Interdomain interactions reveal the molecular evolution of the orange carotenoid protein

Fernando Muzzopappa, Adjélé Wilson, Diana Kirilovsky

Interdomain interactions reveal the molecular evolution of the Orange Carotenoid Protein

Fernando Muzzopappa[1], Adjélé Wilson[1] and Diana Kirilovsky[1*].

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

[*]Corresponding author: diana.kirilovsky@cea.fr

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[1–3]. One keto-carotenoid molecule (3-hydroxy equinenone) is shared between the two domains[4]. In darkness, the OCP remains in the inactive "orange form" (OCP^O) and adopts a closed conformation stabilized by the carotenoid and by interactions between the two domains[4]. Under strong blue-green light, the OCP is photoactivated and converted into the red active form (OCP^R)[5]. Upon light absorption, the H-bonds between the CTD and the carotenoid are broken and the carotenoid migrates (12Å) into the NTD[6] in about 10 µs[7]. Then the inter-domain interactions are disrupted and the protein adopts an open configuration[6,8,9] with a higher hydrodynamic radius[10,11]. By interacting with the core of the cyanobacterial antenna, the phycobilisome (PBS), the NTD of OCP^R induces the dissipation of the excitation energy as heat[6,12–14]. The OCP^R slowly relaxes back to OCP^O in darkness[5]. Another cyanobacterial protein, called Fluorescence Recovery Protein (FRP), accelerates this process[15,16] by increasing the probability of the proper configuration of domains and carotenoid in which they may interact to form OCP^O [17]. The main binding site of the FRP is in a CTD region that is exposed only in the red active form[16,18,19,20]. FRP binding also helps OCP detachment from PBSs[13,21]. 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[22]. 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[22,23]. 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[4,24,25,26]. The existence of these homologues have suggested a modular model for the evolution of OCP[24,23].

The OCP1 and OCP2 domains can be isolated independently as carotenoid proteins (holo-protein) or apo-proteins (without carotenoid), depending on the expression system [27,23,6,25]. Interaction between holo-CTD and apo-NTD results in the formation of a complex, which is spectrally similar to OCP^O (hereafter, OCP-like complex)[23,27,25,28]. This complex can be photoactived leading to efficient carotenoid translocation from the CTD to the NTD and complex disassembly[23,25,27]. Afterwards, reassembling of the OCP-like complex from holo-NTD and apo-CTD was only observed in the case of the OCP2 domains[23]. This suggested that OCP2 is more primitive than OCP1 and hinted towards an evolutionary role of the linker[23].

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[22,23]), 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[22]. 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

104 *Scytonema* (hereafter OCPX). *Tolypothrix* OCP2 and *Scytonema* OCPX were chosen
105 because their primary sequences are close to the OCP2 and OCPX consensus
106 sequences, respectively (Supplementary Fig 2). Furthermore, we also partially studied
107 the OCP2 and OCPX from *Synechocystis* PCC 7509 to confirm the characteristics of
108 each clade. Overexpression was performed in a canthaxanthin producing *E. coli* strain
109 since echinenone binding in *Tolypothrix* and *Scytonema* OCPs induces lower
110 photoactivity (Supplementary Fig. 3). The spectrum of $OCPX^{O}$ showed less vibrational
111 band resolution compared that of $OCP1^{O}$ and $OCP2^{O}$. The $OCPX^{R}$ and $OCP1^{R}$ spectra
112 were identical, with a maximum at 525 nm. The $OCP2^{R}$ spectrum was red shifted by 10
113 nm (Fig. 2A). At 8°C, the accumulation of *Synechocystis* $OCP1^{R}$ was slightly faster than
114 that of $OCP2^{R}$ and $OCPX^{R}$ (Fig. 2B). However, *Tolypothrix* $OCP1^{R}$ accumulated slower
115 than *Synechocystis* $OCP1^{R}$ and *Tolypothrix* $OCP2^{R}$ (Fig. 1B), in agreement with
116 previous observations [22]. This indicates that variations in photoactivation rates can exist
117 between different OCP1s and not only between OCP1 and OCP2 (OCPX). The back
118 conversion (OCP^{R} to OCP^{O}) of OCPX and OCP2 was faster than that of either OCP1,
119 OCPX displaying the fastest conversion rate (Fig. 2C and Supplementary Fig 3C). As
120 expected OCP1 back conversion was accelerated by the presence of *Synechocystis*
121 FRP (Fig 2C). This FRP accelerated the back reaction of all the OCP1s tested in our
122 laboratory including *Tolypothrix* OCP1 (Supplementary Fig 3). By contrast, FRP did not
123 accelerate the back conversion of OCPX and OCP2 (Fig. 2C and [22]). Cyanobacteria
124 strains containing only OCPX or OCP2 do not contain the *frp* gene[22]. OCPX and OCP2
125 of *Synechocystis sp.* PCC 7509 presented the same photoactivation and recovery
126 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[21] and the negative charge of D220[21] (Supplementary Fig 2). However, other amino acids which are also thought to be involved in the OCP-FRP interaction[16] 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[21]. 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[29]. 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 [30].

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 µM). 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 OCP1dimer. Indeed, the mutated R27L OCP1 remained as a

monomer at all concentrations (Fig. 3A and 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.

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 µM) and apo-NTDX (22µM) than apo-NTD2 (5µM) 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[31,32].

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^R, 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[23]). 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[23]). OCP2-like recovery was slower than that of OCP2 (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 [31]).

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, Δ174-179 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^R even in darkness (approximately 45% of OCP^R in E174K-R185E).

After 5 min illumination with strong white light, all OCP1-linker mutants exhibited the typical spectrum of OCP^R (Supplementary Fig. 7). All OCP1 mutants (except for K167A) accumulated the OCP^R 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^R 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 Δ174-179 presented the slowest kinetics. Furthermore, OCP^R 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.

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 (Ea) of the back

conversion for WT OCP1, 37 kcal/mol (in agreement with the values previously published[10]), 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[10,33]. In the linker mutants (with exception of K167A) the break point of the convex plot is shifted toward lower temperatures (Fig. 5D, 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^R (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^R 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 Δ174-179 OCP caused an extensive loss of OCP quenching activity in spite the fact that this mutation stabilized the OCP^R. 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[7,34]. Analysis of the hydrodynamic volume of photoactivated OCPs showed that the apparent Stoke's radius of the Δ174-179 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 (compared Fig. 6D and E). Even the slowest recovery reactions (OCP E174K-R185E and OCP Δ174-179) 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).

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^R to OCP^O back

conversion compared to their respective WTs (Supplementary Fig. 10). These results indicated that the mutations in the linker increased the stability of OCP^R in OCP2 and OCPX as was the case for OCP1. Like in OCP1, mutation E174K slightly increased the rate of OCP^R 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^R accumulation (Supplementary Fig. 10).

Discussion

According to the current evolutionary model, fusion of primitive *ctdh* and *hcp* genes gave rise to a primitive OCP[23,24,22]. We propose that the ancestral HCP was already able to interact with the PBS like the modern HCP4[36]. Addition of an appropriate linker and a CTD made the NTD-PBS interaction more stable, allowing for more efficient photoprotection[6]. 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

311 only under stress conditions in which cells need more protection[22]. Thus, some specific
312 properties of OCP2 might have an advantage for a better response to “permanent”
313 stress conditions while OCP1 properties must be advantageous under low, medium and
314 fluctuating light conditions.
315 The previous evolutionary model proposed that before the genes fused, an OCP-like
316 complex existed formed by the heterodimer of a primitive NTD and CTD that shared a
317 keto carotenoid. Based on *Tolypothrix* OCP2 results, it was originally proposed that this
318 primitive OCP-like could photocycle even in the absence of the linker loop connecting
319 the NTD and CTD[22,23]. However, here we have shown that this is probably not likely
320 since the ancient OCPX-like complex cannot photocycle: after photoactivation, the
321 isolated holo-NTDX and apo-CTDX were unable to reassemble to form the OCP-like
322 complex. Thus, our results strongly suggest that the addition of the linker loop provided
323 the primitive OCP with two new features: the possibility to photocycle by dark recovery
324 into the inactive form and a more controlled photoactivation. The modern OCPX has
325 kept these features, showing a fast dark recovery and a slightly negative effect of the
326 linker on photoactivation. OCP2 kept the fast recovery present in its ancestor and in
327 OCPX. In OCP1, two point mutations acquired during evolution in the CTD (in amino
328 acids 229 and 262) made the recovery reaction considerably slower. This added one
329 more point of regulation in the OCP photocycle which now required an exterior
330 component, the FRP, in order to accelerate the recovery reaction.
331 The OCP1 clade conserved the slightly negative effect of the linker on photoactivation.
332 This fine regulation allowed the accumulation of high concentrations of OCP^R only at
333 high light intensities, which is particularly important for OCP1 and OCPX since they are

expressed even under low light conditions[12,22,30,37]. 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^R to OCP^O 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^O 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^O (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[38].

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[18,19,39] 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[40]. 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[25] and pCDF-NTD1[6] plasmids containing the sequences encoding the isolated *Synechocystis* CTD and NTD domains were described previously[6,25].

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[40]. 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 previosly[40]. OCP, NTD, CTD holo-proteins and apo-proteins expressed in *E. coli* were isolated as described elsewhere[36,40]. 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 µg 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 $A^{1\%}_{cm}$=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 µmol photons m^{-2} 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^R 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 µM) just before turning off the light.

To study CTD-NTD complex formation, holo-CTDs (2.25 µM) 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 µL of purified proteins (10 µM) 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[13]. 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 µmol photons m^{-2} s^{-1}). The PBS concentration used was 0.012 µM and the ratio of carotenoid to PBS was 40:1. OCP samples were pre-illuminated with 5000 µmol photons m^{-2} s^{-1} of white light.

Phylogenetic analysis

In order to build the OCP phylogenetic tree, we performed a BLAST[41] search, using as seed the *slr1963* protein sequence and an E-value cut-off of $1x10^{-4}$, 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[42]. 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[43] 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[44] 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[45] library (Bio.Phylo module). Sequences logos were built with the online server Skylign[46].

Molecular dynamic simulation

The Apo-OCP structure was modelled using the SWISS-MODEL[47] 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[48] 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.

DATA AVAILABILITY

The sequences of the OCPs isolated and studied in this article are accessible in the GenBank/EMBL data libraries (or the IMG database) under the following accession numbers: *slr1963* (OCP1 *Synechocystis,* IMG ID: 2514153952), Fdi2450 (OCP1 *Tolypothrix,* IMG ID: 2501541336), Fdi7374 (OCP2 *Tolypothrix,* IMG ID: 2501546328) and WA1_RS11680 (OCPX *Scytonema,* IMG ID: 2551963320). The protein sequences and their corresponding identifier for the IMG database used for the phylogenetic analysis can be found in the Supplementary Information (available as an .xls file in the online version of the article). The additional data that support the findings of this study are available from the corresponding author upon request.

The author responsible for correspondence and distribution of materials integral to the findings presented in this article is Diana Kirilovsky (diana.kirilovsky@cea.fr).

ACKNOWLEDGMENTS The authors thank Mrs Sandrine Cot for technical help, Dr François Andre for helping with the phylogenetic analysis, Dr Leandro Tabares for helping with MD simulations and M. Guillaume Sarrailhe for helping in the construction

of three OCP mutants. This work was supported by grants from the Agence Nationale de la Recherche (RECYFUEL project (ANR-16-CE05- 0026)), and from the European Union's Horizon 2020 research and innovation program under grant agreement no. 675006 (SE2B). F.M.'s salary was financed by the European Union's Horizon 2020 (project SE2B). The research was also supported by the Centre National de la Recherche Scientifique (CNRS) and the Commissariat à l'Energie Atomique (CEA). The French Infrastructure for Integrated Structural Biology (FRISBI) ANR-10-INBS-05 also partially supported this research.

AUTHOR CONTRIBUTIONS F.M. performed all the OCP characterization experiments and constructed some mutants; A.W. constructed almost all the OCP mutants; D.K. conceived the project, designed and supervised almost all the experiments and analyzed data; the article was written by F.M. and D.K.

COMPETING INTERESTS

The authors declare no competing financial or non-financial interests.

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. Skylign: 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).

Figure legends

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 $OCP1^{O}$ (green), $OCP2^{O}$ (blue), $OCPX^{O}$ (black), $OCPx^{R}$ (dashed black), $OCP1^{R}$ (dashed green) and $OCP2^{R}$ (dashed blue). The inlet shows the vibrionic structure of the OCP^{O} 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 µE of white light illumination. **(c)** Kinetics of OCP deactivation of in the absence

711 (closed symbols) and presence (open symbols) of *Synechocystis* FRP (ratio 1 FRP: 1
712 OCP) of *Synechocystis* OCP1 (green circle), OCP1-R229E-D262K (rose triangle),
713 OCP2 (blue square) and OCPX (black triangle) at 8°C. **(d)** Kinetics of PBS fluorescence
714 quenching under strong blue light (1000 µmol photons $m^{-2}s^{-1}$) induced by pre-
715 photoactivated OCPX (black triangle), OCP2 (blue square) and OCP1 (green circle) at
716 phosphate 0.5 M (closed symbols) and at phosphate 1M (open symbols) at 23°C. **(e)**
717 Kinetics of fluorescence recovery at 0.5 M phosphate. Same symbols than in (d). **(f)**
718 OCP-related PBS fluorescence quenching induced by strong blue-green light (1000
719 µmol photons $m^{-2}s^{-1}$) measured in *Scytonema hofmanni* PCC 7110 cells (1 and 3: cells
720 under dim blue light illumination of 20 µmol photons $m^{-2}s^{-1}$; 2: cells under high blue light
721 of 1000 µmol photons $m^{-2}s^{-1}$). The curves (b-e) are an average of three independent
722 experiments. The error bars represent the standard deviation of the data shown. The
723 figures (a) and (f) show representative experiments, which were repeated 5 times with
724 similar results.

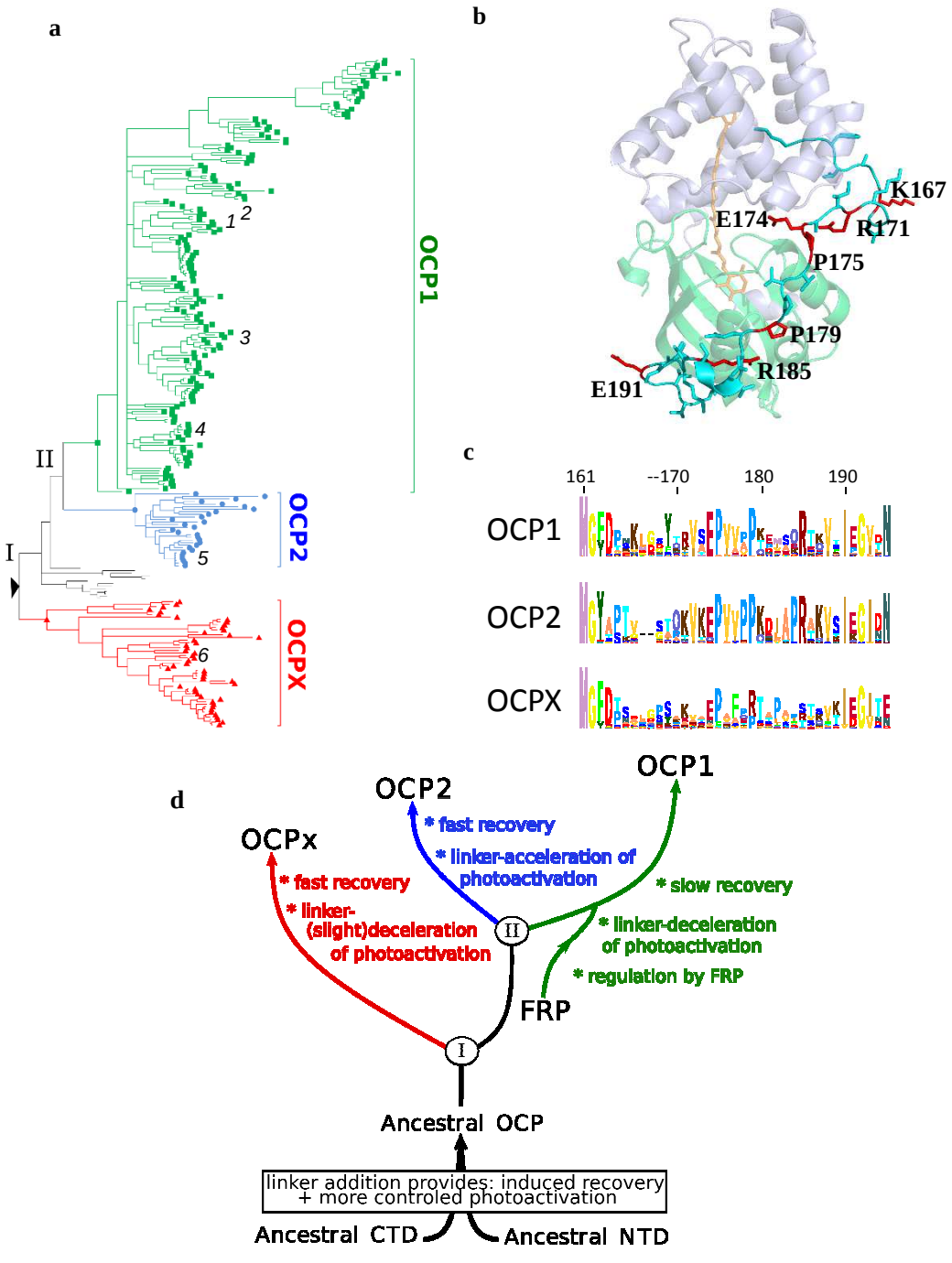
725 **Figure 3 | OCP oligomeric state: Native gel and size exclusion chromatography**.
726 **(a)** Analysis of OCP oligomeric state by Coomassie-Blue stained Native PAGE: OCPX
727 (10 µM), OCP2 (10 µM) and OCP1 (10 µM), two molecular markers (45 kDa and 66
728 kDa) were added (lane 1 and 2). **(b to e)** Comparison of elution profiles in SEC at
729 different concentrations: 120 µM (blue), 80 µM (orange), 40 µM (red), 23 µM (green), 4
730 µM (pink) and 2 µM (sky blue) of holo-WT-$OCP1^O$ **(b)**, Holo-WT-$OCPX^O$ **(c)**, Holo-WT-
731 $OCP2^O$ **(d)** and Holo-R27L-$OCP1^O$ **(e).** Absorbance was measured at 510 nm. The
732 figures (a-e) show representative experiments, which were repeated 3 times with similar
733 results.

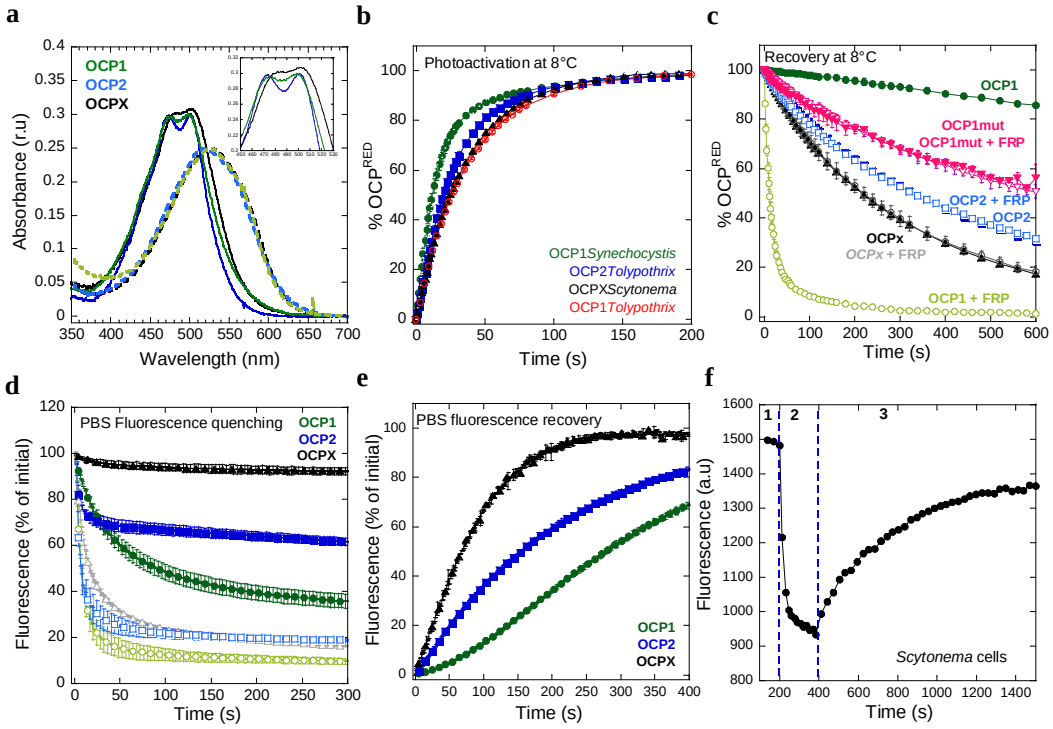
Figure 4 | CTD-NTD interactions. (a) Saturation curve of OCP-like formation from isolated CAN-CTD and apo-NTD. 2.25 µM CAN-CTD was incubated with increasing concentrations of apo-NTD **(b)** Kinetics of OCP-like formation using CAN-CTD (2.25 µM) and apo-NTD (18 µM) from OCP1 (green circle), OCP2 (blue 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^{-2} s^{-1}: **(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 experiments. The error bars represent the standard deviation of the data shown.

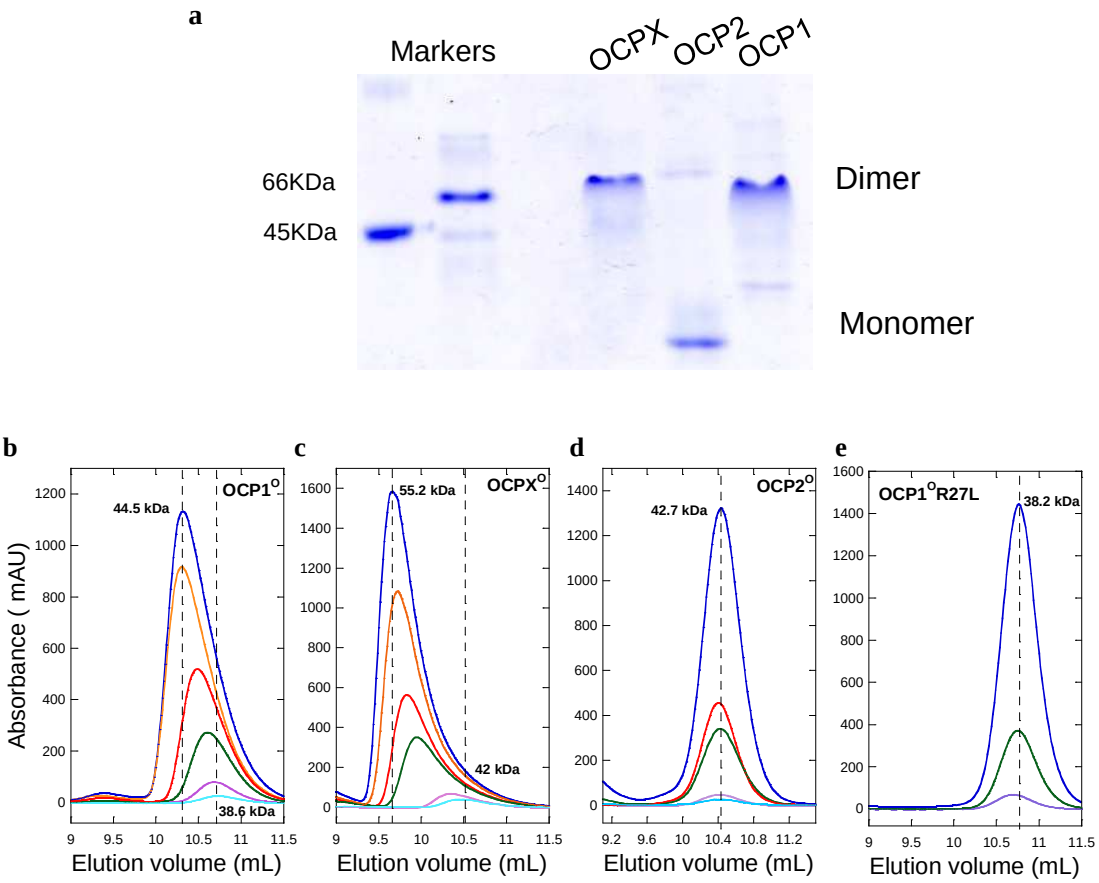
Figure 5 | Photoactivity of OCP1 linker-mutants. **(a)** Accumulation of photoactivated OCP^R and **(b)** OCP^R back-conversion to OCP^O at 15°C of OCP WT (black closed circle), OCP E174K (fuchsia closed triangle), OCP E191K (orange closed square), OCP E174K-R185E (violet open triangle), OCP R185A (blue close diamond), OCP K167A (green open square) **(c)** Percentage of OCP^R 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

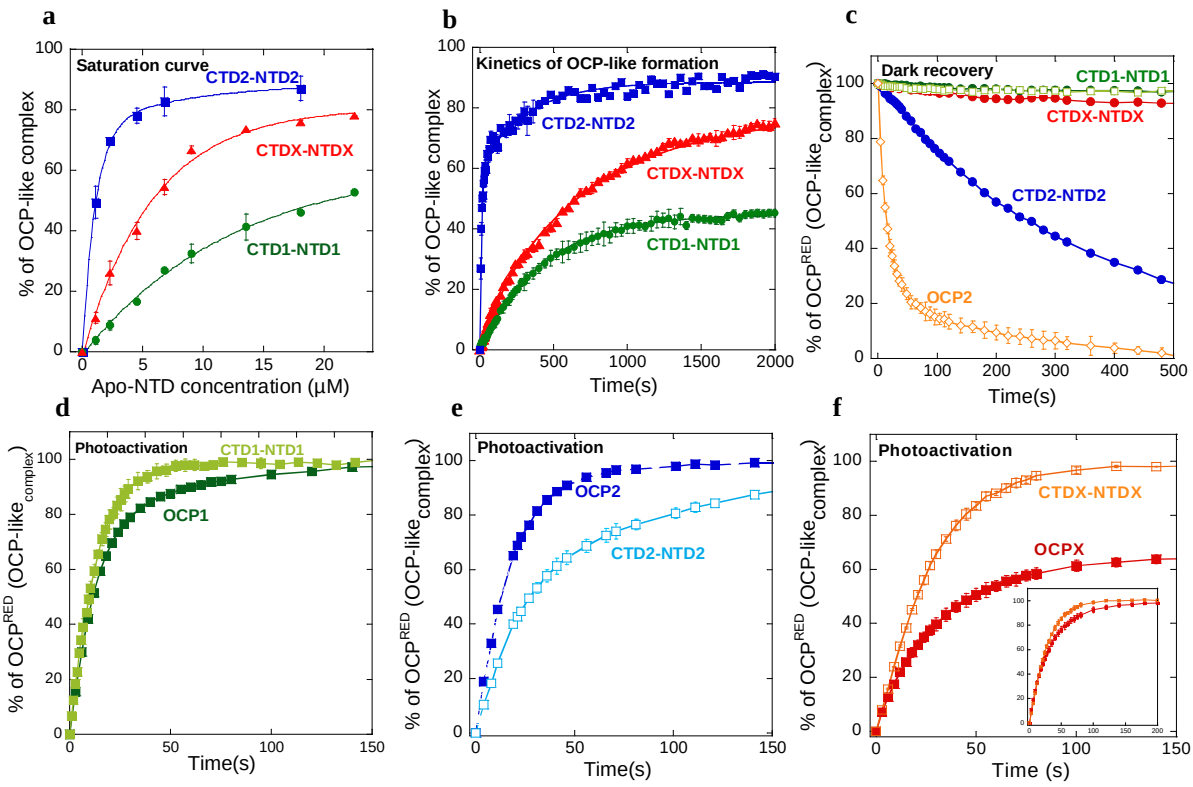
convex curves is shown with dotted lines for the WT OCP (I) and the linker mutants (II) **(e)** Arrhenius plot of back-conversion OCP^R to OCP^O of OCP WT (black), OCP E174K-R185E (dark red), OCP R185A (sky blue), OCP K167A (green) and OCP R171A (orange). The curves are an average of three independent experiments. The error bars represent the standard deviation of the data shown.

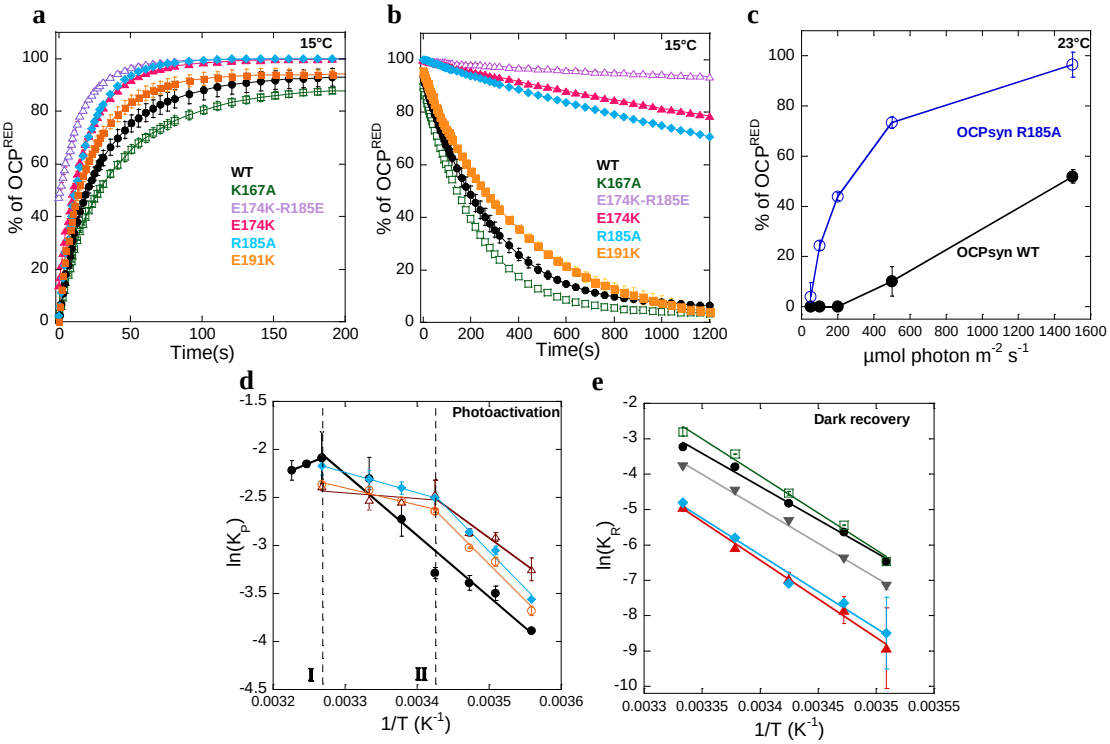
Figure 6 | OCP1 linker mutants: interaction with PBS and FRP. (a) Kinetics of PBS fluorescence quenching under strong blue light induced by previously photoactivated WT OCP^R (black), OCP^R E174K (red), OCP^R E174K-R185E (dark red), OCP^R R185A (sky blue), OCP^R R185E (blue), OCP^R K167A (green), OCP^R R171A (orange), OCP^R P175A-P179A (grey) and OCP^R Δ174-179 (pink). **(b and c)** Kinetics of dark recovery of PBS fluorescence in absence **(b)** or presence of *Synechocystis* FRP **(c)**. Back-conversion OCP^R to OCP^O at 8°C; in presence **(e)** or absence **(d)** of *Synechocystis* FRP (1 FRP per OCP) of OCP^R WT (black), OCP^R E174K (red), OCP^R E174K-R185E (dark red), OCP^R R185A (sky blue), OCP^R R185E (blue), OCP^R K167A (green), OCP^R R171A (orange), OCP^R P175A-P179A (grey) and OCP^R Δ174-179 (pink). The curves are an average of three independent experiments. The error bars represent the standard deviation of the data shown.

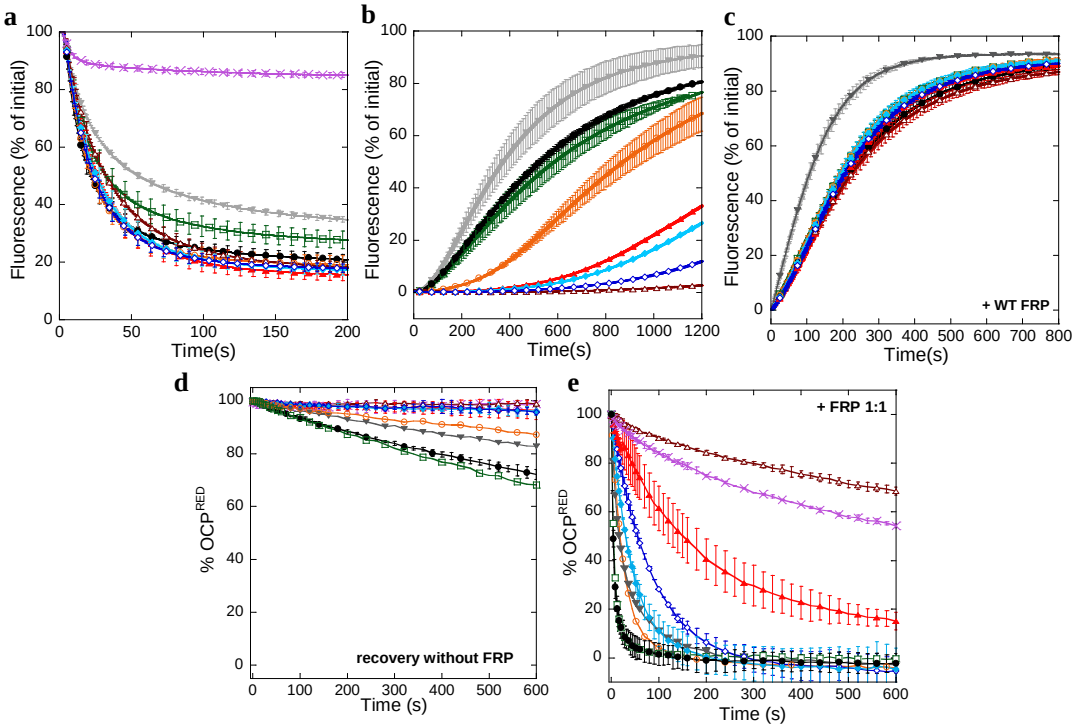

a
OCP1
OCP2
OCPX
I
II
1
2
3
4
5
6
b
K167
E174
R171
P175
P179
R185
E191
c
161
--170
180
190
OCP1
OCP2
OCPX
d
OCP2
OCP1
OCPx
* fast recovery
* linker-acceleration of photoactivation
* slow recovery
* fast recovery
* linker-(slight)deceleration of photoactivation
* linker-deceleration of photoactivation
* regulation by FRP
II
FRP
I
Ancestral OCP
linker addition provides: induced recovery + more controled photoactivation
Ancestral CTD
Ancestral NTD


a
OCP1
OCP2
OCPX
Absorbance (r.u)
Wavelength (nm)
b
Photoactivation at 8°C
% OCP^RED
Time (s)
OCP1Synechocystis
OCP2Tolypothrix
OCPXScytonema
OCP1Tolypothrix
c
Recovery at 8°C
% OCP^RED
Time (s)
OCP1
OCP1mut
OCP1mut + FRP
OCP2 + FRP
OCP2
OCPx
OCPx + FRP
OCP1 + FRP
d
PBS Fluorescence quenching
OCP1
OCP2
OCPX
Fluorescence (% of initial)
Time (s)
e
PBS fluorescence recovery
Fluorescence (% of initial)
Time (s)
OCP1
OCP2
OCPX
f
1
2
3
Fluorescence (a.u)
Time (s)
Scytonema cells


a
Markers
OCPX
OCP2
OCP1
66KDa
45KDa
Dimer
Monomer
b
c
d
e
Absorbance (mAU)
OCP1O
44.5 kDa
38.6 kDa
OCPXO
55.2 kDa
42 kDa
OCP2O
42.7 kDa
OCP1OR27L
38.2 kDa
Elution volume (mL)


a
Saturation curve
CTD2-NTD2
CTDX-NTDX
CTD1-NTD1
% of OCP-like complex
Apo-NTD concentration (µM)
b
Kinetics of OCP-like formation
CTD2-NTD2
CTDX-NTDX
CTD1-NTD1
% of OCP-like complex
Time(s)
c
Dark recovery
CTD1-NTD1
CTDX-NTDX
CTD2-NTD2
OCP2
% of OCP^RED (OCP-like complex)
Time(s)
d
Photoactivation
CTD1-NTD1
OCP1
% of OCP^RED (OCP-like complex)
Time(s)
e
Photoactivation
OCP2
CTD2-NTD2
% of OCP^RED (OCP-like complex)
Time(s)
f
Photoactivation
CTDX-NTDX
OCPX
% of OCP^RED (OCP-like complex)
Time (s)


a
15°C
% of OCP^RED
Time(s)
WT
K167A
E174K-R185E
E174K
R185A
E191K
b
15°C
% of OCP^RED
Time(s)
WT
K167A
E174K-R185E
E174K
R185A
E191K
c
23°C
% of OCP^RED
OCPsyn R185A
OCPsyn WT
µmol photon m-2 s-1
d
Photoactivation
ln(K_P)
I
II
1/T (K-1)
e
Dark recovery
ln(K_R)
1/T (K-1)


a
Fluorescence (% of initial)
Time(s)
b
Fluorescence (% of initial)
Time(s)
c
Fluorescence (% of initial)
+ WT FRP
Time(s)
d
% OCP^RED
recovery without FRP
Time (s)
e
% OCP^RED
+ FRP 1:1
Time (s)