graphics System. <http://www.pymol.org> (2002).
52. Arnold K, Bordoli L, Kopp J, Schwede T. The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. *Bioinformatics.* 2006;22:195–201. doi: 10.1093/bioinformatics/bti770.
53. Blanchet CE, et al. Versatile sample environments and automation for biological solution X-ray scattering experiments at the P12 beamline (PETRA III, DESY) J. *Appl. Crystallogr.* 2015;48:431–443. doi: 10.1107/S160057671500254X.
54. Konarev PV, Volkov VV, Sokolova AV, Koch MHJ, Svergun DI. PRIMUS: a Windows PC-based system for small-angle scattering data analysis. *J. Appl. Crystallogr.* 2003;36:1277–1282. doi: 10.1107/S0021889803012779.
55. Svergun DI. Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *J. Appl. Crystallogr.* 1992;25:495–503. doi: 10.1107/S0021889892001663.
56. Svergun D, Barberato C, Koch MHJ. CRY SOL—a program to evaluate X-ray solution scattering of biological macromolecules from atomic coordinates. *J. Appl. Crystallogr.* 1995;28:768–773. doi: 10.1107/S0021889895007047.
57. Petoukhov MV, et al. New developments in the ATSAS program package for small-angle scattering data analysis. *J. Appl. Crystallogr.* 2012;45:342–350. doi: 10.1107/S0021889812007662.
58. Trott, O. & Olson, A. J. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* 31, 455–461 (2010).
59. Gasteiger, J. & Marsili, M. Iterative partial equalization of orbital electronegativity—a rapid access to atomic charges. *Tetrahedron* 36, 3219–3228 (1980).
60. Wang, J., Wang, W., Kollman, P. A. & Case, D. A. Antechamber: an accessory software package for molecular mechanical calculations. *J. Am. Chem. Soc.* 123, U403 (2001).
61. Wang, J., Wolf, R. M., Caldwell, J. W., Kollman, P. A. & Case, D. A. Development and testing of a general amber force field. *J. Comput. Chem.* 25, 1157–1174 (2004).
62. Case, D. A. et al. The Amber biomolecular simulation programs. *J. Comput. Chem.* 26, 1668–1688 (2005).

63. Nguyen, H., Maier, J., Huang, H., Perrone, V. & Simmerling, C. Folding simulations for proteins with diverse topologies are accessible in days with a physics-based force field and implicit solvent. *J. Am. Chem. Soc.* 136, 13959–13962 (2014).
64. Roe, D. R. & Cheatham III, T. E. PTRAJ and CPPTRAJ: software for processing and analysis of molecular dynamics trajectory data. *J. Chem. Theory Comput.* 9, 3084–3095 (2013).
65. Humphrey, W., Dalke, A. & Schulten, K. VMD: visual molecular dynamics. *J. Mol. Graph.* 14, 33–38 (1996).
66. Srinivasan, J., Cheatham, T. E., Cieplak, P., Kollman, P. A. & Case, D. A. Continuum solvent studies of the stability of DNA, RNA, and phosphoramidate–DNA helices. *J. Am. Chem. Soc.* 120, 9401–9409 (1998).
67. Kollman, P. A. et al. Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models. *Acc. Chem. Res.* 33, 889–897 (2000).
68. Genheden, S. & Ryde, U. The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities. *Expert Opin. Drug Discov.* 10, 449–461 (2015).

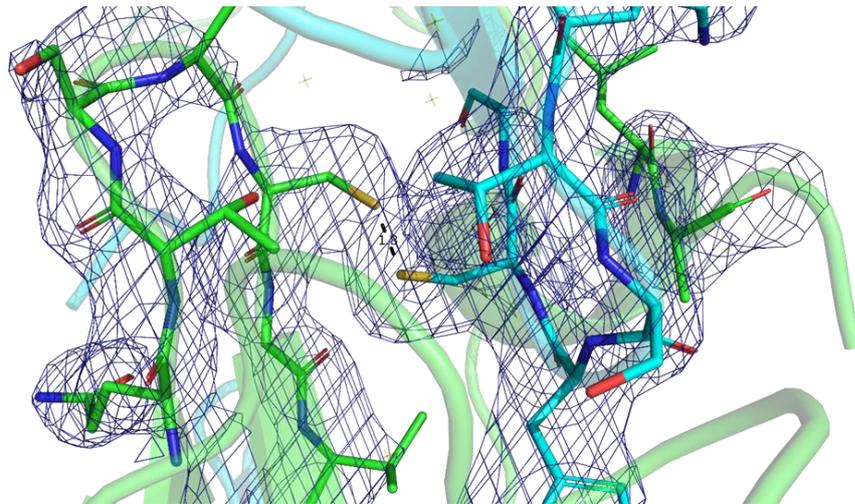
## Appendix: Supplementary figures



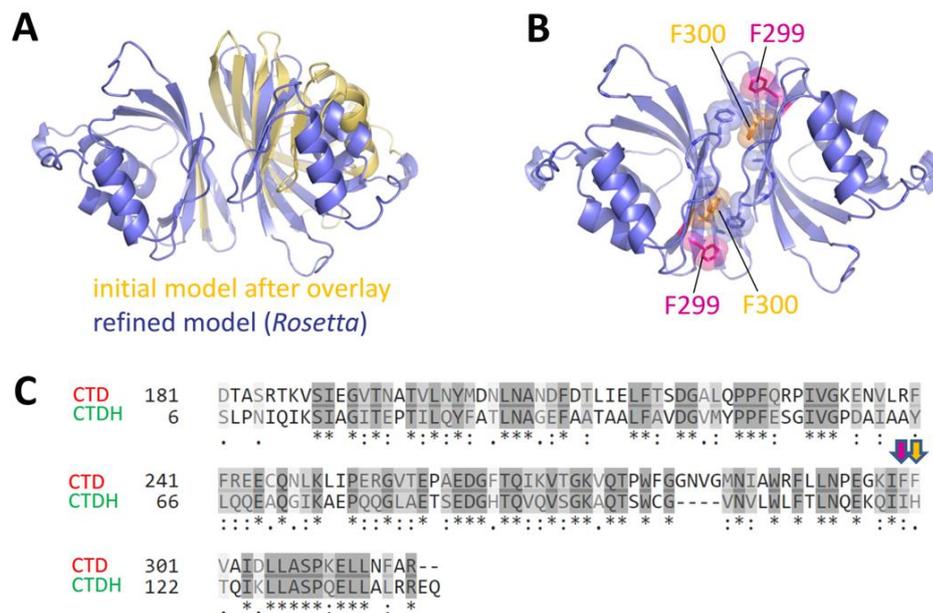
**Supplementary Figure 1 | Assessment of the oligomeric state of apo-AnaCTDH by native-PAGE and SDS-PAGE.** (A) A typical run of an apo-AnaCTDH dimer in native-PAGE (lane 3) with markers (lanes 1, 2). (B) SDS-PAGE reveals that the reduced form of apo-AnaCTDH is in good accordance with a monomeric MW (15kDa).



**Supplementary Figure 2 | Effect of urea on the oligomeric state of CTDH observed by size exclusion chromatography.** Crude CTDH mixture elution profile (dashed bold line); following introduction of 2M urea (solid bold line), a partial shift into a smaller MW assembly is observed, matching the predicted mobility of an AnaCTDH dimer, by column calibration.

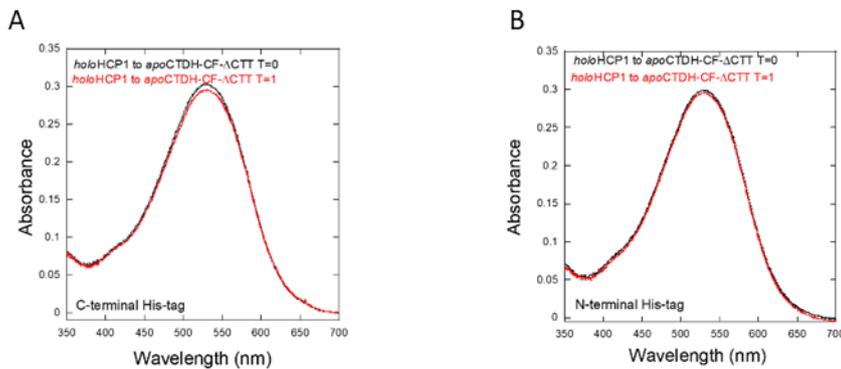


**Supplementary Figure 3 | Representative example of CTDH structure electron density.** Composite Fo-Fc omit contoured at  $1\sigma$  overlaid onto disulfide bond between of monomer A (green cartoon and carbons) and monomer B (cyan cartoon and carbons) in the adjacent asymmetric unit. The sulfur atoms are in yellow stick, and the disulfide bond is 1.8Å in length.



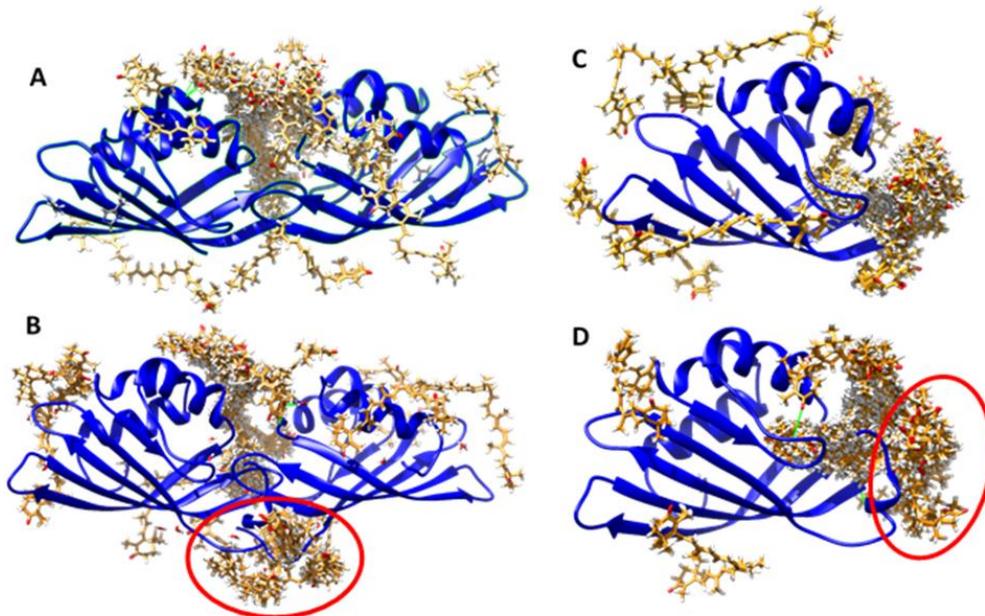
**Supplementary Figure 4 | Modeling of the Apo-CTD dimer in a “back-to-back” conformation.** (A) The A-type of Apo-CTD dimer was obtained by a spatial overlay with the crystal structure of Apo-CTDH dimer (yellow) and then refined by using local protein-protein docking in *RosettaDock* (violet). (B) The refined model was devoid of steric clashes and revealed the potential role of the residues F299 and F300 (numbering of the full-length OCP from *Synechocystis*) in the formation of Apo-CTD dimer in contrast to Apo-CTDH, which lacks these phenylalanine residues. The absence of these residues in AnaCTDH may explain the difference in the ability of AnaCTDH and OCP-CTD to dimerize in the absence

of carotenoid. (C) The alignment of primary structures of OCP-CTD and AnaCTDH showing sequence similarity (greyscale) and the difference in positions 299 and 300 of *Synechocystis* OCP-CTD. Note the 4 amino acid gap in AnaCTDH corresponding to the loop between  $\beta 5$  and  $\beta 6$  strands (see text for more details).



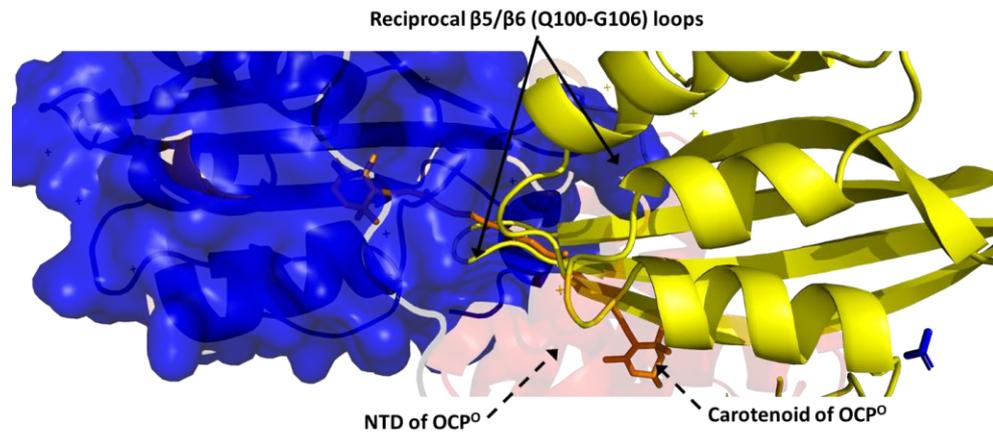
**Supplementary Figure 5 | Carotenoid transfer from HCPs to Ana-CTDH-C103F- $\Delta$ CTT.**

(A) Absorbance spectra at time=0 and after a one hour incubation of holo-HCP1 and apo-CTDH-C103F- $\Delta$ CTT-terminal tail-Cter. (F) Absorbance spectra at time=0 and after a one hour incubation of holo-HCP1 and apo-CTDH-C103F- $\Delta$ C-terminal tail-Nter.

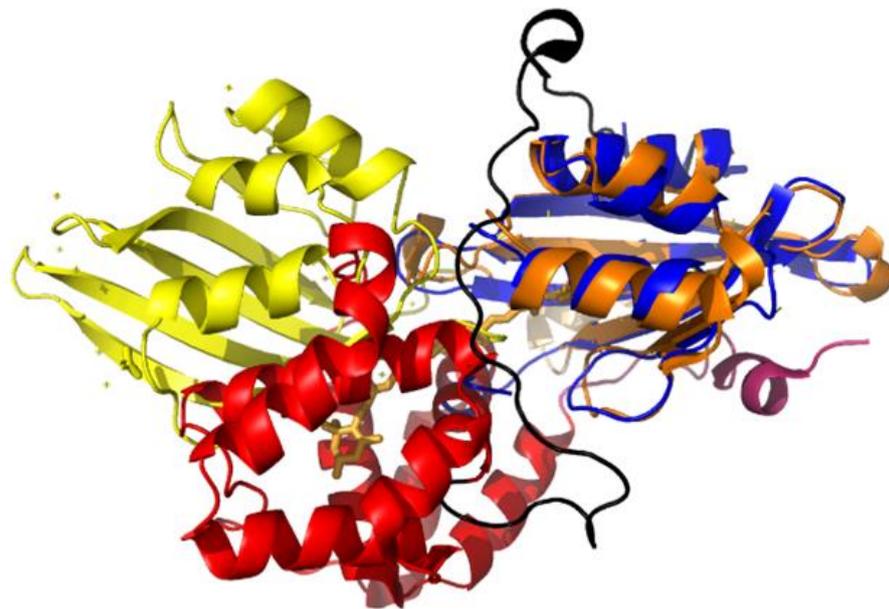


**Supplementary Figure 6 | Docking of canthaxanthin chromophore to apo-AnaCTDH**

**monomers and dimers (with/out C-terminal tail).** (A) Cartoon representation of CTDH dimer without C-terminal tail with all the carotenoid docking clusters super-positioned onto the protein (B) Same as A, but with C-terminal tail. (C) same as A, but with a monomer. (D) same as C, but with C-terminal tail. Circles denote the putative carotenoid binding site entrance.



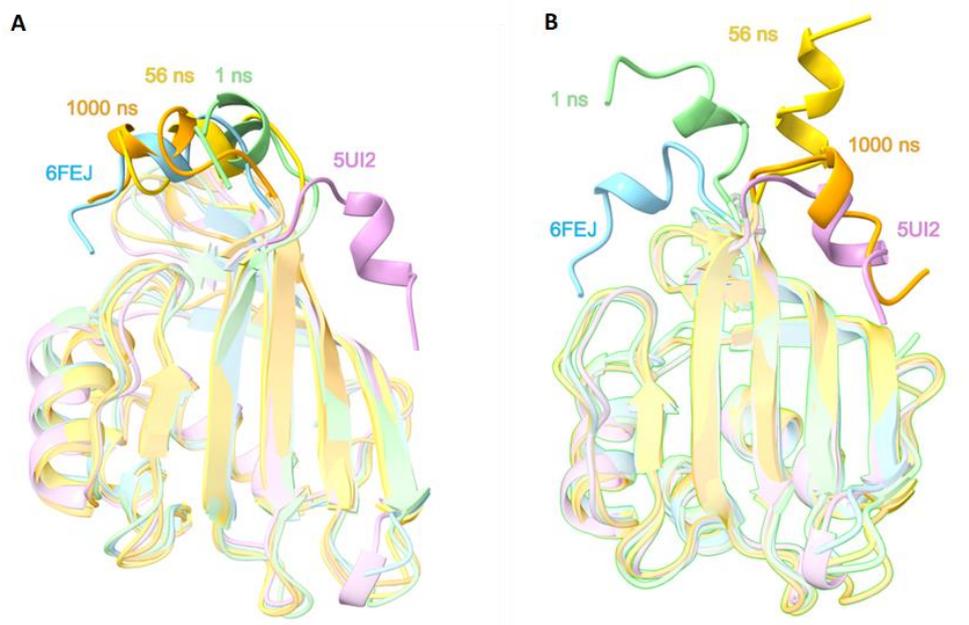
**Supplementary Figure 7 |  $\beta 5/\beta 6$  loops (Q100-G106) occupy carotenoid designated cavities in apo-AnaCTDH dimer.** CTDH monomer surface (blue) super-positioned with CTD of OCP<sup>0</sup>. NTD of OCP<sup>0</sup> (transparent cartoon in the back, indicated using a dashed arrow) shares carotenoid (orange sticks, also indicated using a dashed arrow) with the CTD of OCP<sup>0</sup>. Second CTDH monomer (yellow cartoon) forms a homodimer in a different orientation with respect to NTD. Reciprocal Q100-G106 (QTSWCG) loops blocks the cavity where carotenoid should be situated in the holo form, as indicated by black full arrows (one points to the yellow loop penetrating the blue surface, and the second points to the second loop' surface in blue, penetrating the yellow cartoon).



**Supplementary Figure 8 | CTDH dimer planarity versus OCP<sup>0</sup> bend.** Apo-AnaCTDH dimer is superimposed with OCP<sup>0</sup> by overlapping an apo-AnaCTDH monomer (blue cartoon) with CTD (orange cartoon). NTD (red) facilitates bent carotenoid encapsulation (bright orange sticks) and connected to CTD through the flexible loop (black). The second apo-AnaCTDH monomer (yellow) is nearly on the same axis (horizontal) to the first monomer.



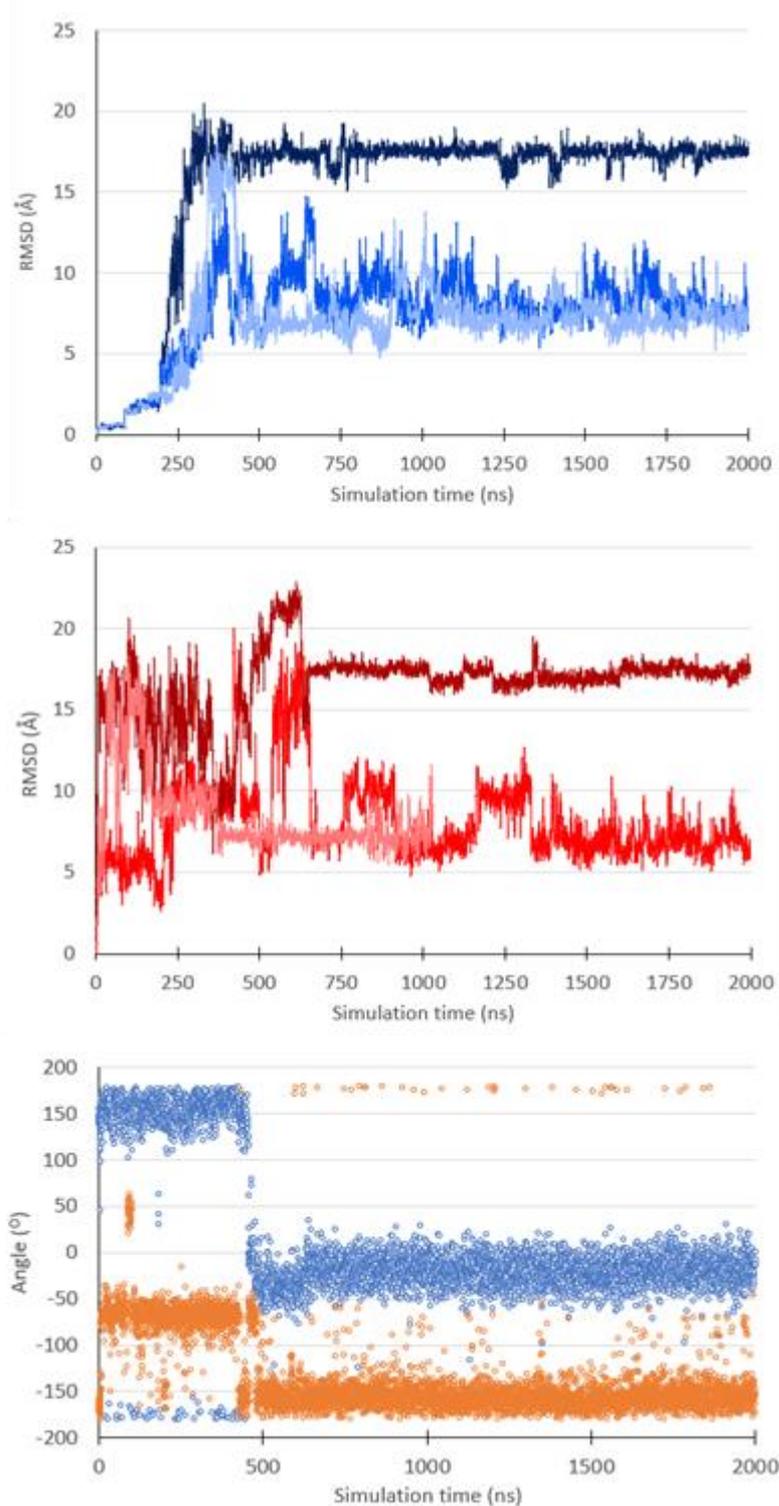
**Supplementary figure 9 | apoCTDH crystal structure under native conditions.** Following MR, twelve protomers were successfully modelled into the electron density comprising the asymmetric unit.



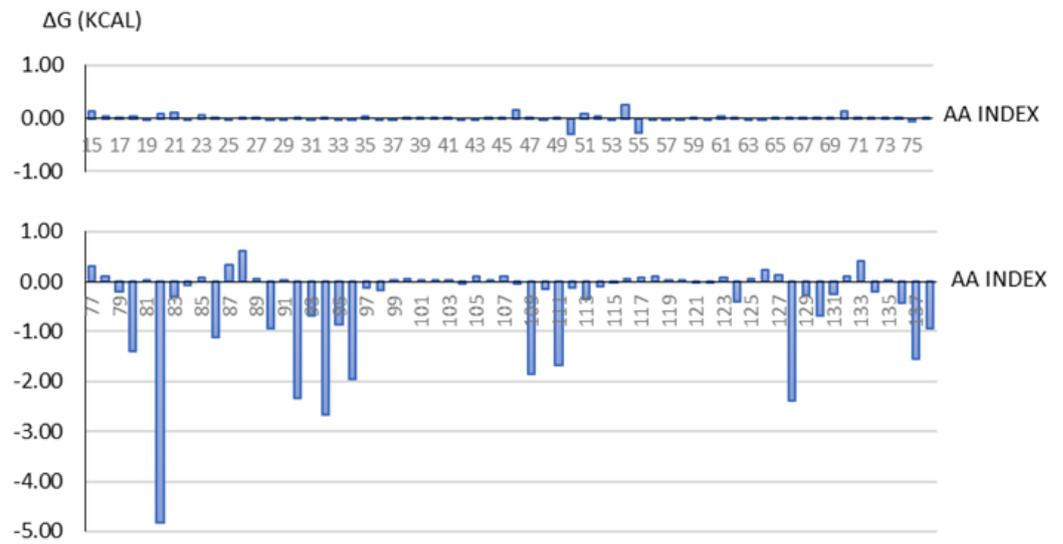
**Supplementary figure 10 | CTT conformational change for apo- and docked-holo-CTDH during MD.** All panels show the protein in cartoon representation. (A) ApoCTDH and (B) CAN docked-holoCTDH. In both A and B, frames presented were sampled at  $t = 1$  ns (green),  $t = 56$  ns (gold) and  $t = 1000$  ns (orange). OCP<sup>O</sup> CTD from PDB 5UI2 (violet) and CTDH structure PDB 6FEJ (cyan), are also shown for comparison. The carotenoid in panel B was omitted for clarity.

	Native	2 M urea
<b>Data collection</b>		
Space group	P212121	P6522
<b>Cell dimensions</b>		
a, b, c (Å)	65.65, 112.28, 318.54	81.08, 81.08, 162.94
$\alpha, \beta, \gamma$ (°)	90.0, 90.0, 90.0	90.0, 90.0, 120.0
Resolution (Å)	47.99–2.90 (3.06–2.90) <sup>a</sup>	53.19–2.75 (2.85–2.75) <sup>a</sup>
R merge	0.06 (0.42) <sup>a</sup>	0.07 (0.33) <sup>a</sup>
$\langle I/\sigma(I) \rangle$	16.7 (3.3) <sup>a</sup>	11.7 (5.0) <sup>a</sup>
Completeness (%)	99.9 (100) <sup>a</sup>	99.8 (100) <sup>a</sup>
Redundancy	3.7 (3.9) <sup>a</sup>	5.6 (6.1) <sup>a</sup>
<b>Refinement</b>		
Resolution (Å)	47.99–2.90 (3.06–2.90) <sup>*</sup>	53.19–2.75
No. of reflections	50566	8742
Rwork/Rfree	0.213/0.249	0.241/0.292
Completeness (%)	99.8	99.58
No. of non-hydrogen atoms	11426	1855
Protein	11351	1828
Ligands	-	12
Water	75	15
<b>B factors (Å<sup>2</sup>)</b>		
Protein	65.67	67.31
Ligands	-	78.36
Water	46.73	45.16
<b>R.m.s. deviations</b>		
Bond lengths (Å)	0.014	0.012
Bond angles (°)	1.712	1.47

**Table S1** . Data-collection and final refinement statistics of both dimeric (UREA 2M) and dodecamer apoCTDH structures. Values in parentheses are for the outer shell.



**Supplementary figure 11 | Raw, MD-derived structural changes of the CTT.** (A) Triplicate of MD runs of docked holoCTDH were examined via the displacement of the CTT from the starting position, as a function of simulation time. Each blue tone represents one run out of the triplicate. (B) Same as A, but for apoCTDH. (C) Change in psi (blue) and phi (red) angles around A129-S130 as a function of simulation time.



**Supplementary figure 12** | Amino acids facilitating back-to-back interaction as calculated by MM-PBSA. MMPBSA analysis (decomposition by residues) of the back-to-back apoCTDH dimer.



## **Part II: results**

---

### **Chapter IV**

# **Molecular evolution of the Orange Carotenoid Protein.**

## *Summary of work*

---

This chapter is based on:

Muzzopappa F, Wilson A and Kirilovsky D. Interdomain interactions reveal the molecular evolution of the orange carotenoid protein. *Nature Plants*. 2019 (in press)

## *Contribution to this work*

In this work, as first author I performed all the biochemical and functional characterization of the different OCPs, the mutant OCPs and the isolated domains (interaction experiments). Most of the plasmids for protein expression were constructed by Adjele Wilson, I prepared the plasmid for overexpression of the *Synechocystis* OCP1-E191K, *Tolypothrix* OCP2-R185E, *Scytonema* OCPX, *Scytonema* OCPX-E174K and OCP2/OCPX from *Synechocystis* PCC 7509. I also performed the bioinformatic analysis.

## *Resume of the article*

In chapter 2, we showed that the OCP-domains (CTD and NTD) and their paralogs (CTDH and HCP) have evolved very different interactions. CTDH and HCP can transiently interact leading to an efficient carotenoid transfer between the proteins. On the other hand, *in vitro* interaction of isolated OCP-domains results in a stable OCP-like complex with similar spectral and functional properties than the OCP. OCP-like complex formed by OCP1 (chapter II and Moldenhauer et al. 2017) and by OCP2 (Lechno-Yossef et al. 2017) domains were previously analyzed, both are capable of photoactivation but only the resultant free domains from OCP2 were able to reconstitute the complex. This raised questions about the evolution of the OCP. *Is the linker required for photoactivation or back-conversion? Since OCP2 was proposed as ancestral OCP, are its domains similar to the ancestors of CTD and NTD that formed the first OCP?* Thus, we address in this chapter the role of the flexible linker and the inter-domain interactions in the modern OCPs and we inferred from these data, complemented by a new phylogenetic analysis of the OCP sequences, the evolutionary history of the OCP.

First, we re-analyzed the sequences of the OCP protein family. We constructed a new phylogenetic tree that we rooted by using an outgroup. This tree suggests that there was a first divergence that gave rise to OCPX and the ancestor of OCP2 and OCP1. Then, the latter one diverges into OCP1 and OCP2, contrary to the previous hypothesis that OCP2 was originated first. Secondly, we characterized the third subclade of OCP, OCPX (uncharacterized) and compared it with the other subclades. We found that similar to OCP2, OCPX back-conversion to OCP<sup>0</sup> is faster than OCP1 and is not accelerated by

FRP, indicating that FRP-interaction is a trait of OCP1 acquired later in evolution. We also found that similar to OCP1, OCPX in solution is in equilibrium between the monomeric and dimeric form. On the other hand, OCP2 is monomeric at all the concentrations tested.

To study the role of the linker we utilized two approaches. The first one was to reconstitute the OCP-like complex with domains from each subclade. In these complexes the linker is absent. Our results indicate that OCP2-domains have the strongest interaction and OCP1-domains the weakest. We found that photoactivation of OCP1-like and OCPX-like complexes is faster than the respective OCPs, in which the linker is present. On the other hand, the absence of linker in OCP2-like complex slowed down the photoactivation. This result indicates that the linker plays a role on photoactivation and it is different in OCP1/OCPX and OCP2. On the other hand, reconstitution of the complex after photoactivation is not possible for OCPX-domains or OCP1-domains, only OCP2-domains are able to such reaction, possibly due to their strong interaction.

The second approach was destined to decipher the role of specific amino acids of the linker. Using OCP1 from *Synechocystis* as model, we introduced mutations in the most conserved residues. Overall, almost all the mutations showed a faster photoactivation than the OCP-WT and a slower back conversion. In addition, they allowed photoactivation at lower light intensities than the WT. The most affected mutants were those in which the residues E174 and R185 were modified. In addition, the deletion of residues from E174 to P179 also largely affected the back conversion and the interaction with the PBS indicating that the length of the linker is important to allow proper binding to the PBS and proper interaction between the domains. In addition, double mutant P175A-P179A also showed a lower quenching activity than OCP-WT, suggesting that those residues stabilize the OCP-PBS complex. FRP could accelerate the back-conversion in all the mutants, indicating that the mutations did not affect OCP-FRP interaction. Then, we introduced the mutations that most affected the OCP1 photocycle in OCP2 and OCPX. Both OCPX-E174K and OCP2-R185E showed slower back-conversion than the respective WTs. In the case of OCPX, the mutation in the linker increases the photoactivation rate, similar to OCP1. On the other hand, linker mutation in OCP2 slowed down the photoactivation rate.

Since the phylogenetic analysis assigned OCPX as the most ancestral OCP and in this case the domains were not able to reconstitute the OCP-like complex after photoactivation, we hypothesize that addition of the linker during gene-fusion enable the back-conversion to OCP<sup>0</sup> and the photocycle. This is also supported by the fact that the presence of conserved amino acids in the linker of the three modern OCPs accelerates this back-conversion. Moreover, the linker provided an extra regulation of the photoactivation. In OCP1 and OCPX, which are reported to be expressed constitutively, this is important because it inhibits photoactivation at low light intensity. On the other

hand, the linker evolved a positive regulation of the photoactivation in OCP2, probably to cope with the strong inter-domain interaction. In general, our results present a complete picture of the evolution of the OCP and support our hypothesis that OCPX was originated first and later in the evolution, OCP1 and OCP2 appeared.

## ***Interdomain interactions reveal the molecular evolution of the Orange Carotenoid Protein.***

---

Fernando Muzzopappa<sup>1</sup>, Adjélé Wilson<sup>1</sup> and Diana Kirilovsky<sup>1</sup>.

### **Affiliations**

<sup>1</sup>Institute for Integrative Biology of the Cell (I2BC), CEA, CNRS, Université Paris-Sud, Université Paris-Saclay, 91198 Gif sur Yvette, France.

### ***Abstract***

The photoactive Orange Carotenoid Protein (OCP) is a blue-light intensity sensor involved in cyanobacterial photoprotection. Three OCP families co-exist (OCPX, OCP1 and OCP2) which originated from the fusion of ancestral domain genes. Here we report the first characterization of an OCPX and the evolutionary characterization of OCP paralogues focusing on the role of the linker connecting the domains. The addition of the linker with specific amino acids enabled the photocycle of the OCP ancestor. OCPX is the paralog closest to this ancestor. A second diversification gave rise to OCP1 and OCP2. OCPX and OCP2 present fast deactivation and weak antenna interaction. In OCP1, the OCP deactivation became slower and the interaction with the antenna became stronger requiring another protein to detach the OCP from the antenna and accelerate its deactivation. OCP2 lost the tendency to dimerize present in OCPX and OCP1 and the role of its linker is slightly different giving a less controlled photoactivation.

### ***Introduction***

---

The Orange Carotenoid Protein, the light intensity sensor involved in cyanobacterial photoprotection, is a modular photoactive protein composed of an N-terminal domain (NTD) and a C-terminal domain (CTD) connected by a long flexible linker loop<sup>1-3</sup>. One keto-carotenoid molecule (3-hydroxy equinenone) is shared between the two domains<sup>4</sup>. In darkness, the OCP remains in the inactive “orange form” (OCP<sup>0</sup>) and adopts a closed conformation stabilized by the carotenoid and by interactions between the two domains<sup>4</sup>. Under strong blue-green light, the OCP is photoactivated and converted into the red active form (OCP<sup>R</sup>)<sup>5</sup>. Upon light absorption, the H-bonds between the CTD and the carotenoid are broken, and the carotenoid migrates (12Å) into the NTD<sup>6</sup> in about 10 μs<sup>7</sup>. Then the inter-domain interactions are disrupted and the protein adopts an open configuration<sup>6,8,9</sup> with a higher hydrodynamic radius<sup>10,11</sup>. By interacting with the core of the cyanobacterial

antenna, the phycobilisome (PBS), the NTD of OCP<sup>R</sup> induces the dissipation of the excitation energy as heat<sup>6,12–14</sup>. The OCP<sup>R</sup> slowly relaxes back to OCP<sup>O</sup> in darkness<sup>5</sup>. Another cyanobacterial protein, called Fluorescence Recovery Protein (FRP), accelerates this process<sup>15,16</sup> by increasing the probability of the proper configuration of domains and carotenoid in which they may interact to form OCP<sup>O</sup><sup>17</sup>. The main binding site of the FRP is in a CTD region that is exposed only in the red active form<sup>16,18,19,20</sup>. FRP binding also helps OCP detachment from PBSs<sup>13,21</sup>. Both FRP activities are essential to recover the full photosynthetic capacity under low light conditions.

Phylogenetic analysis of the OCP family has shown the existence of three different OCP clades: OCP1, OCP2 and OCPX<sup>22</sup>. All previous studies (except one) characterizing the activity and photoactivation of the OCP were done using OCP1 proteins. Recently, the OCP2, from *Tolypothrix* sp. PCC 7601 (hereafter *Tolypothrix*), was characterized<sup>22,23</sup>. No OCPX was characterized in the past. In addition to the OCP genes, genes of homologues of OCP domains are present in many cyanobacterial genomes<sup>4,24,25,26</sup>. The existence of these homologues have suggested a modular model for the evolution of OCP<sup>24,23</sup>.

The OCP1 and OCP2 domains can be isolated independently as carotenoid proteins (holo-protein) or apo-proteins (without carotenoid), depending on the expression system<sup>27,23,25</sup>. Interaction between holo-CTD and apo-NTD results in the formation of a complex, which is spectrally similar to OCPO (hereafter, OCP-like complex)<sup>23,27,25,28</sup>. This complex can be photoactivated leading to efficient carotenoid translocation from the CTD to the NTD and complex disassembly<sup>23,25,27</sup>. Afterwards, reassembling of the OCP-like complex from holo-NTD and apo-CTD was only observed in the case of the OCP2 domains<sup>23</sup>. This suggested that OCP2 is more primitive than OCP1 and hinted towards an evolutionary role of the linker<sup>23</sup>.

The OCP characteristics as a light-regulated modular protein raise the possibility of using the OCP as a photoswitch for an optogenetic toolbox, adding a new class of chromophore and a new type of triggering mechanism in which initially coupled sensor and effector modules reversibly separate upon light absorption. To this end, an understanding of the inter-domain interactions and of the role of the linker connecting the two domains during the photocycle is required for all three OCP types. In this work we did an evolutionary and functional study of the three OCP clades and report the first characterization of two OCPXs, from *Scytonema hofmanni* PCC 7110 (hereafter *Scytonema*) and *Synechocystis* PCC 7509. Our results show the essential role of the linker and specifically of some of its amino acids, in the OCP photocycle. We propose a new evolutionary model for OCP in which a primitive OCP diversified into OCPX and the ancestor of the modern OCP2 and OCP1, which arose as the result of a second diversification.

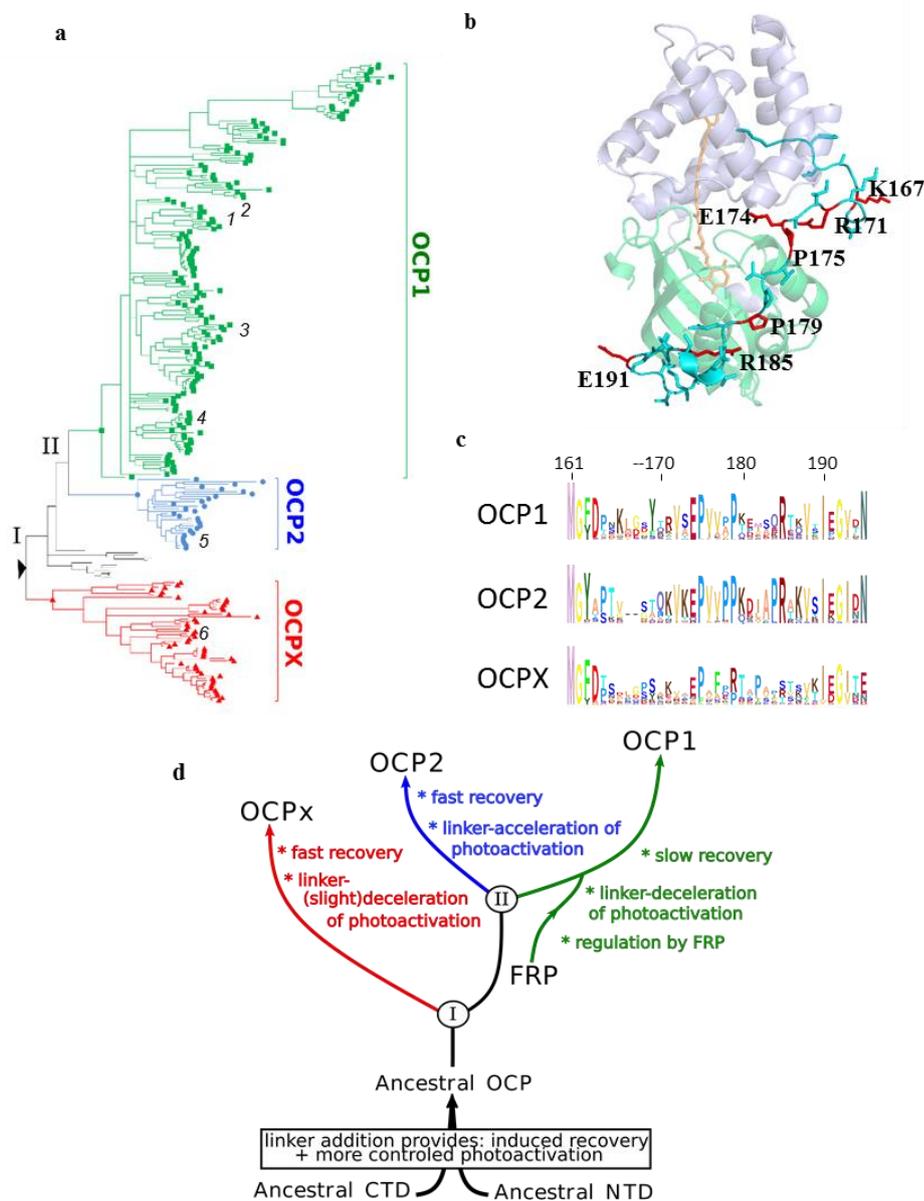
## Results

---

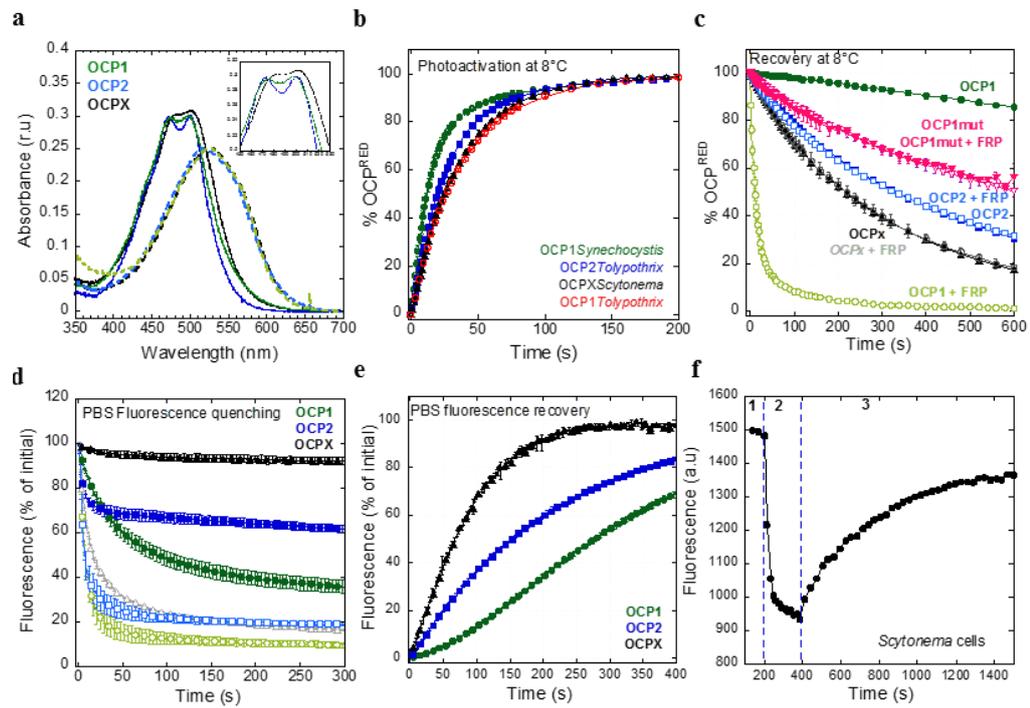
### *OCP evolution and photocycle*

Most of the OCP-containing cyanobacteria strains have only OCP1 (139). A large number contain only OCPX (35). By contrast, most of strains containing OCP2 also harbor OCP1 or OCPX (15); OCP2 is found alone in only 5 strains. In order to better understand the evolutionary history of OCP, we re-analyzed the phylogenetic history of OCP. This time we rooted the tree using CTDH and NTF2 as outgroups (see Materials and Methods and Supplementary Fig. 1). The analysis of the rooted tree strongly suggests that there was a first diversification (Fig. 1A, I) that gave rise to an ancestral OCPX and to the common ancestor of OCP1 and OCP2. Diversification of this ancestor (Fig. 1A, II) led to the modern OCP1 and OCP2 proteins. Our model shows that the OCP2 clade is not evolutionarily older than OCP1 (as it was previously proposed<sup>22,23</sup>), instead the common OCP1 and OCP2 ancestor diverged along two different pathways that probably led to different functional specializations. In *Tolypothrix*, OCP2 is expressed only under stress conditions and in a mutant lacking chromatic adaptation while OCP1 is constitutive<sup>22</sup>. It is important to note that different subgroups may exist within the OCP1 clade, in particular a small subgroup containing marine *Synechococcus* strains seems to be the most divergent OCP1 (Supplementary Fig. 1). Boundaries between OCPX and the rest of the tree are unclear, and a small population seems to be in an intermediary position between OCPX and OCP2/OCP1 (Fig. 1A).

To study the different properties of OCP clades and understand their evolution, we overexpressed and isolated an OCP from each clade: OCP1 from *Synechocystis* sp. PCC 6803 (hereafter OCP1), OCP2 from *Tolypothrix* (hereafter OCP2) and OCPX from *Scytonema* (hereafter OCPX). *Tolypothrix* OCP2 and *Scytonema* OCPX were chosen because their primary sequences are close to the OCP2 and OCPX consensus sequences, respectively (Supplementary Fig 2). Furthermore, we also partially studied the OCP2 and OCPX from *Synechocystis* PCC 7509 to confirm the characteristics of each clade. Overexpression was performed in a canthaxanthin producing *E. coli* strain since echinenone binding in *Tolypothrix* and *Scytonema* OCPs induces lower photoactivity (Supplementary Fig. 3).



**Figure 1 | OCP evolution and OCP linker.** (a) The phylogenetic tree from the Bayesian analysis of the OCP protein family was rooted using All4940 (*Anabaena* PCC 7120 CTDH) as outgroup. Representative cyanobacteria are marked in the tree: 1 (*Synechocystis* OCP1), 2 (*Arthrospira Maxima* OCP1), 3 (*Nostoc* PCC 7120 OCP1), 4 (*Tolypothrix* OCP1), 5 (*Tolypothrix* OCP2) and 6 (*Scytonema* OCPX). A complete tree with taxon names and branch support values is provide in Supplementary Fig. 1. (b) OCP1 from *Synechocystis* structure (PDB ID: 3MG1). NTD is colored in grey, CTD is colored in green and the linker is cyan. The linker amino acids targeted for mutational analysis are shown in red. (c) Sequence logos representation of multiple alignment of OCP subclades linker sequences. (d) Evolutionary model of the OCP. After gene fusion by addition of the linker OCP diverged (divergence point I) into OCPX and the ancestor of OCP1 and OCP2. The last diverged (divergence point II) into the modern OCP1 and OCP2



**Figure 2 | OCP 1, OCP2 and OCPX comparison.** (a) Absorbance spectra of OCP1<sup>O</sup> (green), OCP2<sup>O</sup> (blue), OCPX<sup>O</sup> (black), OCPx<sup>R</sup> (dashed black), OCP1<sup>R</sup> (dashed green) and OCP2<sup>R</sup> (dashed blue). The inset shows the vibrionic structure of the OCP<sup>O</sup> spectra between 425 nm and 525 nm. (b) Kinetics of OCP photoactivation of *Synechocystis* OCP1 (green), OCP2 (blue), OCPX (black) and *Tolypothrix* OCP1 (red) at 8°C under 5000  $\mu\text{E}$  of white light illumination. (c) Kinetics of OCP deactivation of in the absence (closed symbols) and presence (open symbols) of *Synechocystis* FRP (ratio 1 FRP: 1 OCP) of *Synechocystis* OCP1 (green circle), OCP1-R229E-D262K (rose triangle), OCP2 (blue square) and OCPX (black triangle) at 8°C. (d) Kinetics of PBS fluorescence quenching under strong blue light ( $1000\mu\text{mol photons m}^{-2}\text{s}^{-1}$ ) induced by previously photoactivated pre-photoactivated OCPX (black triangle), OCP2 (blue square) and OCP1 (green circle) at phosphate 0.5M (closed symbols) and at phosphate 1M (open symbols) at 23°C. (e) Kinetics of fluorescence recovery at 0.5 M phosphate. Same symbols than in (d). (f) OCP-related PBS fluorescence quenching induced by strong blue-green light ( $1000\mu\text{mol photons m}^{-2}\text{s}^{-1}$ ) measured in *Scytonema hofmanni* PCC 7110 cells (1 and 3: cells under dim blue light illumination of  $20\mu\text{mol photons m}^{-2}\text{s}^{-1}$ ; 2: cells under high blue light of  $1000\mu\text{mol photons m}^{-2}\text{s}^{-1}$ ). The curves are an average of three independent measurement. The error bars represent the standard deviation of the data shown.

The spectrum of OCPX<sup>O</sup> showed less vibrational band resolution compared that of OCP1<sup>O</sup> and OCP2<sup>O</sup>. The OCPx<sup>R</sup> and OCP1<sup>R</sup> spectra were identical, with a maximum at 525 nm. The OCP2<sup>R</sup> spectrum was red shifted by 10 nm (Fig. 2A). At 8°C, the accumulation of *Synechocystis* OCP1<sup>R</sup> was slightly faster than that of OCP2<sup>R</sup> and OCPx<sup>R</sup> (Fig. 2B). However, *Tolypothrix* OCP1<sup>R</sup> accumulated slower than *Synechocystis* OCP1<sup>R</sup> and *Tolypothrix* OCP2<sup>R</sup> (Fig. 1B), in agreement with previous observations<sup>22</sup>. This indicates that variations in photoactivation rates can exist between different OCP1s and not only between OCP1 and OCP2 (OCPX). The back conversion (OCP<sup>R</sup> to OCP<sup>O</sup>) of OCPX and OCP2 was faster than that of either OCP1, OCPX displaying the fastest

conversion rate (Fig. 2C and Supplementary Fig 3C). As expected OCP1 back conversion was accelerated by the presence of *Synechocystis* FRP (Fig 2C). This FRP accelerated the back reaction of all the OCP1s tested in our laboratory including *Tolypothrix* OCP1 (Supplementary Fig 3). By contrast, FRP did not accelerate the back conversion of OCPX and OCP2 (Fig. 2C and <sup>22</sup>). Cyanobacteria strains containing only OCPX or OCP2 do not contain the *frp* gene<sup>22</sup>. OCPX and OCP2 of *Synechocystis* sp. PCC 7509 presented the same photoactivation and recovery characteristics of the *Scytonema* OCPX and *Tolypothrix* OCP2 respectively (Supplementary Fig 4).

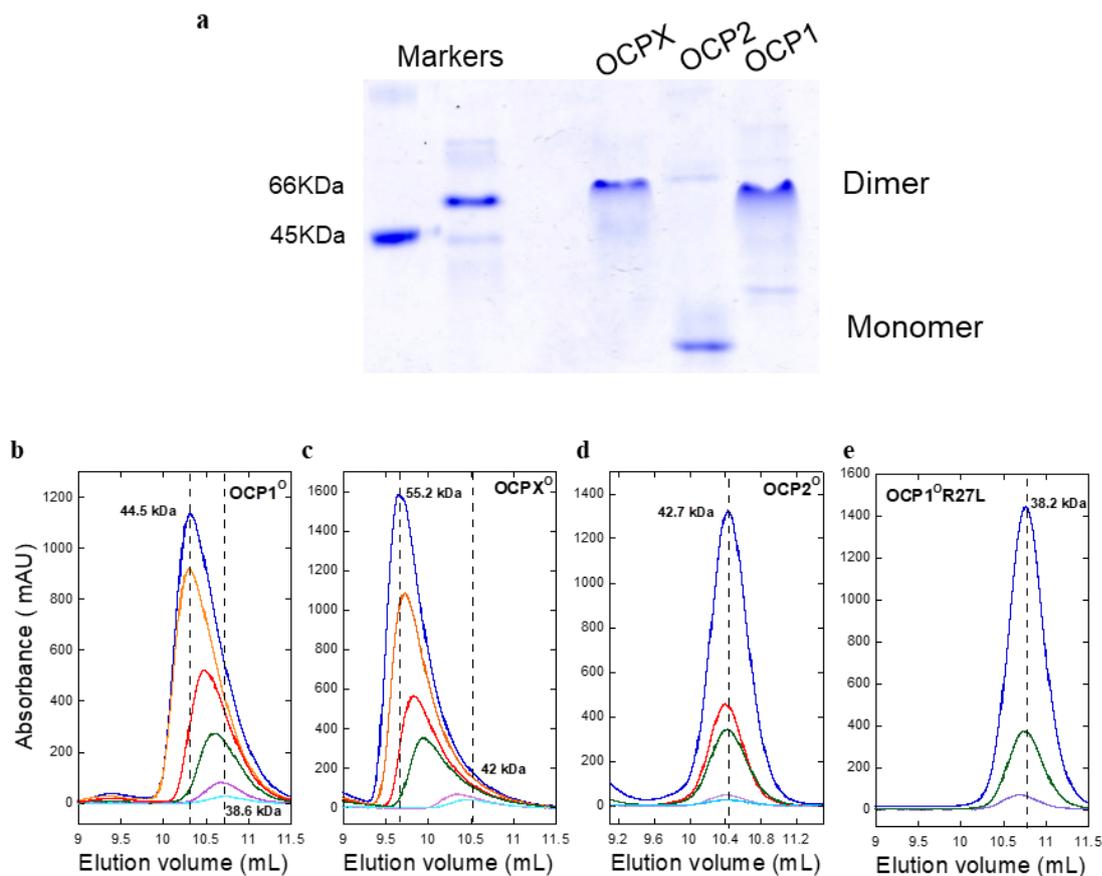
Essential amino acids involved in the FRP-OCP interaction are 100% conserved in the three OCP families, in particular F299<sup>21</sup> and the negative charge of D220<sup>21</sup> (Supplementary Fig 2). However, other amino acids which are also thought to be involved in the OCP-FRP interaction<sup>16</sup> are not conserved: changes concern R229 to E(OCP2) or K(OCPX) and D262 to G(OCP2) or N(OCPX) (Supplementary Fig. 2). In the past, we demonstrated that simple mutations of R229 or D262 do not hinder the OCP-FRP interaction<sup>21</sup>. To further investigate the effect of these mutations on the OCP-FRP interaction, we created a *Synechocystis* OCP1 R229E/D262K double mutant. In this case, FRP was unable to accelerate OCP1 back conversion, suggesting that FRP cannot interact with OCP2 and OCPX due to these amino acid exchanges (Fig. 2C). Interestingly, these mutations also accelerated the back conversion even without FRP, making OCP1 more similar to OCP2 and OCPX.

Next, we studied the OCP-PBS interaction by measuring the PBS fluorescence quenching induced by the three different OCPs. *Synechocystis* PBS were used since it was previously demonstrated that these are the most suitable PBS for in vitro experiments<sup>29</sup>. The three photoactivated OCPs were able to induce large fluorescence quenching in isolated PBS at 1 M phosphate, a concentration which favors OCP-PBS interaction (Fig. 2D). By decreasing the concentration of phosphate from 1 M to 0.5 M differences between the OCPs became measurable: OCP1 induced the largest quenching and OCPX the lowest. This was due to weaker attachment of OCP2 and OCPX to PBS, indicated by a faster recovery of OCP-induced PBS fluorescence quenching (Fig. 2E). In vivo, in *Scytonema* cells containing only OCPX and lacking FRP, strong blue-green light induced a large fluorescence quenching (Fig. 2F), indicating that OCPX binding to *Scytonema* PBSs in the cells can induce PBS fluorescence quenching. Furthermore, in these cells, OCPX was able to detach from PBS and allow fluorescence recovery in the absence of FRP (Fig. 2F). This was also observed in *Gloeobacter violaceus* cells which also contain only OCPXs and no FRP<sup>30</sup>.

### *Oligomeric state of the different OCPs*

In a native gel, OCP1 and both OCPXs appeared as dimers while both OCP2s appeared as a monomer (Fig. 3A and Supplementary Fig 5). In size-exclusion chromatography experiments, the elution volume of OCP1 and both OCPXs exhibited a strong dependency on protein concentration indicating a tendency to dimerize (Fig. 3B and C and Supplementary Fig. 5). OCP1 and OCPX were in monomer-dimer equilibrium with an exchange rate faster than the experiment timescale, giving only one elution peak. Under our experimental conditions, no complete dimerization was observed even for the highest OCP concentrations (120  $\mu$ M). The change in the elution volume was larger for OCPX than OCP1, suggesting a higher dimerization propensity of OCPX. On the other hand, the elution peak of both OCP2s corresponding to its monomeric form was not affected by protein concentration confirming that the OCP2-monomer is stable (Fig. 3D and Supplementary Fig 5). Even though the three OCPs have similar molecular weights, OCP1 had a slightly larger elution volume than monomeric OCPX and OCP2 suggesting differences in monomer structure (Fig. 3D and Supplementary Fig 5).

It has previously been proposed that the hydrogen bond R27-D19 between two OCP1 monomers stabilizes the OCP1 dimer. Indeed, the mutated R27L OCP1 remained as a monomer at all concentrations (Fig. 3E). In OCP2, R27 which is required to stabilize the dimer, is replaced by S and this could explain the stability of OCP2 monomer. However, R27 is also absent in OCPX, which forms dimers, implying that other amino acids are involved in monomer interactions in OCPX.

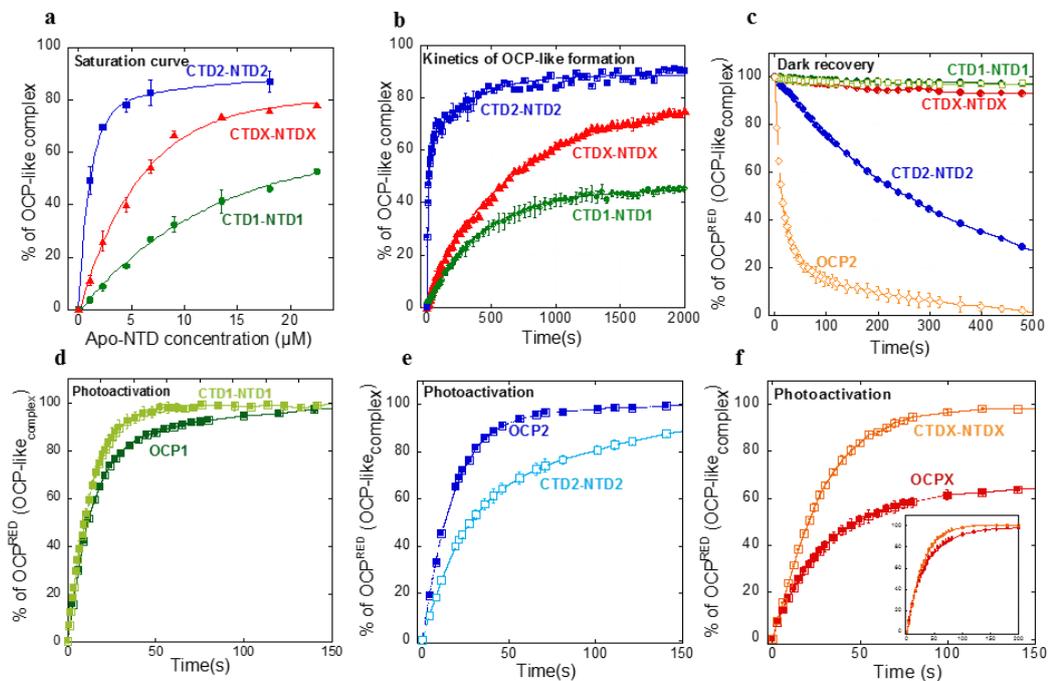


**Figure 3 | OCP oligomeric state: Native gel and size exclusion chromatography.** (a) Analysis of OCP oligomeric state by Coomassie-Blue stained Native PAGE: OCP1 (10  $\mu$ M), OCP2 (10  $\mu$ M), OCPX (10  $\mu$ M) and OCP1-R27L (10  $\mu$ M). (b to e) Comparison of elution profiles in SEC at different concentrations: 120  $\mu$ M (blue curve), 80  $\mu$ M (orange curve), 40  $\mu$ M (red curve), 23  $\mu$ M (green curve), 4  $\mu$ M (pink curve) and 2  $\mu$ M (sky blue curve) of holo-WT-OCP1<sup>O</sup> (b), Holo-WT-OCPX<sup>O</sup> (c), Holo-WT-OCP2<sup>O</sup> (d) and Holo-R27L-OCP1<sup>O</sup> (e). Absorbance was measured at 510 nm.

### *Interdomain interaction in the different OCP clades*

We then studied the interaction between isolated OCP-CTD (CTD) and isolated OCP-NTD (NTD) from the three OCPs. We analyzed the binding affinity of different pairs of NTD-CTD by titrating a fixed canthaxanthin binding CTD (CAN-CTD) concentration with increasing amounts of apo-NTD. We measured the formation of OCP-like complexes (hereafter, OCP-like) (Supplementary Fig. 6). The resulting saturating curves showed that significantly more apo-NTD1 (more than 25  $\mu$ M) and apo-NTDX (22 $\mu$ M) than apo-NTD2 (5 $\mu$ M) were needed to form maximal concentrations of OCP-like complexes (Fig 4A). Furthermore, the formation of OCP2-like was considerably faster than that of the OCPX-like and OCP1-like (Fig. 4B) and its photoactivation was the slowest while that of OCP1-like was the fastest (Fig. 4D, E and F). Taken together all

these results suggest a stronger interaction of CTD2-NTD2 than CTDX-NTDX and NTD1-CTD1, the latter displaying the weakest interaction; CTDX-NTDX interaction is intermediary. However, at this stage we cannot discard that differences in carotenoid affinity between the different NTDs could also influence OCP-like formation<sup>31,32</sup>.



**Figure 4 | CTD-NTD interactions.** (a) Saturation curve of OCP-like formation from isolated CAN-CTD and apo-NTD. 2.25 μM CAN-CTD was incubated with increasing concentrations of apo-NTD (b) Kinetics of OCP-like formation using CAN-CTD (2.25 μM) and apo-NTD (18 μM) from OCP1 (green circle), OCP2 (blue square) and OCPX (red triangle). (c) Back-conversion from free domains to the OCP-like (CTD-NTD) complex at 23°C: CTD1-NTD1 (closed dark green circles), CTD1-NTD1 + *Synechocystis* FRP (open light green circles), CTD2-NTD2 (closed blue circles) and CTDx-NTDx (closed red circles), OCP2 (open orange diamond). (d-f) Comparison of photoactivation kinetics of full OCPs (closed symbols) and OCP-like complexes (open symbols) at 12°C under constant illumination of 1500 μmol photons m<sup>-2</sup> s<sup>-1</sup>: (d) OCP1 and OCP1-like, (e) OCP2 and OCP2-like and (f) OCPX and OCPX-like. Inset in (f): photoactivation kinetics of OCPX and OCPX-like normalized. The curves are an average of three independent measurements. The error bars represent the standard deviation of the data shown.

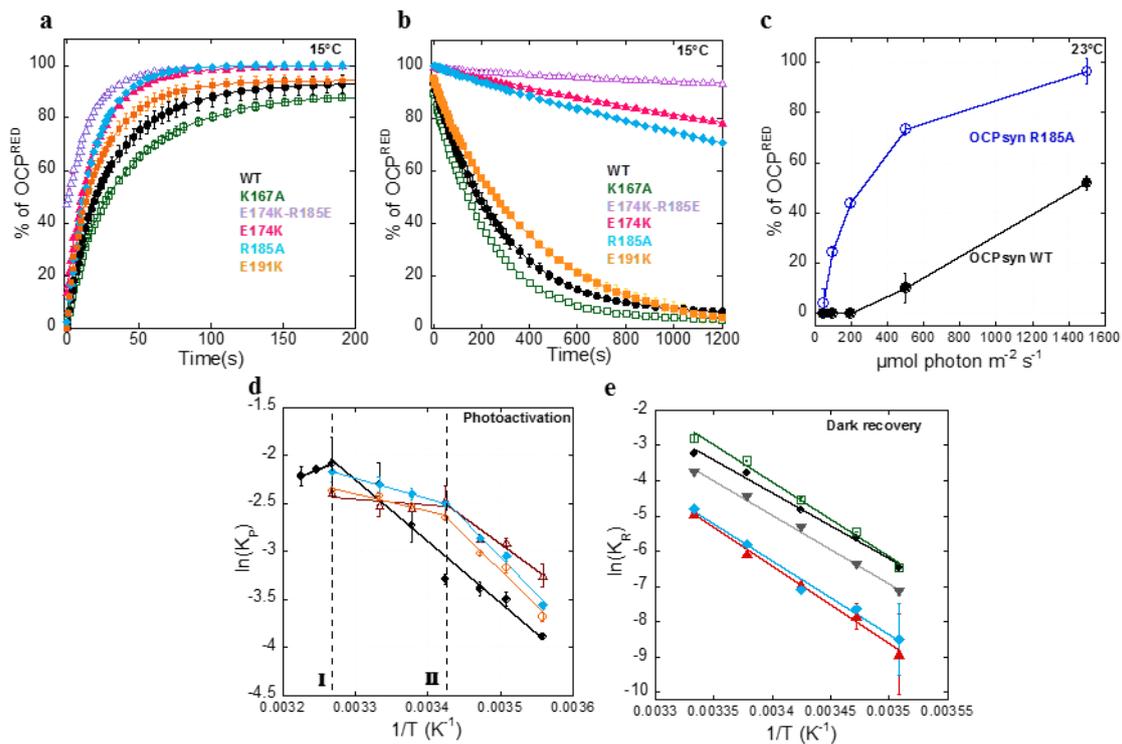
To study the role of the linker connecting the two OCP domains in OCP photoactivation we compared the photoactivation rate of the OCP-like complexes (lacking the linker) with that of the original OCPs. OCP1-like and OCPX-like complexes showed a faster photoactivation than their respective OCPs (Fig. 4D and F). Hence, the presence of the linker slowed down photoactivation. By contrast, photoactivation of OCP2-like was slower than that of OCP2 showing that the presence of the linker accelerated the photoactivation in this case (Fig. 4E). In the photoactivated OCP and OCP-like complexes, the domains were fully separated and the carotenoid was in the NTD. While in the OCP<sup>R</sup>, the domains were still attached via the linker, in the

photoactivated OCP-like, however, the domains were completely free without connections between them. In darkness, holo-NTD1 was not able to form again an OCP-like complex by interaction with apo-CTD1, even in the presence of FRP (Fig. 4C and<sup>23</sup>). OCPX-like complexes recovered only 10 % of their initial state and OCP2-like considerably re-assembled from holo-NTD2 and apo-CTD2 indicating that in this case the domains were reassembled in the proper orientation to allow the carotenoid to move towards the CTD (in agreement with previous report<sup>23</sup>). OCP2-like recovery was slower than that of OCP1 (Fig. 4C). Interestingly, when independently isolated holo-NTD2 was incubated with apo-CTD2, no formation of OCP2-like was observed (Supplementary Fig. 6C). These results suggest that upon photoactivation NTD2 and CTD2 remain somewhat interacting and/or that when the domains are isolated separately apo-CTD can adopt a conformation that hinders the uptake of the carotenoid (see<sup>31</sup>).

### *The role of the linker in photoactivity and recovery of OCP1*

The linker loop connecting the OCP domains is a 25-28 amino acid unstructured region with only a few residues conserved (Fig.1B and 1C). In order to study the role of the linker loop, highly conserved amino acids were mutated in *Synechocystis* OCP1: K167A, R171A, E191K, E174K, R185A, R185E, E174K-R185E,  $\Delta$ 174-179 and P175A-P179A. All these mutated OCP1s were isolated from an echinenone (ECN) producing *E. coli*. In darkness, the OCP1-linker mutants (with the exception of those containing the E174K mutation) showed the same orange absorbance spectrum as WT OCP1 (Supplementary Fig 7). The mutated OCP1s carrying the E174K mutation showed a decrease in the 470 nm peak and an increase at 550 nm, especially the OCP1 E174K-R185E mutant (Supplementary Fig. 7). Spectral deconvolution showed that this change was related to the accumulation of OCP<sup>R</sup> even in darkness (approximately 45% of OCP<sup>R</sup> in E174K-R185E).

After 5 min illumination with strong white light, all OCP1-linker mutants exhibited the typical spectrum of OCP<sup>R</sup> (Supplementary Fig. 7). All OCP1 mutants (except for K167A) accumulated the OCP<sup>R</sup> form faster than the WT OCP1 (Fig. 5A and Supplementary Fig. 8). This affected the photoactivation light dependency: the mutations in the linker induced higher OCP<sup>R</sup> accumulation at low light intensities (Fig. 5C). The back-conversion kinetics to the orange state were slower in the OCP1 mutants than in WT OCP1 (Fig. 5B and Supplementary Fig. 8). E174K and R185A mutations largely affected the rate of recovery; OCP1 E174K-R185E and OCP1  $\Delta$ 174-179 presented the slowest kinetics. Furthermore, OCP<sup>R</sup> was more stable at all the temperatures tested in the linker mutants than in the WT OCP1 (Supplementary Fig. 8). The only exception was OCP1 K167A, which showed the opposite effect: a slower photoconversion and a faster recovery than WT OCP1.



**Figure 5 | Photoactivity of OCP1 linker-mutants.** (a) Accumulation of photoactivated OCP<sup>R</sup> and (b) OCP<sup>R</sup> back-conversion to OCP<sup>O</sup> at 15°C of OCP WT (black closed circle), OCP E174K (fuchsia closed triangle), OCP E191K (orange closed square), OCP E174K-R185E (violet open triangle), OCP R185A (blue close diamond), OCP K167A (green open square) (c) Percentage of OCP<sup>R</sup> accumulated under different light intensities for WT OCP (black circles) and OCP-R185A (blue square) at 23°C. (d) Arrhenius plot of photoactivation of OCP WT (black), OCP E174K (red), OCP K167A (green), OCP R185A (sky blue) and OCP P175A-P179A (grey). Break point of the convex curves is shown with dotted lines for the WT OCP (I) and the linker mutants (II) (e) Arrhenius plot of back-conversion OCP<sup>R</sup> to OCP<sup>O</sup> of OCP WT (black), OCP E174K-R185E (dark red), OCP R185A (sky blue), OCP K167A (green) and OCP R171A (orange). The curves are an average of three independent measurement. The error bars represent the standard deviation of the data shown.

To determine the activation energy of photoactivation and back conversion, the kinetics were measured at different temperatures and Arrhenius plots were generated (Fig. 5D, 5E and Supplementary Fig. 8). The calculated activation energy ( $E_a$ ) of the back conversion for WT OCP1, 37 kcal/mol (in agreement with the values previously published<sup>10</sup>), is statistically the same as in the mutants (Supplementary Fig. 8). By contrast, the pre-exponential factor which is related to the probability of a collision occurring in the proper orientation was different in mutants and WT. The Arrhenius plot of WT OCP1 photoactivation showed a positive convexity that can be explained by a decrease in activation energy of photoactivation at high temperatures (higher than 33°C), possibly due to the destabilization of protein interactions at these high temperatures as previously suggested<sup>10,33</sup>. In the linker mutants (with exception of K167A) the break point of the convex plot is shifted toward lower temperatures (Fig. 5B, from 33°C (I) to 20°C (II)). This indicated that the temperature required to destabilize the OCP structure is lower for the linker mutants.

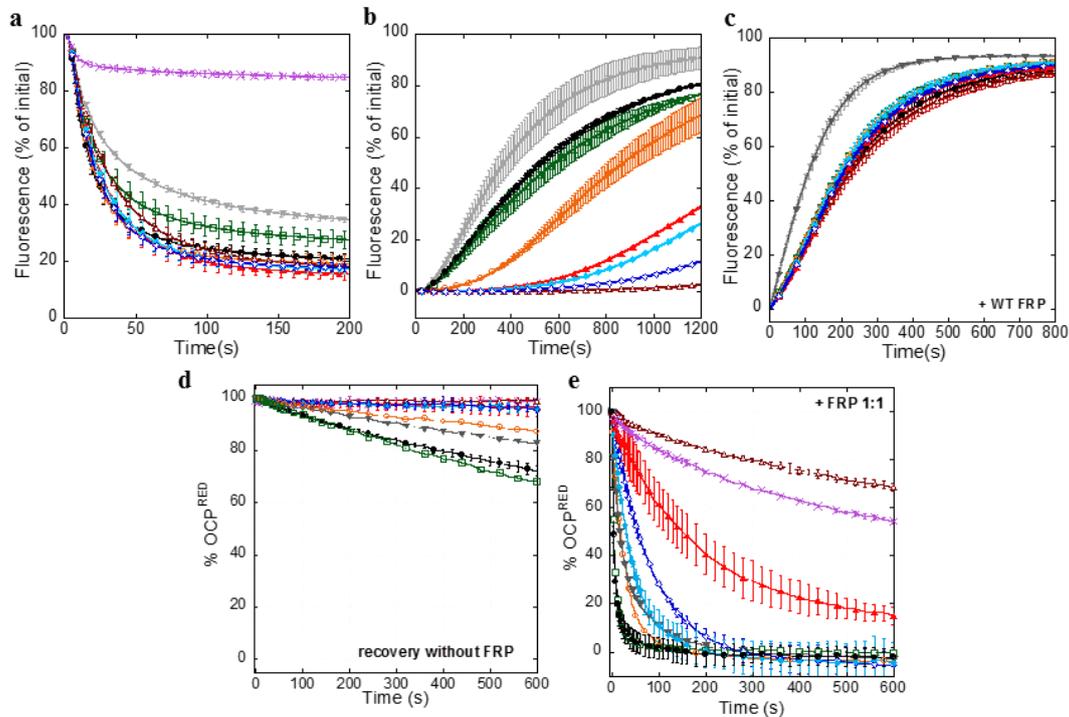
### ***Role of the linker in the OCP1 interaction with phycobilisomes and FRP***

In order to study the role of the mutated amino acids on the OCP-PBS interaction, *in vitro* OCP induced PBS fluorescence quenching was assessed. Most of the linker-mutants induced similar amplitudes of PBS fluorescence quenching as OCP WT (Fig. 6A). PBS fluorescence recovery was slower when the fluorescence quenching was induced by most of these OCP mutants due to a higher stability of their OCP<sup>R</sup> (Fig. 6B). By contrast, OCP P175A/P179A induced less fluorescence quenching and the recovery was faster than that of OCP WT despite its high OCP<sup>R</sup> stability (Fig 5), indicating that the interaction with PBS was affected in this mutant (Fig. 6A and B).

Deletion of 6 amino acids in the  $\Delta$ 174-179 OCP caused an extensive loss of OCP quenching activity in spite the fact that this mutation stabilized the OCP<sup>R</sup>. This could be due to a partial opening of the domains although the protein is red upon illumination. It was recently demonstrated that the protein can become red even without opening<sup>7,34</sup>. Analysis of the hydrodynamic volume of photoactivated OCPs showed that the apparent Stoke's radius of the  $\Delta$ 174-179 OCP was slightly smaller to that of the photoactivated WT OCP suggesting that although in the mutated photoactivated OCP the domains were separated, the complete opening of the protein was not fully reached (Supplementary Fig. 9). Nevertheless, we cannot discard that the lack of these specific amino acids also might have affected the OCP-PBS interaction.

The addition of FRP to photoactivated OCPs accelerated the recovery in the WT OCP and all the OCP-linker mutants (Fig. 6D and E). Even the slowest recovery reactions (OCP E174K-R185E and OCP  $\Delta$ 174-179) were accelerated by FRP addition.

Furthermore, addition of FRP largely accelerated the recovery of PBS fluorescence in the PBSs quenched by all the OCP linker mutants (Fig. 6C).



**Figure 6 | OCP1 linker mutants: interaction with PBS and FRP.** (a) Kinetics of PBS fluorescence quenching under strong blue light induced by previously photoactivated WT OCP<sup>R</sup> (black), OCP<sup>R</sup> E174K (red), OCP<sup>R</sup> E174K-R185E (dark red), OCP<sup>R</sup> R185A (sky blue), OCP<sup>R</sup> R185E (blue), OCP<sup>R</sup> K167A (green), OCP<sup>R</sup> R171A (orange), OCP<sup>R</sup> P175A-P179A (grey) and OCP<sup>R</sup>  $\Delta$ 174-179 (pink). (b and c) Kinetics of dark recovery of PBS fluorescence in absence (b) or presence of *Synechocystis* FRP (c). Back-conversion OCP<sup>R</sup> to OCP<sup>O</sup> at 8°C; in presence (e) or absence (d) of *Synechocystis* FRP (1 FRP per OCP) of OCP<sup>R</sup> WT (black), OCP<sup>R</sup> E174K (red), OCP<sup>R</sup> E174K-R185E (dark red), OCP<sup>R</sup> R185A (sky blue), OCP<sup>R</sup> R185E (blue), OCP<sup>R</sup> K167A (green), OCP<sup>R</sup> R171A (orange), OCP<sup>R</sup> P175A-P179A (grey) and OCP<sup>R</sup>  $\Delta$ 174-179 (pink). The curves are an average of at least three independent measurements. The error bars represent the standard deviation of the data shown.

### Role of the linker in the OCP2 and OCPX clades

In OCP2 and OCPX the number of conserved residues in the linker is significantly lower than in OCP1 (Fig. 1C). In OCP2, E174 is not highly conserved in contrast to R185, which is highly conserved (Fig.1C). In OCPX, a higher diversity of amino acid composition is found in the linker region, where however E174 is highly conserved. Thus, to study whether the role of these amino acids in OCP2 and OCPX is similar to their role in OCP1 we created OCP2 R185E-CAN and OCPX E174K-CAN mutants for comparison with their respective WT OCP-CAN.

Both OCP2 and OCPX linker mutants showed a decreased rate of OCP<sup>R</sup> to OCP<sup>O</sup> back conversion compared to their respective WTs (Supplementary Fig. 10). These results indicated that the mutations in the linker increased the stability of OCP<sup>R</sup> in OCP2 and OCPX as was the case for OCP1. Like in OCP1, mutation E174K slightly increased the rate of OCP<sup>R</sup> accumulation in OCPX (Supplementary Fig. 10) and decreases the light intensity required for complete photoconversion (Supplementary Fig. 11). On the other hand, the R185E mutation in OCP2 slightly decreased the rate of OCP<sup>R</sup> accumulation (Supplementary Fig. 10).

## Discussion

---

According to the current evolutionary model, fusion of primitive *ctdh* and *hcp* genes gave rise to a primitive OCP<sup>23,24,22</sup>. We propose that the ancestral HCP was already able to interact with the PBS like the modern HCP4<sup>36</sup>. Addition of an appropriate linker and a CTD made the NTD-PBS interaction more stable, allowing for more efficient photoprotection<sup>6</sup>. We suppose that the primitive OCP already had the activities found in modern OCPs and that the diversification merely brought a more sophisticated regulation of OCP photoactivation and recovery.

The phylogenetic analysis of a rooted tree we present here shows that the primitive OCP diversified into OCPX and the ancestor of the modern OCP2s and OCP1s, which are the result of a second diversification. OCP1 and OCP2, which co-exist in many strains, seem to have diverged and acquired different specializations by maintaining or changing different properties present in their common ancestor which are still found in modern OCPX (Figure 1D). OCP2 is mostly present in strains which also constitutively express OCPX or OCP1. It is known that at least in *Tolypothrix*, OCP2 is expressed only under stress conditions in which cells need more protection<sup>22</sup>. Thus, some specific properties of OCP2 might have an advantage for a better response to “permanent” stress conditions while OCP1 properties must be advantageous under low, medium and fluctuating light conditions.

The previous evolutionary model proposed that before the genes fused, an OCP-like complex existed formed by the heterodimer of a primitive NTD and CTD that shared a keto carotenoid. Based on *Tolypothrix* OCP2 results, it was originally proposed that this primitive OCP-like could photocycle even in the absence of the linker loop connecting the NTD and CTD<sup>22,23</sup>. However, here we have shown that this is probably not likely since the ancient OCPX-like complex cannot photocycle: after photoactivation, the isolated holo-NTDX and apo-CTDX were unable to reassemble to form the OCP-like complex. Thus, our results strongly suggest that the addition of the linker loop provided the primitive OCP with two new features: the possibility to photocycle by dark recovery into the inactive form and a more controlled photoactivation. The modern OCPX has kept

these features, showing a fast dark recovery and a slightly negative effect of the linker on photoactivation. OCP2 kept the fast recovery present in its ancestor and in OCPX. In OCP1, two point mutations acquired during evolution in the CTD (in amino acids 229 and 262) made the recovery reaction considerably slower. This added one more point of regulation in the OCP photocycle which now required an exterior component, the FRP, in order to accelerate the recovery reaction.

The OCP1 clade conserved the slightly negative effect of the linker on photoactivation. This fine regulation allowed the accumulation of high concentrations of OCP<sup>R</sup> only at high light intensities, which is particularly important for OCP1 and OCPX since they are expressed even under low light conditions<sup>12,22,30,37</sup>. In OCP2, the strong affinity between the domains renders photoactivation too slow for effective photoprotection in the absence of the linker which in this case, functions as an accelerator of photoactivation. Thus, the linker has opposite roles in OCP1 and OCP2 with respect to photoactivation. By contrast, it has the same accelerating role in the OCP<sup>R</sup> to OCP<sup>O</sup> conversion for all OCPs.

Specific linker amino acids are essential to the photocycle: E174K and R185E mutations greatly decreased the recovery rates in the three types of OCPs. The length of the linker is also important to facilitate the recovery to OCP<sup>O</sup> since the deletion of 6 amino acids largely slowed down recovery. Thermodynamic analysis strongly suggested that the role of these amino-acids in the unstructured linker is to provide a specific environment, in which the OCP-domains are reoriented into the proper configuration allowing the inter-domain interaction that favors the formation of carotenoid tunnel and the recovery to OCP<sup>O</sup> (see Supplementary Fig 13). Interestingly, the linker area around the amino acids which most affected OCP recovery is the less flexible part of the linker according to MD simulations (Supplementary Fig. 12).

On the other hand, specific amino acids have only a minor effect on the OCP-PBS interaction. Only the simultaneous mutation of prolines 175 and 179 slightly decreased the strength of the OCP-PBS interaction. Most of the other linker mutations did not affect OCP-induced PBS fluorescence quenching. By contrast, the size of the linker was crucial. The deletion of 6 linker amino acids completely inhibited PBS fluorescence quenching strongly suggesting that a long linker is needed to allow a sufficiently wide separation of the domains allowing the NTD to be buried between two APC trimers in the PBS core<sup>38</sup>.

The FRP was unable to induce recovery after photoactivation of OCP1-like complex indicating that the presence of the linker is essential for FRP activity. However, none of the mutated amino acids in the linker were shown to be essential for FRP interaction. It is possible that after the primary binding of FRP to the CTD, a high local concentration of NTD is required for its secondary binding to the OCP<sup>18,19,39</sup> and acceleration of back conversion. This high concentration is successfully achieved by connecting both domains through the linker.

A monomer to dimer transition could provide an additional regulation of the OCP function. Despite the fact that R27, which is essential for OCP1 dimerization, is not conserved in OCPX, OCPX is able to dimerize. This indicates that another interaction must stabilize the OCPX dimer. Further regulation by a dimer-to-monomer transition might not have been important for carrying out the OCP2 role under stress conditions, leading to its loss during evolution.

Taking all our results together, it is clear that the linker loop played an essential role during OCP evolution, along with inter-domain interaction, by allowing fine tuning of the OCP behavior and better photoprotection and functional divergence between the different families of OCPs. In addition, the understanding of the various molecular determinants controlling the OCP photoswitching could help in the design of optogenetic toolkit for synthetic biology.

## ***Materials and Methods***

---

### ***Construction of OCP, NTD and CTD Plasmids for expression in E. coli cells***

Construction of plasmid pCDF-OCPsynNtag carrying the *Synechocystis* slr1963 gene with a His-tag encoding sequence in the 3'-terminus was described previously<sup>40</sup>. The point mutations were introduced by directed mutagenesis, using the pCDF-OCPSyn-Ntag plasmid as a template and mismatching primers (Eurofins) (Supplementary Table S1). The constructions of the pCDF-CTD1<sup>25</sup> and pCDF-NTD1<sup>6</sup> plasmids containing the sequences encoding the isolated *Synechocystis* CTD and NTD domains were described previously<sup>6,25</sup>.

#### **Construction of pCDF-OCP2 and pCDF-OCP1 from *Tolypothrix***

The *Tolypothrix* sp. PCC 7601 *ocp1* and *ocp2* genes were amplified by PCR using genomic DNA of *Tolypothrix* sp. PCC7601 as template and the primers OCP1-pDuet (F and R) or OCP2-pDuet (F and R). The resulting PCR products were cloned into the BamHI and NotI site of pCDFDuet-1 creating the plasmids *pCDF-NtagOCP2* and *pCDF-NtagOCP1*. With this construction, a His-tag will be present in the N-terminus of the protein. The point mutation R185E was inserted in the OCP2 gene by direct mutagenesis (supplementary Table SI).

#### Construction of the *pCDF-NTD2 Ctag*

To construct the *pCDF-NTD2 Ctag* plasmid, the sequence coding the N-terminus His-tag in the *pCDF-NtagOCP2* plasmid was excised by mutagenesis using F-ocp2-BamHI and R-Duet to create the plasmid *pCDF-OCP2*. Then, a deletion of the nucleotides of the *ocp2* encoding the last 154 amino acids (CTD) and an addition of a sequence encoding a C-terminal His-tag were realized simultaneously by mutagenesis using the plasmid *pCDF-OCP2* as template and synthetic primers (F-DuetOCP2 Ctag and R-NTD2 Ctag).

#### Construction of the *pCDF-CTD2 Ctag*

To construct the *pCDF-CTD2 Ctag* plasmid, first, the nucleotides encoding for 6 His tag were added in the 3'-terminus of the *ocp2* gene by site-directed mutagenesis using the *pCDF-OCP2* and the synthetic primers (F-DuetOCP2 Ctag and R-DuetOCP2 Ctag). Then, the nucleotides encoding the NTD (1-189) of the *ocp2* gene were deleted by site-directed mutagenesis using the plasmid *pCDF-OCP2 Ctag* as template and synthetic primers (F-DNTD2 and R-Duet).

#### Construction of the *pCDF-NtagOCPx from Scytonema hofmanni PCC 7110*

The synthetic gene (Eurofins) of the OCPX from *Scytonema hofmanni* PCC7110 (DNA fragment from genome positions 2,712,095 to 2,713,054) was cloned into the BamHI and NotI site of pCDFDuet-1. With this construction, a His-tag will be present in the N-terminus of the protein. The point mutation E174K was inserted in the OCPX gene by directed mutagenesis (supplementary Table SI).

#### Construction of the *pCDF-NTDX Ctag*

To construct the *pCDF-NTDX Ctag* plasmid, first, the N-terminal extension containing the His-tag initially present in the *pCDF-NtagOCPX* plasmid was excised by mutagenesis using F-OCPX-BamHI and R-Duet to create the plasmid *pCDF-OCPX*. Then, a deletion of the last 154 amino acids (CTD) of the *OCPX* gene and an addition of C-terminal His-tagged were realized simultaneously by mutagenesis using the plasmid *pCDF-OCPX* as template and synthetic primers (F-DuetOCPX Ctag and R-NTDX Ctag).

#### Construction of the *pCDF-CTDx Ctag*

To construct the *pCDF-CTDx Ctag* plasmid, the nucleotides encoding for 6 His tag were added in the 3'-terminus of the *ocpx* gene by site-directed mutagenesis using the *pCDF-OCPx* (previously described in this work) and the synthetic primers (F-DuetOCPx Ctag and R-DuetOCPx Ctag) to create the *pCDF-OCPx Ctag* plasmid. Then, the nucleotides encoding the N-terminal domain (1-189) of the *ocp2* gene were deleted by site-directed mutagenesis using the plasmid *pCDF-OCPx Ctag* as template and synthetic primers (F-NTDx and R-Duet).

### Construction of pCDF-OCP1 (linker mutants)

The specific punctual mutations in the linker region of pCDF-OCP1-Ntag were introduced by site-directed mutagenesis using the plasmid pCDF-OCP1-Ntag as template and the corresponding synthetic primers (Supplementary table 2A-B).

### Construction of the *pCDF-NtagOCPx* from *Synechocystis* sp. PCC 7509

The synthetic gene (Eurofins) of the OCPX from *Synechocystis* sp. PCC 7509 (IMG ID : 2517697059) was cloned into the BamHI and NotI site of pCDFDuet-1. With this construction, a His-tag will be present in the N-terminus of the protein.

### Construction of the *pCDF-NtagOCP2* from *Synechocystis* sp. PCC 7509

The synthetic gene (Eurofins) of the OCP2 from *Synechocystis* sp. PCC 7509 (IMG ID: 2517700424) was cloned into the BamHI and NotI site of pCDFDuet-1. With this construction, a His-tag will be present in the N-terminus of the protein.

## ***Holoprotein production and purification in carotenoid-producing E. coli***

BL21 (DE3) cells from Agilent Technologies were transformed simultaneously with three plasmids: (1) pACBETA, (2) pBAD-CrtW (for canthaxanthin production) or pBAD-CrtO (for echinenone production), and (3) a pCDF-OCP plasmid (carrying either OCP1, OCP2, OCPX, the respective linker mutants, CTD1, NTD1, CTD2, NTD2, CTDX or NTDX). The construction of pACBETA and pBAD-CrtW(-CrtO) was described elsewhere<sup>40</sup>. The *crtBEIY* operon in pACBETA was constitutively expressed under the control of the *crtE* promoter, whereas the *crtW* and *crtO* gene were under the control of the arabinose-inducible promoter *araBAD* and the *OCP* genes were under the control of a T7 RNA polymerase promoter. The expression method used to obtain the holo-proteins was described previously<sup>40</sup>. OCP, NTD, CTD holo-proteins and apo-proteins expressed in *E. coli* were isolated as described elsewhere<sup>36,40</sup>. Briefly, cells were resuspended in lysis buffer (40 mM Tris pH 8/10% glycerol/300mM NaCl/1mM EDTA/1mM PMSF, 1mM caproic acid/1mM benzamidic acid/ 50  $\mu\text{g mL}^{-1}$  DNase) then broken in dim light using a French press. The membranes were pelleted, and the supernatant was loaded on a nickel affinity column (Ni-Probond resin, Invitrogen). Proteins were eluted with 250 mM imidazole and then dialyzed against 40 mM Tris-HCl pH 8. A final purification step was performed by hydrophobic interaction chromatography (HiTrap Phenyl HP column, GE Healthcare). Total OCP concentration was measured using the Bradford method and the carotenoid absorbance at 496 nm using  $A_{\text{cm}}^{1\%}=2158$ . The construction of the *E. coli* strain

overexpressing the *Synechocystis* FRP and the isolation of FRP was described in (Boulay et al, 2010).

### ***Size Exclusion Chromatography***

Purified proteins were injected into a Superdex Increase 75 10/300 column (GE Healthcare) equilibrated with 40mM Tris-HCl pH 8, 150mM NaCl. The Akta FLPC system was equipped with a UV detector at 280 nm, 510nm and 560 nm and was set at a flow rate of 0.8 mL/min. The Gel Filtration Calibration Kit LMW (GE Healthcare) was used to calibrate the column.

### ***Absorbance measurements and kinetic analysis***

Absorbance spectra, kinetics of photoactivation (illumination with 5000  $\mu\text{mol photons m}^{-2} \text{ s}^{-1}$  of white light) and dark recovery of the OCP were measured in a Specord S600 spectrophotometer (Analyticjena) at 8, 12, 15, 19, 23, 27, 33 or 37 °C in 1 cm pathlength cuvette. Spectra were acquired from 250nm to 700nm for each time point. The percentage of OCP<sup>R</sup> formed was calculated from the absorbance changes at 550nm and from spectral deconvolution. To study the effect of FRP, we added *Synechocystis FRP* (2.6  $\mu\text{M}$ ) just before turning off the light.

To study CTD-NTD complex formation, holo-CTDs (2.25  $\mu\text{M}$ ) were incubated with different concentrations of apo-NTD at 12°C for one hour in darkness. The percentage of OCP-like complexes was determined after spectral deconvolution (Supplementary Fig. 6).

### ***Protein Native Gel Electrophoresis***

Nondenaturing gel electrophoresis was performed to determine the oligomeric state of OCPs. 10  $\mu\text{L}$  of purified proteins (10  $\mu\text{M}$ ) were applied to 15% native polyacrylamide gels (pH 8). Electrophoresis was carried out using 25 mM Tris/192 mM Gly buffer, pH 8. Bovine serum albumin (65 kD) and ovalbumin (45 kD) were used as marker proteins.

### ***Isolation of PBS and Fluorescence measurements***

The purification of PBS from *Synechocystis* PCC 6803 was performed as previously described<sup>13</sup>. Fluorescence yield quenching was monitored using a pulse amplitude modulated fluorimeter (101/102/103-PAM, Walz). Measurements were made in a 1 cm path length stirred cuvette. The PBS quenching induced by holo-OCPs was measured in

0.5 M or 1 M potassium phosphate buffer (pH=7.5) at 23°C in strong blue-green light (650  $\mu\text{mol photons m}^{-2} \text{ s}^{-1}$ ). The PBS concentration used was 0.012  $\mu\text{M}$  and the ratio of carotenoid to PBS was 40:1. OCP samples were pre-illuminated with 5000  $\mu\text{mol photons m}^{-2} \text{ s}^{-1}$  of white light.

### *Phylogenetic analysis*

In order to build the OCP phylogenetic tree, we performed a BLAST<sup>41</sup> search, using as seed the *slr1963* protein sequence and an E-value cut-off of  $1 \times 10^{-4}$ , to retrieve protein sequences from IMG. For tree rooting, two outgroups (CTDH *all4940* and NTF2 CP024957.1 *Streptomyces cavourensis*) were added independently giving the same result. Sequences obtained were then aligned using default ClustalOmega<sup>42</sup>. Different trees were constructed using different methods to confirm the result: Maximum likelihood (ML), Bayesian inference and Neighbor-Joining. ML analysis was done using PhyML 3.1<sup>43</sup> in order to build a phylogenetic tree, with the following parameters: LG substitution model, NNI topology search and aBayes algorithm for branch support. Bootstrapped Neighbor-Joining trees were constructed using the Kimura-2-parameter model and 1,000 bootstrap repeats were built to give statistical support. Bayesian analysis was performed using MrBayes 3.1.2<sup>44</sup> software. Metropolis-coupled Markov Chain Monte Carlo (MCMC) (using 4 chains) with 5,000,000 generations were run. Analysis of the distances between nodes was done using the biopython<sup>45</sup> library (Bio.Phylo module). Sequences logos were built with the online server Skyline<sup>46</sup>.

### *Molecular dynamic simulation*

The Apo-OCP structure was modelled using the SWISS-MODEL<sup>47</sup> server, using the OCP structure (PDB: 3MG1) as template in order to model the missing part of the flexible linker. The simulation was performed using GROMACS<sup>48</sup> simulation package version 5.1.1. The coordinate and topology for the simulations were generated using the GROMACS pdb2gmx protocol using the OPLSAA/L force field. Solvation was obtained using the spc216 explicit water model. Simulations were carried out at constant temperature and pressure (300 K, 1 atm), in NaCl 150 mM for 100 ns.

## *References*

---

1. Kirilovsky, D. & Kerfeld, C. A. Cyanobacterial photoprotection by the orange carotenoid protein. *Nat. Plants* 2, 16180 (2016).

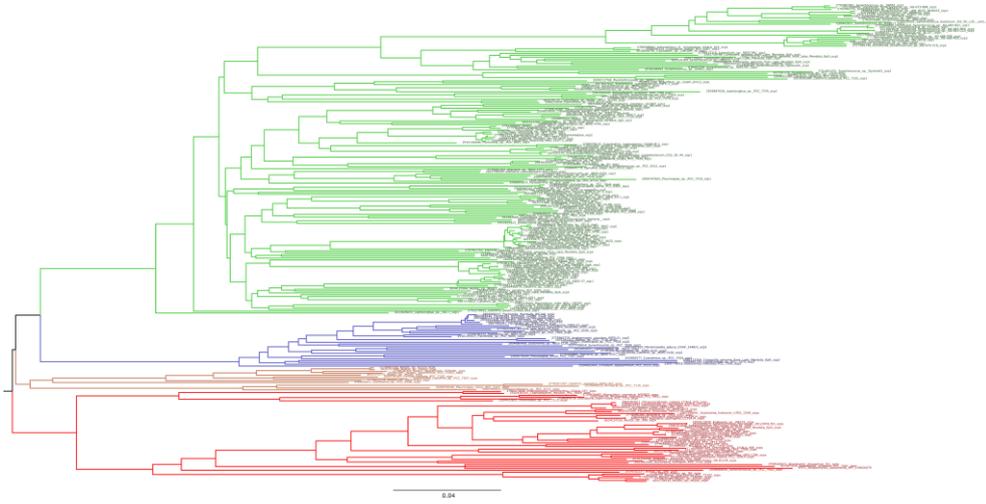
2. Kerfeld, C. A., Melnicki, M. R., Sutter, M. & Dominguez-Martin, M. A. Structure, function and evolution of the cyanobacterial orange carotenoid protein and its homologs. *New Phytol.* 215, 937–951 (2017).
3. Sluchanko, N. N., Slonimskiy, Y. B. & Maksimov, E. G. Features of protein–protein interactions in the cyanobacterial photoprotection mechanism. *Biochem.* 82, 1592–1614 (2017).
4. Kerfeld, C. A. et al. The crystal structure of a cyanobacterial water-soluble carotenoid binding protein. *Structure* 11, 55–65 (2003).
5. Wilson, A. et al. A photoactive carotenoid protein acting as light intensity sensor. *Proc. Natl. Acad. Sci.* 105, 12075–12080 (2008).
6. Leverenz, R. L. et al. A 12 Å carotenoid translocation in a photoswitch associated with cyanobacterial photoprotection. *Science* (80-. ). 348, 1463–1466 (2015).
7. Konold, P. E. et al. Photoactivation mechanism , timing of protein secondary structure dynamics and carotenoid translocation in the Orange Carotenoid. *J. Am. Chem. Soc.* 141, 520–530 (2019).
8. Gupta, S. et al. Local and global structural drivers for the photoactivation of the orange carotenoid protein. *Proc. Natl. Acad. Sci.* 112, E5567–E5574 (2015).
9. Liu, H. et al. Dramatic Domain Rearrangements of the Cyanobacterial Orange Carotenoid Protein upon Photoactivation. *Biochemistry* 55, 1003–1009 (2016).
10. Maksimov, E. G. et al. The Signaling State of Orange Carotenoid Protein. *Biophys. J.* 109, 595–607 (2015).
11. Maksimov, E. G. et al. A comparative study of three signaling forms of the orange carotenoid protein. *Photosynth. Res.* 130, 389–401 (2016).
12. Wilson A, Ajlani G, Verbavatz JM, Vass I, Kerfeld CA, K. D. A Soluble Carotenoid Protein Involved in Phycobilisome-Related Energy Dissipation in Cyanobacteria. *Plant Cell Online* 18, 992–1007 (2006).
13. Gwizdala, M., Wilson, A. & Kirilovsky, D. In vitro reconstitution of the cyanobacterial photoprotective mechanism mediated by the Orange Carotenoid Protein in *Synechocystis* PCC 6803. *Plant Cell* 23, 2631–43 (2011).
14. Wilson, C. W. A. et al. The Essential Role of the N-Terminal Domain of the Orange Carotenoid Protein in Cyanobacterial Photoprotection: Importance of a Positive Charge for Phycobilisome Binding. *Plant Cell* 24, 1972–1983 (2012).

15. Boulay, C., Wilson, A., D'haene, S. & Kirilovsky, D. Identification of a protein required for recovery of full antenna capacity in OCP-related photoprotective mechanism in cyanobacteria. *Proc. Natl. Acad. Sci.* 107, 11620–11625 (2010).
16. Sutter, M. et al. Crystal structure of the FRP and identification of the active site for modulation of OCP-mediated photoprotection in cyanobacteria. *Proc. Natl. Acad. Sci.* 110, 10022–10027 (2013).
17. Sluchanko, N. N. et al. The purple Trp288Ala mutant of *Synechocystis* OCP persistently quenches phycobilisome fluorescence and tightly interacts with FRP. *Biochim. Biophys. Acta - Bioenerg.* 1858, 1–11 (2017).
18. Sluchanko, N. N., Slonimskiy, Y. B., Moldenhauer, M., Friedrich, T. & Maksimov, E. G. Deletion of the short N-terminal extension in OCP reveals the main site for FRP binding. *FEBS Letters* 591, 1667–1676 (2017).
19. Lu, Y. et al. Native mass spectrometry analysis of oligomerization states of fluorescence recovery protein and orange carotenoid protein: Two proteins involved in the cyanobacterial photoprotection cycle. *Biochemistry* 56, 160–166 (2017).
20. Moldenhauer, M. et al. Interaction of the signaling state analog and the apoprotein form of the orange carotenoid protein with the fluorescence recovery protein. *Photosynth. Res.* 135, 125–139 (2018).
21. Thurotte, A. et al. The cyanobacterial Fluorescence Recovery Protein has two distinct activities: Orange Carotenoid Protein amino acids involved in FRP interaction. *Biochim. Biophys. Acta - Bioenerg.* 1858, 308–317 (2017).
22. Bao, H. et al. Additional families of orange carotenoid proteins in the photoprotective system of cyanobacteria. *Nat. Plants* 3, 17089 (2017).
23. Lechno-Yossef, S., Melnicki, M. R., Bao, H., Montgomery, B. L. & Kerfeld, C. A. Synthetic OCP heterodimers are photoactive and recapitulate the fusion of two primitive carotenoproteins in the evolution of cyanobacterial photoprotection. *Plant J.* 91, 646–656 (2017).
24. Melnicki, M. R. et al. Structure, Diversity, and Evolution of a New Family of Soluble Carotenoid-Binding Proteins in Cyanobacteria. *Mol. Plant* 9, 1379–1394 (2016).
25. Muzzopappa, F. et al. The paralogs to the C-terminal domain of the cyanobacterial OCP are carotenoid donors to HCPs. *Plant Physiol.* 175, 1283–1303 (2017).
26. Kerfeld, C. A. Structure and Function of the Water-Soluble Carotenoid-Binding Proteins of Cyanobacteria. *Photosynth. Res.* 81, 215–225 (2004).

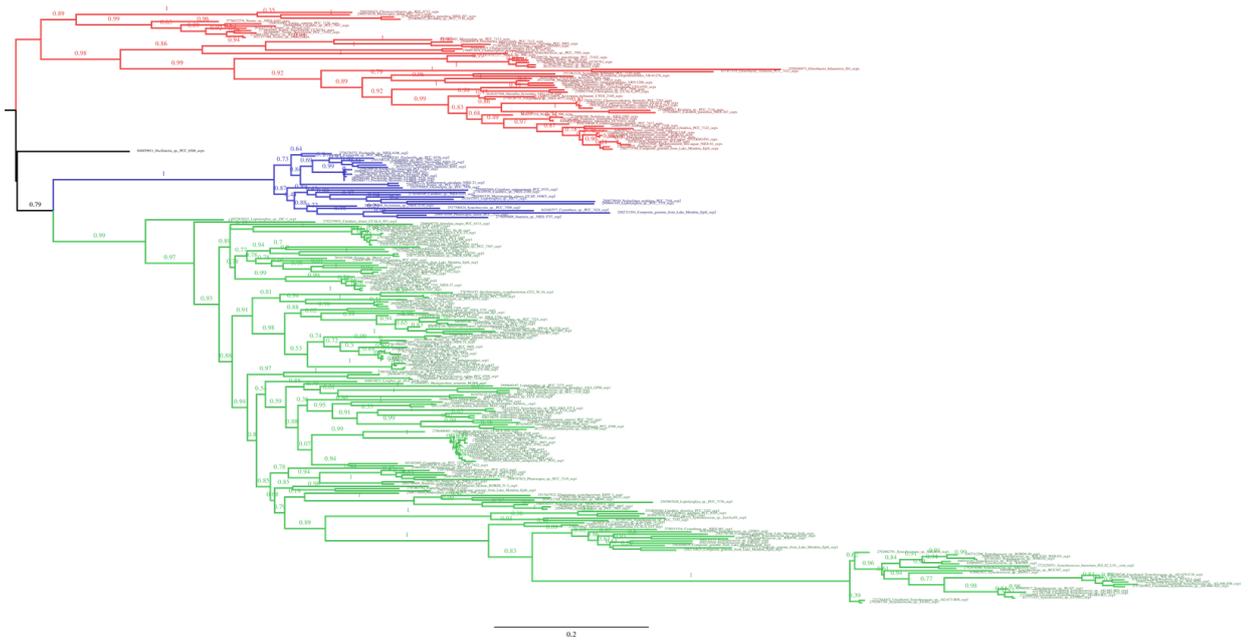
27. Maksimov, E. G. et al. The Unique Protein-to-Protein Carotenoid Transfer Mechanism. *Biophys. J.* 113, 402–414 (2017).
28. Moldenhauer, M. et al. Assembly of photoactive orange carotenoid protein from its domains unravels a carotenoid shuttle mechanism. *Photosynth. Res.* (2017). doi:10.1007/s11120-017-0353-3
29. Jallet, D. et al. Specificity of the Cyanobacterial Orange Carotenoid Protein: Influences of Orange Carotenoid Protein and Phycobilisome Structures. *Plant Physiol.* 164, 790–804 (2014).
30. Bernát, G. et al. Unique properties vs. Common themes: The atypical cyanobacterium *gloeobacter violaceus* PCC 7421 is capable of state transitions and blue-light-induced fluorescence quenching. *Plant Cell Physiol.* 53, 528–542 (2012).
31. Harris, D. et al. Structural rearrangements in the C-terminal domain homolog of Orange Carotenoid Protein are crucial for carotenoid transfer. *Commun. Biol.* 1, 125 (2018).
32. Slonimskiy, Y. B. et al. Light-controlled carotenoid transfer between water-soluble proteins related to cyanobacterial photoprotection. *FEBS J.* 286, 1908–1924 (2019).
33. Maksimov, E. G. et al. Fluorescent Labeling Preserving OCP Photoactivity Reveals Its Reorganization during the Photocycle. *Biophys. J.* 112, 827 (2017).
34. Maksimov, E. G. et al. The photocycle of orange carotenoid protein conceals distinct intermediates and asynchronous changes in the carotenoid and protein components. *Sci. Rep.* 7, 15548 (2017).
35. Wilson, A. et al. Structural determinants underlying photoprotection in the photoactive orange carotenoid protein of cyanobacteria. *J. Biol. Chem.* 285, 18364–18375 (2010).
36. López-Igual, R. et al. Different Functions of the Paralogs to the N-Terminal Domain of the Orange Carotenoid Protein in the Cyanobacterium *Anabaena* sp. PCC 7120. *Plant Physiol.* 171, 1852–1866 (2016).
37. Wen, Y. et al. Orange and red carotenoid proteins are involved in the adaptation of the terrestrial cyanobacterium *Nostoc flagelliforme* to desiccation. *Photosynth. Res.* 1, 1–11 (2019).
38. Harris, D. et al. Orange carotenoid protein burrows into the phycobilisome to provide photoprotection. *Proc. Natl. Acad. Sci.* (2016). doi:10.1073/pnas.1523680113

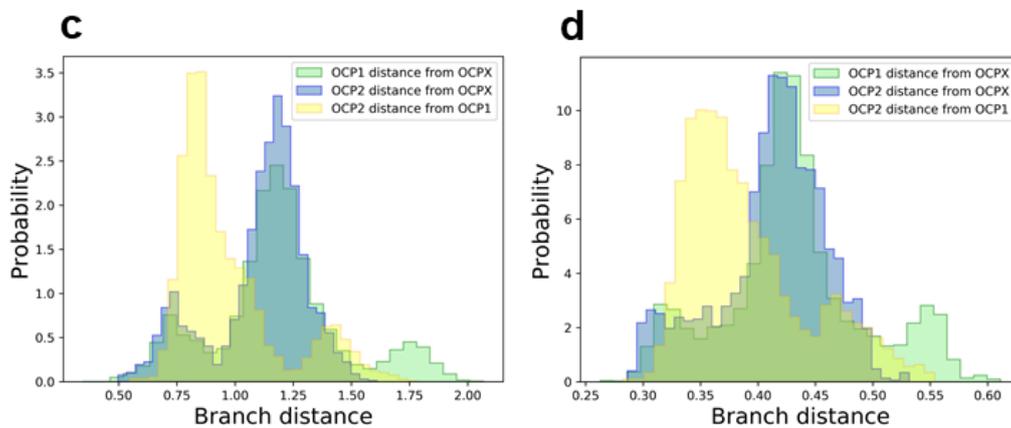
39. Sluchanko, N. N. et al. OCP–FRP protein complex topologies suggest a mechanism for controlling high light tolerance in cyanobacteria. *Nat. Commun.* 9, 1–15 (2018).
40. Bourcier De Carbon, C., Thurotte, A., Wilson, A., Perreau, F. & Kirilovsky, D. Biosynthesis of soluble carotenoid holoproteins in *Escherichia coli*. *Sci. Rep.* 5, 9085 (2015).
41. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *J. Mol. Biol.* 215, 403–410 (1990).
42. Sievers, F. et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* 7, 539 (2011).
43. Guindon, S. et al. New Algorithms and Methods to Estimate Maximum-Likelihood Phylogenies: Assessing the Performance of PhyML 3.0. *Syst. Biol.* 59, 307–321 (2010).
44. Ronquist, F. & Huelsenbeck, J. P. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19, 1572–1574 (2003).
45. Cock, P. J. A. et al. Biopython: freely available Python tools for computational molecular biology and bioinformatics. *Bioinformatics* 25, 1422–1423 (2009).
46. Wheeler, T. J., Clements, J. & Finn, R. D. Skyline: a tool for creating informative, interactive logos representing sequence alignments and profile hidden Markov models. *BMC Bioinformatics* 15, 7 (2014).
47. Waterhouse, A. et al. SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res.* 46, W296–W303 (2018).
48. Berendsen, H. J. C., van der Spoel, D. & van Drunen, R. GROMACS: A message-passing parallel molecular dynamics implementation. *Comput. Phys. Commun.* 91, 43–56 (1995).

**Appendix: Supplementary figures**

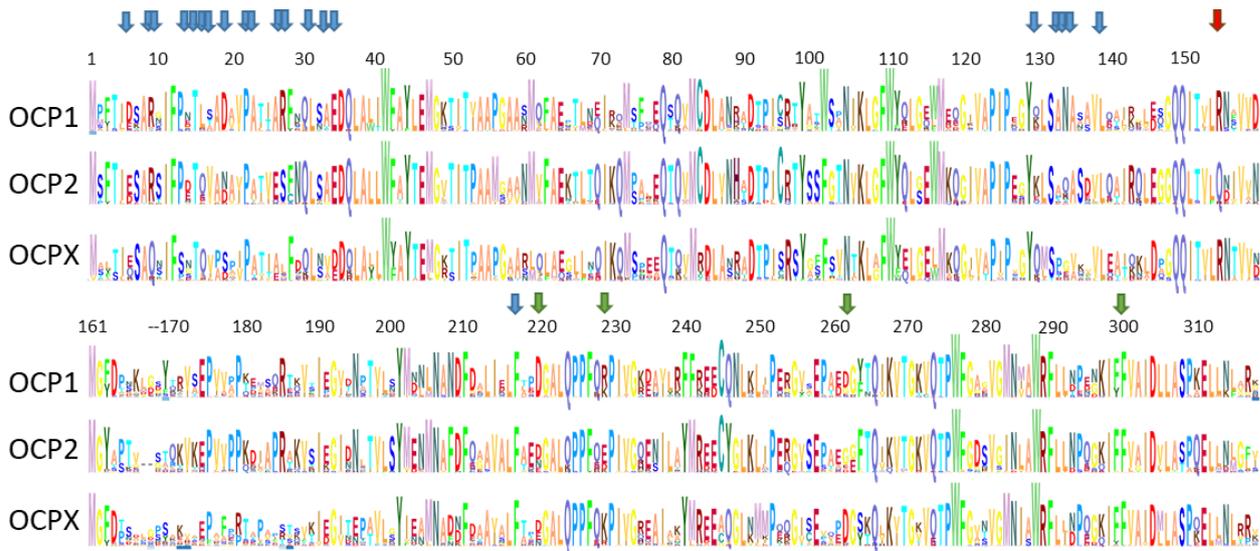


**b**

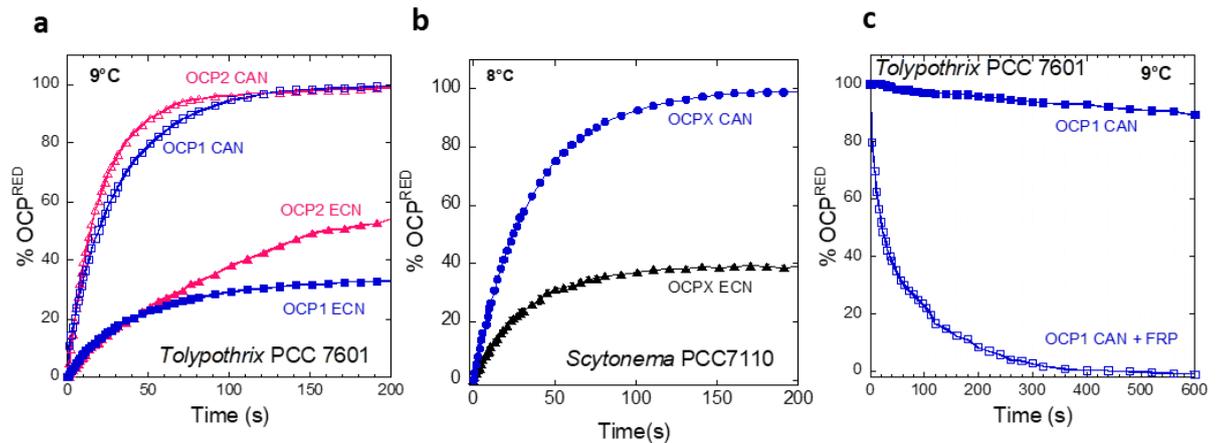




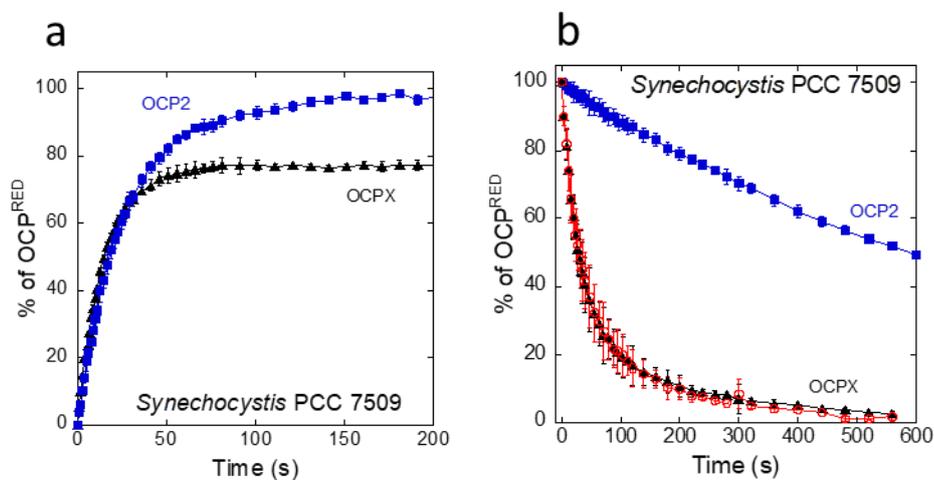
**Supplementary Figure 1 | Phylogenetic tree of OCPs:** Phylogenetic tree of the OCPs protein sequences reconstructed using (a) Neighbor-joining methods, and (b) maximum likelihood phylogenetic. *Streptomyces* NTF2 or *all4940* were used as outgroup for rooting. Supporting this tree topology, histogram of branch distances between OCP1 or OCP2 and OCPX, and between OCP1 and OCP2, extracted from Neighbor-joining tree (c) Bayesian tree (d), shows that OCP1 and OCP2 are equally distant from OCPX, and relatively close between them.



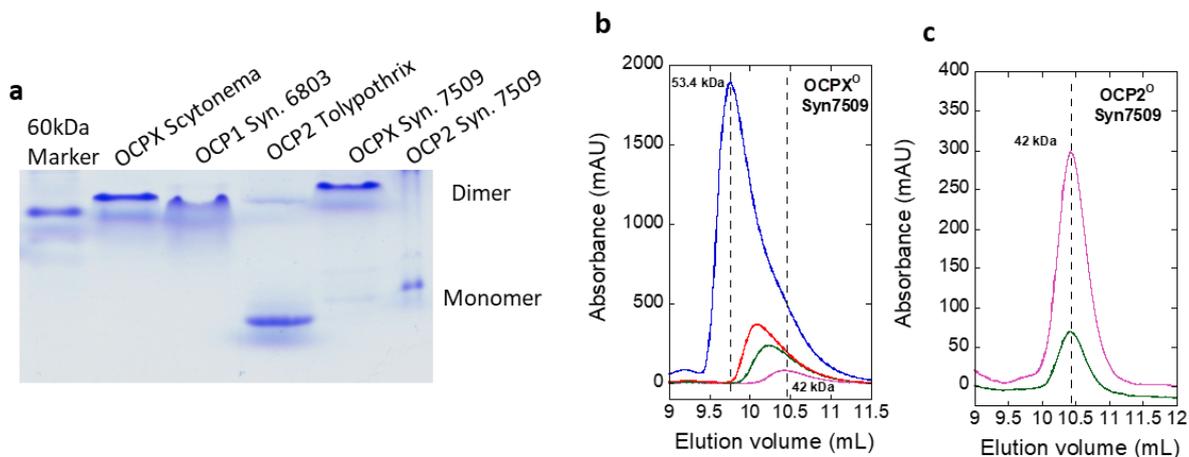
**Supplementary Figure 2 | OCPs sequence alignment.** Logos constructed from OCP1, OCP2 and OCPX clades alignments using Skylign. Numbering corresponds to the positions in *Synechocystis* OCP1. The residues involved in the dimerization interactions<sup>3</sup> are marked by blue arrows. The amino acids proposed to be involved in OCP-FRP<sup>19</sup> interaction are marked by green arrows. The position corresponding to the R155 of *Synechocystis* OCP1, involved in OCP-PB interaction, is marked with a red arrow.



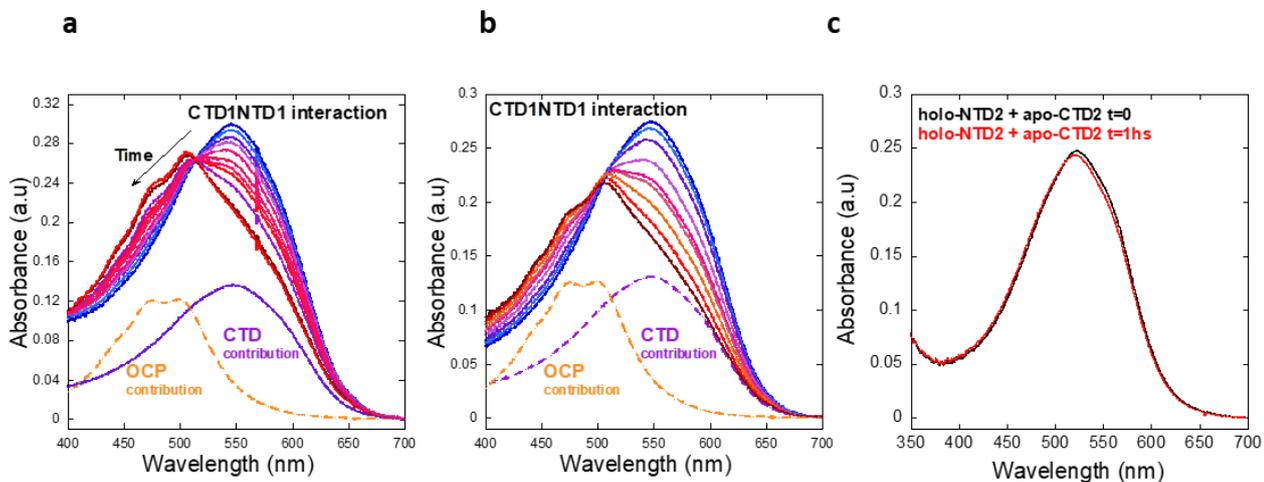
**Supplementary Figure 3 | Tolypothrix and Scytonema OCPs: photoactivation.** (a) Photoactivation kinetics of Tolypothrix OCPs: OCP1CAN (open blue square), OCP1ECN (closed blue square), OCP2 CAN (open fuchsia triangle) and OCP2 CAN (closed fuchsia triangle). (b) Photoactivation kinetics of Scytonema OCPX CAN (blue circles) and OCPX ECN (black triangles) (c) Dark recovery kinetics of OCP1Tolypothrix-CAN in absence(closed square) or presence of Synechocystis FRP (open square). The figures (a-c) show representative experiments, which were repeated 3 times with similar results.



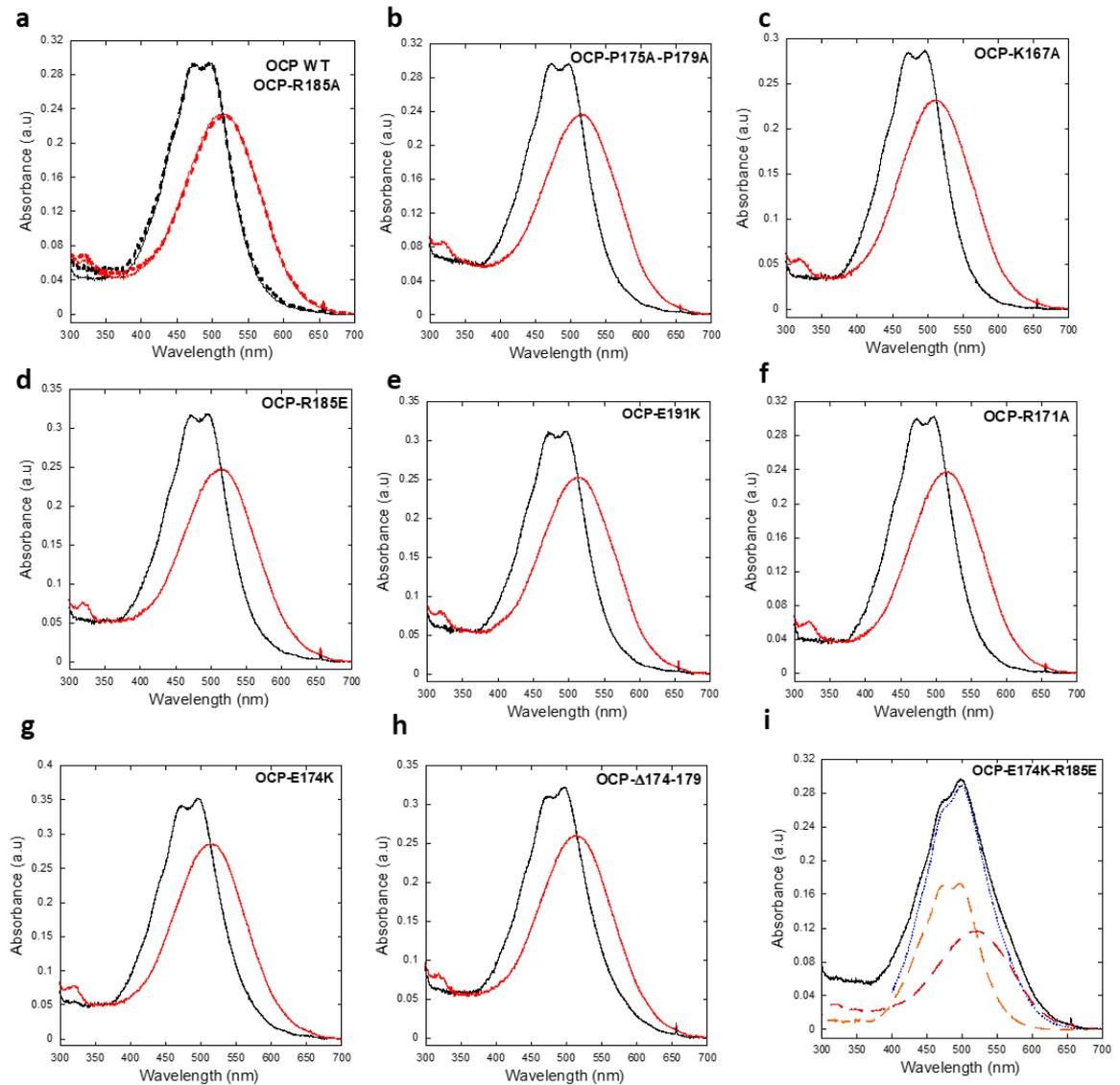
**Supplementary Figure 4 | OCP2 and OCPX from Synechocystis PCC 7509: photoactivation and recovery.** (a) Photoactivation kinetics of OCP2 CAN (blue square) and OCPX CAN (black triangle) from Synechocystis sp. PCC 7509 at 8°C under 5000  $\mu$ E of white light illumination. (b) Back-conversion kinetics of Synechocystis sp. PCC 7509 OCP2 CAN (blue square) and OCPX CAN in the absence (black triangle) and in presence of FRP (red circle) at 8°C. The curves are an average of three independent measurements. The error bars represent the standard deviation of the data shown



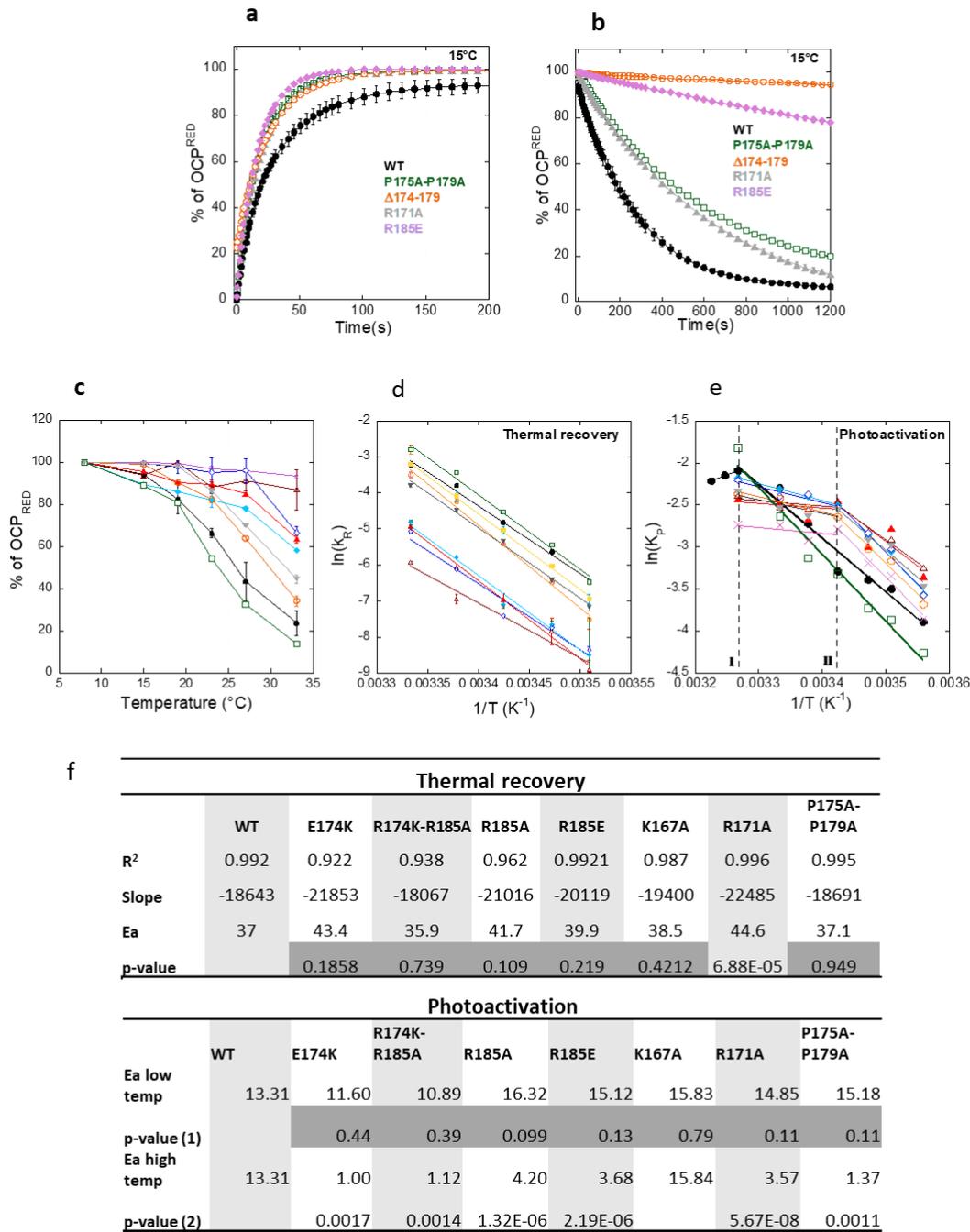
**Supplementary Figure 5 | Oligomeric state of Synechocystis PCC 7509 OCPs.** (a) Analysis of OCP oligomeric state by Coomassie-Blue stained Native PAGE : Scytonema OCPX (10  $\mu$ M), Synechocystis PCC 6803 OCP1 (10  $\mu$ M), Tolypothrix OCP2 (10  $\mu$ M) and Synechocystis 7509 OCPX and OCP2 (10  $\mu$ M). (b-c) Comparison of elution profiles of Synechocystis PCC 7509 OCPX<sup>0</sup> (b) and OCP2<sup>0</sup> (c) at different concentrations: 120  $\mu$ M (blue curve), 40  $\mu$ M (red curve), 23  $\mu$ M (green curve) and 6  $\mu$ M (pink curve). Absorbance was measured at 510 nm. The figures (a-c) show representative experiments, which were repeated 3 times with similar results.



**Supplementary Figure 6 | Domain interaction. Following OCP-like complex formation.** (a) Absorbance spectrum changes during OCP1-like formation when Holo-CTD1 (2.25  $\mu$ M) was incubated with Apo-NTD1 (20  $\mu$ M) at 12°C for 30 minutes. Contribution of holo-OCP1 (dotted orange line) and holo-CTD1 (violet line) in the final spectrum (30 min) is shown. (b) Titration of holo-CTD1 (2.25  $\mu$ M) (blue) with increasing concentrations of apo-NTD1 from 1.25  $\mu$ M (sky blue) to 27.6  $\mu$ M (dark red) at 12°C. They were incubated during 30 min. Contribution of holo-OCP1 (dotted orange line) and holo-CTD1 (dotted violet line) in the spectrum obtained after 30 min incubation of holo-CTD1 (2.25  $\mu$ M) and apo-NTD1 (27.6  $\mu$ M) is shown. (c) Incubation of holo-NTD2 (2.25  $\mu$ M) with apo-CTD2 (11.25  $\mu$ M) at 23°C in darkness. The lack of spectral change indicate that no OCP-like complex had been formed. The figures (a-c) show representative experiments, which were repeated 3 times with similar results.

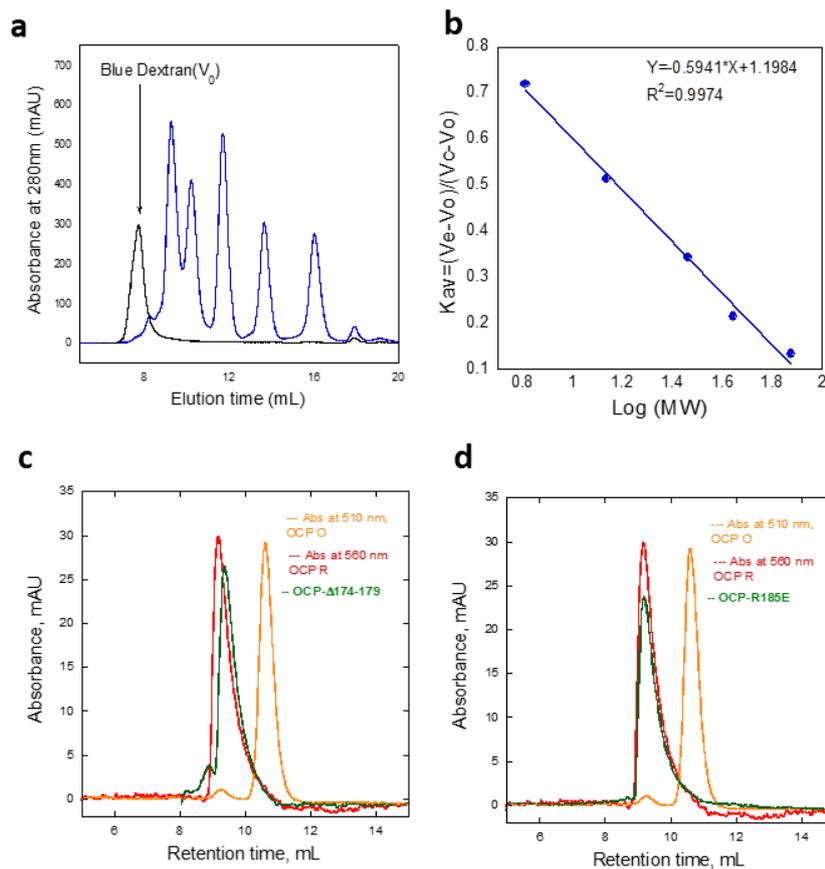


**Supplementary Figure 7 | Absorbance spectra of OCP1 linker mutants.** Dark (black) and photoactivated (red) spectrum of WT OCP and OCP R185A (dashed lines) (a), OCP-P175A-P179A (b), OCP-K167A (c), OCP-R185E (d), OCP-E191K (e), OCP-R171A (f), OCP-E174K (g), OCP-  $\Delta$ 174-179 (h), and dark spectra of OCP-e174K-R185E (black curve), holo-OCP<sup>O</sup> contribution (dotted orange curve) and holo-OCP<sup>R</sup> contribution (dotted red curve) obtained from spectral decomposition (dotted blue curve) (i). The figures (a-i) show representative experiments, which were repeated 3 times with similar results.

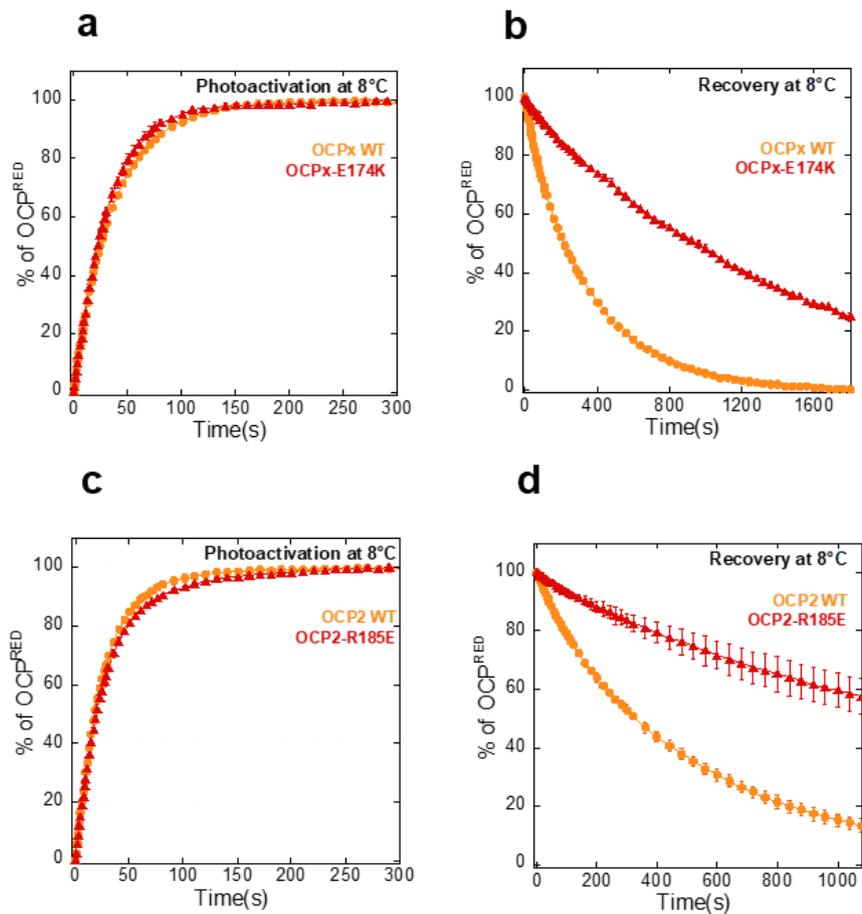


**Supplementary Figure 8 | OCP1 linker mutants photoactivation at different temperatures.** (a) Accumulation of photoactivated  $OCP^R$  and (b)  $OCP^R$  back-conversion to  $OCP^O$  at 15°C of OCP WT (black closed circle), OCP R185E (violet diamond), OCP R171A (grey triangle), OCP P175A-P179A (green square) and OCP  $\Delta$ 174-179 (orange open circle). (c)  $OCP^R$  stability dependence on the temperature. (d) Arrhenius plot of dark recovery and (e) Arrhenius plot of photoactivation of : WT OCP (black close circle), OCP E174K (red close triangle), OCP E191K (yellow close square), OCP E174K-R185E (dark red open triangle), OCP R185A (sky blue close diamond), OCP R185E (blue open diamond), OCP K167A (green open square), OCP R171A (orange open circle), OCP P175A-P179A (grey close inverted triangle) and OCP  $\Delta$ 174-179 (pink cross). Break point of the convex curves is shown with dotted lines for the WT OCP (I) and the linker mutants (II). The curves are an average of three independent measurement. The error bars represent the standard deviation of the data shown.

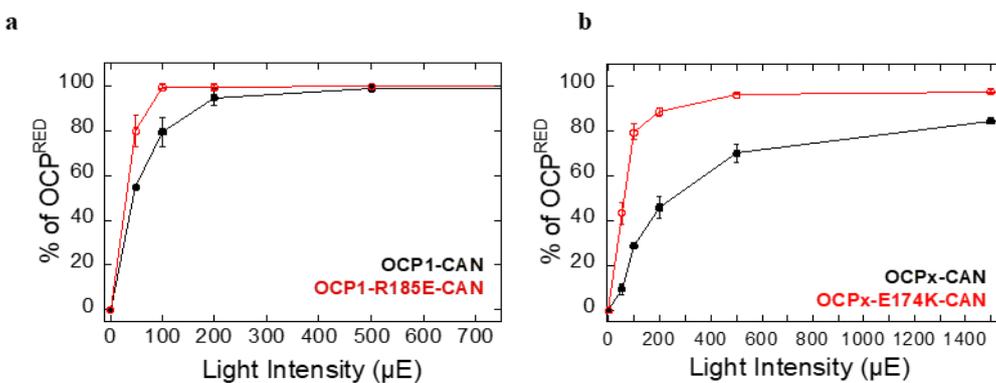
**f) Parameters from data fitting using Arrhenius equation.** Activation energy ( $E_a$ ) was calculated from the slope of Arrhenius equation; values are presented in kcal/mol. We performed T-test to test the null hypothesis that there is no statistical difference between each Arrhenius slope of OCP-mutants and the Arrhenius slope of the WT OCP. We cannot reject the null hypothesis for the p-values that are highlighted in dark-grey (with a 95% confidence interval and 15 degrees of freedom). In photoactivation sub-table, p-value(1) correspond to t-test comparing slopes at low temperatures (8°C-19°C) and p-value(2) correspond to t-test comparing slopes at high temperatures (19°C-27°C).



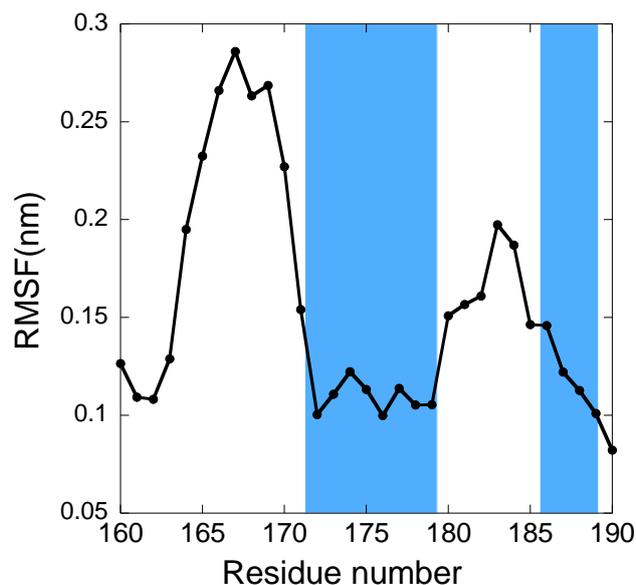
**Supplementary Figure 9 | OCP1 linker mutants: size exclusion chromatography.** (a) Elution profile of proteins used for calibration: Aprotinin, Ribonuclease A, Carbonic anhydrase, Ovalbumin and Conalbumin. (b) Calibration curve plotted using the gel-phase distribution coefficient ( $K_{av}$ ) versus logarithm of the molecular weight (Log Mw).  $K_{av} = (V_e - V_o)/(V_c - V_o)$  where  $V_e$  = elution volume,  $V_o$  = column void volume. (c-d) Comparison of elution profiles of holo-WT-OCP1<sup>O</sup>-ECN 6μM (orange curve) and holo-WT-OCP1<sup>R</sup>-ECN 6μM (red curve) with photoactivated holo-OCP1-Δ174-179-ECN 6μM (green curve) (c) and holo-OCP1-R185E-ECN 6μM (green curve) (d). The figures (a,c,d) show representative experiments, which were repeated 3 times with similar results.



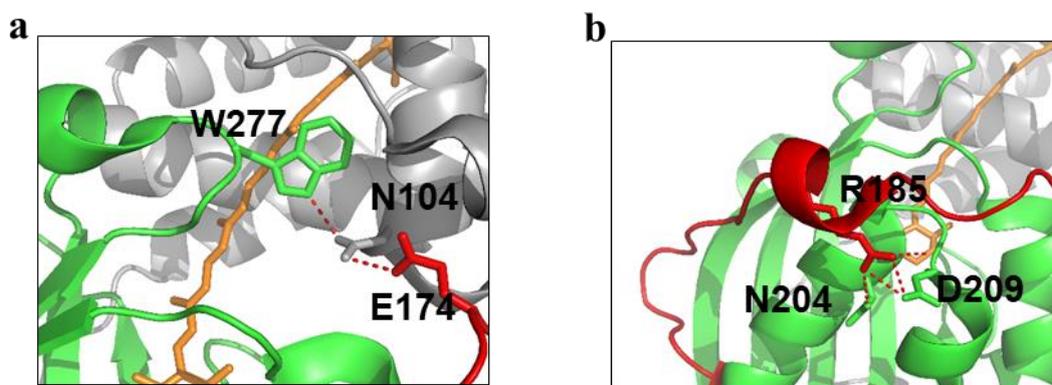
**Supplementary Figure 10 | Photoactivity and recovery of OCP2 and OCPX linker-mutants.** (a and c) Kinetics of photoactivation (under 500  $\mu\text{E}$  of white light illumination) and (b and d) back-conversion (OCP<sup>R</sup> to OCP<sup>O</sup>) of OCPX WT or OCP2 WT (orange circles) and OCPX E174K or OCP2 R185E (red triangles). Experiments were carried out at 8°C. The curves are an average of three independent measurement. The error bars represent the standard deviation of the data shown.



**Supplementary Figure 11 | Effect of linker mutations on photoactivation at different light intensities.** Percentage of OCP that is photoactivated after illumination at different light intensities of OCP1-CAN (black)/ OCP1-R185E-CAN(red) (a) and OCPX-CAN (black) / OCPX-E174K-CAN(red) (b). The curves are an average of three independent measurements. The error bars represent the standard deviation of the data shown.



**Supplementary Figure 12 | Role of linker amino acids in the OCP.** Root mean square fluctuation (RMSF) of protein backbone as function of apo-OCP1 residue number, only the linker region is shown and the less flexible regions are colored in sky blue.



**Supplementary Figure 13 | Structure, function of the OCP linker.** (a and b) Zoom into the key interactions between E174 (a) or R185 (b) and the domains of the OCP (CTD: green; NTD: grey; linker: red).

(a) In OCP1<sup>o</sup>, the E174 is close to the principal interface NTD-CTD forming an H-bond with N104, in the NTD, which interacts with W277 in the CTD. The N104-W277 interaction largely stabilizes the closed orange form. Thus, E174 could have a role in the proper positioning of N104 to allow its interaction with W277.

(b) R185 which is involved in an H-bond with N204 and D209 in the exterior surface of the CTD also could be important to stabilize the closure of the domains during back conversion.

Primer names		Primer sequences
<b>A- Synechocystis OCP Mutants</b>		
E174K	F	5'-GTAAGGACGGCAAACGCATCGCTAAAACCGTAGTTCTCCCAAG-3'
	R	5'-AGCGATGCGTTTGCCGTCTTACCAGCGGTGAAGCC-3
ΔE174-P179	F	5'-GTAAGGACGGCAAACGCATCGCTCAAGATACCGCCAGCCGCACCAAAG-3'
	R	5'-AGCGATGCGTTTGCCGTCTTACCAGCGGTGAAGCC-3
R185E	F	5'-CAAGATACCGCCAGCGAGACCAAAGTTCCATTG-3'
	R	5'-CAATGGAACTTTGGTCTCGCTGGCGGTATCTTG-3'
R185A	F	5'-CAA GAT ACC GCC AGC GCC ACC AAA GTT TCC ATT G-3'
	R	5'-C AAT GGA AAC TTT GGT GGC GCT GGC GGT ATC TTG-3'
K167A	F	5'-GGC TTC ACC GCT GGT GCG GAC GGC AAA CGC ATC-3'
	R	5'-GAT GCG TTT GCC GTC CGC ACC AGC GGT GAA GCC-3'
R171A	F	5'-GGT AAG GAC GGC AAA GCC ATC GCT GAG CCC GTA G-3'
	R	5'-C TAC GGG CTC AGC GAT GGC TTT GCC GTC CTT ACC-3'
P175A-P179A	F	5'-CGC ATC GCT GAG GCC GTA GTT CCT GCC CAA GAT ACC GCC-3'
	R	5'-GGC GGT ATC TTG GGC AGG AAC TAC GGC CTC AGC GAT GCG -3'
OCP1-R27L	F	5'-CCCGCTACCATCGCCCTTTTGTAGCCAACTCAATGC-3
	R	5'-GCATTGAGTTGGCTAAAAGGGCGATGGTAGCGGG-3
E191K	F	5'-CCAAAGTTTCCATTAAGGGGTGACCAATGCC-3
	R	5'-GGCATTGGTCACCCCTTAATGAAACTTTGG-3
R229E-D262K	F	5'-GAA CCC GCT GAA AAG GGT TTC ACC CAA ATT AAA -3 5'- CC TTC CAA GAG CCA ATT GTC GGT AAA G-3
	R	5'-TTG GGT GAA ACC CTT TTC AGC GGG TTC AGT AAC-3 5'-CTT TAC CGA CAA TTG GCT CTT GGA AGG-3
<b>B- Construction of the pCDF-NtagOCP1, pCDF-NtagOCP2 and mutants plasmids</b>		
OCP1-pDuet	F	5'-AACAGAGGATCCGCCGTTTACCATCGATTGAGCT-3'
	R	5'-TCTGTTGCGGCGCGTTAACGGTTGACAGTTCACAG-3'
OCP2-pDuet	F	5'-AACAGAGGATCCGCCGTTTACTATCAAGTCTGCA-3'
	R	5'-TCTGTTGCGGCGCGCAACATATTGCACCCAAGGTTTC-3'
OCP2-R185E	F	5'-CAATGCTGACCTTAGCCTCGGGTGCATATCTTTAGG-3'
	R	5'-CCTAAAGATATCGCACCCGAGGCTAAGGTCAGCATTG-3'
<b>C- Construction of the pCDF-NTD2 Ctag plasmid</b>		
F-ocp2-ΔBamHI		5'-TTAATA AGG AGA TAT ACC ATG CCG TTT ACT ATC AAG TCTGCACGCTCC-3'
R-Duet		5'-CAT GGT ATA TCT CCT TAT TAA AGT TAA ACA AAA TTA TTT CTA-3'
F-DuetOCP2 Ctag		5'-CAC CAC CAC CAC CAC TAG AGC AAGTTTGGGTAACAGTGTCTGG-3'
R-NTD2 Ctag		5'-GTG GTG GTG GTG GTG GTG TGG GGC ATA CCC CAT ATT GAC AAC-3'

**Supplementary Table 1 A | Sequences of primers used in this study.** Listed primer include those used for construction of plasmid for protein expression in *E. coli* as described in the Material and Methods section.

Primer names	Primer sequences	
<b>D- Construction of the pCDF-CTD2 Ctag plasmid</b>		
F-DuetOCP2 Ctag	5'-CAC CAC CAC CAC CAC CAC TAGAGCAAGTTTGGGGTAAACAGTGTCTGG-3'	
R-DuetOCP2 Ctag	5'-GTG GTG GTG GTG GTG GTGCTGGCTAAAGCCCATGTTAAGTAGTTCCCTG-3'	
F-ΔNTD2	5'-TTAATA AGG AGA TAT ACC ATGATTGAGGGCATCACCAACTTGACAGTC-3'	
R-Duet	5'-CAT GGT ATA TCT CCT TAT TAA AGT TAA ACA AAA TTA TTT CTA-3'	
<b>E- Construction of the pCDF-NtagOCPx and mutants plasmids</b>		
OCPx-pDuet	F	5' ATCAACCACAGCCAGGATCCGACTTTTACTATCGAGTCC--3'
	R	5'-CTTAAGCATTATGCGGCCGCCACTACTACACAGGACG-3
OCPx-E174K	F	5'-GTTCAAACAGATGCCAAACCTTTGTTCCCAACGC-3'
	R	5'-GCGTGGGAACAAAGGTTTGGCATCTGTTTGAAC-3'
<b>F- Construction of the pCDF-NTDx Ctag plasmid</b>		
F-ocpx-ΔBamHI	5'-TTAATA AGG AGA TAT ACC ATG ACT TTT ACT ATC GAGTCCGCCCAAAG-3'	
R-Duet	5'-CAT GGT ATA TCT CCT TAT TAA AGT TAA ACA AAA TTA TTT CTA-3'	
F-DuetOCPx Ctag	5'-CAC CAC CAC CAC CAC CAC TAG TAG TGG CGG CCG CAA TG-3'	
R-NTDx Ctag	5'-GTG GTG GTGGTGGTGGTGGCTCCC GAATCCCATGTCTACTACAGCATTG-3'	
<b>G- Construction of the pCDF-CTDx Ctag plasmid</b>		
F-DuetOCPx Ctag	5'-CAC CAC CAC CAC CAC CAC TAG TAG TGG CGG CCG CAA TG-3'	
R-DuetOCPx Ctag	5'-GTG GTG GTG GTG GTG GTG CAC AGG ACG CAG ATT TAG TAG TTCTTG-3'	
F-ΔNTDx	5'-TTAATA AGG AGA TAT ACC ATG GCT GTC GAGGGC ATCACAGACTCTAC-3'	
R-Duet	5'-CAT GGT ATA TCT CCT TAT TAA AGT TAA ACA AAA TTA TTT CTA-3'	

**Supplementary Table 1 B | Sequences of primers used in this study.** Listed primer including those used for construction of plasmid for protein expression in *E. coli* as described in the Material and Methods section.



## **Part III**

---

# **Conclusions and Perspectives**

## ***General discussion***

---

In the last 13 years, the OCP-related NPQ mechanism has been extensively characterized. Nevertheless, the role/mechanism of the paralogs of the OCP, the CTDH and HCP, was less clear when I began my thesis. At that time, we knew that, at least in vitro, HPC4 can induce PBS fluorescence quenching and HCP2 and HCP3 are good singlet oxygen quenchers. These results suggested that HCPs are involved in photoprotection mechanisms in cyanobacteria cells. By contrast, we knew almost nothing about CTDH. Since CTDH and HCP4 usually are located in adjacent loci on the genome, it was proposed that they could form a heterodimer with similar characteristics to the OCP, resembling the interaction between the ancestors of the OCP domains. Moreover, it was hypothesized that the fusion of their ancestors created the modern OCPs. This thesis work was undertaken to study the uncharacterized CTDH and its interaction with the HCP. One of the main findings presented in this thesis is that CTDHs are transient carotenoid proteins involved in carotenoid transfer to other proteins. Then, I explored the different subclades of OCP by studying the role of domain interactions and of the flexible linker on the evolution of the OCPs. Overall, all our results have shed light on the evolution and different roles of carotenoid proteins in cyanobacteria. In this section, I discuss the main findings and their implications.

### ***Role of CTDH and HCP proteins in cyanobacteria***

The paralogs of the CTD domain of the OCP, called CTDHs, are capable to bind carotenoid. The carotenoid binding stabilizes the dimeric form of the CTDH in which the carotenoid is shared by the two monomers. The low amount of holo-CTDH isolated from canthaxanthin producing *Synechocystis* cells strongly suggested that CTDHs are transient carotenoid proteins. The results of my thesis strongly suggest that the main role of CTDHs is the carotenoid delivery to HCPs. In addition, holo-CTDH could have a supplementary and important role as  $^1\text{O}_2$  quencher. The main experimental results behind this proposition were the carotenoid transfer from holo-CTDH to apo-HCP and the fact that only apo-CTDHs are able to uptake carotenoids from membranes. Further support is provided by other results: 1) when isolated from carotenoid producing *E. coli*, HCPs bind less carotenoid than CTDHs (in absence of any carotenoid acceptor, CTDH remains in the holoform) due to its lower capacity of carotenoid uptake from membranes; 2)

when they are co-expressed or when their apoforms are mixed with carotenoid containing membranes, the yield of holo-HCP is increased (unpublished results) confirming the requirement of CTDH as carotenoid provider; 3) In all the genomes containing *hcp* genes also one *ctdh* gene is present. Nevertheless, the fact that some holo-HCP can be isolated from *E. coli* or cyanobacteria suggested that during protein synthesis and folding, carotenoids can be incorporated to these proteins but with low efficiency.

Moreover, in *Nostoc flagelliforme*<sup>1</sup>, transcription of both HCP and CTDH is up-regulated during desiccation. However, the increase of transcription of CTDH is less pronounced than that of HCP1/HCP6. This can be interpreted as the following: since the CTDH may serve as catalyzer of the carotenoid loading into different HCPs, it is required in a lesser amount than the HCPs, although a synchronization of the expression pattern would increase the effectiveness of the system. By contrast, the levels of OCP remain almost constant during desiccation. Thus, it is tempting to hypothesize that OCP-paralogs (CTDH and HCP) have evolved in parallel to the OCP in order to respond to different stress conditions and different photoprotective mechanisms. It could be that the principal role of the upregulated HCPs is mostly antioxidant while OCP is mostly useful to decrease the energy arriving to the photochemical centers.

### ***Carotenoid translocation: CTD-NTD vs CTDH-HCP***

In the past, it was demonstrated that carotenoid translocation has an essential role in the OCP photocycle and activity. In darkness the carotenoid is shared between the two domains, CTD and NTD. However, during OCP (or OCP-like) photoactivation the carotenoid is translocated to the NTD. This results in the formation of the active OCP<sup>R</sup> that can induce energy dissipation at the level of PBSs. In this thesis, I show that carotenoid translocation also plays an important role for HCP activity: HCP receives the carotenoid, which is essential for their activity from another protein: the CTDH. The parallel study of the mechanism of carotenoid transfer between isolated free OCP domains and between HCPs and CTDHs allow us to conclude that the carotenoid translocation and its direction are determined by the carotenoid-affinity of each component and the strength of interaction between them. Nevertheless, there are important differences between carotenoid translocation and protein interaction between CTDH/HCP and OCP-CTD/OCP-NTD couples.

First, the interaction between the isolated domains of the OCP results in a stable complex with OCP-like features (when holo-CTD interacts with apo-NTD). On the other hand, the interaction between holo-CTDH and apo-HCP is transient and no complex is formed. This may be explained as a stronger protein interaction between CTD and NTD than between CTDH and HCP. Nevertheless, the interaction with the carotenoid could also affect the complex formation. For example, if the difference of the affinity for the carotenoid of the two domains in the OCP is lower than between the CTDH and HCP. This balance could allow the carotenoid to be in an intermediary position between the two domains and stabilize the complex.

Second, during the CTDH-HCP interaction the carotenoid is completely translocated to the HCP. This carotenoid translocation is spontaneous and light-independent while in the OCP (and OCP-like) the translocation of the carotenoid to the NTD takes place only upon strong illumination (during photoactivation). The molecular basis behind the light-induced carotenoid translocation in the OCP are unknown. In the same way is unclear why the carotenoid transfer in the CTDH-HCP is light independent. A possible explanation for this difference is that the carotenoid channel formed in the CTD-NTD complex is different than the one formed in the CTDH-HCP transient complex. In the first case, it forces the carotenoid to be in a bent and highly energetic position between the two domains. While in the second case there are no structural constraints and the carotenoid can freely move to the HCP. Some factors can contribute to this difference: **I-** it is possible that the NTD of the OCP could adopt two different conformations to bind the carotenoid: one of them stabilizes the carotenoid in the OCP<sup>O</sup> and the other one allows the complete incorporation of the carotenoid into the NTD during photoactivation. In this case, the HCP could adopt only the latter conformation and provoke the direct carotenoid transfer without stable complex formation. **II-** As it was mentioned before protein interaction might be stronger in the CTD-NTD couple and this could stabilize enough the complex to allow the correct positioning of the carotenoid between the two proteins. In the case of CTDH-HCP, if the interaction is weaker the complex is not stable and the carotenoid channel is not properly formed leading to the direct translocation of the carotenoid.

### *Multidirectional carotenoid transfer*

In this thesis, it is shown that AnaCTDH (under reducing conditions) can deliver the carotenoid to HCP4 and TeHCP4 but not retrieve it from them.

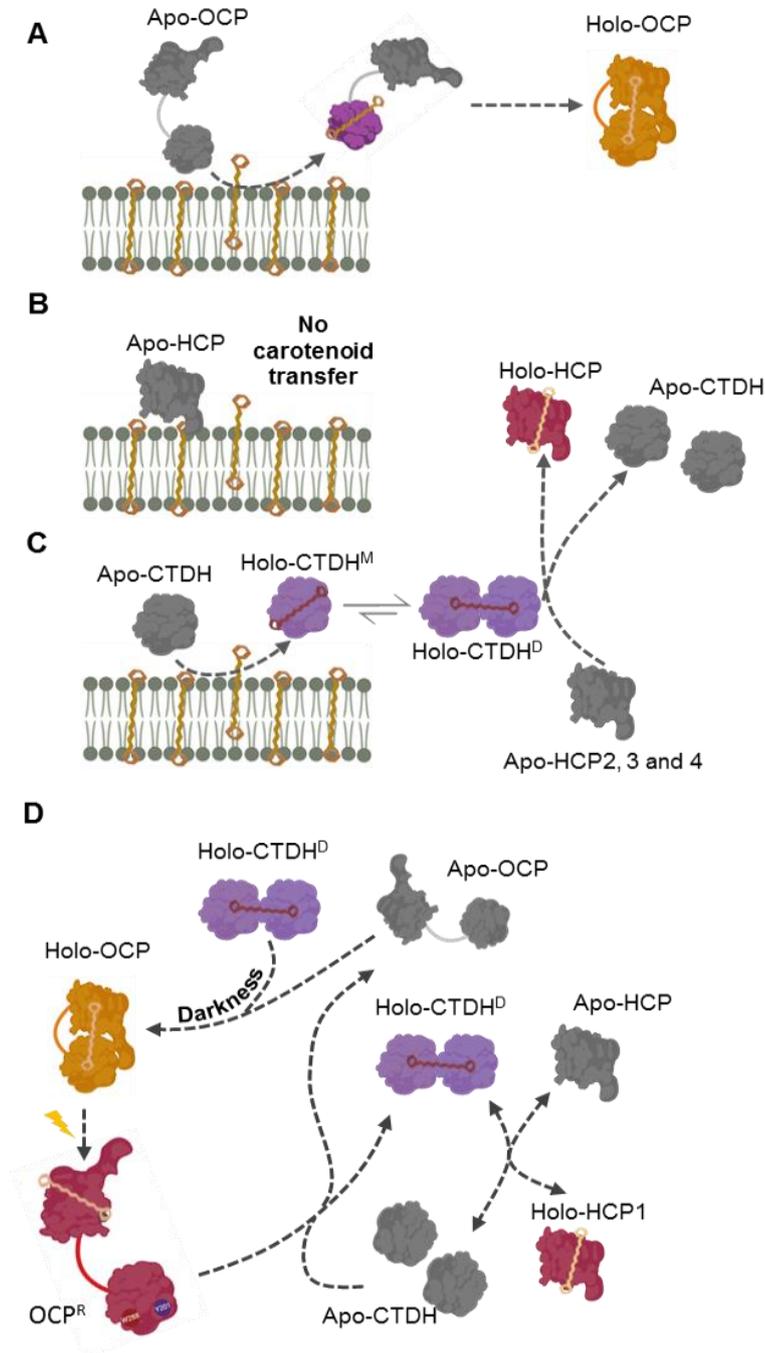
Interestingly, TeCTDH could deliver the carotenoid only to TeHCP. This high specificity of TeCTDH may be explained by the fact that there is only one HCP in this strain and both, TeCTDH and TeHCP, could have co-evolved to improve their interaction. On the other hand, in *Anabaena* there are four HCPs, and most likely the AnaCTDH have evolved to be able to interact with different HCPs.

Neither holo-AnaCTDH nor holo-TeCTDH can transfer the carotenoid to the isolated apo-NTD. However, AnaCTDH is able to partially transfer the carotenoid to the OCP. This unexpected result strongly suggests that the carotenoid transfer takes place from CTDH to the CTD domain of the OCP. In agreement with this, CTDH can transfer the carotenoid to isolated CTD. In addition, AnaCTDH can retrieve carotenoid from OCP<sup>R</sup>, in which the carotenoid is in the NTD. This result showed that an “inverse” carotenoid transfer from NTD to CTDH also exists. In agreement with this, isolated NTD can transfer the carotenoid to apo-CTDH. It was suggested that this “inverse” carotenoid transfer could play the FRP role by deactivating OCP<sup>R</sup> in strains lacking FRP (these strains usually contain OCPX). Interaction between apo-CTDH and OCP<sup>R</sup> would lead to carotenoid transfer to CTDH and formation of apo-OCP, which is inactive. Then holo-OCP could be regenerated in darkness by carotenoid transfer from the same CTDH to the apo-OCP. It is important to bear in mind that this is just an *in vitro* observation and it will require further *in vivo* experiments to prove this hypothesis.

Finally, the AnaCTDH can retrieve carotenoid from the enigmatic holo-HCP1. Despite the fact that this protein is expressed *in vivo*, its role is unclear because HCP1 cannot quench neither <sup>1</sup>O<sub>2</sub> nor PBS fluorescence. Although our results suggested that HCP1 could be a carotenoid carrier, this is only a hypothesis and further *in vivo* experiments are required to get a better understanding about this protein. In the case of CTDH-HCP1 interaction, the carotenoid translocation can take place in both directions and the steady state is determined by the equilibrium of those reactions. It is important to note that HCP1 still requires the presence of CTDH to retrieve carotenoid from the membrane.

Overall, the paralogs of the OCP-domains are capable of multiple interactions between them leading to a very intricate pattern of carotenoid transfer. Carotenoid translocation among these proteins, which is summarized in the Figure 1, can have deep implications in the photoprotection of cyanobacteria. In this hypothetical model, CTDH can serve as both <sup>1</sup>O<sub>2</sub> quencher and carotenoid donor to HCPs in the cell. In this case, the carotenoid

transfer from CTDH to HCP would ensure the presence of holo-HCP which also can quench  $^1\text{O}_2$  and, in more extreme cases, induce dissipation of excess of energy in the PBS.



**Figure 1 | Carotenoid translocation among OCP-paralogs.** Apo-OCP (A) and apo-CTDH (C) are capable of carotenoid uptake from membranes while apo-HCP (B) is not able to do it. After holo-CTDH formation (C), it interacts with Apo-HCPs and transfers the carotenoid to them to form holo-HCPs. Model based on results described in [65]. (D) Multidirectional carotenoid transfer. Ana-HCP1 in addition to be able to receive the carotenoid from the CTDH is the only HCP that is also able to give the carotenoid to CTDH. Under strong illumination, photoactivated OCP<sup>R</sup> is able to transfer the carotenoid to apo-CTDH forming

holo-CTDH and apo-OCP. In darkness, the holo-CTDH can transfer the carotenoid back to the OCP. Figure adapted from Muzzopappa and Kirilovsky 2019.

### *Mechanism of carotenoid transfer in OCP-related proteins*

The structural and mechanistic basis behind the CTDH-HCP interaction and associated carotenoid transfer are not easy to decipher due to the fact that this interaction is transient. Nevertheless, I showed in this work that by solving the tridimensional structure of apo-AnaCTDH (in collaboration with Dvir Harris), some aspects of the carotenoid transfer can be unraveled. The apoform of the CTDH can form two types of dimers (back-to-back dimer and F-type dimer). Although the structure of the holoform was not solved, it is clear that it is a dimer binding a single carotenoid molecule. This dimeric form probably is the F-type dimer, since in this case the carotenoid channels of each monomer face the other one, and the carotenoid can be completely buried between them. Thus, the formation of the F-type dimer is stabilized by carotenoid binding.

The oxidized holo-AnaCTDH is unable to transfer the carotenoid. In this CTDH, as well as in all the CTDH belonging to the subclade CTDH2, there is a 100% conserved cysteine in the (F-type) dimerization surface. This Cys can form a disulfide bond that hinders monomerization and carotenoid transfer. This confirms the F-type as the main dimeric form of the holo-CTDH and demonstrated that monomerization of the CTDH is required for the carotenoid translocation to HCPs. In addition, the presence of the Cys could allow a further regulation of the system in the CTDH2-containing strains. The local cellular redox-state could determine the formation of the disulfide bond and regulate the carotenoid transfer from CTDH to HCP. However, this is just a hypothesis and further *in vivo* experimentations will be required to confirm it.

The most remarkable feature of the apo-CTDH structure is the high mobility of the CTT which can populate two main flexible conformations in the apoform (#1 and #2 in figure 5A, chapter III). In the holoform, the CTT adopts an open position interacting with the  $\beta$ -sheet. In this way, the carotenoid channel is unblocked (as it is observed in the OCP structure). These differences between *holo*- and *apo*- CTDHs suggested a role for the CTT in carotenoid uptake/delivery. Mutational analysis showed that deletion of the CTT inhibits carotenoid transfer and uptake. The hydrophobic patch at the end of the CTT may help to extract the carotenoid from the membrane and to translocate it during HCP-interaction. Mutation of L137 which is at the

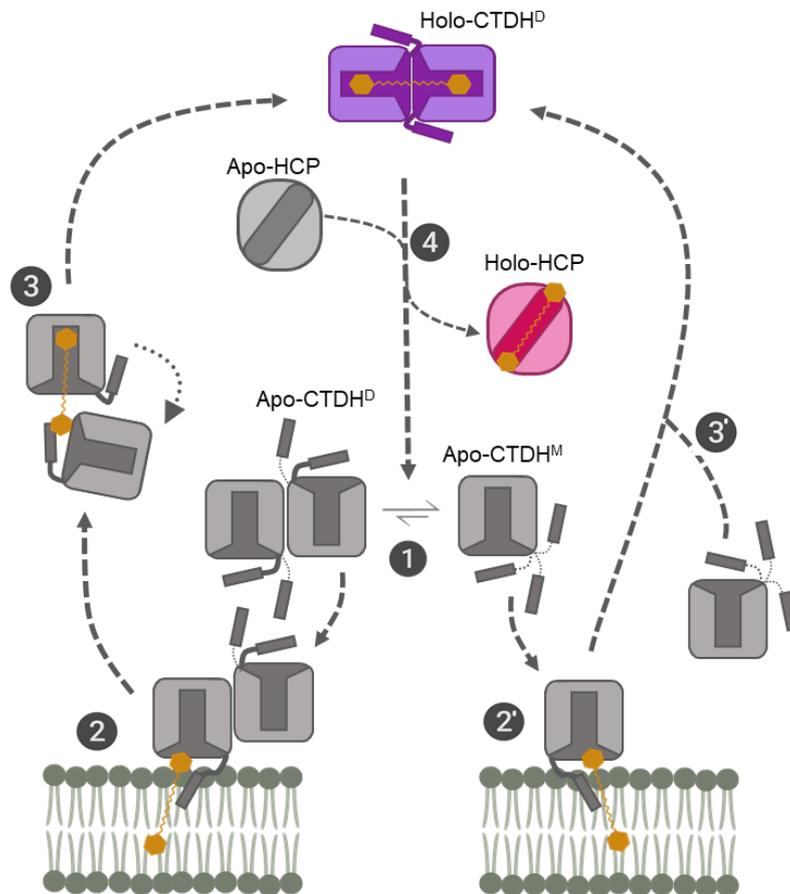
end of the hydrophobic patch strongly decreased the rate of carotenoid uptake/transfer, confirming its role on facilitating the movement of the carotenoid. MD simulations suggested that in addition, L137 is involved in anchoring the CTT to the  $\beta$ -sheets during A-type to F-type dimer transitions. MD simulations showed that the L137 can interact with the hydrophobic part of the  $\beta$ -sheet in the open position of the CTT. Replacement of L137 by a charged residue destabilized this interaction and increased the flexibility of the CTT. In agreement with this, the volume of the CTDH-L137D monomeric is higher than that of the WT. Thus, the CTT also participates on carotenoid uptake by promoting the switch between A-type and F-type dimers. On the other hand, L128 was found to stabilize the back-to-back dimer probably by stabilizing the close conformation. Supporting this, the amplitude of the carotenoid uptake (from membranes or HCP1) was reduced when this residue was replaced by aspartate. Taking together all the results presented in account, we can propose a model for the mechanism of carotenoid uptake of CTDH, which is illustrated in the figure 2.

### ***Role of the OCP flexible linker and evolution of the OCP***

The presence of *ctdh* and *hcp* genes in the same OCP-containing genomes prompted the researchers to suggest that the *ocp* gene is the result of a gene fusion between the genes coding for the ancestors of the OCP domains. The isolated domains of the modern OCPs can interact and form an OCP-like complex, as it was shown in the chapter I and IV and in previous publications<sup>2,3</sup>. In the case of the domains of OCP2, this OCP-like complex can be photoactivated and then relax to the initial state while OCP1-like complex can be irreversibly photoactivated. This observation together with biochemical differences between OCP1 and OCP2 were the basis to proposed that OCP2 represent an ancestral OCP and that their domains resemble the features of the ancestors of the OCP. In this model the linker plays no or almost no role in the OCP mechanisms and there is no explicit explanation why the OCP have evolved as a single protein instead of two interacting proteins and the evolutionary advantage that this represents. Our results from chapter IV largely contributed to the reformulation of this paradigm.

It was first suggested that the OCP was originated by a gene fusion event of ancestors of CTDH/HCP4. Another alternative explanation is that the *ctdh* and *hcp4* are the products of the fission of the *ocp* gene. This could be an intermediary step during the genetic deletion of *ocp*. In both situations, the primary event leading to the presence of both of them (*ctdh/hcp* and *ocp*) in

the same genome was gene duplication (before the gene fusion or fission event). This defines the *ctdh* and *hcp* as paralogs of the *ocp* domains. In agreement with the findings presented on this thesis, paralogs usually have different functions.



**Figure 2 | Carotenoid uptake mechanism.** (1) In solution, the apoform of CTDH is in equilibrium between the monomeric form and the dimeric (A-type) form. In the dimer, the CTT adopt the close conformation and cover the empty carotenoid channel. On the other hand, in the monomer the CTT is free to move and populate its three main conformations (open, close and extended). (2) One subunit of the dimer can interact with the membrane through the hydrophobic patch of the CTT, which by burrowing in the membrane can extract the carotenoid. (2') The apo-CTDH monomer also is able to interact with the membrane via the CTT, which facilitates carotenoid extraction. (3) In the monomer of the A-type dimer, which incorporated the carotenoid, the CTT is no longer required to cover the channel, now occupied by the carotenoid, and moves to the open position disrupting the dimer interaction and provokes the formation of the F-type dimer which by facing both channels stabilize the carotenoid binding. (3') In the case of the monomeric form, once the carotenoid was incorporated, the CTT moves to the open position as well. Interaction with a second monomeric CTDH lead to the formation of the F-type dimer, the main holo-form. (4) The holo-CTDH can interact with apo-HCP and transfer the carotenoid forming holo-HCP4 and apo-CTDH.

Up to now, only unrooted trees had been used to classify the subclades of OCPs. I used outgroups to root the OCP phylogenetic tree. Two different outgroups (I used independently both CTDH and NTF2) and three different approaches to infer the phylogenetic tree gave the same tree topology. This tree indicated that OCPX diverged first and represent an “evolutionary old” subclade of OCP. Resulting from a second divergence event, OCP1 and OCP2 are equally modern. OCPX is characterized by a fast back-conversion, a lack of FRP-interaction (similar to OCP2) and a strong propensity to dimerize (like OCP1). Surprisingly, the OCPX-like complex formed by interaction of the OCPX domains was not able of back-conversion after photoactivation. Since this is the ancestral OCP, this result suggested that the ancestors of OCP domains were not able to complete the photocycle. We hypothesize that gene fusion and linker addition to the OCP allowed to complete the photocycle by the back-conversion to the inactive state. If photoactivation rates of the OCP-like complex and OCP are compared, it can be concluded that the linker slows-down the photoactivation in OCP1 and OCPX. This is important to control the photoactivation at low light intensities. In OCP2 the flexible linker seems to enhance the photoactivation, probably to cope with a strong inter-domain interaction.

Specific amino acids, which are highly conserved, are very important for the role of the linker. Mutation in E174 and R185 in OCP1 from *Synechocystis* largely decreased the rate of back-conversion and increased the rate of photoactivation. This indicates that both amino acids are essential for the role of the linker. The structure of OCP1 provides a possible explanation of the importance of these amino acids (supplementary figure 13, chapter IV). In the structure E174 is found forming an H-bond with the N104 (from the NTD), which is interacting with W277 (from the CTD). The E174-N102 interaction could help to orient the domains during back-conversion and stabilize the OCP<sup>O</sup>. The R185 is interacting with N204 and D209 from the CTD, probably stabilizing the close conformation of the OCP. In addition to the composition of the linker, its length also seems to be important for its role. Deletion of 6 amino acids also strongly affects the back-conversion. Another effect of a shorter linker is the complete inhibition of the PBS quenching activity of the OCP. The current model of OCP-PBS interaction indicates that the NTD need to be buried into the PBS while the CTD remains outside of the PBS. In this interaction the length of the linker is important to allow this configuration. In addition to this, the specific prolines 175 and 179 (in OCP1) stabilize the OCP binding to the PBS.

Overall, the characterization of the three OCP subclades and the role of the linker supported the evolutionary model suggested by our rooted phylogenetic tree. If we assume that the main features of the first ancestral OCP are kept by the OCPX, we can easily observe how some of these traits were kept in the OCP2 (fast back-conversion) or in the OCP1 (negative regulation of the photoactivation by the linker and dimerization). In addition, the modern OCP1 and OCP2 have acquired new unique traits (OCP1 evolved to interact with the FRP and OCP2 stabilized the monomeric form). These slight changes on the different OCPs, could tune in different ways the activity of OCPs and improve them to respond to various stress situations. For example, since OCPX and OCP1 were shown to be expressed constitutively, the regulation of the linker could avoid unnecessary photoactivation and loss of the photosynthetic capacity at low light intensity. On the other hand, since OCP2 is expressed only in stress conditions (at least in *Tolypothrix*) this is not required.

## *Perspectives*

---

### *In vivo function of CTDH-HCP system*

In this thesis I presented the first characterization *in vitro* of the CTDH-HCP system. However, further experimentation is required to assess the exact function of the HCP proteins (at the moment there is information about only 4 out of 9 subclades of HCPs<sup>4</sup>). It was reported that some of them are up-regulated in stress conditions such as dehydration<sup>1</sup>. I anticipate that future studies of gene transcription or protein expression will shed light on the conditions in which the CTDH-HCP system is required. This, together with CTDH or HCP deletion mutants, will help to elucidate the role *in vivo* of these proteins and corroborate if it corresponds with the proposed one based on *in vitro* results. *In vitro*, most of studied HCPs can quench <sup>1</sup>O<sub>2</sub> with high efficiency and only one sub-clade can induce energy dissipation in the PBS. The <sup>1</sup>O<sub>2</sub> quenching can be studied by His-mediated chemical trapping<sup>5</sup> *in vivo*. An important question is whether HCP4 is involved in permanent PBS energy quenching. A proper model to study this is *Thermosynechococcus elongatus* since it contains one CTDH and one HCP4/5 but not OCP. First, the conditions in which these proteins (CTDH/HCP) are expressed have to be determined. Possible conditions could be high-light, low CO<sub>2</sub> or nutrient starvation. Then the quenching of the PBS in these conditions can be studied by time-resolved fluorescence spectroscopy (since the primary quenching site

and the rate of quenching should be similar to those determined for the OCP). *T. elongatus*  $\Delta$ HCP mutant must be constructed and used as control to confirm that the PBS quenching induced in those conditions (if exists) is related to the HCP activity.

Another question that raises from the multidirectional carotenoid transfer mechanism is whether the CTDH can have the same role that FRP in the strains containing OCPX, HCP and CTDH (like *Scytonema*). Since most of the OCPX-containing cyanobacteria are not model organisms, a previous step is to develop a method to genetically transform them. Although preliminary experiments *in vitro* indicated that the presence of CTDH does not affect the quenching capacity of OCP,  $\Delta$ CTDH mutant can be constructed in this kind of strain to investigate this potential role of the CTDH.

In addition, the working model of carotenoid transfer mechanism that we proposed need to be further studied. One approach that can be used is the construction of a CTDH mutant that stabilizes A- or F- type dimers or the monomer. Then, these mutants can be used to explore in detail the involvement of each one of these oligomeric states in the mechanism. Another interesting option is to use protein-crosslinking coupled to mass spectrometry to study the transient CTDH-HCP complex and to identify interacting residues.

### *Open questions about the OCP*

The most interesting feature of the OCP is its photocycle and the structural changes induced by light. As I described in the introduction of this thesis, some advances have been made by using time-resolved spectroscopy. These experiments pointed out the existence of red intermediary states during photoactivation and orange intermediary states during the back-conversion. It will be very interesting to know the structural features of those intermediary states which will contribute to the understanding of the OCP photocycle. One approach to study this could be to construct mutants in which the OCP is possibly locked in one these states. The amino acids involved in the carotenoid channel, in which the carotenoid is translocated during photoactivation, could be potential target residues for these mutations. In addition, serial femtosecond crystallography can shed light on the first events during photoactivation and time-resolved SAXS can be used to study the global structural changes during photocycle. However, the most important tridimensional structure that is missing is the structure of the OCP<sup>R</sup>, which is

hard to crystalize probably due to its high flexibility. This structure is not only very important for the study of the photocycle, but it is also important to understand how the OCP binds to the PBS.

Although the spectral properties of the carotenoid bound to the OCP indicates that it is capable of direct quenching of the PBS, until now there is no evidence of that process. The alternative option is that after OCP binding to the PBS, structural changes in the PBS induce energy dissipation at the level of bilins. One experiment that could help to solve this open question is femtosecond transient absorption. In this technique, it is possible to follow the energy transfer and the spectral signature of the specie that is populated by measuring the fast changes in the absorption spectrum after excitation of the sample with a short laser pulse. Selective excitation of the PBS in PBS-OCP (or PBS-NTD) complexes could show (if the carotenoid is involved in the quenching) that the energy is transferred to the carotenoid.

So far, little attention has been paid to the role of the oligomeric state of the OCP on its (photo) activity. It was reported that the OCP monomerizes upon photoactivation<sup>6</sup>. Thus, it is possible that the equilibrium between monomer and dimer is involved in an additional regulation. In agreement with this, dimerization seems to be conserved in OCP1 and OCPX that are constitutively expressed and require fine regulation to be photoactivated only at high light intensity. I constructed an OCP1 mutant which is stable as monomer; it does not dimerize. This mutant can be used in photoactivation experiments to investigate if the dimerization of OCP can regulate its activity. In addition to this, OCP can be labeled with fluorescent dyes and the dimerization can be studied by FRET (Förster resonance energy transfer) that is usually used to measure distances. The dyes will be attached to cysteines in the OCP, thus a previous requirement is to mutate all the Cys in the OCP and introduce two new Cys in positions in which FRET is only possible in the dimeric form. If the label is a fluorescent protein genetically fused to the OCP this experiment can be performed *in vivo* (however, photosynthetic cells are not the best models for these experiments due to the high autofluorescence). In this case, *Synechocystis* can be engineered to overexpress two different OCPs, one labeled with GFP and the other one with mCherry (two fluorescent proteins usually used for FRET essays).

Another question about an enigmatic feature of the OCP<sup>0</sup> that have received scant attention is its heterogeneity in solution. It has been observed, by Raman and pump-probe spectroscopy, the presence of two OCP<sup>0</sup> populations<sup>7</sup>. This was initially explained as two different positions of the

carotenoid. In this case, single molecule spectroscopy measurements (SMS) are suitable to evaluate the different properties of these population. For example, by labeling with fluorescent dyes in a way in which it allows to follow the photoactivation by FRET<sup>8</sup>, SMS can provide information about the kinetics of photoactivation in both populations. Additionally, it can be investigated if the photoactivation of these two populations have the same light dependence. In a different approach, the OCP can be labeled with acceptor/donor pairs in different parts and structural information can be obtained about the two different population (in the inactive and active states).

The last open question about the OCP that I would like to discuss here is related to the chapter IV of this thesis. It was reported that the flexible linker of the *Synechocystis* OCP can be post-transcriptionally modified by phosphorylation<sup>9</sup> (also: Natalia Battchikova, personal communication). The flexible linker of the OCP is intrinsically disordered. Disordered proteins (or domains) are known to be regulated by phosphorylation<sup>10</sup>. For example, phosphorylation can stabilize or destabilize secondary structures like  $\alpha$ -helices inducing conformational changes. Thus, it is possible that the phosphorylation of the linker induces conformational changes that affect the regulatory role of the linker. Phosphomimic mutations can be used to assess *in vitro* and *in vivo* the possible role of the OCP phosphorylation. In addition, it would be also important to find in which conditions this OCP modification takes place.

### *Application of the CTDH-HCP and OCP systems*

A potential application of the OCP is in the topic of optogenetics. Optogenetics is related to the use of light to control genetic expression. Here, the modularity and photoactivity of the OCP can be exploited to generate new light-controlled molecular switches of gene expression. For example, the isolated domains of the OCP2 can be used. In this case, the CTD2 can be fused to a DNA-binding domain in order to target the system to specific genes. The NTD2 can be fused to a transcriptional factor that acts as repressor of gene expression. Thus, if the cell also produces carotenoids, in darkness the CTD2 and NTD2 will form the OCP2-like complex and through the DNA-binding domain it will bind the promotor of specific genes. The action of the fused transcriptional repressor will inhibit the expression of the gene target. Illumination of the cells will lead to photoactivation of the OCP2-like complex, domain separation and re-activation of the gene target expression (by detachment of the repressor). The election of the OCP2 domains for this

system allows to switch off again the gene expression by OCP2-like complex reassembling in darkness. In addition, this system can be used not only to control gene expression but also to control the activity of other different kind of enzymes.

Our large knowledge about OCP and OCP-related NPQ mechanism could also serve to construct cyanobacteria mutants producing a higher biomass. Due to the increase of atmospheric CO<sub>2</sub> concentrations and the global warming, humanity is facing one of the most dangerous climate emergencies. An imperative requirement to overcome this situation is the transition to clean and renewable energies. One option, the most pertinent to this thesis, is the biofuel. This fuel can be produced directly from biomass. There are different classifications of biofuel depending on the biomass source but, here, I will focus on the biofuel produced from microalgae or cyanobacteria. The use of cyanobacteria presents extra benefits compared with plant biomass sources: it allows the lands to be used for crops and it grows faster. In general, biofuel production in plants, algae and cyanobacteria has another advantage: photosynthetic organisms consume atmospheric CO<sub>2</sub>. Nevertheless, the growth conditions of cyanobacteria are expensive (photobioreactors, water supply) compared with the yield of biofuel produced from the biomass. Thus, the cultivation of cyanobacteria needs to be improved. Different strategies can be used to increase the CO<sub>2</sub> assimilation by cyanobacteria and the biofuel production. Here, I focus only on how to increase the biomass production by overcoming the photosynthetic limitation at high light intensities. Due to ROS production, photoinhibition and the activation of photoprotective mechanisms, the photosynthetic efficiency reaches a plateau at high light and does not increase further. It was shown that in *Tobacco* plants<sup>11</sup>, the genetic modification of the NPQ mechanism that accelerates the carotenoid interconversion in the xanthophyll cycle leads to an increase of 15% in biomass production in natural field. In this mutant, the relaxation of the NPQ is faster than in the WT. Depending on the mixing conditions of the bioreactor used to grow the cyanobacteria, the cells are exposed to fluctuating light conditions. Thus, it is possible that cyanobacteria bioengineered to induce fast relaxations of the OCP-related NPQ will increase the growth yield. One approach to achieve this could be introducing mutations in the OCP, which alter the PBS binding inducing a faster detachment in darkness and relaxation of the NPQ. Good candidates for these mutations are the R155<sup>12</sup> and the P175/P179 that were shown to be involved in PBS-OCP stabilization in the chapter IV. Another approach could be to replace the OCP1 of cyanobacteria used for biomass productions (like *Synechocystis* PCC 6803) by OCPX,

which I showed in chapter IV to have a faster back-conversion to the inactive state and yet conserve the fine regulation of the photoactivation like OCP1. Moreover, amino acids identified to be involved in FRP interaction could be introduced in OCPX in order to generate an OCPX, in which the back-conversion could be accelerated even more by the FRP.

## *Conclusions*

---

In conclusion, the aim of this thesis was to study the role of the CTDH family and its interaction with the HCP. This study has identified a **novel carotenoid transfer mechanism between CTDH and HCP**. One of the more significant findings is that **the main role of CTDH is to transfer the carotenoid to HCP** and ensure the formation of holo-HCP. The C-terminal tail of the CTDH plays an important role in this mechanism by facilitating the movement of the carotenoid. The CTDH and HCP are paralogs of the OCP domains that have evolved in parallel to them. Modern OCPs were originated by gene fusion of ancestors of CTDH/CTD and NTD/HCP. The analysis of the role of the linker connecting the OCP domains, presented in this thesis, revealed that **the addition of a linker provided the first OCP the possibility of completing the photocycle**. In addition, it provided a more regulated photoactivation, which is different depending on the OCP subfamily. This result complemented with biochemical characterization of the different OCP subclades and a bioinformatic analysis provided insight for the evolution of the OCP. In our model, the first OCP diverged into OCPX and the precursor of OCP1 and OCP2, which have emerged later in the evolution.

## *References*

---

1. Wen, Y. et al. Orange and red carotenoid proteins are involved in the adaptation of the terrestrial cyanobacterium *Nostoc flagelliforme* to desiccation. *Photosynth. Res.* 1, 1–11 (2019).
2. Moldenhauer, M. et al. Assembly of photoactive orange carotenoid protein from its domains unravels a carotenoid shuttle mechanism. *Photosynth. Res.* 133, 327–341 (2017).
3. Lechno-Yossef, S., Melnicki, M. R., Bao, H., Montgomery, B. L. & Kerfeld, C. A. Synthetic OCP heterodimers are photoactive and recapitulate the fusion of two primitive carotenoproteins in the evolution of cyanobacterial photoprotection. *Plant J.* 91, 646–656 (2017).

4. López-Igual, R. *et al.* Different Functions of the Paralogs to the N-Terminal Domain of the Orange Carotenoid Protein in the Cyanobacterium *Anabaena* sp. PCC 7120. *Plant Physiol.* **171**, 1852–1866 (2016).
5. Rehman, A. U., Cser, K., Sass, L. & Vass, I. Characterization of singlet oxygen production and its involvement in photodamage of Photosystem II in the cyanobacterium *Synechocystis* PCC 6803 by histidine-mediated chemical trapping. *Biochim. Biophys. Acta - Bioenerg.* **1827**, 689–698 (2013).
6. Zhang, H. *et al.* Molecular mechanism of photoactivation and structural location of the cyanobacterial orange carotenoid protein. *Biochemistry* **53**, 13–19 (2014).
7. Slouf, V. *et al.* Ultrafast spectroscopy tracks carotenoid configurations in the orange and red carotenoid proteins from cyanobacteria. *Photosynth. Res.* 105–117 (2017).
8. Maksimov, E. G. *et al.* Fluorescent Labeling Preserving OCP Photoactivity Reveals Its Reorganization during the Photocycle. *Biophys. J.* **112**, 827 (2017).
9. Chen, Z. *et al.* Effects of Phosphorylation of  $\beta$  Subunits of Phycocyanins on State Transition in the Model Cyanobacterium *Synechocystis* sp. PCC 6803. *Plant Cell Physiol.* **56**, 1997–2013 (2015).
10. Bah, A. & Forman-Kay, J. D. Modulation of Intrinsically Disordered Protein Function by Post-translational Modifications. *J. Biol. Chem.* **291**, 6696–705 (2016).
11. Kromdijk, J. *et al.* Improving photosynthesis and crop productivity by accelerating recovery from photoprotection. *Science* **354**, 857–861 (2016).
12. Adjélé Wilson, Michal Gwizdala, Alberto Mezzetti, Maxime Alexandre, Cheryl A. Kerfeld, D. K. The Essential Role of the N-Terminal Domain of the Orange Carotenoid Protein in Cyanobacterial Photoprotection: Importance of a Positive Charge for Phycobilisome Binding. *Plant Cell* **24**, 1972–1983 (2012).

## Résumé en français

---

### *Pour information*

*This is a translated version of the general discussion of this thesis.*

Au cours des 13 dernières années, le mécanisme d'*OCP-related NPQ* a été largement caractérisé. L'OCP est photoactivé en réponse à de fortes intensités lumineuses. Au cours de cette photoactivation, l'OCP subit des changements de conformation globaux qui lui permettent de rejoindre l'antenne complexe de la cyanobactérie (PBS). De cette manière, l'OCP est capable d'induire une dissipation d'énergie, d'éviter une réduction excessive de la chaîne photosynthétique et la production d'espèces réactives de l'oxygène (ROS).

Nonobstant, le rôle et mécanisme des paralogues de l'OCP, CTDH et HCP, était moins clair lorsque j'ai commencé ma thèse. À ce moment-là, nous savions que HPC4 pouvait induire une atténuation de la fluorescence du PBS et que HCP2 et HCP3 étaient de bons inhibiteurs de l'oxygène singulet (*in vitro*). Ces résultats suggèrent que les HCP sont impliqués dans les mécanismes de photoprotection dans les cellules de cyanobactéries. De l'autre côté, nous ne savions presque rien sur l'CTDH. CTDH et HCP4 étant généralement situés dans des *loci* adjacents du génome, il a été proposé de former un hétérodimère présentant des caractéristiques similaires à celles de l'OCP, ressemblant à l'interaction entre les ancêtres des domaines d'OCP. De plus, il a été supposé que la fusion de leurs ancêtres a créé les OCP modernes. Ce travail de thèse a été entrepris pour étudier la CTDH non caractérisée et son interaction avec le HCP. L'une des principales conclusions présentées dans cette thèse est que les CTDH sont des protéines caroténoïdes transitoires impliquées dans le transfert du caroténoïde vers d'autres protéines. Ensuite, j'ai exploré les différents *subclades* d'OCP en étudiant le rôle des interactions de domaine et de l'éditeur de liens flexibles sur l'évolution des OCP. Dans l'ensemble, tous nos résultats ont mis en lumière l'évolution et les différents rôles des protéines caroténoïdes dans les cyanobactéries.

## ***Rôle des protéines CTDH et HCP dans les cyanobactéries***

---

Les paralogues du domaine CTD de l'OCP, appelés CTDH, sont capables de se lier au caroténoïde. La liaison des caroténoïdes stabilise la forme dimère de la CTDH dans laquelle le caroténoïde est partagé par les deux monomères. La faible quantité d'holo-CTDH isolée à partir de cellules de *Synechocystis* produisant de la canthaxanthine suggère que les CTDH sont des protéines caroténoïdes transitoires. Les résultats de ma thèse suggèrent fortement que le rôle principal des CTDH est la délivrance de caroténoïdes aux HCPs. En outre, l'holo-CTDH pourrait jouer un rôle supplémentaire important sur le *quenching* du  $^1\text{O}_2$ . Les principaux résultats expérimentaux à l'origine de cette proposition sont le transfert de caroténoïde de l'holo-CTDH à l'apo-HCP et le fait que seules les apo-CTDH sont capables d'absorber les caroténoïdes des membranes. D'autres résultats sont fournis: 1) lorsqu'ils sont isolés d'*E. coli* producteur de caroténoïde, les HCP se lient moins au caroténoïde que les CTDH (en l'absence de tout accepteur de caroténoïde, la CTDH reste dans l'holo-forme) en raison de sa plus faible capacité d'absorption des caroténoïdes par les membranes; 2) lorsqu'ils sont co-exprimés ou lorsque leurs apo-formes sont mélangées avec des membranes contenant des caroténoïdes, la quantité de holo-HCP est augmentée (résultats non publiés), ce qui confirme l'exigence de CTDH en tant que fournisseur de caroténoïde; 3) Dans tous les génomes contenant les gènes *hcp*, un gène *ctdh* est également présent. Nonobstant, le fait que certains holo-HCP puissent être isolés d'*E. coli* ou de cyanobactéries suggèrent que, lors du repliement des protéines, des caroténoïdes peuvent être incorporés à ces protéines mais avec une faible efficacité.

### ***Translocation des caroténoïdes: CTD-NTD vs CTDH-HCP***

Dans le passé, il a été démontré que la translocation des caroténoïdes joue un rôle essentiel dans le photocycle et l'activité de l'OCP. Dans l'obscurité, le caroténoïde est partagé entre les deux domaines, CTD et NTD. Cependant, lors de la photoactivation OCP (ou d'OCP-like), le caroténoïde est transféré au NTD. Ceci aboutit à la formation de l'OCP<sup>R</sup> actif pouvant induire la dissipation d'énergie du PBS. Dans cette thèse, je montre que la translocation des caroténoïdes joue également un rôle important dans l'activité des HCP: les HCP reçoivent le caroténoïde, qui est essentiel pour leur activité à partir d'une autre protéine: la CTDH. L'étude parallèle du mécanisme de transfert

des caroténoïdes entre les domaines OCP libres isolés et entre les HCP et les CTDH nous permet de conclure que la translocation des caroténoïdes et sa direction sont déterminées par l'affinité caroténoïde de chaque composant et par la force de ses interactions. Nonobstant, il existe des différences importantes entre la translocation des caroténoïdes et l'interaction des protéines entre les couples CTDH / HCP et OCP-CTD / OCP-NTD.

Tout d'abord, l'interaction entre les domaines isolés de l'OCP résulte en un complexe stable avec des caractéristiques similaires à l'OCP (lorsque l'holo-CTD interagit avec l'apo-NTD). D'autre part, l'interaction entre l'holo-CTDH et l'apo-HCP est transitoire et aucun complexe n'est formé. Ceci peut s'expliquer par une interaction protéique plus forte entre le CTD et le NTD qu'entre le CTDH et le HCP. Néanmoins, l'interaction avec le caroténoïde pourrait également affecter la formation du complexe. Par exemple, si la différence d'affinité pour le caroténoïde des deux domaines de l'OCP est inférieure à celle entre CTDH et HCP. Cet équilibre pourrait permettre au caroténoïde de se placer dans une position intermédiaire entre les deux domaines et de stabiliser le complexe.

Deuxièmement, pendant l'interaction CTDH-HCP, le caroténoïde est complètement transféré au HCP. Cette translocation caroténoïde est spontanée et indépendante de la lumière, alors que dans l'OCP (et l'OCP-like), la translocation du caroténoïde vers la NTD n'a lieu que sous fort éclairage (pendant la photoactivation). La base moléculaire derrière la translocation des caroténoïdes induite par la lumière dans l'OCP est inconnue. De même, on ignore pourquoi le transfert de caroténoïde dans le CTDH-HCP est indépendant de la lumière. Une explication possible de cette différence est que le canal caroténoïde formé dans le complexe CTD-NTD est différent de celui formé dans le complexe transitoire CTDH-HCP. Dans le premier cas, il oblige le caroténoïde à se trouver dans une position courbée et hautement énergétique entre les deux domaines. Dans le second cas, il n'y a pas de restriction structurelle et le caroténoïde peut se déplacer librement vers le HCP. Certains facteurs peuvent contribuer à cette différence: **I-** il est possible que le NTD de l'OCP adopte deux conformations différentes pour lier le caroténoïde: l'un stabilise le caroténoïde dans l'OCP<sup>0</sup> et l'autre permet l'incorporation complète du caroténoïde dans les NTD pendant la photoactivation. Dans ce cas, le HCP pourrait adopter uniquement cette dernière conformation et provoquer le transfert direct de caroténoïde sans formation de complexe stable. **II-** Comme il a été mentionné précédemment, l'interaction des protéines pourrait être plus forte dans le couple CTD-NTD et pourrait stabiliser suffisamment le complexe pour permettre le

positionnement correct du caroténoïde entre les deux protéines. Dans le cas de CTDH-HCP, si l'interaction est plus faible, le complexe n'est pas stable et le canal caroténoïde n'est pas correctement formé et la caroténoïde transloque directement.

### *Transfert de caroténoïde multidirectionnel*

Dans cette thèse, il est démontré qu'AnaCTDH (dans des conditions réductrices) peut transférer le caroténoïde à HCP2, HCP3 et HCP4 et même à TeHCP4 sans le récupérer. Il est intéressant de noter que TeCTDH pourrait administrer le caroténoïde uniquement à TeHCP. Cette haute spécificité de TeCTDH peut s'expliquer par le fait qu'il n'y a qu'un seul HCP dans cette souche et que les deux TeCTDH et TeHCP pourraient avoir co-évolué pour améliorer leur interaction. Par ailleurs, à *Anabaena*, il existe quatre HCPs et très probablement l'AnaCTDH a évolué pour pouvoir interagir avec différents HCPs.

Ni holo-AnaCTDH ni holo-TeCTDH ne peuvent transférer le caroténoïde à l'apo-NTD isolé. Cependant, AnaCTDH est en mesure de transférer partiellement le caroténoïde à l'OCP. Ce résultat inattendu suggère fortement que le transfert de caroténoïde s'effectue de CTDH au domaine CTD de l'OCP. En accord avec cela, CTDH peut transférer le caroténoïde en CTD isolé. En outre, AnaCTDH peut extraire le caroténoïde de l'OCP<sup>R</sup>, dans lequel le caroténoïde se trouve dans le NTD. Ce résultat a montré qu'il existe également un transfert «inverse» de caroténoïde des NTD vers les CTDH. En accord avec cela, un NTD isolé peut transférer le caroténoïde à l'apo-CTDH. Il a été suggéré que ce transfert «inverse» de caroténoïde pourrait jouer le rôle de FRP en désactivant l'OCP<sup>R</sup> dans les souches dépourvues de FRP (ces souches contiennent généralement de l'OCPX). L'interaction entre apo-CTDH et OCP<sup>R</sup> entraînerait un transfert de caroténoïde vers CTDH et la formation d'apo-OCP, qui est inactif. Ensuite, l'holo-OCP pourrait être régénéré dans l'obscurité par transfert de caroténoïde du même CTDH vers l'apo-OCP. Il est important de noter qu'il s'agit simplement d'une observation *in vitro* et que de nouvelles expériences *in vivo* seront nécessaires pour prouver cette hypothèse.

Enfin, l'AnaCTDH peut récupérer le caroténoïde de l'énigmatique holo-HCP1. Malgré le fait que cette protéine est exprimée *in vivo*, son rôle n'est pas clair car HCP1 ne peut pas *quencher* ni le <sup>1</sup>O<sub>2</sub> ni la fluorescence du PBS. Bien que nos résultats suggèrent que HCP1 pourrait être un transporteur du

caroténoïde, il ne s'agit que d'une hypothèse et de nouvelles *expériences in vivo* sont nécessaires pour mieux comprendre cette protéine. Dans le cas de l'interaction CTDH-HCP1, la translocation des caroténoïdes peut avoir lieu dans les deux sens et l'état d'équilibre est déterminé par l'équilibre de ces réactions. Il est important de noter que le HCP1 nécessite toujours la présence de CTDH pour récupérer le caroténoïde de la membrane.

Globalement, les paralogues des domaines OCP sont capables de multiples interactions entre eux, ce qui conduit à un schéma très complexe de transfert de caroténoïde. La translocation des caroténoïdes parmi ces protéines peut avoir des implications profondes dans la photoprotection des cyanobactéries. Dans ce modèle hypothétique, la CTDH peut servir à la fois de neutralisateur de  $^1\text{O}_2$  et de donneur de caroténoïde aux HCP dans la cellule. Dans ce cas, le transfert de caroténoïde de CTDH à HCP garantirait la présence d'holo-HCP qui peut également neutraliser  $^1\text{O}_2$  et, dans des cas plus extrêmes, induire une dissipation de l'excès d'énergie dans le PBS.

### *Mécanisme de transfert de caroténoïde dans les protéines apparentées à l'OCP*

Les bases structurelles et mécanistes de l'interaction CTDH-HCP et du transfert de caroténoïde associé ne sont pas faciles à déchiffrer, cette interaction étant transitoire. Néanmoins, j'ai montré dans ce travail qu'en résolvant la structure tridimensionnelle d'AnaCTDH (apo) (en collaboration avec Dvir Harris), certains aspects du transfert de caroténoïde peuvent être résolus. L'apoforme de la CTDH peut former deux types de dimères (dimère de type A et dimère de type F). Bien que la structure de l'holoforme n'ait pas été résolue, il est clair qu'il s'agit d'un dimère se liant à une seule molécule de caroténoïde. Cette forme dimérique est probablement le dimère de type F, car dans ce cas les canaux caroténoïdes de chaque monomère se font face, et le caroténoïde peut être complètement enterré entre eux. Ainsi, la formation du dimère de type F est stabilisée par liaison de caroténoïde.

L'holo-AnaCTDH oxydée est incapable de transférer le caroténoïde. Dans cette CTDH, ainsi que dans toutes les CTDH appartenant au sous-groupe CTDH2, il existe une cysteine conservée à 100% dans la surface de dimérisation de type (F). Cette Cys peut former une liaison disulfure qui empêche la monomérisation et le transfert de caroténoïde. Cela confirme le type F en tant que forme dimérique principale de l'holo-CTDH et démontre que la monomérisation de la CTDH est nécessaire pour la translocation du

caroténoïde en HCP. De plus, la présence du Cys pourrait permettre une régulation supplémentaire du système dans les souches contenant CTDH2. L'état redox cellulaire local pourrait déterminer la formation de la liaison disulfure et réguler le transfert de caroténoïde de la CTDH au HCP. Toutefois, il ne s'agit que d'une hypothèse et de nouvelles expérimentations *in vivo* seront nécessaires pour la confirmer.

La caractéristique la plus remarquable de la structure de l'apo-CTDH est la grande mobilité du CTT, qui peut peupler deux conformations flexibles principales de l'apoforme (n° 1 et n° 2 de la figure 5A, chapitre III). Dans le holoforme, le CTT adopte une position ouverte en interaction avec la feuille  $\beta$ . De cette façon, le canal caroténoïde est débloqué (comme il est observé dans la structure de l'OCP). Ces différences entre les CTDH holo- et apo ont suggéré un rôle du CTT dans la capture de caroténoïdes. L'analyse de mutation a montré que la suppression du CTT inhibe le transfert et l'absorption des caroténoïdes. Le patch hydrophobe à la fin du CTT peut aider à extraire le caroténoïde de la membrane et à le transloquer lors de l'interaction HCP. La mutation de L137D qui est à la fin du patch hydrophobe a fortement diminué le *rate* d'absorption / transfert de caroténoïde, confirmant ainsi son rôle dans la facilitation du mouvement du caroténoïde. Les simulations de MD suggèrent en outre que L137 participe à l'ancrage de la CTT aux  $\beta$ -sheets lors des transitions de dimères de type A à type F. Les simulations de MD ont montré que le L137 pouvait interagir avec la partie hydrophobe de la  $\beta$ -sheets en la position ouverte du CTT. Le remplacement de L137 par un résidu chargé a déstabilisé cette interaction et accru la flexibilité du CTT. En accord avec cela, le volume du monomère CTDH-L137D est supérieur à celui du WT. Ainsi, le CTT participe également à la capture de caroténoïde en favorisant la commutation entre les dimères de type A et de type F. D'autre part, il a été constaté que L128 stabilisait le dimère type A, probablement en stabilisant la conformation proche du CTT. À cet effet, l'amplitude de l'absorption de caroténoïde (par les membranes ou HCP1) a été réduite lorsque ce résidu a été remplacé par de l'aspartate.

### ***Rôle de le linker d'OCP et évolution de l'OCP***

La présence de gènes *ctdh* et *hcp* dans les mêmes génomes contenant OCP a incité les chercheurs à suggérer que le gène *ocp* est le résultat d'une fusion de gènes entre les gènes codant pour les ancêtres des domaines OCP. Les domaines isolés des OCP modernes peuvent interagir et former un complexe OCP-like, comme il a été montré aux chapitres I et IV et dans les publications

précédentes. Dans le cas des domaines de OCP2, ce complexe de type OCP peut être photoactivés puis se relâchent à l'état initial tandis que les complexes de type OCP1 peuvent être photoactivés de manière irréversible. Cette observation, ainsi que les différences biochimiques entre OCP1 et OCP2, ont été à la base de la proposition selon laquelle OCP2 représente un OCP ancestral et que leurs domaines ressemblent aux caractéristiques des ancêtres de l'OCP. Dans ce modèle, le *linker* ne joue aucun rôle ou presque pas dans les mécanismes de l'OCP et il n'y a aucune explication explicite de la raison pour laquelle l'OCP a évolué en tant que protéine unique au lieu de deux protéines en interaction et de l'avantage évolutif que cela représentait. Nos résultats du chapitre IV ont largement contribué à la reformulation de ce paradigme.

Il a d'abord été suggéré que l'OCP avait été créé par un événement de fusion de gènes d'ancêtres de CTDH / HCP4. Une autre explication alternative est que *ctdh* et *hcp4* sont les produits de la fission du gène *ocp*. Cela pourrait être une étape intermédiaire lors de la suppression génétique d'*ocp*. Dans les deux situations, le principal événement ayant conduit à la présence des deux (*ctdh* / *hcp* et *ocp*) dans le même génome était la duplication de gènes (avant l'événement de fusion ou de fission de gènes). Cela définit *ctdh* et *hcp* comme des paralogues des domaines *ocp*. En accord avec les résultats présentés sur cette thèse, les paralogues ont généralement des fonctions différentes.

Jusqu'à présent, seuls les arbres phylogénétique non racinés avaient été utilisés pour classer les *subclades* des OCP. J'ai utilisé des groupes externes pour enracer l'arbre phylogénétique OCP. Deux groupes externes différents (j'ai utilisé indépendamment CTDH et NTF2) et trois approches différentes pour déduire l'arbre phylogénétique donnaient la même topologie d'arbre. Cet arbre indique que OCPX a divergé en premier et représente un *subclade* «plus ancien évolutif» d'OCP. Issue d'un deuxième événement de divergence, OCP1 et OCP2 sont également modernes. OCPX se caractérise par une conversion arrière rapide, un manque d'interaction FRP (similaire à OCP2) et une forte propension à dimériser (comme OCP1). Remarquablement, le complexe d'OCPX-like formé par l'interaction des domaines OCPX n'était pas capable de conversion arrière après la photoactivation. S'agissant d'un OCP ancestral, ce résultat suggère que les ancêtres des domaines OCP n'étaient pas en mesure de compléter le photocycle. Nous émettons l'hypothèse que la fusion de gènes et l'addition de *linker* à l'OCP permettaient de compléter le photocycle par conversion en retour à l'état inactif. Si l'on compare les taux de photoactivation du complexe de type OCP et d'OCP, on peut en conclure que le *linker* ralentit la photoactivation dans OCP1 et OCPX.

Ceci est important pour contrôler la photoactivation à de faibles intensités lumineuses. Dans OCP2, le *linker* semble améliorer la photoactivation, probablement pour faire face à une interaction forte entre les domaines.

Les acides aminés spécifiques, qui sont hautement conservés, sont très importants pour le rôle du linker. La mutation dans E174 et R185 dans OCP1 de *Synechocystis* a largement diminué le *rate* de désactivation d'OCP et augmenté le *rate* de photoactivation. Cela indique que les deux acides aminés sont essentiels pour le rôle du linker. La structure d'OCP1 fournit une explication possible de l'importance de ces acides aminés (Figure supplémentaire 13, chapitre IV). Dans la structure, E174 forme une liaison d'hydrogène avec le N104 (provenant du NTD), qui interagit avec le W277 (CTD). L'interaction E174-N102 pourrait aider à orienter les domaines lors de la reconversion et à stabiliser l'OCP<sup>0</sup>. Le R185 interagit avec N204 et D209 du CTD, stabilisant probablement la conformation inactive de l'OCP. Outre la composition du *linker*, sa longueur semble également être importante pour son rôle. La suppression de 6 acides aminés affecte également fortement la désactivation. Un autre effet d'un *linker* plus court est l'inhibition complète de l'activité de *quencher* du PBS de l'OCP. Le modèle actuel d'interaction OCP-PBS indique que les NTD doivent être enfouis dans le PBS tandis que le CTD reste en dehors de celui-ci. Dans cette interaction, la longueur de la *linker* est importante pour permettre cette configuration. De plus, les prolines spécifiques (175 et 179 dans OCP1) stabilisent la liaison de l'OCP au PBS.

Globalement, la caractérisation des trois sous-clades OCP et le rôle du linker ont conforté le modèle évolutif suggéré par notre arbre phylogénétique enraciné. Si nous supposons que les caractéristiques principales du premier OCP ancestral sont conservées par OCPX, nous pouvons facilement observer comment certaines de ces caractéristiques ont été conservées dans l'OCP2 (conversion rapide en retour) ou dans l'OCP1 (régulation négative de la photoactivation par le *linker* et dimérisation). De plus, les OCP1 et OCP2 modernes ont acquis de nouveaux *traits* uniques (OCP1 a évolué pour interagir avec le FRP et OCP2 a stabilisé la forme monomère). Ces légers changements sur les différents OCP pourraient accorder de différentes manières l'activité des OCP et les améliorer pour répondre à diverses situations de stress. Par exemple, comme il a été prouvé qu'OCPX et OCP1 sont exprimés de manière constitutive, la régulation du *linker* pourrait éviter une photoactivation inutile et une perte de la capacité photosynthétique à faible intensité lumineuse. D'autre part, puisque OCP2 n'est exprimé que dans des conditions de stress (du moins chez *Tolypothrix*), cela n'est pas obligatoire.

## Conclusions

En conclusion, le objectif de cette thèse était d'étudier le rôle de la famille CTDH et son interaction avec le HCP. Cette étude a identifié un nouveau mécanisme de transfert de caroténoïde entre CTDH et HCP. L'une des découvertes les plus significatives est que le rôle principal de CTDH est de transférer le caroténoïde en HCP et d'assurer la formation d'holo-HCP. La queue C-terminale de la CTDH joue un rôle important dans ce mécanisme en facilitant le mouvement du caroténoïde. CTDH et HCP sont des paralogues des domaines OCP qui ont évolué en parallèle. Les OCP modernes ont été créés par la fusion de gènes d'ancêtres CTDH / CTD et NTD / HCP. L'analyse du rôle du *linker* reliant les domaines OCP, présentée dans cette thèse, a révélé que l'ajout d'un *linker* offrait au premier OCP la possibilité de compléter le photocycle. En outre, il fournissait une photoactivation plus régulée, qui varie en fonction du *subclade* d'OCP. Ce résultat, complété par une caractérisation biochimique des différents *subclades* de l'OCP et une analyse bioinformatique ont permis de mieux comprendre l'évolution de l'OCP. Dans notre modèle, le premier OCP a divergé entre OCPX et le précurseur d'OCP1 et OCP2, apparus plus tard dans l'évolution.

**Titre:** Translocation des caroténoïdes et évolution des protéines dans la photoprotection des cyanobactéries.

**Mots clés :** cyanobactéries, évolution des protéines, photoprotection, caroténoïde, photosynthèse

**Abstract :** Les cyanobactéries sont des organismes photosynthétiques capables de convertir le CO<sub>2</sub> en composés organiques et de produire de l'oxygène en utilisant l'énergie lumineuse. Néanmoins, de fortes intensités lumineuses saturent l'appareil photosynthétique, ce qui conduit à la production d'espèces réactives de l'oxygène, dangereuses pour la cellule. Pour y faire face, la photoactive *orange carotenoid protein* (OCP) induit une dissipation thermique de l'énergie excédentaire récoltée par le complexe d'antennes, le phycobilisome (PBS), afin de diminuer l'énergie arrivant aux centres photochimiques. L'OCP est composé de deux domaines, le domaine C-terminal (CTD) et le domaine N-terminal (NTD), reliés par un domaine de liaison flexible (*linker*). Pendant la photoactivation, le caroténoïde est transféré vers le NTD, les domaines se séparent et le NTD peut interagir avec le PBS. Trois familles d'OCP coexistent (OCPX, OCP1 et OCP2) dans les cyanobactéries modernes. Outre l'OCP, de nombreuses cyanobactéries contiennent également des homologues des domaines OCP, le CTDH et HCP. Les HCP sont une famille de protéines caroténoïdes présentant différents traits photoprotecteurs. La plupart d'entre eux sont de très bons quenchers d'oxygène singulet, et un *subclade* est capable d'interagir avec le PBS et d'induire une dissipation de l'énergie thermique comme l'OCP. Le rôle de CTDH était inconnu. La présence de ces homologues parallèlement à l'OCP a conforté l'idée générale que l'OCP a une origine évolutive modulaire et que la CTDH et HCP pourraient interagir pour former un complexe *OCP-like* ayant des caractéristiques et une fonction similaires à celles de l'OCP.

Dans cette thèse, je présente la première caractérisation des protéines CTDH. Les CTDH sont des dimères se liant à une molécule de caroténoïde. Le rôle principal de la CTDH est de transférer son caroténoïde au HCP. De plus, les CTDH sont capables de récupérer les caroténoïdes des membranes contrairement à les HCP. Ces résultats suggèrent fortement que les CTDH sont des transporteurs de caroténoïde qui assurent le chargement en caroténoïde sur les HCP. Ce nouveau mécanisme de translocation des caroténoïdes pourrait être multidirectionnel. La résolution de deux structures tridimensionnelles de l'ApoCTDH d'*Anabaena* a montré que la queue C-terminale du CTDH (CTT) peut adopter différentes conformations. De plus, l'analyse de mutation a démontré que le CTT joue un rôle essentiel dans la translocation des caroténoïdes. Enfin, je rapporte une caractérisation moléculaire du *linker* reliant les domaines de différents OCP modernes et son rôle au cours de l'évolution de l'OCP. Tout d'abord, j'ai caractérisé les OCP des trois *subclades*, y compris l'OCPX non caractérisé. OCPX et OCP2 présentent une désactivation rapide par rapport à OCP1. Alors que OCP1 et OCPX peuvent dimériser, OCP2 est stable en tant que monomère. Enfin, j'ai constaté que le *linker* est essentiel pour la désactivation de l'OCP et qu'il régule la photoactivation. Dans OCP1 et OCPX, le *linker* ralentit la photoactivation, tandis que dans OCP2, il augmente le taux de photoactivation. L'analyse bioinformatique complète cette caractérisation et fournit une image claire de l'évolution de l'OCP pour répondre efficacement aux conditions de stress.

**Title:** Carotenoid translocation and protein evolution in cyanobacterial photoprotection.

**Keywords :** cyanobacteria, protein evolution, photoprotection, carotenoid, photosynthesis.

**Abstract:** Cyanobacteria are photosynthetic organisms capable of CO<sub>2</sub> conversion into organic compounds and production of O<sub>2</sub> by using light energy. Nevertheless, high light intensities saturate the photosynthetic apparatus leading to production of reactive oxygen species, which are dangerous for the cell. To cope with this, the photoactive Orange Carotenoid Protein (OCP) induces thermal dissipation of the excess energy harvested by the antenna complex, the phycobilisome (PBS) to decrease the energy arriving at the photochemical centers. The OCP is composed of two domains connected by a flexible linker, the C-terminal domain (CTD) and the N-terminal domain (NTD). During photoactivation, the carotenoid is translocated to the NTD, the domains separate and the NTD is able to interact with the PBS. Three OCP families co-exist (OCPX, OCP1 and OCP2) in modern cyanobacteria. In addition to the OCP, many cyanobacteria also contain homologs of OCP domains, the CTDH and HCP. The HCPs are a family of carotenoid proteins with different photoprotective traits. Most of them are very good singlet oxygen quenchers, and one sub-clade is able to interact with the PBS and to induce thermal energy dissipation like OCP. The role of CTDH was unknown. The presence of these homologs in parallel to the OCP supported the general idea that the OCP has a modular evolutionary origin and that the CTDH and HCP can interact forming an OCP-like complex with similar characteristics and function than the OCP.

In this thesis, I present the first characterization of the CTDH proteins. CTDHs are dimers binding a carotenoid molecule. The main role of the CTDH is to transfer its carotenoid to the HCP. In addition, CTDHs are able to uptake carotenoids from membranes but not HCPs. These results strongly suggested that the CTDHs are carotenoid carriers that ensure the proper carotenoid loading into HCPs. This novel carotenoid translocation mechanism could be multidirectional. The resolution of two tridimensional structures of the ApoCTDH from *Anabaena* showed that the C-terminal tail of the CTDH (CTT) can populate different conformations. Moreover, mutational analysis demonstrated that the CTT has an essential role in carotenoid translocation. Finally, I report a molecular characterization of the flexible linker connecting the domains of different modern OCPs and its role during the evolution of the OCP. First, I characterized OCPs from the three subclades, including the uncharacterized OCPX. OCPX and OCP2 present a fast deactivation compared with OCP1. While OCP1 and OCPX can dimerize, OCP2 is stable as monomer. Finally, I found that the linker is essential for the OCP deactivation and it regulates the photoactivation. In OCP1 and OCPX the linker slows down the photoactivation, while in OCP2 it increases the photoactivation rate. Bioinformatic analysis complements this characterization and provides a clear picture of the evolution of the OCP to respond efficiently to stress conditions.