Structural and spectroscopic characterization of HCP2

Maria Agustina Dominguez-Martin[1], Tomáš Polívka[4], Markus Sutter[1,2], Bryan Ferlez[1], Sigal Lechno-Yossef[1], Beronda L. Montgomery[1,3,5] & Cheryl A. Kerfeld[1,2,3]

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

[2] Environmental Genomics and Systems Biology and Molecular Biophysics and Integrated Bioimaging Divisions, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA

[3] Department of Biochemistry and Molecular Biology, Michigan State University, East Lansing, MI 48824, USA

[4] Institute of Physics, Faculty of Science, University of South Bohemia, Branišovská 31, 370 05 České Budějovice, Czech Republic

[5] Department of Microbiology and Molecular Genetics, Michigan State University, East Lansing, MI 48824, USA

- Address correspondence to: ckerfeld@lbl.gov

The author responsible for distribution of materials integral to the findings presented in this article is: Cheryl A. Kerfeld (ckerfeld@lbl.gov).

Address: MSU-Department of Energy Plant Research Laboratory, Michigan State University, East Lansing, Michigan 48824, USA

Abstract

The Helical Carotenoid Proteins (HCPs) are a large group of newly identified carotenoid-binding proteins found in ecophysiologically diverse cyanobacteria. They likely evolved before becoming the effector (quenching) domain of the modular Orange Carotenoid Protein (OCP). The number of discrete HCP families—at least nine—suggests they are involved in multiple distinct functions. Here we report the 1.7 Å crystal structure of HCP2, one of the most widespread HCPs found in nature, from the chromatically acclimating cyanobacterium *Tolypothrix* sp. PCC 7601. By purifying HCP2 from the native source we are able to identify its natively-bound carotenoid, which is exclusively canthaxanthin. In solution, HCP2 is a monomer with an absorbance maximum of 530 nm. However, the HCP2 crystals have a maximum absorbance at 548 nm, which is accounted by the stacking of the β1 rings of the carotenoid in the two molecules in the asymmetric unit. Our results demonstrate how HCPs provide a valuable system to study carotenoid-protein interactions and their spectroscopic implications, and contribute to efforts to understand the functional roles of this large, newly discovered family of pigment proteins, which to-date remain enigmatic.

Keywords: Helical Carotenoid Protein, photoprotection, cyanobacteria, spectroscopic, crystal structure, carotenoid aggregates.

Introduction

Carotenoids are ancient pigments that carry out a diverse range of biological functions such as providing membrane structural support, contributing to light harvesting or producing ecologically relevant colorations (*1-4*). Despite their highly hydrophobic nature, carotenoids are ubiquitous in aqueous cellular environments. To enhance solubility, carotenoids form complexes with proteins. The binding of carotenoids to proteins also extends their functional repertoire, for example by holding them in defined orientations proximal to other chromophores to serve as accessory pigments in photosynthetic complexes (*5-7*). Carotenoproteins also confer photoprotective functions by absorbing and/or dissipating excess absorbed light energy or by 'quenching' reactive oxygen species (ROS) (*8, 9*).

In cyanobacteria, the Orange Carotenoid Protein (OCP) is a 35 kDa water-soluble photoactive protein (recently reviewed in (*10, 11*)) responsible for a non-photochemical quenching mechanism that enables cells to avoid photodamage and growth inhibition caused by high light or nutrient stresses (*12*). The OCP is structurally and functionally modular (*10, 11, 13*), consisting of a sensor and an effector domain. The first crystal structure of the OCP (*14*) revealed two discrete structural domains: an all-helical carotenoid-binding N-terminal domain (NTD), composed of two discontinuous four-helix bundles, and a C-terminal domain (CTD) with a mixed α-β fold. The NTD (pfam09150) is found only in cyanobacteria, whereas the CTD is a member of the ubiquitous Nuclear Transport Factor-2 (NTF2) superfamily (pfam02136). The carotenoid spans the two protein domains, through the largest interacting surface (the major interface) which is stabilized by a salt bridge. The OCP has been shown to bind various keto-carotenoids when purified from cyanobacteria (3'-hydroxy-echinenone (hECN), echinenone (ECN) and canthaxanthin (CAN)) (*14, 15*).

Upon absorption of blue-light, the OCP converts from a dark-stable orange form, OCP^O, to a light-activated red form, OCP^R (*16*). The color is related to the spectroscopic properties of the absorbing state, S_2, while the lowest excited state S_1 is forbidden for one-photon transitions from the ground state and thereby invisible in absorption spectra (*1*). Yet, the S_1 state is quickly populated from the S_2 state, and in many photosynthetic systems the S_1 state is important for both light-harvesting and photoprotection. Photoactivation is also characterized by the dissociation of the interaction of the NTD and CTD (*17, 18*) and the carotenoid translocates 12 Å to the NTD

(*19*). Thus, the carotenoid of the OCP occupies two distinct positions in the protein depending on the state of photoactivation: resting (OCP^O) and activated (OCP^R). The residues defining these two distinct positions are referred to as carotenoid-protein configurations cpcO and cpcR, respectively. In the OCP^O, the carotenoid is only sparingly solvent accessible, with ca. 96% of its surface area interacting with the protein. However, the solvent accessibility of the carotenoid is increased in OCP^R, with the β1 ring becoming much more solvent exposed after domain separation at the major interface (Figure 1 in (*19*)). The NTD of the OCP^R form mediates excitation quenching of phycobilisomes (PBS) (*13, 20*).

The more than doubling of the number of available cyanobacterial genomes in recent years has enabled the identification of new families of the OCP and homologs to its constituent domains (*20, 21*). At least nine different clades of NTD homologs have been identified across ecophysiologically diverse cyanobacteria (*10, 11*). These paralogs have been named Helical Carotenoid Proteins (HCPs), as they are all predicted to conserve both the all-helical fold of the NTD as well as the residues specific for binding carotenoid (*21*). Homologs to the CTD (CTDHs) have also been found in nearly every genome encoding an HCP (*21*). The principles of protein evolution by gene fusion combined with the taxonomic species distribution among the HCP/NTD phylogeny have led to the suggestion that the NTD of the OCP was derived from an HCP, and was likely to have combined with a CTDH into a single polypeptide as a result of a domain fusion event (*10, 14, 22, 23*).

The first systematic attempt to determine the function of paralogous HCPs focused on the four HCPs (HCP1-4) that are found in *Nostoc* PCC 7120 (hereafter Nostoc) (*24*). Carotenoid binding was confirmed when the genes were overexpressed in a CAN-producing *E. coli* strain. Functionally, HCP4 (*all4941*) was the only paralog shown to bind and quench the PBS, where it induced constitutive fluorescence quenching. Notably, the primary structure of the HCP4 is most similar to the OCP-NTD, consistent with the proposal that it is the HCP most closely related to the ancestor of the NTD (*21*). Nostoc HCP2 and HCP3 (*all3221* and *alr4783*, respectively) were shown to be effective $^1O_2$ quenchers, similar to the OCP-NTD (*24*) and the Red Carotenoid Protein (RCP)(*14*). However, singlet oxygen quenching is not unexpected for carotenoproteins. Ecophysiologically, the reason for the multiplicity of HCP paralogs in a single organism, each

presumably with a specific function, their relative carotenoid binding specificity, and their distinct functional roles in general, remains unclear.

The genome of the filamentous, chromatically acclimating cyanobacterium *Tolypothrix* sp. PCC 7601 (also known as *Fremyella diplosiphon* (UTEX481), hereafter Tolypothrix) encodes an OCP1, an OCP2, three HCPs (HCP1, HCP2 and HCP3) and one CTDH. Here, we present a structural and biophysical characterization of Tolypothrix HCP2 (gene ID: 2501541399). We developed a system to overexpress the carotenoproteins in Tolypothrix, enabling identification of the natively bound pigment. Holo-HCP2 is a monomer in solution and quenches singlet oxygen. The 1.7 Å resolution crystal structure of HCP2 revealed a putative dimer with a stacking of the carotenoid β1 rings; this stacking is strikingly manifested spectroscopically, the visible absorbance peak is shifted 18 nm in the HCP2 crystal relative to solution. Our results provide an ideal model system to spectrally characterize and probe carotenoid-protein interactions.

Materials and methods

1. Overexpression and purification of HCP2 in Tolypothrix

To construct an overexpression vector of HCP2-10x His-tagged, the gene was PCR-amplified from genomic DNA of Tolypothrix, cloned between the constitutively expressed *apcA* promoter (*25*) and a sequence coding for HisTag, and expressed on a pDU1-based replicating plasmid (*26*). The resulting plasmid (pSL88) was introduced into WT Tolypothrix strain SF33, a shortened filament strain (*27*), by conjugation, using the conjugative plasmid pRL443 (*28*) and the methylating plasmid pSL17 (Lechno-Yossef *et al*., unpublished). The HCP2 overexpression strain was grown in buffered BG-11 pH 8.0 medium supplemented with 25 μg/ml of kanamycin.

Cell cultures were grown in liquid BG-11 media at 30 °C, bubbled with air containing 3% CO_2 with stirring and continuous illumination with red light. Cells were harvested at 14,900 *g* for 20 min. The cells pellet was resuspended in 100 ml of 50 mM Tris pH 8.0 /200 mM NaCl containing *DNase*I (*Sigma*) and protease inhibitor cocktail (*Sigma*). The cells were lysed using a French Pressure cell (SLM/Aminco model FA-079) at 1,000 PSI. The lysate was cleared by centrifugation (45 min at 41,600 *g* at 4 °C). Purification of HCP2 was performed by Ni-NTA affinity chromatography (HisTrap Affinity column, *GE Healthcare*) followed by size exclusion chromatography (Superdex 75 pg 16/60 GL, *GE Healthcare*). The isolated HCP2 holoprotein was identified by SDS-PAGE and immunodetection with antibodies against the HisTag.

2. Expression and purification of apo-HCP2 in *E. coli*

To produce Apo-HCP2 in *E. coli*, N-terminal 6x His-tagged Tolypothrix *hcp2* was cloned in a pET28 vector (*Novagene*) and expressed in BL21(DE3) cells. Cells were grown at 37 °C to an OD_{600} ca. 0.6-0.8, followed by induction with 100 μM isopropyl β-D-1-thiogalactopyranoside (IPTG), and incubated in a shaker overnight at 30 °C. The harvested cells were resuspended in 50 mM Tris pH 8.0/200 mM NaCl, containing protease inhibitor cocktail (*Sigma*) and DNase I (*Sigma*) and lysed through a cell disruptor (*Constant Systems* Aberdeenshire, UK) at 35 kPSI. The clarified lysate was applied to a Ni-NTA affinity chromatography followed by a size exclusion chromatography as explained above.

3. Native polyacrylamide gel electrophoresis analysis

The HCP2 samples were subjected to electrophoretic separation in a 15% polyacrylamide gel (*Bio-Rad*) and run at a voltage of 200 V for 60 min at 4 ºC. Gels were stained with Coomassie Brilliant Blue G250 for protein visualization.

4. Extraction and analysis of carotenoids by LC-MS/MS

Samples of purified HCP2 were extracted with acetone at -20 °C for 30 min. After centrifugation (21,000 *g* for 10 min at 4 °C), the supernatant was incubated at -20 °C for 30 min and then centrifuged. This procedure was repeated several times until a white pellet was obtained, which indicated the absence of pigment in the protein sample. The combined supernatants were evaporated to dryness, dissolved in 70 μl of solvent A (methanol-acetonitrile-water [42:33:25 by volume]) and 30 μl of solvent B (methanol-acetonitrile-ethyl acetate [50:20:30 by volume]). Carotenoids were identified by liquid chromatography-mass spectrometry using a *Waters* Xevo G2-XS QTof mass spectrometer interfaced with a *Waters* Acquity UPLC system. 5 µl of sample were injected onto a *Waters* Acquity BEH C18 column (2.1 x 100 mm) and separated using the following gradient: 30% solvent B at the time of injection, linear increase to 100% solvent B over 10 min, 100% solvent B for 2 min, and linear decrease to 30% solvent B for 3 min (total time = 15 min) using a flow rate of 0.3 ml/min and column temperature of 40 °C. Pigments were ionized by APCI (atmospheric pressure chemical ionization) in negative ion mode and identified by a combination of elution time and accurate mass analysis.

5. Analytical Size Exclusion chromatography (SEC)

Analytical SEC to estimate the oligomeric state of HCP2 holoproteins was carried out at 4 °C on a Superdex 75 10/300 GL column (*GE Healthcare*) with 50 mM Tris-HCl, pH 8.0, and 200 mM NaCl as the running buffer. 100 µl of sample was injected onto the column. The elution volume was used for molecular mass estimation using standard curves for column calibration with *Bio-Rad* molecular mass standards.

6. Measurement of ultraviolet-visible spectra in solution

Samples were buffer-exchanged into 50 mM Tris-HCl, pH 8.0, and 200 mM NaCl before spectroscopic measurements. Ultraviolet–visible absorption spectra were collected with a Cary 60 spectrophotometer (*Agilent*).

7. Single crystal microspectrophotometric measurements

Single crystal UV-visible absorption spectra were measured at beam line 9-1 at the Stanford Synchrotron Radiation Lightsource (SSRL) with a modified 4DX microspectrophotometer (*29*). A Hamamatsu light source with deuterium and halogen lamps (model DH-2000-BAL, Ocean Optics Inc.) was used to illuminate the sample with an output from 250 nm to 900 nm. The UV solarization-resistant optical fibers (50 micron and 450 micron) were employed to deliver the light to the sample and transmitting the light through the sample to the detector. The UV-visible absorption spectra were collected on a spectrum analyzer (model QE65000, *Ocean Optics* Inc.) under PC computer control with SpectraSuite (O*cean Optics*, Inc.) software.

8. Crystallization of HCP2

Crystals of Tolypothrix HCP2 were obtained by vapor diffusion in sitting drop experiments at room temperature. 3 μl of protein solution (40 μM, 1 mg/ml in 10 mM Tris-HCl pH 8.0/ 50 mM NaCl) was mixed with 1 μl of reservoir solution containing 0.2 M ammonium iodide and 20% PEG 3,350. Crystals were stabilized by adding 0.2 M ammonium iodide and 20% PEG 3,350 (in reservoir solution) to the drop, mounted on a nylon loop (CrystalCap ALS HT, *Hampton Research*), and then frozen in liquid nitrogen. X-ray diffraction was measured at beam line 5.0.2 of the Advanced Light Source (ALS) at Lawrence Berkeley National Lab.

Diffraction data were integrated with XDS (*30*) and scaled with SCALA (CCP4) (*31*). The HCP2 structure was solved by phenix.MR_Rosetta (*32*) starting with a homology model generated with Phyre (*33*). The resulting solution was refined and rebuilt using phenix.refine (*34*) and COOT (*35*). Statistics for diffraction data collection, structure determination and refinement are summarized in Table 1.

Further analysis of the structure was performed using the following software: PDBePISA at the EBI (http://www.ebi.ac.uk/pdbe/pisa/) and Profunc (https://www.ebi.ac.uk/thornton-srv/databases/profunc/) (*31*).

9. Singlet Oxygen quenching assay by Electron Paramagnetic Resonance (EPR)

A 1 M solution of 2,2,6,6-tetramethyl-4-piperidone hydrochloride (TEMPD-HCl) (*Sigma*) was prepared fresh in Tris buffered saline (100 mM Tris pH 8.0, 400 mM NaCl) before each assay. Each quenching reaction was prepared such that the final solution contained 50 mM Tris pH 8.0, 200 mM NaCl, 0.1 mM methylene blue as a photosensitizer, 100 mM TEMPD-HCl, and a variable amount of HCP2 protein. Reactions were prepared in quartz EPR tubes (2 mm OD) and illuminated with ~1,000 μE/m^{2}s of red light (575 – 725 nm) from an LED source for 3 min. EPR spectra were measured in the dark immediately following illumination. EPR spectra were collected using a Bruker E680 spectrometer at X-Band (9.8 GHz) equipped with a TE 011 mode cylindrical resonator at room temperature (*Bruker* Biospin Corp.). After illumination with red light, EPR samples were placed in the spectrometer in the dark and spectra were collected using 100 kHz modulation frequency, 0.1 mT modulation amplitude, and 40 mW microwave power. The time constant was 82 ms and total scan duration was 335.5 sec. 1024 points were collected for each scan and two scans were averaged for each sample. Built-in Xepr software (*Bruker* BioSpin Corp.) was used to quantify the magnitude of the low field feature of the three-line spectrum of the TEMPD nitroxide radical for each sample by manual double integration. The difference between the area calculated from double integration of each illuminated sample and a control reaction kept in the dark and not excited with red light (containing no protein) was used to quantify the amount of TEMPD radical formed as a function of illumination and HCP2 concentration. A Stern-Volmer plot was generated by dividing the amount of TEMPD radical produced after illumination in the absence of HCP2 by the amount of TEMPD radical produced after illumination in the presence of different concentrations of HCP2. Where present, the error bars represent the standard deviation of the quotient of their respective arithmetic means from three technical replicates. The Stern-Volmer data was fit with a linear regression ($Y_{-HCP2}/Y_{+HCP2} = 1 + K_{SV}[HCP2]$; where Y_{-HCP2} is the EPR signal intensity without HCP2, Y_{+HCP} is the EPR signal intensity with HCP2, and the Stern-

263 Volmer constant K_{SV} is the slope) using IgorPro (*Wavemetrics Inc*, Lake Oswego, OR). and the
264 I50 estimated by calculating the HCP2 concentration for which the TEMPD radical signal is
265 reduced to 50% of its initial value (i.e. when $Y_{-HCP2}/Y_{+HCP2} = 2$).
266
267 **10. Isolation of PBS and *in vitro* PBS fluorescence quenching assays**

268 The Tolypothrix WT cell cultures were grown in liquid BG-11 media at 30 °C, bubbled with
269 air containing 3% CO_2 with stirring and continuous illumination with red light, where no
270 phycoerythrin is expressed (*36*). PBS were isolated as described previously (*36, 37*). PBS were
271 kept in 0.75 M potassium phosphate buffer, pH 7.5. Isolated PBS protein composition was assessed
272 by SDS–PAGE and ultraviolet–visible spectra to confirm its integrity. For fluorescence quenching
273 measurement, HCP2 was added to the isolated PBS at different protein ratios. The fluorescence of
274 PBS was excited at 580 and 660 nm and fluorescence emission spectra were recorded at room
275 temperature (22 °C) from 600 to 800 nm in a fluorimeter (SpectraMax M2, *Molecular Devices*).

276
277 **11. Accession numbers**

278 Sequence data from this article can be accessed from the IMG database
279 (https://img.jgi.doe.gov/) using the gene ID 2501541399 listed in the Supplementary Table 1.
280 Structural coordinates have been deposited in the Research Collaboratory for Structure
281 Bioinformatics (RSCB) Protein Data Bank (http://www.rcsb.org/pdb) under the accession code
282 6MCJ.

283

Results

1. Purification, spectroscopy of HCP2 in solution and identification of the native-bound carotenoid

HCP2 was overexpressed in Tolypothrix from a plasmid with a C-terminal 10x His-Tag using the constitutive promoter *apcA*. The overexpression strain was grown under red-enriched light with 3% CO_2, and it yielded cell extracts that were used to purify the His-tagged HCP2 holoprotein by affinity chromatography followed by size exclusion chromatography (SEC) to yield pure protein (Figure 1A).

The purified HCP2 was pink-violet confirming the binding of the carotenoid with an absorption maximum of 530 nm (A_{530}/A_{280} = 2.7). The absorption spectra of the isolated HCP2 did not exhibit vibrational features and no spectral bands at long wavelengths were identified (Figure 1B). Furthermore, analysis by LC-MS/MS showed that HCP2 bound 100% of the ketocarotenoid CAN.

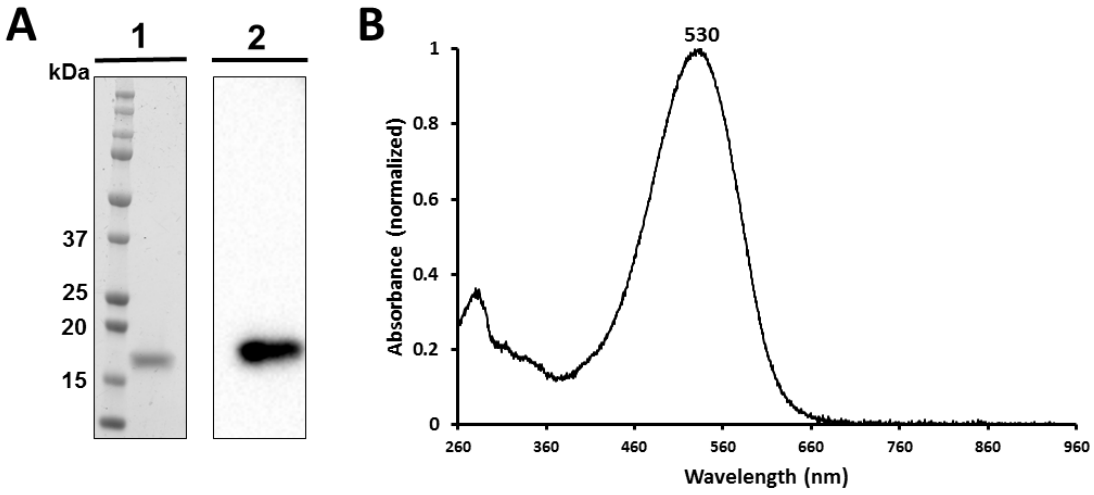
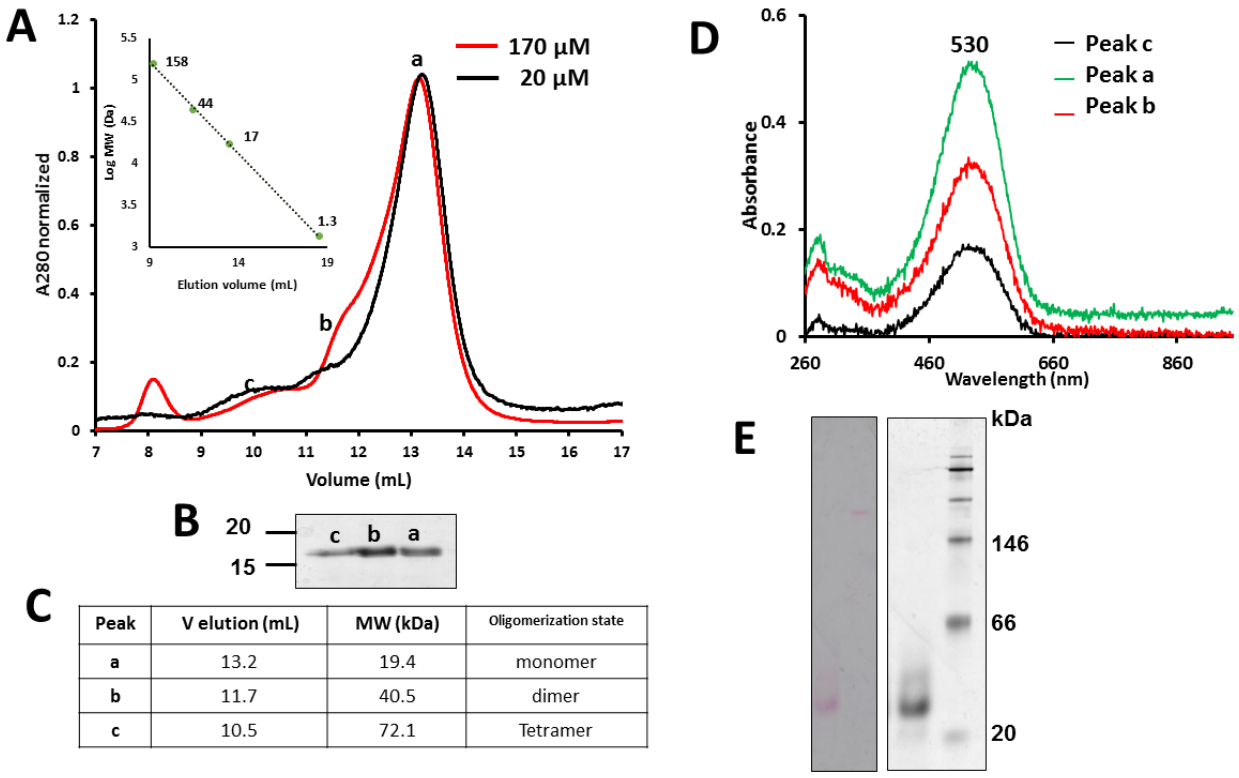


Figure 1: Biochemical and spectroscopic properties of the HCP2 overexpressed in Tolypothrix. (A) (1) Coomassie blue-stained SDS-PAGE. (2) Anti-His Immunodetection. 15 µl of the protein were loaded. (B) UV-visible absorbance spectrum (260-960 nm).

2. Characterization of the quaternary structure of HCP2 in solution

To characterize the quaternary structure of the HCP2 in solution, two different concentrations of protein (20 μM and 170 μM) were subjected to analytical size exclusion chromatography (SEC)

(Figure 2A). The estimated mass for the main peak at both concentrations was 19.4 kDa, which matches the calculated size for a monomeric HCP2 (18.2 kDa). In the chromatogram, two shoulders are observed in the higher concentrated sample (peak b and c in Figure 2A). The calculated sizes for those indicate potential dimer and tetramer states of HCP2, although the main oligomeric state is a monomer. The absorption spectra for all oligomeric forms were similar, exhibiting a maximum peak at 530 nm (Figure 2D). In addition, the HCP2 (20 μM) migrated as a single band in a native gel (Figure 2E).



C

Peak	V elution (mL)	MW (kDa)	Oligomerization state
a	13.2	19.4	monomer
b	11.7	40.5	dimer
c	10.5	72.1	Tetramer

Figure 2: Quaternary structure of HCP2 in solution. (A) size-exclusion chromatography for HCP2 at two concentrations, 20 and 170 μM. (B) Coomassie stained SDS-PAGE of the fractions collected from the chromatography. (C) Summary of the analytical SEC data (D) UV-visible spectra of the fractions from the peaks in the chromatogram. (E) Unstained (left) and Coomassie blue stained (right) native gel. 15 µl of protein were loaded.

3. Crystal Structure of the HCP2

The structure of the HCP2 was determined at a resolution of 1.7 Å. The asymmetric unit contains two HCP2 molecules. Clear electron density is observed for the non-covalently bound CAN for both chains of the HCP2 structure (Supplementary Figure 1). The HCP2 is composed of eight α-helixes (A-H) organized as an all-helical domain (Figure 3, 4). The helical domain can be divided into two four-helix bundles (Figure 4). The two bundles are composed of discontinuous segments of the protein. Bundle 1 is formed by helices A, B, G and H and bundle 2 is formed by helices C, D, E and F. Both bundles are joined through a fourteen residues interbundle loop (between helices f and g). The two bundles form the binding site of the CAN between helices H and B (from bundle 1) and F and D (from bundle 2). The two bundles are structurally similar and can be superimposed with a root-mean-square (RMSD) C_α of 0.940 Å (228 aligned atoms). The carotenoid is largely buried within the protein, with only 15% of its surface exposed to solvent. The areas of solvent exposure are concentrated near the two terminal rings β1 and β2. The protein surface surrounding these rings are of distinctly different charge: the β1-adjacent surface (β1 surface) is predominantly positively charged, whereas the charge on the protein surrounding the β2 ring (β2 surface) is mostly negative (Figures 4D and Supplementary Figure 7). Furthermore, the dimerization in the crystal reduces the solvent accessibility to 5.4% (Figure 4A).

The two HCP2 molecules in the asymmetric unit interact across the β1 surface (Figure 4A). The interaction along β1 surface includes a coplanar stacking of the 4-keto rings of the CAN molecules, the distance between the two rings is about 5 Å (Figure 4B). Each monomer has approximately 4.6% of its surface involved in the formation of the putative dimer (871 $Å^2$ buried area from the total surface 18,800 $Å^2$). 10 out of 18 residues involved in the dimerization (Figure 3) are conserved in the HCP2 clade (Supplementary Figure 2). Of the 13 hydrogen bonds, only one (Asn138-Asn138) involves residues highly conserved in HCP2 (Figure 3 and Supplementary Figure 2). The calculated ΔG of interaction is ca. - 1.6 kcal/mol indicative of a weak interaction.

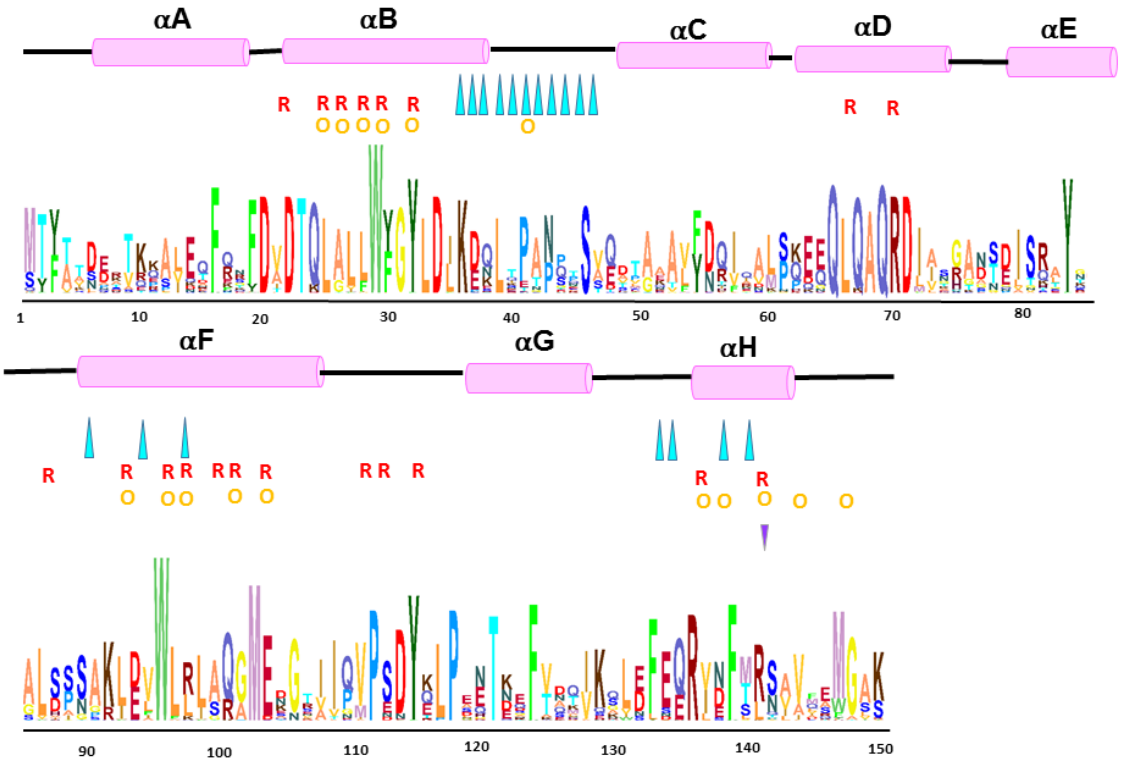


Figure 3: Secondary Structure and Sequence Conservation of the HCP2. Secondary structure cartoon showing the conserved residues within 4 Å of the carotenoid that correspond to the cpcR (R) and cpcO (O) configuration. Positions of residues that are involved in the putative dimerization in the crystal are highlighted with turquoise triangles. The purple triangle indicates the conserved residue R141 (R155). HMM is adapted from (*21*).

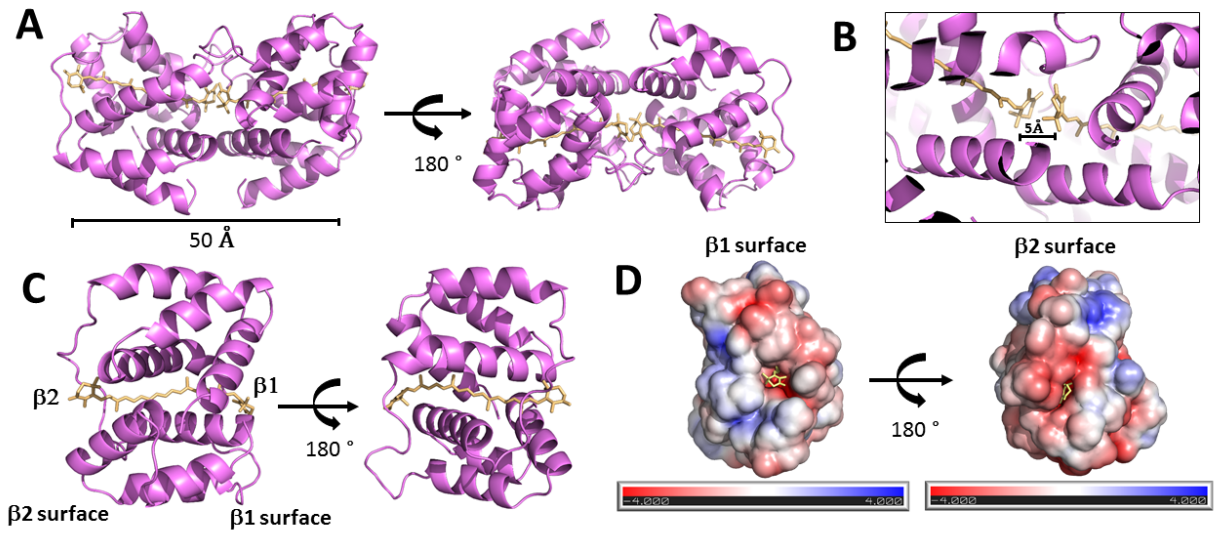


Figure 4: Structure of the HCP2 Dimer (A) Ribbon diagram of the dimer in the HCP2 asymmetric unit. The carotenoid is shown in yellow sticks. (B) Close-up view of the interaction between β1 rings in the dimer. (C). Structure of the HCP2 monomer, showing the β1 and B2 surfaces. (D) Charge distribution of HCP2 with the surface colored by electrostatic potential from –4 kT/e (red) to 4 kT/e (blue). HCP2 is oriented as 90° from panel c.

The 24 residues within 4 Å of the carotenoid are listed in Supplementary Table 2. Of these 20 residues are highly conserved in the HCP2 clade (Figure 5). The carotenoid in HCP2 is surrounded by seven aromatic residues (six are absolutely conserved) (Supplementary Figure 3B). In HCP2, the intraprotein aromatic-sulfur interactions between Trp28-Met103, Trp96-Met147 and Tyr115-Met111 flank the central portion of the polyene. The Trp28 and Met147 are in bundle 1 and Met103 and Trp96 are in bundle 2, therefore aromatic-sulfur interactions connect the two bundles, while the Tyr115-Met111 interaction is in the interbundle linker.

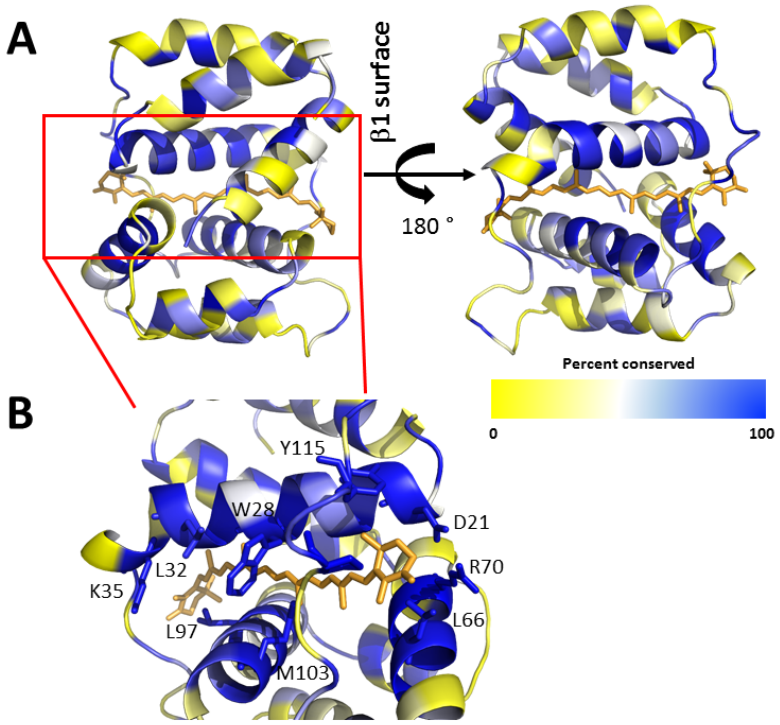


Figure 5: Amino acid sequence conservation of HCP2 mapped on Tolypothrix HCP2 structure. (A) Cartoon model of secondary structure elements colored by amino acid conservation, from least conserved (yellow) to most conserved (blue) in accordance with the color bar. (B) Zoomed view of the most conserved region with highly conserved residues labeled.

4. Spectroscopic properties of crystalline HCP2

Figure 6 shows the absorption spectra of CAN in benzene, of HCP2 in solution and in the HCP2 crystal. The absorption maximum of HCP2 in solution is at 530 nm, which is 50 nm red-shifted from the absorption maximum of CAN in benzene (480 nm). The origin of this significant red-shift is partially due to protein-induced change of CAN configuration. In solution the β1 and β2 rings of CAN take the s-cis configuration, which represents the lowest energy configuration of carotenoids with conjugation extended to the end rings (*38*). Binding to the HCP2 twists the β1 ring to the s-trans configuration (Figure 4), effectively prolonging the conjugation length, and resulting in the observed red shift of absorption spectrum.

The absorption spectrum of the HCP2 crystal exhibits significant differences from the absorption spectrum of HCP2 in buffer (Figure 6). The spectrum was corrected for scattering by subtracting the scattering curve, as shown in Supplementary Figure 4. The main absorption band

of the HCP2 crystal has a peak at 548 nm, further red-shifted relative to HCP2 in buffer. The major difference between HCP2 absorption spectrum in buffer and in the crystal are the spectral bands in the 650-960 nm region that are observed exclusively in HCP2 crystals. For carotenoids in solution as well as in proteins in a buffer, this region is typically free from any spectral bands, as the lowest energy transition is forbidden for one photon processes, thus being invisible in absorption spectra (*39*).

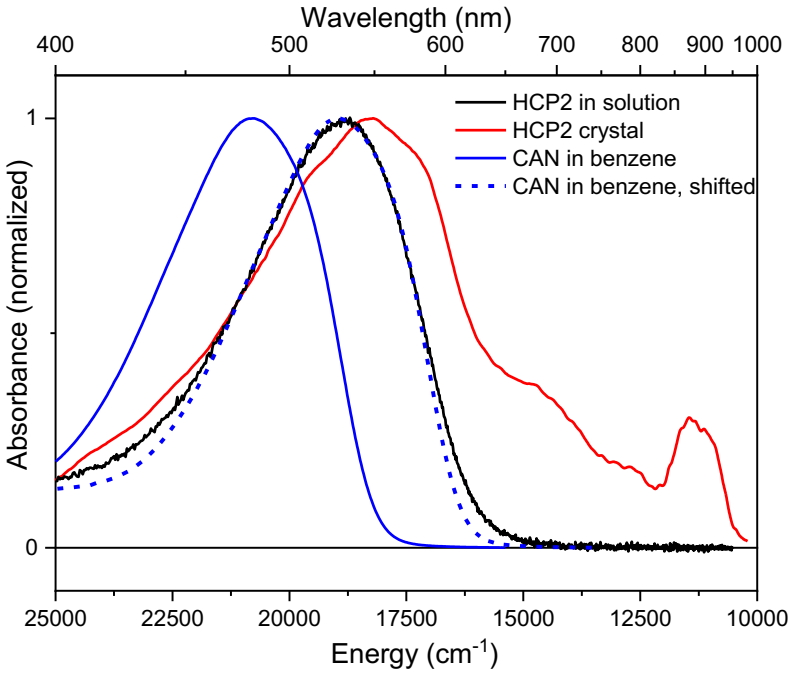


Figure 6. **Absorption spectra of HCP2**. Absorption spectra of HCP2 in buffer (black) and HCP2 crystal (red) compared to the absorption spectrum of CAN in benzene (blue). The blue dashed line represents the absorption spectrum of CAN in benzene shifted to overlap with absorption of HCP2 in buffer. All spectra are normalized to absorption maximum. The absorption spectra are presented with a linear energy scale (cm^{-1}) to be able to directly compare the widths of various spectral bands.

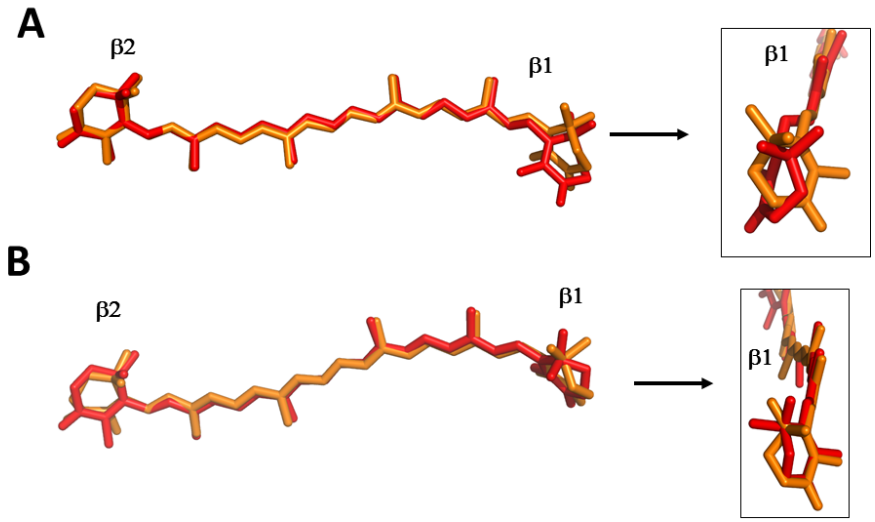


Figure 7: Superposition between CAN carotenoids from different structures. (A) Superposition of the CAN molecule from the structures HCP1 (pdb: 5FCX) and HCP2 (pdb: 6MCJ). (B) Superposition of the CAN from the structures of RCP (pdb: 4BX4) and HCP2 (pdb: 6MCJ).

5. Does HCP2 function similarly to the OCP NTD?

We tested if HCP2 could quench PBS fluorescence in vitro using different PBS:HCP2 ratios (from 1:40 to 1: 1,640) to establish if the function of the HCP2 was in non-photochemical quenching. We tested concentration ranges that reflect less avid quenching (Supplementary Figure 5) but even across this concentration range HCP2 is not able to quench PBS. We also measured the singlet oxygen ($^1O_2$) quenching activity of HCP2 since this has been observed as a secondary function of the OCP (*14, 40*) Electron paramagnetic resonance (EPR) spin trapping was applied for $^1O_2$ detection using 2,2,6,6-tetramethyl 4-piperidone hydrochloride (TEMPD-HCl). When this nitrone reacts with $^1O_2$, it is converted into the stable nitroxide radical which is paramagnetic and detectable by EPR spectroscopy. The production of $^1O_2$ was induced with the illumination of the photosensitizer methylene blue. Figure 8A shows the typical EPR signal of the nitroxide radical obtained after 3 min of illumination (1000 μmol quanta m^{-2} s^{-1}) with or without illumination of a solution containing methylene blue and TEMPD-HCl. When this reaction was incubated in presence of increasing concentrations of purified HCP2, a decrease in the EPR signal was detected

(Figure 8A), which indicates that HCP2 acts as a $^1O_2$ quencher. The Stern-Volmer plot shows a direct relation between the concentration of HCP2 and the $^1O_2$ quenching activity (Figure 8B) with a calculated I_{50} of 5.4 μM.

To investigate the relative contribution of the carotenoid and protein binding to the quenching of $^1O_2$, we tested the HCP2 apoprotein under the same conditions. The apoprotein quenches $^1O_2$ only sparingly (I_{50} of 53.1 µM) (Figure 8B).

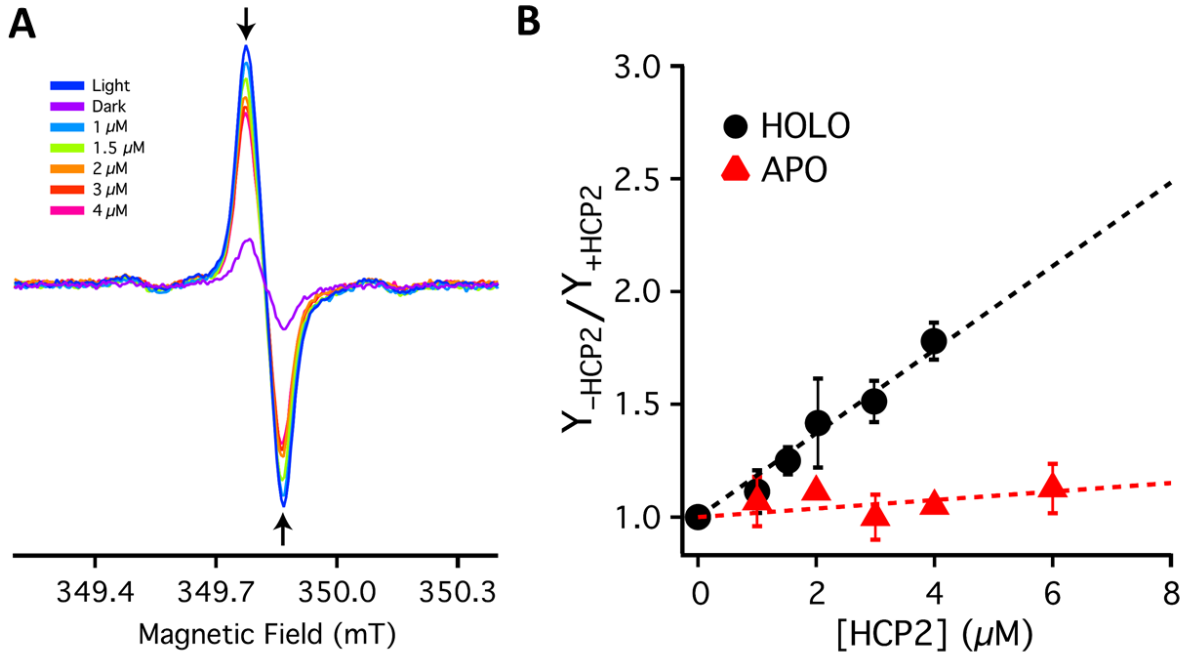


Figure 8: $^1O_2$ quenching activity of HCP2. (A) Representative EPR spectra of the low-field TEMPD radical signal following illumination in the presence of different concentrations of Holo-HCP2. (B) Stern-Volmer plots of singlet oxygen quenching by holo (black squares and dotted line) and apo (red triangles and dotted line) HCP2 measured using the yield of TEMPD radical in the absence (Y_{-HCP2}) or presence (Y_{+HCP2}) of HCP2. Error bars represent the standard deviation from the mean of three technical replicates. Fits of the data with a straight line with a Y-intercept of 1 are plotted as dashed lines for samples containing holo (black) and apo (red) HCP2.

Discussion

HCP2 is one of the most widespread subtype of HCPs in cyanobacteria. It typically co-occurs with three paralogs (HCP1, HCP3 and HCP4). This hints at critical functions for these proteins, and possible distinctive roles that rely on subtle differences in the carotenoprotein interactions. There are seven cyanobacterial strains that, like Tolypothrix, encode HCP1-HCP3 and CTDH but lack HCP4 (*Nostoc* ATCC 29133, *Scytonema* PCC 7110, *Calothrix* PCC 7102, *Synechocystis* PCC 7509, *Leptolyngbya* 2LT21S03, *Leptolyngbya* JSC-1 and Tolypothrix PCC 7601). Notably, all of these genomes, except *Leptolyngbya* 2LT21S03, encode an OCP gene (either OCP1, OCP2 or both). In addition, in the Tolypothrix genome, HCP2 is adjacent to HCP3, and proximal to the CTDH. This arrangement however is not generally conserved among cyanobacteria. Based on RNAseq analysis in Tolypothrix, full-length OCPs (OCP1 and OCP2) and the HCP paralogs are differentially regulated. For instance, OCP1 and OCP2 are upregulated in the ΔrcaE mutant that cannot chromatically acclimate, whereas the expression of the HCP paralogs is either down or not significantly impacted (*41*). We therefore conclude that HCPs and the OCPs have distinct roles.

Previous functional investigation of HCPs used proteins prepared in an *E. coli* expression system that only produced CAN (and also contained large amounts of co-purified apoprotein), therefore the identity of the natively bound carotenoid was unknown. There is a precedent for HCPs to bind a range of carotenoids: OCP and HCP1 have both been shown to bind various carotenoids including CAN, ECN, β-carotene, 3'-hECN and the carotenoid glycoside deoxymyxoxanthophyll (*15, 21*). The production of HCP2 in Tolypothrix enabled us to purify 100% holoprotein and to identify the carotenoid it selectively binds in vivo; HCP2 binds exclusively CAN. The selectivity of CAN by HCP2 is likely functionally relevant. Tolypothrix produces β-carotene, ECN, CAN and myxoxanthophyll (*42*). Interestingly, when the cells were grown under white or red light CAN was only about 10% of the total carotenoid (*42*). Our cells were grown under enriched red light; therefore, we can exclude the idea that the selectivity of HCP2 for CAN is due to the abundance of the pigment in the cell. In general, CAN levels in cyanobacteria are highly increased under high light or UV-B stress conditions (*43-45*). We suggest

that the selectivity of CAN by HCP2 is directly related to the function of the protein but that function remains enigmatic.

HCPs are homologs to the NTD of the OCP, and comprise at least nine distinct families of HCPs(*21*). While our study is focused on structural and spectroscopic characterization of this carotenoprotein, we tested if our HCP2 preparation, containing the natively bound carotenoid, functions similarly to the NTD of the OCP. Unlike the OCP, HCP2 does not quench PBS in the range 1:40 to 1:1,640 (Supplementary Figure 5). The OCP has a secondary protective function, the quenching of reactive oxygen species (*14, 40*). The Nostoc HCPs prepared in *E. coli* were shown to quench singlet oxygen variously (*24*). In that study the relative contribution of the protein and the carotenoid to singlet quenching was not evaluated. Given that the HCPs are enriched in aromatic amino acids, which are known to be important for quenching, we measured the Tolypothrix HCP2 $^1O_2$ quenching activity in vitro, comparing the apo- and holoprotein. The holoprotein exhibited significantly higher $^1O_2$ quenching activity than apoprotein, underlining the importance of CAN, yet Tolypothrix HCP2 showed a lower $^1O_2$ quenching activity than the *E. coli* derived Nostoc HCP2 (I_{50} of 5.4 μM versus 1.2-1.5 μM, respectively) (*24*). Considering both HCP2s contain the same carotenoid, we could not attribute the difference in I_{50} to the pigment content. Among carotenoids, CAN is a relatively good quencher of singlet oxygen (*46*) due to its increased conjugated double-bond system. Moreover since CAN is a 4,4'-diketo β-carotene derivative, it is also more stable against oxidation than carotenoids lacking the keto groups (*47*). It must be noted though that other Nostoc HCPs (expressed in the CAN-producing *E.coli* strain) showed a lower capacity for singlet oxygen quenching than Tolypothrix HCP2, indicating that the presence of CAN is not the key factor determining the efficiency of singlet oxygen quenching.

To date, the only structure of an HCP1 determined is 2.5 Å resolution from Nostoc (*21*), which contained a mixture of carotenoids that precluded a precise analysis of carotenoid-protein interactions. Notably, both HCP2 and HCP1 contain dimers in the asymmetric unit. Moreover, in both dimers there is a coplanar stacking of the β1-ring of the CAN molecules ~5 Å apart. Furthermore, the structure of the NTD with bound carotenoid (also known as RCP (pdb: 4XB4), also contains a similarly oriented (across the β1 face) dimer in the asymmetric unit. The HCP1 and HCP2 monomers are structurally similar and can be superimposed with a C-α RMSD of 0.88 Å

(over 714 aligned atoms) (Supplementary Figure 6). HCP1 and HCP2 dimers share some features, but also key differences: both bury approximately the same amount of surface area (876 $Å^2$ in HCP1) but the dimerization is mediated by six hydrogen bonds in HCP1 but only two of them involve highly conserved residues (Thr50-Asn111 and Ala54-Glu57) (*21*). Moreover, in HCP1, the loop connecting the α-C and α-D helices is proteolyzed. Given the weak binding interface, we question the physiological relevance of the HCP2 dimer. An electrostatic rendering of the β1 surface shows a sizeable nonpolar region (Figure 4 and Supplementary Figure 7), consistent with the hypothesis that the protein buries the β1 surface by interacting with some other binding partner in vivo. The most obvious prospective interaction partners for HCP2 are the CTDH or the PBS. However, we were unable to show an interaction between the HCP2 and CTDH (data not shown) nor between HCP2 and the PBS under the conditions that we tested. Therefore, we speculate that HCP2 has an unknown binding partner, which most likely will be related to one of its functions.

In addition to the striking observation of the carotenoid ring stacking in HCPs, the HCP2 structure and the absorption spectra of the crystals provide a valuable model for probing carotenoid-protein interactions. HCP2 in buffer exhibited a maximum peak at 530 nm (Figure 1A). Similar absorption maxima (525 nm) are reported for RCP(*13, 19*) or HCP1(*21*) binding CAN, suggesting comparable effects of the binding site on spectroscopic properties of CAN in these three proteins. Small differences in the absorption maxima of HCP1 and HCP2 could be related to slightly different rotation of the end ring in s-trans configuration (β1) that are rotated in respect to the main conjugation plane by 60° and 34° degrees for HCP1 and HCP2, respectively (Figure 7). The more planar orientation of the β1 ring in HCP2 may account for the slightly larger red-shift of HCP2 absorption spectrum.

However, the spectroscopic properties of HCP2 in solution differ markedly from those in the crystal. Although the binding cleft induces a significant red shift of the absorption maximum of CAN, the width and overall shape of the CAN absorption band remains nearly the same in solution and in HCP2 as evidenced by the artificially shifted absorption spectrum of CAN in solution to match the absorption maximum of CAN in HCP2 (dashed blue line in Figure 6). Since the width and shape of the carotenoid absorption band is related to a conformational disorder of terminal rings(*1*), it implies that the terminal rings of CAN in HCP2 still have some freedom to

rotate as they do in solution. Thus, although the mean values of the ring torsions obtained from crystal structure are different than in solution (*38*), the conformational disorder remains comparable to that in solution. This is clearly caused by the fact that the terminal rings of CAN in HCP2 are exposed to buffer (Figure 4), making it comparable to solution. This is in striking contrast to OCP structure, in which the terminal rings are deep in the binding cleft and locked in specific configuration by hydrogen bonds (*14*). Therefore, OCP exhibits a narrower absorption spectrum with clearly resolved vibrational bands indicating significantly diminished conformational disorder compared to HCP2. The loss of vibrational structure in HCP2 (and also in RCP and HCP1) is a clear spectroscopic sign of less constrained terminal rings in comparison to OCP.

Previously, different spectroscopic features have been reported for OCP crystals (*48, 49*). A red-shift, albeit smaller, of crystal vs. buffer absorption spectrum has been recently reported also by Bandara et al. for OCP (*48*). The even more red-shifted spectrum of the HCP2 crystal may be related to more constrained terminal rings in the crystal, which may take slightly different orientation as in the buffer as both terminal rings are at the surface of the protein (Figure 4) and thus exposed to buffer. Thus, the packing of HCP2 in crystal may mildly affect the terminal ring orientation resulting in the observed red-shift. The effect of the crystal packing on the CAN terminal rings is also demonstrated by the clearly-resolved vibrational bands of the crystal absorption spectrum. The energy distance between the sub-bands in the main absorption bands is around 1200 cm^{-1}, nicely matching the vibrational separation of carotenoid vibrational bands in solution (*1*). To restore the resolution of vibrational bands, the conformational disorder of the terminal rings must be diminished, implying that the packing of HCP2 in the crystal put some constraints on the movement of the terminal rings, locking them in a certain defined orientation which eventually results in the observed vibrational bands in the absorption spectrum. When HCP2 is dissolved in a buffer, the constraints are released, resulting in the featureless absorption spectrum as shown in Figure 6.

Apart from a single example of a peridinin deoxyderivative (*50*), the lowest energy transition has never been observed in absorption spectra of carotenoids. Recently, however, Bandara et al. reported absorption spectrum of an OCP binding echinenone crystal with a pronounced spectral

band at 800 nm (*48*) and they suspected this band may be related to the formally forbidden S_0-S_1 band of echinenone. Here, the spectrum of HCP2 crystal is even more complicated, as there are a few spectral bands within the 650-950 nm spectral region. The possible origin of the bands can be, at least partially, traced back to the arrangement of CAN molecules in the HCP2 crystal. The crystal packing exhibits a clear dimerization of CAN molecules, in the form of the head-to-tail (J-type) aggregate. J-type aggregates (or dimers in this case) are typically composed of molecules forming a chain. In the arrangement in the HCP2 crystal, the β-ring of one CAN molecule (head) is only 5 Å from the β-ring of the other CAN molecule (tail). Due to this very short distance, the two carotenoid molecules do not behave as individual entities, but rather form a "supermolecule" with spectroscopic properties that differ from those of individual molecules. This is a well-known effect described for carotenoids in hydrated solvents, in which the carotenoids tend to form aggregates characterized by new spectral bands, with positions dependent on the type of aggregation. J-aggregates of carotenoids generate spectral bands that are significantly red-shifted from the main absorption band (*51*). There are no studies of CAN aggregates, but aggregation of astaxanthin (which differs from CAN only by the presence of hydroxyl groups at the terminal rings) in hydrated solvents has been extensively studied (*52, 53*). It is known that astaxanthin J-aggregates can generate spectral bands red-shifted by nearly 3000 cm^{-1} from the lowest main absorption band. Here, the shoulder peaking around 680 nm is shifted from the main absorption maximum by ~3500 cm^{-1}, which is a realistic aggregation-induced shift in a well-defined J-type dimer that appears in the crystal packing. The two CANs in the HCP2 dimer are arranged in the head-to-tail orientation with ring-to-ring distance of 5 Å, which can also induce some π-π stacking of the ring, further enhancing the interaction between the CANs in the dimer. Thus, we propose that the 680 nm band is due to dimerization of HCP2 in the crystal that brings two CAN molecules into a close contact, forming J-type dimer. This dimer exists only in the crystal, so no such band is observed when HCP2 is dissolved in a buffer.

The most red absorption band peaking at 890 nm can, however, can hardly be explained by the CAN aggregation in the crystal. It has a certain resemblance to the spectral band reported by Bandara et al. (*48*), but there are arguments against assigning it to the S_0-S_1 transition. First, the spectral band is far too narrow to be associated with such a transition. From weak S_1 fluorescence data that were reported for a few carotenoids we know that the spectral profile of the S_0-S_1 transition should have a width comparable to that of the main absorption band (S_0-S_2 transition)

603 (*54*). It is evident from Figure 5 that it is clearly not the case here as the 890 nm band is much
604 narrower than the main absorption band. Further, the 890 nm absorption band has most likely too
605 low energy to be assigned to the S_0-S_1 transition. We know from the fluorescence measurements
606 that the S_0-S_1 emission band peaks at the second (0-2) vibrational band (*54*). Then, assuming the
607 mirror image between the S_1-S_0 emission and hypothetical S_0-S_1 absorption band, it would place
608 the S_1 energy significantly below 11000 cm^{-1}, which is nearly 4000 cm^{-1} lower than the S_1 energy
609 experimentally observed for hydroxyechinenone in OCP (*55*). Even though the expected S_1 energy
610 of CAN should be lower than that of hydroxyechinenone due to the one additional conjugated keto
611 oxygen of CAN, based on the measured S_1 energies of various carotenoids, the downshift of the
612 S_1 energy when going from hydroxyechinenone to CAN should not exceed 500 cm^{-1} (*1*). Thus, the
613 890 nm spectral band cannot be associated with the S_0-S_1 transition. Instead, this band could be
614 due to long-range arrangement of carotenoids in the HCP2 crystal as it is observed solely for
615 crystals, never for HCP2 in buffer.
616
617 Our data showing that absorption spectrum of HCP2 crystal differs from that measured for
618 HCP2 in buffer is in line with recent theoretical investigation of another carotenoid-binding
619 protein, crustacyanin (*56*). This protein binds two molecules of carotenoid astaxanthin. There is a
620 very large (>100 nm) red shift of absorption maximum upon carotenoid binding to crustacyanin
621 apoprotein (*57*); the origin of this shift has been a matter of considerable debate. Recent
622 calculations showed that to explain the color tuning by the protein it is necessary to include
623 dynamical effects and temperature-dependent fluctuations (*56*). These authors showed that certain
624 key parameters determining spectroscopic properties, such as bond length changes in astaxanthin,
625 may differ significantly in crystal and in a less constrained environment that allows for
626 temperature-induced fluctuations. We suggest it is likely that similar effects occur also for HCP2
627 resulting in the observed difference between absorption spectra of HCP2 crystal and HCP2 in
628 solution.
629
630 **Conclusions**
631 HCP2, indeed potentially the entire family of the HCPs, offer a useful model system to probe
632 carotenoid protein interactions in different environments; the subtle differences among the HCPs
633 undoubtedly reflects distinctive functions, in perhaps distinct subcellular environmental

conditions. Our structural and biophysical characterization of HCP2 provides an important step in understanding the structural basis of function in the large, poorly characterized family of water soluble carotenoid-binding proteins.

643 **Table 1:** Data collection and refinement statistics
644

Data collection	**HCP2**
Resolution range (Å)	44.26 - 1.71 (1.80 - 1.71)
Space group	P 21
Unit cell dimensions	35.2 71.6 56.3 Å 90 92.3 90 °
Total reflections	98580 (13974)
Unique reflections	29788 (4255)
Multiplicity	3.3 (3.3)
Completeness (%)	98.8 (97.6)
Mean I/sigma(I)	15.7 (4.1)
R-merge	0.036 (0.273)
R-meas	0.052 (0.341)
$CC^{1/2}$	0.999 (0.893)
Refinement	
Number of reflections	29761 (2931)
Number of reflections used for R-free	2009 (191)
R-work (%)	16.7 (25.7)
R-free (%)	19.9 (30.4)
Number of non-hydrogen atoms	2576
macromolecules	2300
ligands	108
solvent	290
Protein residues	290
RMS (bonds, Å)	0.005
RMS (angles, °)	1.00
Ramachandran favored (%)	99.3
Ramachandran allowed (%)	0.7
Ramachandran outliers (%)	0
Clashscore	2.78
Average B-factor ($Å^2$)	27.6

645 Statistics for the highest-resolution shell are shown in parentheses.
646
647
648

REFERENCES

1. Polivka, T., and Sundstrom, V. (2004) Ultrafast dynamics of carotenoid excited States-from solution to natural and artificial systems, *Chem Rev 104*, 2021-2071.
2. Frank, H. A., and Cogdell, R. J. (1996) Carotenoids in photosynthesis, *Photochem Photobiol 63*, 257-264.
3. El-Agamey, A., Lowe, G. M., McGarvey, D. J., Mortensen, A., Phillip, D. M., Truscott, T. G., and Young, A. J. (2004) Carotenoid radical chemistry and antioxidant/pro-oxidant properties, *Arch Biochem Biophys 430*, 37-48.
4. Landrum, J. T., and Bone, R. A. (2001) Lutein, zeaxanthin, and the macular pigment, *Arch Biochem Biophys 385*, 28-40.
5. Hickman, M., and Schweger, C. E. (1991) Oscillaxanthin and myxoxanthophyll in two cores from Lake Wabamun, Alberta, Canada, *Journal of Paleolimnology 5*, 127-137.
6. Hickman, M., and Schweger, C. E. (1996) The Late Quaternary palaeoenvironmental history of a presently deep freshwater lake in east-central Alberta, Canada and palaeoclimate implications, *Palaeogeography Palaeoclimatology Palaeoecology 123*, 161-178.
7. Züllig, H. (1989) Role of carotenoids in lake sediments for reconstructing trophic history during the late Quaternary, *Journal of Paleolimnology 2*, 23-40.
8. Steiger, S., Schäfer, L., and Sandmann, G. (1999) High-light-dependent upregulation of carotenoids and their antioxidative properties in the cyanobacterium *Synechocystis* PCC 6803, *Journal of Photochemistry and Photobiology 52*, 14-18.
9. Zhu, Y., Graham, J., Ludwig, M., Xiong, W., Alvey, R., Shen, G., and Bryant, D. (2010) Roles of xanthophyll carotenoids in protection against photoinhibition and oxidative stress in the cyanobacterium *Synechococcus* sp. strain PCC 7002, *Arch Biochem Biophys 504(1)*, 86-99.
10. Kerfeld, C. A., Melnicki, M. R., Sutter, M., and Dominguez-Martin, M. A. (2017) Structure, function and evolution of the cyanobacterial orange carotenoid protein and its homologs, *New Phytol 215*, 937-951.
11. Bao, H., Melnicki, M. R., and Kerfeld, C. A. (2017) Structure and functions of Orange Carotenoid Protein homologs in cyanobacteria, *Curr Opin Plant Biol 37*, 1-9.
12. Wilson, A., Punginelli, C., Gall, A., Bonetti, C., Alexandre, M., Routaboul, J., Kerfeld, C., van Grondelle, R., Robert, B., Kennis, J., and Kirilovsky, D. (2008) A photoactive carotenoid protein acting as light intensity sensor, *Proc Natl Acad Sci USA 105*, 12075-12080.
13. Leverenz, R. L., Jallet, D., Li, M. D., Mathies, R. A., Kirilovsky, D., and Kerfeld, C. A. (2014) Structural and functional modularity of the orange carotenoid protein: distinct roles for the N- and C-terminal domains in cyanobacterial photoprotection, *The Plant cell 26*, 426-437.
14. Kerfeld, C., Sawaya, M., Brahmandam, V., Cascio, D., Ho, K., Trevithick-Sutton, C., Krogmann, D., and Yeates, T. (2003) The crystal structure of a cyanobacterial water-soluble carotenoid binding protein, *Structure 11*, 55-65.
15. Punginelli, C., Wilson, A., Routaboul, J., and Kirilovsky, D. (2009) Influence of zeaxanthin and echinenone binding on the activity of the Orange Carotenoid Protein, *Biochimica et Biophysica Acta - Bioenergetics 1787*, 280-288.

16. Wilson, A., Punginelli, C., Gall, A., Bonetti, C., Alexandre, M., Routaboul, J.-M., Kerfeld, C. a., Grondelle, R. V., Robert, B., Kennis, J. T. M., Kirilovsky, D., van Grondelle, R., Robert, B., Kennis, J. T. M., and Kirilovsky, D. (2008) A photoactive carotenoid protein acting as light intensity sensor, *Proc Natl Acad Sci U S A 105*, 12075-12080.
17. Gupta, S., Guttman, M., Leverenz, R., Zhumadilova, K., Pawlowski, E., Petzold, C., Lee, K., Ralston, C., and Kerfeld, C. (2015) Local and global structural drivers for the photoactivation of the orange carotenoid protein, *Proc Natl Acad Sci USA 112*, E5567-5574.
18. Liu, H., Zhang, H., Orf, G., Lu, Y., Jiang, J., King, J., Wolf, N., Gross, M., and Blankenship, R. (2016) Dramatic Domain Rearrangements of the Cyanobacterial Orange Carotenoid Protein upon Photoactivation, *Biochemistry 55*, 1003-1009.
19. Leverenz, R. L., Sutter, M., Wilson, A., Gupta, S., Thurotte, A., Bourcier de Carbon, C., Petzold, C. J., Ralston, C., Perreau, F., Kirilovsky, D., Kerfeld, C. a., Bourcier De Carbon, C., Petzold, C. J., Ralston, C., Perreau, F., Kirilovsky, D., and Kerfeld, C. A. (2015) A 12 Å carotenoid translocation in a photoswitch associated with cyanobacterial photoprotection, *Science 348*, 1463-1466.
20. Bao, H., Melnicki, M., Pawlowski, E. G., Sutter, M., Lechno-Yossef, S., Montgomery, B. L., and Kerfeld, C. A. (2017) Functional characterization of a new family of Orange Carotenoid Proteins, *Nature plants*.
21. Melnicki, M., Leverenz, R., Sutter, M., Lopez-Igual, R., Wilson, A., Pawlowski, E., Perreau, F., Kirilovsky, D., and Kerfeld, C. (2016) Structure, Diversity, and Evolution of a New Family of Soluble Carotenoid-Binding Proteins in Cyanobacteria, *Mol Plant 9*, 1379-1394.
22. Kerfeld, C. (2004) Structure and function of the water-soluble carotenoid-binding proteins of cyanobacteria, *Photosynthesis Research* *81*, 215-225.
23. Kerfeld, C. (2004) Water-soluble carotenoid proteins of cyanobacteria, *Archives of biochemistry and biophysics 430*, 2-9.
24. Lopez-Igual, R., Wilson, A., Leverenz, R., Melnicki, M., Bourcier de Carbon, C., Sutter, M., Turmo, A., Perreau, F., Kerfeld, C., and Kirilovsky, D. (2016) Different Functions of the Paralogs to the N-Terminal Domain of the Orange Carotenoid Protein in the Cyanobacterium *Anabaena* sp. PCC 7120, *Plant physiology* *171*, 1852-1866.
25. Singh, S. P., and Montgomery, B. L. (2015) Regulation of BolA abundance mediates morphogenesis in Fremyella diplosiphon, *Front Microbiol 6*, 1215.
26. Wolk, C. P., Fan, Q., Zhou, R., Huang, G., Lechno-Yossef, S., Kuritz, T., and Wojciuch, E. (2007) Paired cloning vectors for complementation of mutations in the cyanobacterium Anabaena sp. strain PCC 7120, *Arch Microbiol 188*, 551-563.
27. Cobley, J. G., Zerweck, E., Reyes, R., Mody, A., Seludo-Unson, J. R., Jaeger, H., Weerasuriya, S., and Navankasattusas, S. (1993) Construction of shuttle plasmids which can be efficiently mobilized from Escherichia coli into the chromatically adapting cyanobacterium, Fremyella diplosiphon, *Plasmid 30*, 90-105.
28. Elhai, J., Vepritskiy, A., Muro-Pastor, A. M., Flores, E., and Wolk, C. P. (1997) Reduction of conjugal transfer efficiency by three restriction activities of Anabaena sp. strain PCC 7120, *J Bacteriol 179*, 1998-2005.

29. Hadfield, A., and Hajdu, J. (1993) A Fast and Portable Microspectrophotometer for Protein Crystallography, *J Appl Crystallogr 26*, 839-842.
30. Kabsch, W. (2010) Integration, scaling, space-group assignment and post-refinement, *Acta Crystallogr D Biol Crystallogr 66*, 133-144.
31. Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., Keegan, R. M., Krissinel, E. B., Leslie, A. G., McCoy, A., McNicholas, S. J., Murshudov, G. N., Pannu, N. S., Potterton, E. A., Powell, H. R., Read, R. J., Vagin, A., and Wilson, K. S. (2011) Overview of the CCP4 suite and current developments, *Acta Crystallogr D Biol Crystallogr 67*, 235-242.
32. Terwilliger, T. C., Dimaio, F., Read, R. J., Baker, D., Bunkoczi, G., Adams, P. D., Grosse-Kunstleve, R. W., Afonine, P. V., and Echols, N. (2012) phenix.mr_rosetta: molecular replacement and model rebuilding with Phenix and Rosetta, *J Struct Funct Genomics 13*, 81-90.
33. Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N., and Sternberg, M. J. (2015) The Phyre2 web portal for protein modeling, prediction and analysis, *Nat Protoc 10*, 845-858.
34. Afonine, P. V., Grosse-Kunstleve, R. W., Echols, N., Headd, J. J., Moriarty, N. W., Mustyakimov, M., Terwilliger, T. C., Urzhumtsev, A., Zwart, P. H., and Adams, P. D. (2012) Towards automated crystallographic structure refinement with phenix.refine, *Acta Crystallogr D Biol Crystallogr 68*, 352-367.
35. Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics, *Acta Crystallogr D Biol Crystallogr 60*, 2126-2132.
36. Gantt, E., Lipschultz, C. A., Grabowski, J., and Zimmerman, B. K. (1979) Phycobilisomes from blue-green and red algae: isolation criteria and dissociation characteristics, *Plant Physiol 63*, 615-620.
37. Araoz, R., and Hader, D. P. (1997) Ultraviolet radiation induces both degradation and synthesis of phycobilisomes in Nostoc sp.: a spectroscopic and biochemical approach, *Fems Microbiol Ecol 23*, 301-313.
38. Young, A. J., Phillip, D. M., and Hashimoto, H. (2002) Ring-to-chain conformation may be a determining factor in the ability of xanthophylls to bind to the bulk light-harvesting complex of plants, *J Mol Struct 642*, 137-145.
39. Polivka, T., Herek, J. L., Zigmantas, D., Akerlund, H. E., and Sundstrom, V. (1999) Direct observation of the (forbidden) S1 state in carotenoids, *Proc Natl Acad Sci U S A 96*, 4914-4917.
40. Sedoud, A., López-Igual, R., Ur Rehman, A., Wilson, A., Perreau, F., Boulay, C., Vass, I., Krieger-Liszkay, A., Kirilovsky, D., Lopez-Igual, R., Rehman, A. U., Wilson, A., Perreau, F., Boulay, C., Vass, I., Krieger-Liszkay, A., and Kirilovsky, D. (2014) The Cyanobacterial Photoactive Orange Carotenoid Protein Is an Excellent Singlet Oxygen Quencher, *The Plant cell* *26*, 1781-1791.
41. Pattanaik, B., Busch, A. W. U., Hu, P., Chen, J., and Montgomery, B. L. (2014) Responses to iron limitation are impacted by light quality and regulated by RcaE in the chromatically acclimating cyanobacterium Fremyella diplosiphon, *Microbiology 160*, 992-1005.
42. Fiksdahl, A., Foss, P., Liaaenjensen, S., and Siegelman, H. W. (1983) Carotenoids of Blue-Green-Algae .11. Carotenoids of Chromatically-Adapted Cyanobacteria, *Comp Biochem Phys B 76*, 599-601.

43. Lakatos, M., Bilger, W., and Budel, B. (2001) Carotenoid composition of terrestrial Cyanobacteria: response to natural light conditions in open rock habitats in Venezuela, *Eur J Phycol 36*, 367-375.
44. Albrecht, M., Steiger, S., and Sandmann, G. (2001) Expression of a ketolase gene mediates the synthesis of canthaxanthin in Synechococcus leading to tolerance against photoinhibition, pigment degradation and UV-B sensitivity of photosynthesis, *Photochem Photobiol 73*, 551-555.
45. Gotz, T., Windhovel, U., Boger, P., and Sandmann, G. (1999) Protection of photosynthesis against ultraviolet-B radiation by carotenoids in transformants of the cyanobacterium synechococcus PCC7942, *Plant Physiol 120*, 599-604.
46. Di Mascio, P., Kaiser, S., and Sies, H. (1989) Lycopene as the most efficient biological carotenoid singlet oxygen quencher, *Arch Biochem Biophys 274*, 532-538.
47. Woodall, A. A., Lee, S. W., Weesie, R. J., Jackson, M. J., and Britton, G. (1997) Oxidation of carotenoids by free radicals: relationship between structure and reactivity, *Biochim Biophys Acta 1336*, 33-42.
48. Bandara, S., Ren, Z., Lu, L., Zeng, X., Shin, H., Zhao, K. H., and Yang, X. (2017) Photoactivation mechanism of a carotenoid-based photoreceptor, *Proc Natl Acad Sci U S A 114*, 6286-6291.
49. Kerfeld, C. A., Wu, Y. P., Chan, C., Krogmann, D. W., and Yeates, T. O. (1997) Crystals of the carotenoid protein from Arthrospira maxima containing uniformly oriented pigment molecules, *Acta Crystallogr D Biol Crystallogr 53*, 720-723.
50. Greco, J. A., LaFountain, A. M., Kinashi, N., Shinada, T., Sakaguchi, K., Katsumura, S., Magdaong, N. C., Niedzwiedzki, D. M., Birge, R. R., and Frank, H. A. (2016) Spectroscopic Investigation of the Carotenoid Deoxyperidinin: Direct Observation of the Forbidden S0 --> S1 Transition, *J Phys Chem B 120*, 2731-2744.
51. Simonyi, M., Bikadi, Z., Zsila, F., and Deli, J. (2003) Supramolecular exciton chirality of carotenoid aggregates, *Chirality 15*, 680-698.
52. Musser, A. J., Maiuri, M., Brida, D., Cerullo, G., Friend, R. H., and Clark, J. (2015) The Nature of Singlet Exciton Fission in Carotenoid Aggregates, *J Am Chem Soc 137*, 5130-5139.
53. Kopsel, C., Moltgen, H., Schuch, H., Auweter, H., Kleinermanns, K., Martin, H. D., and Bettermann, H. (2005) Structure investigations on assembled astaxanthin molecules, *J Mol Struct 750*, 109-115.
54. Fujii, R., Onaka, K., Kuki, M., Koyama, Y., and Watanabe, Y. (1998) The 2A(g)(-) energies of all-trans-neurosporene and spheroidene as determined by fluorescence spectroscopy, *Chem Phys Lett 288*, 847-853.
55. Polivka, T., Chabera, P., and Kerfeld, C. A. (2013) Carotenoid-protein interaction alters the S(1) energy of hydroxyechinenone in the Orange Carotenoid Protein, *Biochim Biophys Acta 1827*, 248-254.
56. Loco, D., Buda, F., Lugtenburg, J., and Mennucci, B. (2018) The Dynamic Origin of Color Tuning in Proteins Revealed by a Carotenoid Pigment, *J Phys Chem Lett 9*, 2404-2410.
57. Cianci, M., Rizkallah, P. J., Olczak, A., Raftery, J., Chayen, N. E., Zagalsky, P. F., and Helliwell, J. R. (2002) The molecular basis of the coloration mechanism in lobster shell: beta-crustacyanin at 3.2-A resolution, *Proc Natl Acad Sci U S A 99*, 9795-9800.

830
831

Acknowledgement

This work was supported by the National Science Foundation (IOS 1557324) and the Office of Science of the U.S. Department of Energy DE-FG02-91ER20021. This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No. 795070. T.P. thanks the Czech Science Foundation (18-21631S) for financial support. The authors thank members of the Kerfeld Lab for helpful discussions. We thank the Metabolomic facility core at Michigan State University. We also thank Dr. John McCracken for technical help and use of his EPR spectrometer for the spin trapping experiments. This research used resources of the Advanced Light Source, which is a DOE Office of Science User Facility under contract no. DE-AC02-05CH11231 and the Stanford Synchrotron Radiation Lightsource, SLAC National Accelerator Laboratory, is supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences under Contract No. DE-AC02-76SF00515.

Author contribution

M.A.D.M. designed and performed research, analysed and interpreted data, and wrote the manuscript. C.A.K. designed the research, analysed and interpreted data, and wrote the manuscript. T. P. interpreted the spectroscopic data and wrote the manuscript. B. F. performed the EPR experiments. M.S., S.L.-Y. and B.L.M. performed research and contributed to the analysis and interpretation of the data.