Lawrence Berkeley National Laboratory
LBL Publications

Title
Orange carotenoid proteins: structural understanding of evolution and function

Permalink
https://escholarship.org/uc/item/5rf4c5cb

Authors
Kerfeld, Cheryl A
Sutter, Markus

Publication Date
2024-05-01

DOI
10.1016/j.tibs.2024.04.010

Copyright Information
This work is made available under the terms of a Creative Commons Attribution-NonCommercial License, available at https://creativecommons.org/licenses/by-nc/4.0/

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

Orange carotenoid proteins: structural understanding of evolution and function

Cheryl A. Kerfeld[1,2,3], Markus Sutter[1,2,3]

[1] MSU-DOE Plant Research Laboratory, Michigan State University, East Lansing, MI 48824, USA

[2] Environmental Genomics and Systems Biology Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA

[3] Molecular Biophysics and Integrated Bioimaging Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA

Correspondence: ckerfeld@lbl.gov

Kerfeld orcid: 0000-0002-9977-8482

Sutter orcid: 0000-0001-6290-4820

Keywords: Cyanobacteria, Carotenoids, Protein evolution, Protein dynamics, Photoprotection, Photosynthesis

Abstract

Cyanobacteria uniquely contain a primitive water-soluble carotenoprotein, the Orange Carotenoid Protein (OCP). Nearly all extant cyanobacterial genomes contain genes for the OCP or its homologs, implying an evolutionary constraint for cyanobacteria to conserve its function. Genes encoding the OCP and its two constituent structural domains, the N-terminal domain, Helical Carotenoid Proteins (HCPs), and its C-terminal domain (CCPs), are found in the most basal lineages of extant cyanobacteria. These three carotenoproteins exemplify the importance of the protein for carotenoid properties—including protein dynamics--in response to environmental changes in facilitating a photoresponse and energy quenching. Here we review new structural insights for these carotenoproteins and situate the role of the protein in what is currently understood about their functions.

Carotenoids and the Orange Carotenoid Protein**Carotenoids** are considered the oldest, most widespread, and numerous class of pigments and are proposed to be biosignatures for detecting life [1]. They extend the spectral range of light-harvesting antennae of nearly all phototrophic organisms, from bacteria to plants,, where they function as **accessory pigments**, absorbing photons not captured by chlorophylls in photosynthetic complexes. Because carotenoids are fast and effective in deactivating (**quenching**) excited states of nearby **chromophores** through thermal dissipation, they frequently play roles in **photoprotection** mechanisms where they disarm the ROS (reactive oxygen species) inherently produced during oxygenic photosynthesis [2]. It seems likely that the first role of carotenoids in cyanobacteria was as a component of membranes [3], where they provided a structural role and likely influenced permeability to oxygen. The next key event was the complexation of carotenoids to proteins, thereby vastly expanding the potential roles of carotenoids from membrane-structure to light harvesting and its necessary accompaniment, photoprotection.

One of the simplest model systems for the study of how the protein environment tunes carotenoid function is the cyanobacterial Orange Carotenoid Protein (OCP), a 34 kDa protein that noncovalently binds a single carotenoid (Figure 1A) [4, 5]. The OCP regulates photoprotection in cyanobacteria, the organisms that originated oxygenic photosynthesis. The OCP is the only known photoactive protein that uses a carotenoid as its sole chromophore; the absorption of blue-green light causes OCP to convert from a dark stable orange form, OCP^O, to a light-activated red form, OCP^R [6]. The OCP^R directly participates in a photoprotective mechanism by binding to the light-harvesting antenna, the **phycobilisome** (**PBS**) [7], and dissipating excess captured light energy [6]. Recent structural studies of the OCP and of its primitive antecedents, described below, have provided unexpected new insights into how carotenoid properties are tuned by protein structure, and how protein dynamics—of both large and small amplitudes—regulate carotenoid function. This review will focus on the results of these structural studies and place them in the context of the evolution and dynamics of OCP function. It is known that determinants of quenching properties are inherent in the protein sequence, while the spectroscopic properties play a subsidiary role in in understanding OCP function [8]. Therefore, much of the spectroscopy

that has been done on OCP in the past can be reinterpreted in the context of the newly discovered structural and evolutionary aspects described here.

The Modularity of OCP Structure and Function

The cyanobacterial OCP was discovered by David Krogmann as a contaminant in cytochrome purifications [9], and the crystal structure of the OCP (Figure 1A) was determined in 2003, prior to an understanding of its physiological function; however, its singlet oxygen quenching properties and its ability to reversibly change color from orange to red were described with the structure, and a hypothesis that the OCP may be photoresponsive and play a role in photoprotection was suggested [10-12]. Eventually it was shown to indeed be a photoactive protein responsible for both light sensing and energy dissipation in cyanobacteria [13].

Structurally, the OCP consists of two domains (Figure 1A): the all-alpha helical N-terminal domain (NTD, formed by residues 1-164, (*Synechocystis* sp. PCC 6803 (hereafter *Syn* 6803) numbering) and the C-terminal domain (CTD, residues 190-314), formed by alpha helices and beta sheet. While the NTD is found only in cyanobacteria, the CTD is a member of the versatile NTF2 superfamily, widespread among all kingdoms of life [12, 14]. The two domains are joined by a linker (residues 165-190) that in the orange, inactive state (OCP^O) is excluded from the interface between the NTD and CTD and wraps around the surface of the CTD (Figure 1B) [12]. A keto-carotenoid—such as canthaxanthin or echinenone— spans the two domains, with a keto group of the CAR (carotenoid) hydrogen bonded to a conserved Trp and Tyr in the CTD. (Figure 1A, inset). In addition to the large interface for interaction across which the CAR spans, the CTD and NTD also interact through a "latch" formed by the N-terminal extension (NTE), residues 1-19, that nestle against the CTD surface opposite of the linker (Figure 1A, B) [12].

The structural modularity of the OCP was shown to extend to functional modularity: controlled proteolysis showed that photo activated OCP^R is sensitive to protease exposure and rapidly converted all OCP into a permanently red and constitutively active form known as the RCP; this proteolytic protein fragment consists of the NTD and the carotenoid [15]. Krogmann had identified similar red proteolytic fragments in purifications from wild cyanobacterial blooms [9],

suggesting that, in cells, proteolysis could be used regulate quenching activity Recombinant RCP was readily crystallized and the structure subsequently determined [16]. Combined with X-ray footprinting [17], the structural analysis of the RCP in comparison to the full-length OCP to determine the solvent accessible surface area differences between OCP^O to OCP^R, revealing that the NTD and CTD domains remained structurally intact, but dissociated, with the carotenoid migrating entirely into the NTD during the activation [16].

The modular OCP Evolved from the Fusion of Two Carotenoproteins

The biochemical and structural studies described above suggest an analogy between the OCP and other modular photoreceptors that employ sensor and effector domain architectures such as BLUF (blue light using flavin) [18] and LOV (light, oxygen voltage) domains [19]. In this context, the NTD could be considered the effector (quenching) domain, activated when the carotenoid was fully enclosed within, and the CTD acting as a regulatory domain [20]. This is consistent with the proposed evolution of the OCP [5, 21, 22]; it evolved from the fusion of two discrete proteins that formed a heterodimer. Already at the time of the first structural description of the OCP, which coincided with the beginning of the genomics era, it was noted that in the few diverse cyanobacterial genomes available, in addition to the OCP, several had genes encoding homologs to both the NTD and the CTD [12]. The NTD homologs are now known collectively as the **Helical Carotenoid Proteins** (HCPs) and these have been phylogenetically resolved into at least nine clades [23]. The proteins from all of the clades are proposed to bind carotenoid, and the crystal structures of two of these (*Anabaena* sp. PCC 7120 (hereafter *Anabaena* 7120) HCP1 and HCP2) have been determined [23, 24] (Figure 1C), confirming that they bind a carotenoid as observed in the RCP [15]. Many organisms encode multiple HCPs in addition to the OCP. For example, *Anabaena* 7120 contains four HCPs: HCP1-HCP4 in addition to OCP. Among all HCPs, only HCP4 has thus far been verified to quench PBS [25], and among HCP1-4 in *Anabaena* 7120, it has the highest sequence identity (52%) with the NTD of *Anabaena* 7120 OCP. The function of the other HCPs is as-yet unknown (see Box 1), and are likely distinct, given their varying occurrences in expression and proteomics studies (summarized in [23] and, more recently provided in [26-30]).

The role of HCP4 in both quenching and the evolution of the OCP is also supported by the observation that a gene encoding a CTD homolog is frequently adjacent to an HCP4 locus [21, 23]. These proteins, termed CCP (C-terminal domain-like carotenoprotein or CTD homologs (CTDHs)), are now known to bind a single carotenoid as homodimers (Figure 1D), shifting the absorption maximum by 80nm (versus the free carotenoid in solvent) to ~560nm, with a further 10nm red shift in a dimer of dimers state at high concentrations [31]. Additionally, some CCPs have been shown to accept carotenoid from HCPs [27, 32-34]. Results from devolving the OCP into its constituent domains and evaluating photoactivation and recovery [21, 35] are consistent with a model that the CCP, the precursor of the CTD, regulated quenching by the HCP by providing a carotenoid environment within the heterodimer that rendered the carotenoid inactive for quenching.

It also seems likely that this fusion event happened several times in cyanobacterial evolution. This is evident from the multiple discrete families of the OCP that are now known bioinformatically. In 2017 it was shown that there were two clearly resolved families of the OCP, known as OCP1 (the canonical, including the extensively characterized OCP from *Syn* 6803) and OCP2, as well as a third, more mixed group, OCPx [36] which has recently been renamed OCP3 and further resolved into 3 subgroups [37]. OCP1-3 all photoactivate from orange to red, and all quench PBS, but with some differences in quenching and regulatory properties [27]. For example, *Tolypothrix* sp. PCC 7601 (*Tolypothrix* hereafter) OCP2 photoactivates faster and recovers faster, and exhibits lower fluorescence quenching than its paralogous OCP1 [36]; it also appears to have less flexibility [8]. Additionally, the OCP^R to OCP^O recovery for OCP1 is accelerated by the binding of the all-helical Fluorescence Recovery Protein (FRP) [38, 39], whereas OCP2 and OCP3 appear to be insensitive to FRP [36, 37]. Interestingly, the evolution of regulation of OCP1 by FRP has been shown to be the result of complementarity between the CTD dimer and the FRP, which was acquired by horizontal gene transfer, likely from an ancestral delta-proteobacterium [40]. These differences in activity and regulation may account for the co-occurrence of OCP paralogs within the same genome—for example, OCP1 and OCP2 in *Tolypothrix* [36] or the two OCP3 paralogs found in *Nostoc flagelliforme* [27] -- to confer more flexibility in the range of photoprotective

responses. More specifically, in *Tolypothrix*, OCP1 appears to be constitutively expressed with OCP2 being additionally expressed under high irradiance [36, 41] while the *Nostoc flagelliforme* paralogs likewise exhibited differential expression and differences in quenching properties [27].

Large-scale Structural Dynamics Tune OCP Function

While there were differences in the kinetics, as noted above, OCP1, OCP2, and OCP3 all photoactivate and quench PBS. Indeed, the notion that the quenching mechanism is conserved across OCP families was reinforced when Slonimskiy et al demonstrated that OCP3 from *Gloeobacter*, an organism with an atypical, bundle-shaped PBS, could quench the PBS of *Syn* 6803 [37]. However, the precise structural details of how the OCP quenches the PBS and the structure of the OCP^R remained elusive. Elements of the activation and quenching process were known: Biophysical studies of the OCP^R had provided some information; in addition to the required carotenoid translocation into the NTD [16], other events such as detachment of the NTE from the CTD and domain separation had been detected [42, 43]. It required capturing the OCP^R in the act of quenching the PBS to learn its structure [44]. Contrary to all previous models, four activated OCP^Rs, arranged as two dimers, quench the 6.2 MDa PBS with its 396 phycocyanobilin pigments (Figure 2A). Also unexpectedly, quenching of the OCP^R-PBS complex revealed a drastic structural rearrangement relative to the inactive form; the CTD rotated approximately 220 degrees around the NTD and with a net displacement ~60 Å relative to its position in the OCP^O structure (Figure 2B) [44]. The structure of the OCP^R bound to the PBS explains why the obvious experiment, illuminate the OCP^O crystals with light and then collect diffraction data, always failed, as such a large structural displacement upon illumination would be prevented by the close crystal packing interactions. Few such large domain motions are known (e.g. a 90 Å translocation as observed for the nuclear localization signal domain of influenza virus polymerase [45]), though the increasing application of cryo-EM and the development of data analysis software (3Dflex [46] and 3DVA [46]) promise to reveal more such insights from analysis of conformational flexibility exhibited in single particle datasets.

The OCP^R-PBS structure, which has since been extended to a resolution of 1.6-2.1 Å [47], explains how the protein "activates" the carotenoid by providing a snapshot of its environment in the act of quenching. The carotenoid translocation discovered in the RCP structure presented a major advance in understanding of the photoprotective mechanism; however, many details such as the position of the NTE, of the interdomain linker and of the CTD, all of which are not present in the RCP [16], are protein elements that influence the carotenoid properties. The interdomain linker, for example, contributes to the electrostatic environment around the CAR and results in an even more pronounced polarization than in RCP. In the recent high resolution structure [47], modeling the sidechain motions and water molecules permitted quantum mechanical calculations of the transition dipole moment (TDM) conferred by the protein environment. This included the contribution of some PBS residues (in ApcA and ApcB) as well as OCP residues R155, E34 and D35 (conserved in OCP1 and OCP2), and showed that it is not distortions in the CAR, but changes in the protein environment, that are the drivers for potentiating quenching [47].

The OCP^R dimer structure also suggests a potentially new aspect of the regulation conferred by the CTD. The CTD mediates the dimerization of the OCP^R through inter-beta sheet interactions (Figure 2C); the dimer forms a buttress-like structure, with one NTD bound to the T cylinder and the second bound to the B cylinder in the PBS core (Figure 2A) [44]. Such an interface seems to be favored by NTF2 domains generally; a very similar interface is observed in the structure of a protein with unknown function (DUF4783 family protein from *Bacteroides uniformis* ATCC 8492, PDB ID 4Q53) that we can structurally superimpose almost perfectly despite only 9% sequence identity (Figure 2C). The predominantly hydrophobic residues forming the CTD dimerization interface are shielded by the NTE in the OCP^O form [12, 14]. This adds an additional regulatory implication of the oligomeric states of OCP1. A prerequisite to detaching the NTE from the CTD is the dissociation of the $OCP1^O$ dimer (Figure 1). Based on structural arguments, the dimeric resting $OCP1^O$ state found in the crystal (Figure 1A) has always been asserted to be biologically relevant by structural biologists because of the surface area excluded [5, 12]. In addition to the obvious occluding of the requisite protein motions for domain separation, the OCP dimer interface is mediated by conserved residues forming the OCP^O dimer: D6, R9, G10, P13, N14, T15,

L16, A17, F227, R229 (Figure 3). Notably, these residues are not well conserved in OCP3, consistent with the observation that it is predominantly monomeric [37]; although there is discrepant earlier data [48]. Recently, the dissociation of the $OCP1^O$ dimer into monomers, the prerequisite for photoactivation from a structural perspective, was shown to be light-driven [49] and important for both photoactivation and recovery of *Syn* 6803 OCP [50]. Recently, protein motion has also been suggested to play a role in regulating the activity of the OCP^R; the dimerized CTDs are the most flexible within the range of motion exhibited by the OCP^R-PBS complex [47]. The interdomain linker moves in conjunction with the CTD-CTD dimer which fluctuates as a unit, with linker residues E174-V176 making the closest approach to the CAR through the course of the motion. This potentially modifies the TDM of the carotenoid, expanding the regulatory repertoire to the structural dynamics of the CTD.

It was previously demonstrated that $OCP1^R$ is a dimer in solution [36], but the observation of the structure of the dimer bound to the PBS also provided new insight into the regulation of the OCP1 by FRP. While precise mechanism of FRP catalysis of dark conversion is not known, several lines of evidence have implicated that it binds the CTD of OCP^R [39, 51]. Solution studies of the interaction of engineered FRP variants with OCP^R in solution led to the suggestion of a 1:1 complex of OCPR to FRP [52]. However this model required that FRP, which based on structural arguments was suggested to function as a dimer, was required to monomerize to bind the CTD. Now, with the observation of the quenching-active dimer, the FRP dimer likely interacts with the CTD part of the OCP^R dimer, acting as a sort of chaperone to protect the exposed hydrophobic beta sheet until it can reform with its NTD counterpart. Interestingly, the distance between the two head groups of FRP that are thought to interact with OCP are spaced apart similarly (both ~80 Å) to the distance between the two NTDs in OCP^R dimer associated with the PBS. Each FRP head group binding a CTD would bring the CTDs close to their respective NTDs and facilitate recombination (Figure 2D).

Different but Similar—Conserved Features of Quenching by OCP Homologs

FRP binding is not a universal feature of OCP-mediated photoprotective systems; OCP2 and OCP3 are insensitive to it [36, 37]. We now consider if other structural features known from the OCP^R

structure are relevant to quenching by OCP2, OCP3 and HCP4. While the OCP1 has been extensively characterized structurally and functionally, far less data is available for the OCP2 and OCP3 families. However representatives of both paralogous OCP families, (*Tolypothrix* OCP2 and *Gloeobacter kilaueensis* JS1 OCP3), as well as the *Anabaena* 7120 HCP4 have been shown to quench PBS [25, 36, 37]. Likewise, structural data is now available for the orange form of both OCP2 (PDB IDs 8PYH and 8PZK) [53] and OCP3 (8A0H) [37]; they superimpose on the $OCP1^{O}$ structure of *Syn* 6803 with less than 1 Å rmsd. A sequence alignment of these with *Syn* 6803 OCP, with residues known from the *Syn* 6803 OCP^{R} structure to be important for binding and quenching the PBS marked, is shown in Figure 3. The PBS binding residues are largely conserved among the three OCP paralogs. While HCP4 lacks the equivalent of an NTE, all of the other PBS binding residues are conserved relative the OCP NTD (Figure 3), suggesting that the binding sites on the PBS are similar; this is also evident from a cursory survey of their interaction counterparts in these species.

The protein environment of the carotenoid tunes its transition dipole moment to poise it for quenching. Using the $OCP1^{R}$ -PBS complex as a template and placing homology models for the $OCP2^{R}$, $OCP3^{R}$ and HCP4 in that context, the electrostatics of the carotenoid environment for each was visualized (Figure 4). Given the conservation of the PBS residues at the OCP binding site, the conservation of PBS binding residues among the OCPs, and the experimental observation of PBS quenching by heterologous OCP, such as the *Gloeobacter kilaueensis* JS1 OCP3 quenching the *Syn* 6803 PBS [37], we include the influence of the PBS on the electrostatics (Figure 4). The carotenoid environment electrostatics of the OCP3 and HCP4 are similar to that observed of $OCP1^{R}$, while OCP2 is relatively non-polar at the β1 surface, which interacts with the PBS. The recent structure of the OCP^{R}-PBS complex was of a resolution sufficient for quantum mechanical calculations of the transition dipole moment of the carotenoid [47]; residues that play a significant role in setting the TDM, such as R155 (substituted as Q in all OCP2), E34, D35 are conserved across the homologs. While OCP1 and OCP2 have the linker motif EPV (174-176) suggested to be important in affecting either stability or duration of quenching [47], OCP3 has a substantial difference with an E substituted for Pro175 (Figure 3).

Another difference between OCP1 and OCP2 that was documented is that $OCP1^R$ forms dimers, while $OCP2^R$ was predominantly a monomer [36]; for OCP3 the oligomeric state has only been characterized for OCP^O. Considering the conservation of residues that stabilize the CTD dimer interface in the OCP^R-PBS state, and the extent of solvent accessible surface area buried in homology models of the $OCP2^R$ (962 $Å^2$) and $OCP3^R$ (1146 $Å^2$), suggest that both likely form dimers as in the OCP1 quenching complex. Given the data that $OCP2^R$ is predominantly a monomer in solution, we propose that association of two monomers of $OCP2^R$ as they bind the PBS for quenching juxtaposes their CTDs in proximity to facilitate OCP^R dimer formation.

Concluding Remarks

When the ketocarotenoid canthaxanthin becomes associated with protein to form the OCP, its absorption maximum shifts to 495 nm (compared to 490 nm in DMSO). Photoactivation of OCP^O to OCP^R further shifts it another 10 nm, the result of the protein environmental change surrounding the CAR with the protein. These visible changes are a result of the sensitive tuning of the carotenoid environment by the protein. The OCP exemplifies how the lack of covalent bonds to CAR, while making their binding sites in proteins difficult to predict, permits CAR function to be tuned by the dynamics of its protein environment [54], including intra-protein and inter-domain motion. Given advances in structural tools such as cryo-EM and its data analysis promise many more surprises in the fundamental understanding of the mechanism of function of carotenoproteins that have implications for engineering of photoprotection; for example, accelerating recovery from photoprotection has been shown to increase primary productivity in plants [55]. Manipulation of photoprotection in cyanobacteria could be especially valuable in production strains and communities used in biotechnology [56]. Likewise, understanding the molecular architecture of pigment interactions in antenna and how energy transfer and dissipation can be regulated is fundamental knowledge forthe construction of artificial photosynthetic systems [57]. Beyond photosynthesis, photoreceptors like OCP have the potential to become light-activated switches foroptogenetic applications [58, 59].

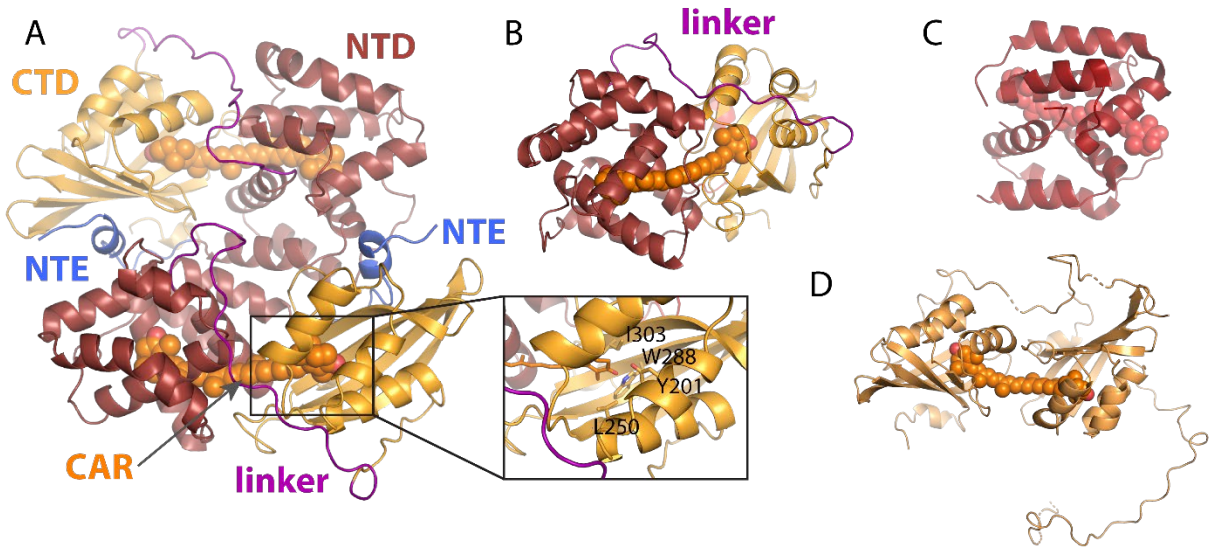

Figure 1. A) OCP^{O} dimer structure observed in crystallographic studies, highlighting the different domains across which the CA spans (space filling representation) (PDB ID 4XB5). B) OCP^{O} monomer, highlighting the linker (PDB ID 4XB5) C) HCP2 structure (PDB ID 5FCX), and D) CCP2 dimer model (SASBDB ID SASDHD6). NTD: N-terminal domain, dark red; CTD: C-terminal domain, orange; NTE: N-terminal extension, blue; CAR: carotenoid, orange; linker: magenta.

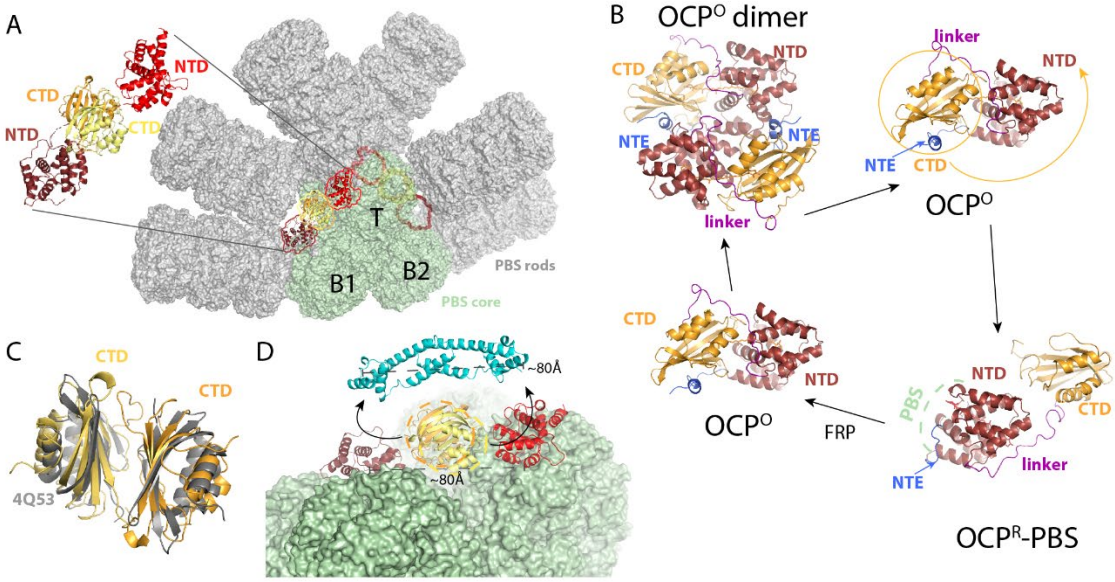

Figure 2. OCP-PBS structure overview. A) Overview of the position of the two OCP dimer pairs, OCP dimer in the back is outlined and dimer in front is shown zoomed in. B) OCP activation and recovery pathway: OCP^{O} dimer to monomer; structural changes to the OCP^{O} monomer lead to the OCP^{R}-PBS form where the NTD and CTD are separated; the FRP returns the OCP^{R} form back

to a monomeric OCP^{O} which can then dimerize again. C) CTD dimer as observed in the OCP^{R}-PBS complex and alignment with a structurally homologous NTF2 domain protein (pdb ID 4Q53). D) Distance between the head groups of FRP match the distance between the NTDs in OCP^{R}-PBS.

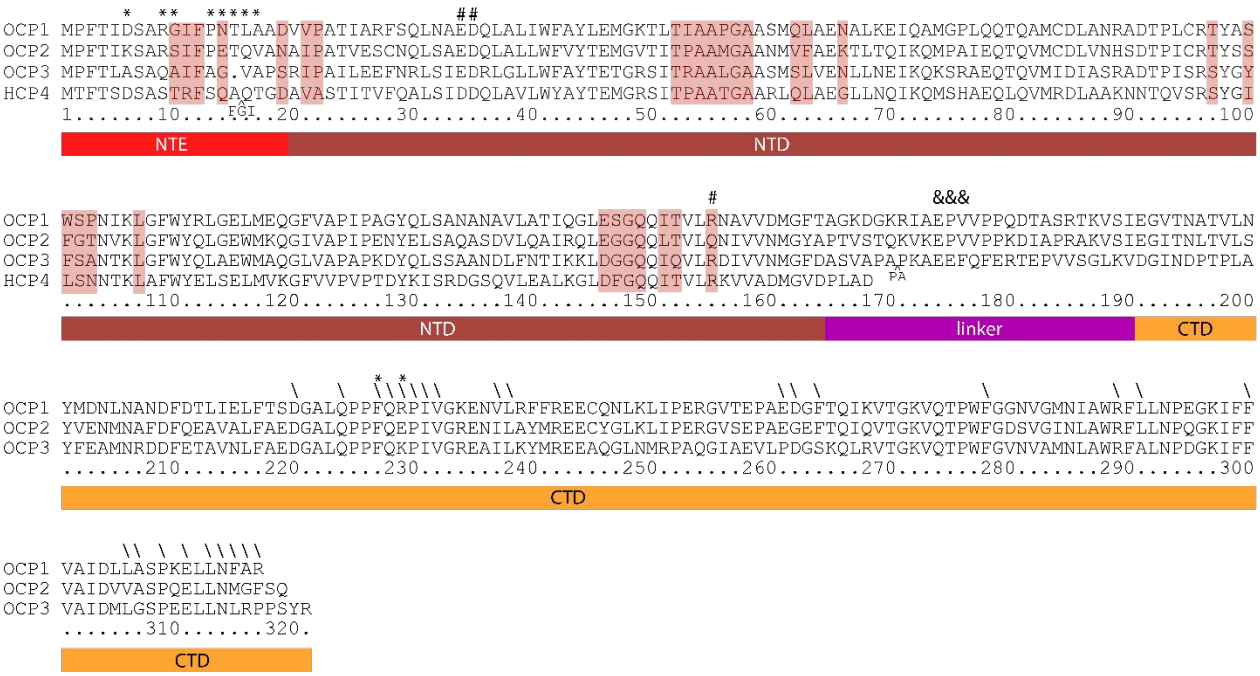

Figure 3. Sequence alignment of *Syn* 6803 OCP1, *Tolypothrix* OCP2, *Gloeobacter* OCP3 and HCP4 from *Anabaena* 7120. PBS binding residues in the *Syn* 6803 OCP-PBS complex are shaded red, OCP^{O} dimer residues marked with an asterisk, OCP^{R} CTD dimer residues with a backslash, TDM residues with a hash and the EPV motif in the linker is denoted with ampersands. Insertion residues for HCP4 (pos. 15) and OCP3 (pos. 170) are shown in small print below the sequence.

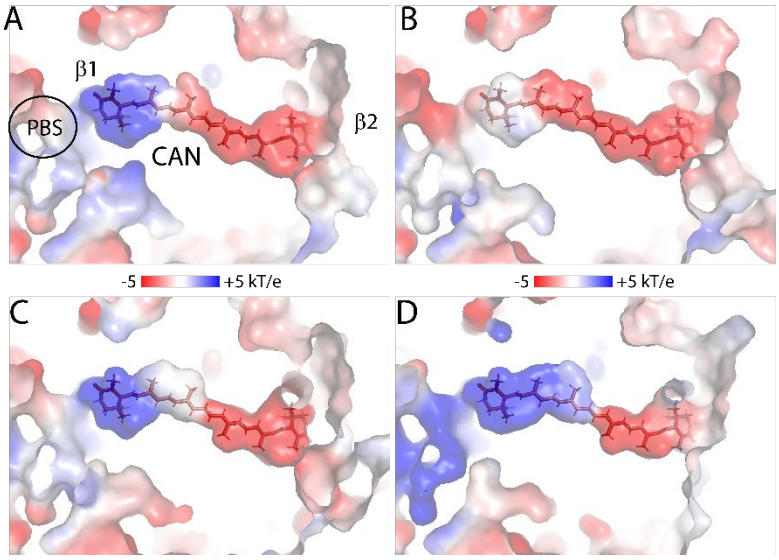

Figure 4. Electrostatic surface representation of the carotenoid environment of (A) the *Synechocystis* sp. PCC 6803 OCP1^{R}-PBS complex cryo-EM structure (B-D) homology models based on the same PBS-OCP complex for (B) *Tolypothrix* sp. PCC 7601 OCP2 (C) *Gloeobacter kilaueensis* JS1 OCP3 and (D) *Anabaena* sp. PCC 7120 HCP4. Figure generated with PyMOL using the APBS plugin.

Acknowledgements

The research in the Kerfeld lab was supported by the Office of Science of the U.S. Department of Energy under awards DE-SC0020606 and DE-FG02-91ER20021, and with support from Michigan AgBio.

Declaration of Interests

The authors declare no competing interests.

REFERENCES

1. Schwieterman, E.W. et al. (2015) Nonphotosynthetic Pigments as Potential Biosignatures. Astrobiology. 15(5): p. 341-361.
2. Hirayama, O. et al. (1994) Singlet oxygen quenching ability of naturally occurring carotenoids. Lipids. 29(2): p. 149-50.
3. Vershinin, A. (1999) Biological functions of carotenoids--diversity and evolution. BioFactors. 10(2-3): p. 99-104.
4. Kirilovsky, D. and C. Kerfeld (2016) Cyanobacterial photoprotection by the orange carotenoid protein. Nature plants. 2(12): p. 16180.
5. Kerfeld, C.A. et al. (2017) Structure, function and evolution of the cyanobacterial orange carotenoid protein and its homologs. New Phytologist. 215(3): p. 937-951.
6. Wilson, A. et al. (2008) A photoactive carotenoid protein acting as light intensity sensor. Proc Natl Acad Sci U S A. 105(33): p. 12075-12080.
7. Adir, N. et al. (2020) The amazing phycobilisome. Biochim Biophys Acta Bioenerg. 1861(4): p. 148047.
8. Kuznetsova, V. et al. (2020) Comparative ultrafast spectroscopy and structural analysis of OCP1 and OCP2 from Tolypothrix. Biochimica Et Biophysica Acta-Bioenergetics. 1861(2).
9. Wu, Y.P. and D.W. Krogmann (1997) The orange carotenoid protein of Synechocystis PCC 6803. Biochimica et Biophysica Acta - Bioenergetics. 1322(1): p. 1-7.
10. Kerfeld, C. (2004) Structure and function of the water-soluble carotenoid-binding proteins of cyanobacteria. Photosynthesis Research. 81(3): p. 215-25.
11. Kerfeld, C. (2004) Water-soluble carotenoid proteins of cyanobacteria. Archives of biochemistry and biophysics. 430(1): p. 2-9.
12. Kerfeld, C. et al. (2003) The crystal structure of a cyanobacterial water-soluble carotenoid binding protein. Structure. 11(1): p. 55-65.
13. Wilson, A. et al. (2006) A soluble carotenoid protein involved in phycobilisome-related energy dissipation in cyanobacteria. The Plant cell. 18(4): p. 992-1007.
14. Wilson, A. et al. (2010) Structural determinants underlying photoprotection in the photoactive orange carotenoid protein of cyanobacteria. The Journal of biological chemistry. 285(24): p. 18364-75.
15. Leverenz, R.L. et al. (2014) Structural and functional modularity of the orange carotenoid protein: distinct roles for the N- and C-terminal domains in cyanobacterial photoprotection. The Plant cell. 26(1): p. 426-37.
16. Leverenz, R.L. et al. (2015) A 12 Å carotenoid translocation in a photoswitch associated with cyanobacterial photoprotection. Science. 348(6242): p. 1463-1466.
17. Gupta, S. et al. (2012) Structure and dynamics of protein waters revealed by radiolysis and mass spectrometry. Proc Natl Acad Sci U S A. 109(37): p. 14882-7.
18. Barends, T.R. et al. (2009) Structure and mechanism of a bacterial light-regulated cyclic nucleotide phosphodiesterase. Nature. 459(7249): p. 1015-8.
19. Swartz, T.E. et al. (2007) Blue-light-activated histidine kinases: two-component sensors in bacteria. Science. 317(5841): p. 1090-3.
20. Kirilovsky, D. and C.A. Kerfeld (2013) The Orange Carotenoid Protein: A blue-green light photoactive protein. Photochemistry and Photobiological Sciences 12: p. 1135-1143.
21. Lechno-Yossef, S. et al. (2017) Synthetic OCP heterodimers are photoactive and recapitulate the fusion of two primitive carotenoproteins in the evolution of cyanobacterial photoprotection. Plant Journal. 91(4): p. 646-656.

22. Bao, H. et al. (2017) Structure and functions of Orange Carotenoid Protein homologs in cyanobacteria. Current Opinion in Plant Biology. 37: p. 1-9.
23. Melnicki, M. et al. (2016) Structure, Diversity, and Evolution of a New Family of Soluble Carotenoid-Binding Proteins in Cyanobacteria. Mol Plant. 9(10): p. 1379-1394.
24. Dominguez-Martin, M.A. et al. (2019) Structural and spectroscopic characterization of HCP2. Biochimica Et Biophysica Acta-Bioenergetics. 1860(5): p. 414-424.
25. Lopez-Igual, R. et al. (2016) Different Functions of the Paralogs to the N-Terminal Domain of the Orange Carotenoid Protein in the Cyanobacterium *Anabaena* sp. PCC 7120. Plant physiology. 171(3): p. 1852-66.
26. Ho, M.Y. et al. (2017) Far-red light photoacclimation (FaRLiP) in Synechococcus sp PCC 7335: I. Regulation of FaRLiP gene expression. Photosynthesis Research. 131(2): p. 173-186.
27. Yang, Y.W. et al. (2023) Functional specialization of expanded orange carotenoid protein paralogs in subaerial Nostoc species. Plant Physiology. 192(4): p. 2640-2655.
28. Yang, Y.W. et al. (2019) Orange and red carotenoid proteins are involved in the adaptation of the terrestrial cyanobacterium Nostoc flagelliforme to desiccation. Photosynthesis Research. 140(1): p. 103-113.
29. Llewellyn, C.A. et al. (2020) Synthesis, Regulation and Degradation of Carotenoids Under Low Level UV-B Radiation in the Filamentous Cyanobacterium Chlorogloeopsis fritschii PCC 6912. Frontiers in Microbiology. 11.
30. Ho, M.Y. and D.A. Bryant (2019) Global Transcriptional Profiling of the Cyanobacterium Synechocystis sp. PCC 9212 in Far-Red Light: Insights Into the Regulation of Chlorophyll Synthesis. Frontiers in Microbiology. 10.
31. Dominguez-Martin, M.A. et al. (2020) Structural analysis of a new carotenoid-binding protein: the C-terminal domain homolog of the OCP. Scientific Reports. 10(1).
32. Muzzopappa, F. et al. (2017) Paralogs of the C-Terminal Domain of the Cyanobacterial Orange Carotenoid Protein Are Carotenoid Donors to Helical Carotenoid Proteins. Plant Physiology. 175(3): p. 1283-1303.
33. Slonimskiy, Y.B. et al. (2019) Light-controlled carotenoid transfer between water-soluble proteins related to cyanobacterial photoprotection. Febs Journal. 286(10): p. 1908-1924.
34. Harris, D. et al. (2018) Structural rearrangements in the C-terminal domain homolog of Orange Carotenoid Protein are crucial for carotenoid transfer. Communications Biology. 1.
35. Moldenhauer, M. et al. (2017) Assembly of photoactive orange carotenoid protein from its domains unravels a carotenoid shuttle mechanism. Photosynthesis Research(doi:10.1007/s11120-017-0353-3): p. 1-15.
36. Bao, H. et al. (2017) Additional families of orange carotenoid proteins in the photoprotective system of cyanobacteria. Nature Plants. 3(8).
37. Slonimskiy, Y.B. et al. (2022) A primordial Orange Carotenoid Protein: Structure, photoswitching activity and evolutionary aspects. International Journal of Biological Macromolecules. 222: p. 167-180.
38. Boulay, C. et al. (2010) Identification of a protein required for recovery of full antenna capacity in OCP-related photoprotective mechanism in cyanobacteria. Proc Natl Acad Sci USA. 107(25): p. 11620-5.
39. Sutter, M. et al. (2013) Crystal structure of the FRP and identification of the active site for modulation of OCP-mediated photoprotection in cyanobacteria. Proc Natl Acad Sci USA. 110(24): p. 10022-10027.
40. Steube, N. et al. (2023) Fortuitously compatible protein surfaces primed allosteric control in cyanobacterial photoprotection. Nature Ecology & Evolution.

41. Petrescu, D.I. et al. (2021) Environmental Tuning of Homologs of the Orange Carotenoid Protein-Encoding Gene in the Cyanobacterium. Frontiers in Microbiology. 12.
42. Bandara, S. et al. (2017) Photoactivation mechanism of a carotenoid-based photoreceptor. Proc Natl Acad Sci USA. 114(24): p. 6286-6291.
43. Gupta, S. et al. (2015) Local and global structural drivers for the photoactivation of the orange carotenoid protein. Proc Natl Acad Sci USA. 112(41): p. E5567-74.
44. Domínguez-Martín, M.A. et al. (2022) Structures of a phycobilisome in light-harvesting and photoprotected states. Nature. 609(7928): p. 835-845.
45. Thierry, E. et al. (2016) Influenza Polymerase Can Adopt an Alternative Configuration Involving a Radical Repacking of PB2 Domains. Molecular Cell. 61(1): p. 125-137.
46. Punjani, A. and D.J. Fleet (2021) 3D variability analysis: Resolving continuous flexibility and discrete heterogeneity from single particle cryo-EM. Journal of Structural Biology. 213(2).
47. Sauer, P.V. et al. (2024) Structural and quantum chemical basis for OCP-mediated quenching of phycobilisomes. Sci Adv.
48. Muzzopappa, F. et al. (2019) Interdomain interactions reveal the molecular evolution of the orange carotenoid protein. Nature Plants. 5(10): p. 1076-1086.
49. Rose, J.B. et al. (2023) Photoactivation of the orange carotenoid protein requires two light-driven reactions mediated by a metastable monomeric intermediate. Physical Chemistry Chemical Physics. 25(48): p. 33000-33012.
50. Andreeva, E.A. et al. (2022) Oligomerization processes limit photoactivation and recovery of the orange carotenoid protein. Biophysical Journal. 121(15): p. 2849-2872.
51. Tsoraev, G.V. et al. (2023) Stages of OCP-FRP Interactions in the Regulation of Photoprotection in Cyanobacteria, Part 1: Time-Resolved Spectroscopy. Journal of Physical Chemistry B.
52. Sluchanko, N.N. et al. (2018) OCP-FRP protein complex topologies suggest a mechanism for controlling high light tolerance in cyanobacteria. Nature Communications. 9.
53. Sluchanko, N.N. et al. (2024) Structural framework for the understanding spectroscopic and functional signatures of the cyanobacterial Orange Carotenoid Protein families. International Journal of Biological Macromolecules. 254.
54. Arcidiacono, A. et al. (2023) How orange carotenoid protein controls the excited state dynamics of canthaxanthin. Chemical Science. 14(40): p. 11158-11169.
55. De Souza, A.P. et al. (2022) Soybean photosynthesis and crop yield are improved by accelerating recovery from photoprotection. Science. 377(6608): p. 851-+.
56. Abed, R.M.M. et al. (2009) Applications of cyanobacteria in biotechnology. Journal of Applied Microbiology. 106(1): p. 1-12.
57. Kloz, M. et al. (2011) Carotenoid Photoprotection in Artificial Photosynthetic Antennas. Journal of the American Chemical Society. 133(18): p. 7007-7015.
58. Piccinini, L. et al. (2022) A synthetic switch based on orange carotenoid protein to control blue-green light responses in chloroplasts. Plant Physiology. 189(2): p. 1153-1168.
59. Dominguez-Martin, M.A. and C.A. Kerfeld (2019) Engineering the orange carotenoid protein for applications in synthetic biology. Current Opinion in Structural Biology. 57: p. 110-117.
60. Chothia, C. and J. Gough (2009) Genomic and structural aspects of protein evolution. Biochemical Journal. 419: p. 15-28.
61. Kondrashov, F.A. et al. (2002) Selection in the evolution of gene duplications. Genome Biology. 3(2).
62. Galhardo, R.S. et al. (2007) Mutation as a stress response and the regulation of evolvability. Critical Reviews in Biochemistry and Molecular Biology. 42(5): p. 399-435.

TEXT BOX—Puzzling Pieces: The HCPs

Protein domains are evolutionary modules/units of structure and function. Once internally duplicated within a genome, a domain is under less selective pressure to remain the same. Phylogenomic analysis suggests that HCP paralogs and the NTD of the OCP arose from gene duplication and subfunctionalization, with PBS quenching just one of the possible functions [23]. Despite structural redundancy, including the binding or carotenoids, the relatively low sequence homology between HCP paralogs within a genome supports the idea that each has a distinct function. For example, *Anabaena* 7120 which contains HCP1-HCP4, with the exception of the HCP1:HCP4 homology (49% identity) all other pairwise alignments are between 23-36% identical. Change in function is typically associated with sequence identity less than 40% [60]. In protein evolution the acquisition of new function is the strongest driver of protein sequence change with paralogs evolving more rapidy than their corresponding orthologs [61]. Furthermore, cellular stress has been shown to correlate with high mutation rates [62]; and the inherent dangers of light harvesting as well as the advent of oxygenic photosynthesis would suggest primitive carotenoproteins like the HCPs strong candidates for mutation and divergence. We suggest that the HCP families evolved discrete functions, with only some, such primitive HCP4, able to bind and quench the PBS. Accordingly, the extant function of, for example, *Anabaena* 7120 HCP1-3 likely reflects their more ancient roles as cyanobacterial carotenoproteins.

Glossary

Carotenoids

Yellow, orange or red pigments that are produced by a variety of organisms and are frequently involved in photosynthesis. The basic building blocks are isoprene units that are linked into chains of conjugated double bonds.

Accessory pigments

Accessory pigments function to extend the range of light that photosynthetic organisms can utilize but are not directly used for the conversion of light energy into cellular energy.

Chromophore

A molecule that can absorb light and emit in the visible spectrum, hence giving it a color. A carotenoid, for example, achieves this property by multiple conjugated double bonds.

Photoprotection

Mechanisms of photosynthetic organisms to prevent damage to the photosystem caused by excess light. In Cyanobacteria, the OCP is the primary means of non-photochemical quenching and dissipates excess energy absorbed by phycobilisomes and prevents damage to the Photosystem II.

Phycobilisome (PBS)

Light harvesting antenna complex of cyanobacteria and certain algae to facilitate light absorption. They are large protein complexes consisting of several hundreds of protein chains that form cylindrical cores and rods. There are two types of proteins, phycobiliproteins that bind chromophores and linker proteins to connect them.

Helical Carotenoid Proteins

Homologs to the N-terminal domain of OCP. They contain a carotenoid in the same position as an activated OCP^{R}.