



## OPEN ACCESS

## EDITED BY

Michael A. Kennedy,  
Miami University, United States

## REVIEWED BY

Esko Oksanen,  
European Spallation Source, Sweden  
Antonio Chaves-Sanjuan,  
University of Milan, Italy

## \*CORRESPONDENCE

Thomas Friedrich,  
✉ friedrich@chem.tu-berlin.de

## SPECIALTY SECTION

This article was submitted  
to Structural Biology,  
a section of the journal  
Frontiers in Molecular Biosciences

RECEIVED 17 October 2022

ACCEPTED 17 January 2023

PUBLISHED 26 January 2023

## CITATION

Moldenhauer M, Tseng H-W, Kraskov A,  
Tavraz NN, Yaroshevich IA, Hildebrandt P,  
Sluchanko NN, Hochberg GA, Essen L-O,  
Budisa N, Korf L, Maksimov EG and  
Friedrich T (2023), Parameterization of a  
single H-bond in Orange Carotenoid  
Protein by atomic mutation reveals  
principles of evolutionary design of  
complex chemical photosystems.  
*Front. Mol. Biosci.* 10:1072606.  
doi: 10.3389/fmolb.2023.1072606

## COPYRIGHT

© 2023 Moldenhauer, Tseng, Kraskov,  
Tavraz, Yaroshevich, Hildebrandt,  
Sluchanko, Hochberg, Essen, Budisa, Korf,  
Maksimov and Friedrich. This is an open-  
access article distributed under the terms  
of the [Creative Commons Attribution  
License \(CC BY\)](#). The use, distribution or  
reproduction in other forums is permitted,  
provided the original author(s) and the  
copyright owner(s) are credited and that  
the original publication in this journal is  
cited, in accordance with accepted  
academic practice. No use, distribution or  
reproduction is permitted which does not  
comply with these terms.

# Parameterization of a single H-bond in Orange Carotenoid Protein by atomic mutation reveals principles of evolutionary design of complex chemical photosystems

Marcus Moldenhauer<sup>1</sup>, Hsueh-Wei Tseng<sup>2</sup>, Anastasia Kraskov<sup>1</sup>, Neslihan N. Tavraz<sup>1</sup>, Igor A. Yaroshevich<sup>3</sup>, Peter Hildebrandt<sup>1</sup>, Nikolai N. Sluchanko<sup>4</sup>, Georg A. Hochberg<sup>5</sup>, Lars-Oliver Essen<sup>6</sup>, Nediljko Budisa<sup>2,7</sup>, Lukas Korf<sup>6</sup>, Eugene G. Maksimov<sup>3</sup> and Thomas Friedrich<sup>1\*</sup>

<sup>1</sup>Department of Bioenergetics, Institute of Chemistry PC 14, Technische Universität Berlin, Berlin, Germany,

<sup>2</sup>Department of Biocatalysis, Institute of Chemistry L1, Technische Universität Berlin, Berlin, Germany,

<sup>3</sup>Department of Biophysics, Faculty of Biology, Lomonosov Moscow State University, Moscow, Russia, <sup>4</sup>A.N. Bach Institute of Biochemistry, Federal Research Center Fundamentals of Biotechnology of Russian Academy of Sciences, Moscow, Russia, <sup>5</sup>Max-Planck-Institute of Terrestrial Microbiology, Evolutionary Biochemistry Group, Marburg, Germany, <sup>6</sup>Department of Chemistry and Unit for Structural Biology, Philipps-Universität Marburg, Marburg, Germany, <sup>7</sup>Department of Chemistry, University of Manitoba, Winnipeg, MB, Canada

**Introduction:** Dissecting the intricate networks of covalent and non-covalent interactions that stabilize complex protein structures is notoriously difficult and requires subtle atomic-level exchanges to precisely affect local chemical functionality. The function of the Orange Carotenoid Protein (OCP), a light-driven photoswitch involved in cyanobacterial photoprotection, depends strongly on two H-bonds between the 4-ketolated xanthophyll cofactor and two highly conserved residues in the C-terminal domain (Trp288 and Tyr201).

**Method:** By orthogonal translation, we replaced Trp288 in *Synechocystis* OCP with 3-benzothienyl-L-alanine (BTA), thereby exchanging the imino nitrogen for a sulphur atom.

**Results:** Although the high-resolution (1.8 Å) crystal structure of the fully photoactive OCP-W288\_BTA protein showed perfect isomorphism to the native structure, the spectroscopic and kinetic properties changed distinctly. We accurately parameterized the effects of the absence of a single H-bond on the spectroscopic and thermodynamic properties of OCP photoconversion and reveal general principles underlying the design of photoreceptors by natural evolution.

**Discussion:** Such “molecular surgery” is superior over trial-and-error methods in hypothesis-driven research of complex chemical systems.

**Abbreviations:** BTA, 3-benzothienyl-L-alanine; CAN, canthaxanthin; ECN, echinenone; ESI-MS, electrospray ionization mass spectrometry; IPTG, isopropyl-β-D-thiogalactopyranoside; MmPylRS, pyrrolysine-tRNA synthetase from *Methanosarcina mazei*; ncAA, non-canonical amino acid; OCP, Orange Carotenoid Protein; OCP<sup>D</sup>, dark-adapted orange form of OCP; OCP<sup>R</sup>, light-adapted (active) red form of OCP; PBS, phycobilisome; RR, resonance Raman; SEC, size-exclusion chromatography; Trp, tryptophan; Tyr, tyrosine.

## KEYWORDS

atomic mutations, hydrogen bond strength/energy, non-canonical amino acids, orthogonal translation, Orange Carotenoid Protein

## 1 Introduction

Maintaining the marginally stable three-dimensional architecture of proteins often requires a complicated network of weak interactions, including hydrogen bonds. The study of individual hydrogen bonds in the context of such complex protein structures is often a formidable task. A kind of “molecular surgery” (Thi To et al., 2022) with “atomic mutations” (Minks et al., 1999) is required to remove or replace individual atoms or groups of atoms in the context of complex proteins containing thousands of atoms with a complicated network of, for example, covalent, non-covalent or hydrogen bonding and van der Waals interactions. It has been suggested that the function of the Orange Carotenoid Protein (OCP), a blue light-driven photoswitch involved in photoprotection of cyanobacteria (Holt and Krogmann, 1981), strongly depends on two critical H-bonds between the ketocarotenoid cofactor and a tryptophan (Trp) and a tyrosine (Tyr) residue in the surrounding protein matrix (Figure 1). However, classical protein mutagenesis does not allow the precise parameterization of the contributions of these H-bonds to a protein’s functional integrity.

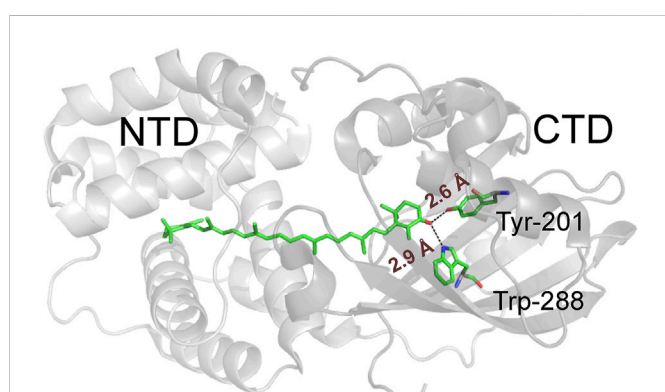
Structurally, the 35 kDa OCP consists of an N-terminal domain (NTD) composed entirely of  $\alpha$ -helices and a mixed  $\alpha/\beta$ -type C-terminal domain (CTD) of about equal size, both encapsulating a single ketocarotenoid molecule in a common central cavity (Kerfeld et al., 2003). Upon absorption of a photon, OCP converts from the dark-adapted, orange OCP<sup>O</sup> state to the red, active OCP<sup>R</sup> state. The embedded ketocarotenoid is at the heart of OCP photoswitching activity and function. In native cyanobacteria this cofactor is 3'-hydroxyechinenone (Muzzopappa and Kirilovsky, 2020); whereas the chemical identity in the equally active recombinant form is a 4-ketolated xanthophyll. In the compact OCP<sup>O</sup> state, for which X-ray crystal structures are available (e.g., PDB codes 3MG1, 4XB5, 5UI2), the 4-keto oxygen

atom of the xanthophyll is involved in an H-bonding network to the two highly conserved residues within the C-terminal domain (Tyr201 and Trp288 according to amino acid numbering of *Synechocystis* OCP). This local network fixes the cofactor in a rigid, slightly bent and twisted conformation (Figure 1) (Kerfeld et al., 2003; Leverenz et al., 2015). The H-bonds in this network break upon photoexcitation, releasing the terminal  $\beta$ -ring of the xanthophyll, leading to a number of detectable spectroscopic and conformational changes, e.g., xanthophyll isomerization and translocation, separation of domains, and a red-shift in absorbance (Leverenz et al., 2015).

The structure of the OCP<sup>R</sup> state has not yet been elucidated because it is very elongated and flexible (Gupta et al., 2015; Sluchanko et al., 2017a). In OCP<sup>R</sup>, the protein is able to interact closely with the water-soluble antenna complexes of cyanobacteria, the phycobilisomes (PBSs). By exposing the chromophore to the terminally emitting pigments of the PBSs, OCP<sup>R</sup> acts as an efficient quencher of PBS fluorescence, protecting the underlying photosystems from the deleterious effects of overexcitation and dissipating the electronic excitation energy into heat.

Mutations of Tyr201 and Trp288 often destabilize the compact OCP<sup>O</sup> form and generate permanently red-shifted OCP forms with separated domains that are - despite not being photoactive - constitutively capable of PBS fluorescence quenching (Maksimov et al., 2016). As an example, the OCP-W288A mutant protein upon heterologous expression yields a mixture of species with a dominant fraction of photo-inactive red-purple protein, which is capable of PBS fluorescence quenching. At the same time, the absorbance spectra of many dark-adapted OCP species show spectral heterogeneity with admixtures of red states in the OCP<sup>O</sup> form (Yaroshevich et al., 2021). However, mutational analyses occasionally revealed OCP variants with markedly reduced contamination by red states in the compact, dark-adapted OCP<sup>O</sup> form. Such properties were observed for the OCP-Y201W mutant, whose spectrally pure, “super-orange” OCP<sup>O</sup> state prove advantageous for femtosecond pump-probe UV-Vis spectroscopy to investigate the photocycle (Yaroshevich et al., 2021). For OCP-Y201W, the analysis of femtosecond transient UV-Vis absorption experiments combined with quantum chemical calculations led to the formulation of a reaction mechanism for the photochemistry in the early events that occur upon light-triggered H-bond breaking and involve the formation of an oxocarbenium ion on the ketocarotenoid (Yaroshevich et al., 2021).

Site-directed mutagenesis studies to probe the role of Trp288 and Tyr201 in OCP suffer from the limited ability to substitute aromatic amino acids within the standard amino acid repertoire. As a result, inaccurate conclusions have been drawn about which amino acid residues or their chemical properties are *necessary* for particular functional aspects, leaving unanswered the question of which chemical functionalities are *sufficient* for function. In this work, we performed a kind of “molecular surgery” in which a canonical amino acid residue was site-specifically replaced with a non-canonical analogue to remove the critical H-bond donor function of Trp288 of OCP with atomic precision. To this end, we used a recently evolved orthogonal pair (Tseng et al., 2020) consisting of a



**FIGURE 1**

Structural representation of wild-type *Synechocystis* OCP with the N- and C-terminal domains depicted in different shades of grey. The critical H-bond-donating residues (Trp288 and Tyr201) and the ECN cofactor are shown as ball-and-stick models, while the H-bonds are indicated by dashed lines.

custom-tailored *Methanosarcina mazei* pyrrolysine tRNA synthetase (*MmPylRS*) and its cognate tRNA to replace Trp288 in OCP with 3-benzothienyl-*L*-alanine (BTA), an isostructural analogue of tryptophan in which the imino nitrogen is substituted by a sulphur atom. The resulting OCP-W288\_BTA protein is photoactive and, according to the X-ray crystal structure solved at 1.8 Å resolution, fully isomorphous to the wild-type protein, allowing us to parametrize the consequences of the absence of a single H-bond by spectroscopic, thermodynamic and computational methods.

## 2 Materials and methods

### 2.1 DNA constructs and molecular cloning

Engineering of the *Methanosarcina mazei* pyrrolysyl-tRNA synthetase (*MmPylRS*) gene towards efficient loading of its cognate tRNA<sup>Pyl</sup> with the non-canonical amino acid (ncAA) 3-benzothienyl-*L*-alanine (BTA) was carried out as described previously (Tseng et al., 2020). In detail, the *MmPylRS* variant used here is denoted as *MmPylRS*-13IYPER\_SMG carrying the following set of mutations: T13I/T56P/H62Y/A100E/S193R/N346S/C348M/V401G conferring high BTA incorporation efficiency, and Y384F, which confers an increased aminoacylation rate of the aaRS enzyme, as known from previous studies (Ko et al., 2013). For further increase of aaRS processivity, the *MmPylRS* cDNA sequence was codon-optimized for enhanced recombinant expression in *E. coli* [see (Tseng et al., 2020) for sequence information]. The cDNA of the *MmPylRS* variant was cloned into plasmid pBU16 (ampicillin resistance) which contained pyrrolysyl-tRNA<sub>CUA</sub> (tRNA<sub>CUA</sub><sup>Pyl</sup>) driven by a *glnS*' promoter (Hauf et al., 2017).

The OCP gene (*slr 1963*) of *Synechocystis* sp. PCC 6803 was optimized for codon usage in *E. coli* by artificial gene synthesis (GeneOptimizer algorithm, GeneArt, Thermo Fisher Scientific, Waltham, MA, United States) and cloned into the first multiple cloning site of the pRSFDuet-1 vector (Novagen, Merck KGa, Darmstadt, Germany) by using *Bam*HI and *Not*I restriction sites. The in-frame amber stop codon mutation Trp-288-TAG was generated using the Q5SDM Kit (New England Biolabs, Ipswich, MA, United States). For ease of eliminating truncated protein products, the 6xHis-Tag at the N-terminus was deleted and a new 6xHis-Tag was added to the C-terminus (resulting plasmid termed pRSFDuet-cTag herein, carrying kanamycin resistance) with an additional peptide linker sequence [the resulting C-terminal sequence was . . . NFARPGSSHHHHHH\* with the linker sequence PGSS in bold, as suggested by (de Carbon et al., 2015)] using the Q5SDM Kit (New England Biolabs, Ipswich, MA, United States). The resulting N-terminal sequence corresponds to the sequence of wild-type OCP from *Synechocystis* sp. PCC 6803 (amino acid sequence MPFTID . . .). All cDNA constructs were verified by DNA sequencing (LGC genomics, Berlin, Germany). Synthesis of the carotenoid cofactor was accomplished using the p25crtO plasmid (carrying chloramphenicol resistance) which did not encode the *crtZ* gene but contained the contiguous gene cluster harboring the genes *crtY*, *crtI*, *crtB*, and *crtE* from *Pantoea ananatis* (formerly *Erwinia uredovora*) and the *crtO* gene from *Synechocystis* sp. PCC 6803 for echinenone/canthaxanthin synthesis (Misawa et al., 1995), and the strategy for expressing OCP variant proteins in echinenone/

canthaxanthin-producing in *E. coli* strains followed the procedures described previously (Maksimov et al., 2016).

### 2.2 Protein expression and purification

To produce OCP with a site-specific ncAA incorporation, the three plasmids described above (pRSFDuet-cTag carrying the cDNA of the desired OCP variant, the p25crtO plasmid for ECN/CAN synthesis and the pBU16 plasmid carrying the components of the orthogonal translation system) were co-transformed into *E. coli* strain BL21 (DE3). Single colonies were picked and pre-cultured in 10 mL of Modified Tyler Medium (MTM: 2.5 g/L NH<sub>4</sub>Cl [46.7 mM], 0.71 g/L Na<sub>2</sub>SO<sub>4</sub> [5 mM], 6.8 g/L KH<sub>2</sub>PO<sub>4</sub> [50 mM], 7.1 g/L Na<sub>2</sub>HPO<sub>4</sub> [50 mM], 6.0 g/L glycerol [65 mM], 2.5 g/L D-glucose [13.9 mM], 5 μM FeCl<sub>2</sub>, 2 μM CaCl<sub>2</sub>, 1 μM MnCl<sub>2</sub>, 1 μM ZnSO<sub>4</sub>, 0.2 μM CoCl<sub>2</sub>, 0.2 μM CuCl<sub>2</sub>, 0.2 μM NiCl<sub>2</sub>, 0.2 μM Na<sub>2</sub>MoO<sub>4</sub>, 0.2 μM Na<sub>2</sub>SeO<sub>3</sub>, 0.2 μM H<sub>3</sub>BO<sub>3</sub>, 0.2 μM vitamin B12, 0.2 μM nicotinic acid, 0.2 μM pyridoxine-HCl, 0.2 μM thiamine-HCl, 0.2 μM *p*-aminobenzoic acid, 5 nM folic acid, 5 nM riboflavin) (Tyler et al., 2005; Hoopes et al., 2015) supplemented with 100 μg/mL ampicillin (Amp), 50 μg/mL kanamycin (Kan) and 3734 μg/mL chloramphenicol (Cm) at 37°C in an orbital shaker at 200 rpm until OD<sub>600</sub> reached 0.8–1.0. The pre-culture was transferred into 100 mL MTM supplemented with 100 μg/mL Amp, 50 μg/mL Kan, and 3734 μg/mL Cm and grown at 37°C in an orbital shaker at 200 rpm until OD<sub>600</sub> reached 0.6–0.8. Then, 2 mM of the ncAA was added to the culture, which subsequently was further incubated at 37°C for 30 min. Subsequently, target gene expression was induced by adding 1 mM isopropyl-β-*D*-thiogalactopyranoside (IPTG) and incubation at 25°C and 250 rpm proceeded for 72 h. Cells were harvested by centrifugation at 10,000 × g at 4°C for 20 min and frozen until use. For the expression of the OCP-W288A variant, the same conditions were used except the steps of non-canonical amino acid-related biochemistry.

The frozen cell pellets were resuspended in phosphate buffered saline (1xPBS: 137 mM NaCl, 2.7 KCl, 12 mM phosphate at pH 7.4) and supplemented with 0.5 mg/mL lysozyme (Ovobest, Neuenkirchen-Vörden, Germany) and protease inhibitor solution (1 mM benzamidine, 1 mM ε-amino-caproic acid). Cell lysis was performed by an OmniRuptor 4,000 sonifier (OMNI International, Kennesaw, GA, United States) with 5 cycles of 3 min sonification and 3 min cooling breaks on ice between each cycle, the tip power was tuned to 80%. To remove the cell debris, the cell solution was centrifuged at 18,000 × g at 4°C and the clarified supernatant was loaded on a 5 mL Co<sup>2+</sup>-HiTrap Talon crude column (GE Healthcare, Chicago, IL, United States) by using a peristaltic pump. The protein was eluted by an imidazole-containing phosphate buffer (1xPBS, 350 mM imidazole pH 7.4). To remove carotenoid-free apoprotein and other contaminants, the protein solution was purified by hydrophobic interaction chromatography (HIC) by using a Hiprep 16/10 Phenyl HP column (GE Healthcare) as follows: The protein solution was dialyzed overnight in HIC buffer A (0.5 M (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 0.1 M urea, 5 mM phosphate pH 7.5), loaded on the HIC buffer A-equilibrated column and finally eluted by a binary gradient of HIC buffer B (0.1 M urea, 5 mM phosphate pH 7.5). Finally, the collected and concentrated protein solution was purified using a Superdex 200 Increase 10/300 column (GE Healthcare) to separate the monomeric orange from the dimeric pink protein species.

## 2.3 Mass spectrometry

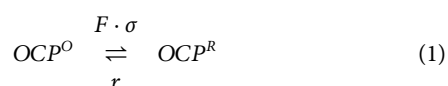
To identify ncAA incorporation into OCP by *MmPylRS*, intact protein masses were analyzed by an electrospray ionization mass spectrometry (ESI-MS) source (Waters H-class instrument). The proteins were eluted through an Acquity UPLC protein BEH C4 column (300 Å, 1.7 µm, 2.1 mm × 50 mm) with a flow rate of 0.3 mL/min. The resulting ions were analyzed by a Waters XEVO G2-XS QToF analyzer (Waters Corp., Milford, MA, United States). Spectra deconvolution was performed using MaxEnt 1 employing the maximum entropy deconvolution algorithm. Observed protein masses were compared to theoretical values determined by the Expsy ProtParam tool (Gasteiger et al., 2005).

## 2.4 Carotenoid extraction and HPLC analysis

For analysis of the carotenoid content of OCP holoprotein variants, 50 µL of concentrated protein samples were mixed with 1 mL acetone and centrifuged at maximum speed at 4°C to spin down precipitated protein. The yellowish supernatant was evaporated in a centrifugal vacuum concentrator (Eppendorf, Hamburg, Germany) until all acetone was removed and the carotenoid sample was precipitated as red crystals. The remaining water solution was removed and the carotenoid crystals were re-dissolved in 50 µL acetone. The carotenoid-rich solution was transferred into a vial for HPLC analysis. HPLC was performed on an UFLC NexeraX2 system (Shimadzu, Kyoto, Japan) by using a Accucore™ C30 column (Thermo Fisher Scientific, 250 mm × 2.1 mm, 2.6 µm particle size, pore size 150 Å). As mobile phase two eluents A (methanol:water, 95:5) and B (methanol:THF, 7:3) were used with the following protocol to elute carotenoids: 0–4.3 min 0% buffer B, 4.3–8.6 min 0–100% buffer B, 8.6–15.6 min 100% buffer B, 15.6–20.1 min 0% buffer B and a flow rate set to 0.4 mL/min. The eluted carotenoids were verified by mass spectrometry to correlate elution times to specific carotenoid species, and by thin-layer chromatography by comparison with reference samples.

## 2.5 UV/VIS spectroscopy and kinetic analysis

Absorption spectra were recorded using a Maya2000Pro spectrometer (Ocean Insight, Orlando, FL, United States) coupled by a fiber to a deuterium tungsten light source (Sarspec, Vila Nova de Gaia, Portugal) and a cuvette holder (CVH100, Thorlabs, Bergkirchen, Germany). For kinetic measurements, a temperature-controlled cuvette holder qpod2e (Quantum Northwest, Liberty Lake, WA, United States) was fiber-coupled to a CCS100/M spectrometer (Thorlabs) and a SLS201 L/M tungsten light source (Thorlabs). For illumination with actinic light, a 3W (Avonec, Wesel, Germany) light-emitting diode (LED) with a maximum emission at 455 nm was used. Analysis of the observed time constants for photoconversion (OCP<sup>O</sup>→OCP<sup>R</sup>) and back-relaxation (OCP<sup>R</sup>→OCP<sup>O</sup>) was performed based on a simple two-state kinetic model described in (Maksimov et al., 2015) as follows:



In a two-state model with first order rate constants, the total (observed) relaxation rate  $k_{obs}$  during photoconversion by light is the sum of the light-dependent forward rate constant (the product of illumination intensity (fluence)  $F$  and the photoconversion cross-section  $\sigma$ ) and the light-independent relaxation rate  $r$ . Thus, whereas the rate constant of the light-independent back-relaxation can be directly measured, the rate-constant of the light-dependent forward reaction ( $F \cdot \sigma$ ) is calculated as:

$$F \cdot \sigma = k_{obs} - r \quad (2)$$

As a prerequisite, the photoswitching reaction always needs to be followed at the same illumination intensity at the sample, which was obeyed in the present experiments.

## 2.6 Resonance Raman (RR) spectroscopy

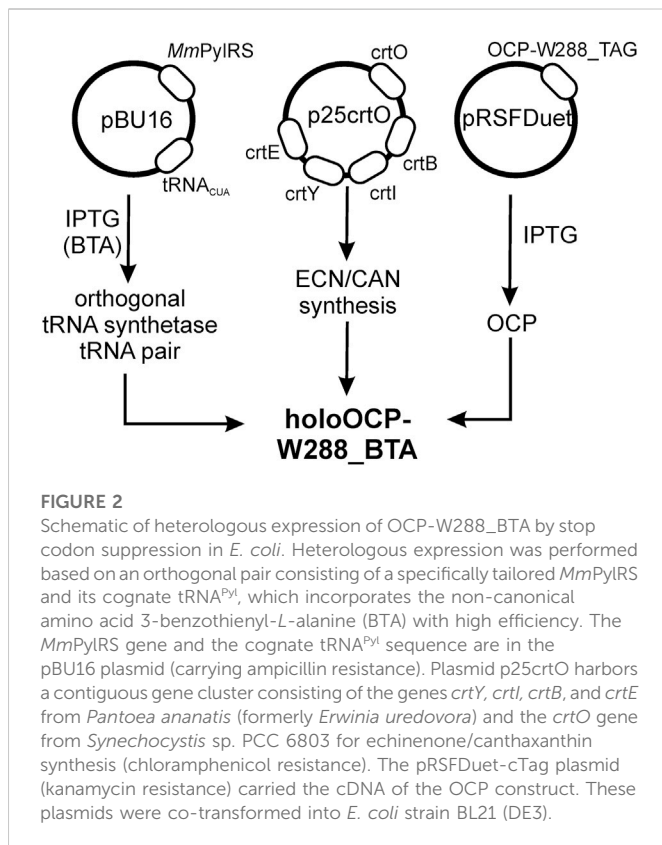
Spectra were recorded at −140°C (Linkam cryostat, Resultec, Illerkirchberg, Germany) using a Fourier-Transform Raman spectrometer RFS-100/S (Bruker, Rosenheim, Germany) equipped with a 1,064 nm cw NdYAG laser (Compass 1064-1500N, Coherent LaserSystems, Santa Clara, CA, United States) as described (Moldenhauer et al., 2017). Although 1,064 nm excitation is not in resonance with chromophore absorption, the chromophore signals are ~1:1,000 enhanced over the protein bands justifying use of the term “resonance Raman” (RR) spectroscopy for our pre-resonant conditions. For each RR spectrum, 1,000 single scans were averaged. The samples were prepared at a concentration of 230 µM (wild-type OCP) and 243 µM (OCP-W288\_BTA) in 1xPBS, pH 7.4. For photoconversion, a 3-W LED (Avonec, Wesel, Germany) was used.

## 2.7 Crystallization and X-ray data collection

Protein crystallization was carried out as sitting drop experiments by mixing 300 nL of protein and well solution each. A thawed aliquot of OCP-W288\_BTA with a concentration of 2 mg/mL stored in 1xPBS Buffer (137 mM NaCl, 2.7 KCl, 12 mM phosphate, pH 7.4) was used for crystallization. The crystal used for structure determination grew in the presence of 0.2 M MgCl<sub>2</sub>, 0.1 M Tris pH 7.0% and 10% (w/v) PEG 8000. Crystals were mounted on micro mounts and flash-frozen in liquid nitrogen before transport. The synchrotron X-ray diffraction data was collected at beamline P14 operated by EMBL Hamburg at the PETRA III storage ring (DESY, Hamburg, Germany).

## 2.8 Crystal structure solution and refinement

Diffraction data were processed using X-ray Detector Software (XDS) (Kabsch, 1993). Subsequent data reduction was performed using the CCP4i2 pipeline running AIMLESS (Evans and Murshudov, 2013) and PHASER (McCoy et al., 2007) for molecular replacement using the OCP wild-type structure (PDB: 5TUW) as search model. Afterwards, the structure was refined by several rounds of manual model building and optimization using Coot (Emsley and Cowtan, 2004) followed by refinement and quality check with phenix.refine (Afonine et al., 2012). The refinement strategy involved rigid body refinement, individual B-factors, TLS parameters and occupancies. OCP-W288\_BTA

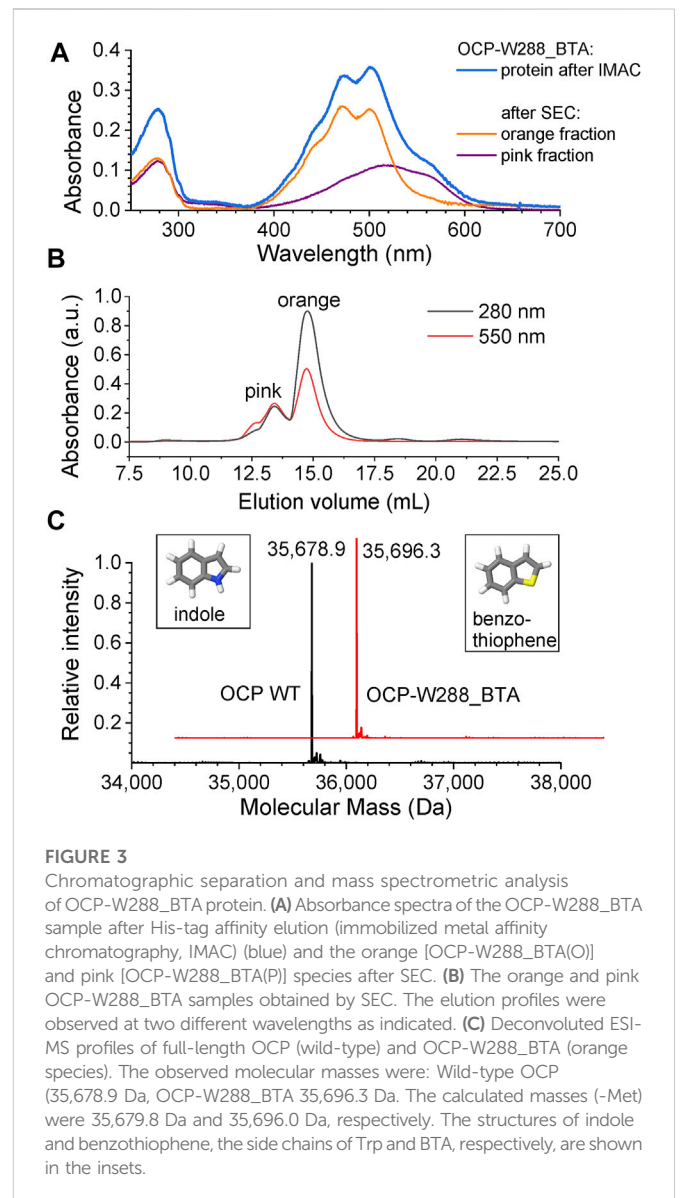


coordinates and structure factors were deposited in the RCSB protein database under the PDB accession code 7ZXV. The structure of the BTA side chain was obtained by quantum chemical (QC) calculations and the corresponding coordinates were used in structure refinement. Crystal structure depictions were produced using PyMOL 2.5.2.

## 2.9 Quantum chemical (QC) calculations

To obtain the equilibrium geometry for BTA, QC calculations with  $\omega$ B97x (Mardirossian and Head-Gordon, 2014)/6-311G\*\* (Ditchfield et al., 1971) theory level were used. Calculations were done using the ORCA 4.2 software package (Neese et al., 2020).

The calculation of the molecular complex between BTA, tryptophan and the xanthophyll was performed with quantum dynamics (QD). Two systems consisting of tryptophan/BTA, tyrosine, and the xanthophyll were taken from the PDB crystal structures of 3MG1 and 7ZXV, respectively. Both complexes were optimized with constraints applied to nine atomic nuclei: three of the heavy atoms for both amino acid backbones, and three methyl groups at the C1 and C5' atoms of ECN. The optimized systems were passed for conducting 10 ps QD calculations with integration time step of 1 fs using the  $\omega$ B97x/def2-SVP (Weigend and Ahlrichs, 2005) method. Simulations were performed at 300 K and 1 bar pressure using the Bussi-Parrinello Langevin dynamics (Bussi and Parrinello, 2007). Calculation was done with TeraChem v1.9 (Seritan et al., 2020) software using the super-computing center at Lomonosov Moscow State University.



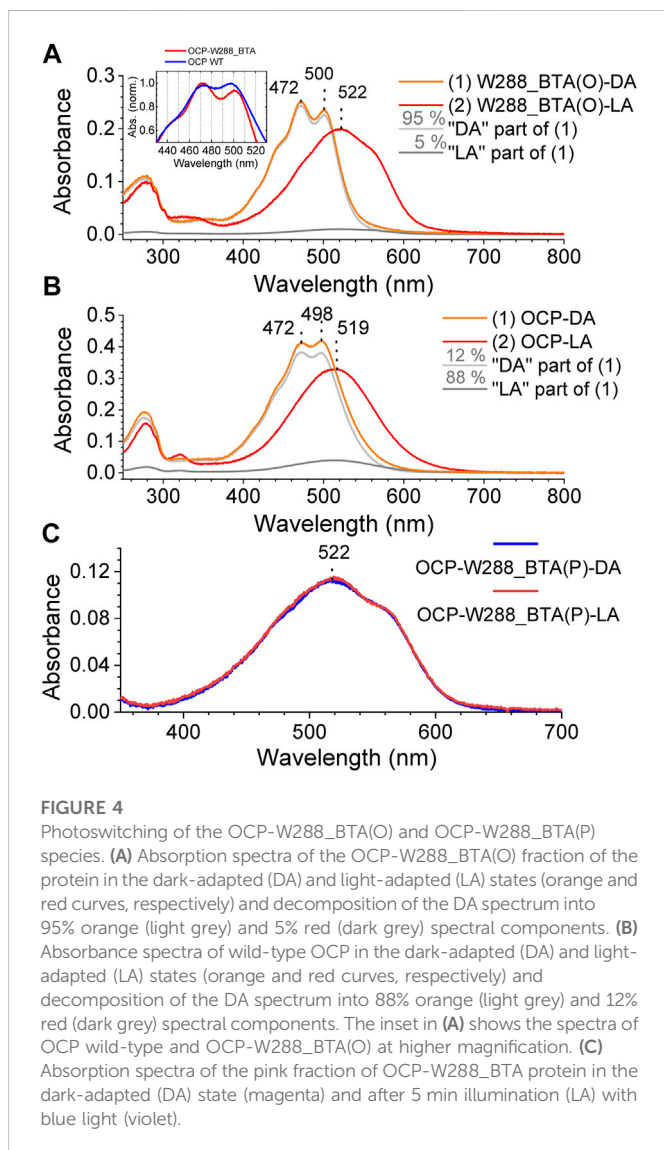
## 3 Results and discussion

We used an *Escherichia coli* expression host for the biosynthesis of the BTA-containing OCP holoprotein variant (OCP-W288\_BTA). Our expression system is based on three plasmids (Figure 2). Affinity chromatography using His-tag yielded an intensely stained protein solution (blue spectrum in Figure 3A) showing characteristic features of dark-adapted OCP with a small amount of a red-shifted species, which shows up as a shoulder at 570 nm (Sluchanko et al., 2017a). After an additional purification step with preparative size-exclusion chromatography (SEC), an orange and a pink protein species could be separated based on their different elution profiles (Figure 3B) (Sluchanko et al., 2017b). While the pink species [OCP-W288\_BTA(P)] eluted earlier, the orange fraction [OCP-W288\_BTA(O)] corresponded to the retention time of an OCP monomer. Figure 3A shows the representative spectra of a protein preparation that contained approximately 35% pink and 65% orange species. The purified proteins of the orange OCP-W288\_BTA(O) variant and

TABLE 1 Carotenoid content determined by HPLC (see chromatograms in Supplementary Figure S1).

Sample	Carotenoid	Retention time (min)	Relative content (mean $\pm$ S.D) in %
OCP wild-type	canthaxanthin	7.6	3.0 $\pm$ 3.1
	<b>echinenone</b>	<b>9.7</b>	<b>91.7 <math>\pm</math> 2.5</b>
	<i>ECN prepeak</i> <sup>[a]</sup>	9.5	5.3 $\pm$ 1.7
OCP-W288_BTA orange	<b>canthaxanthin</b>	<b>7.5</b>	<b>53.9 <math>\pm</math> 4.8</b>
	<b>echinenone</b>	<b>9.7</b>	<b>42.4 <math>\pm</math> 4.0</b>
	<i>ECN prepeak</i> <sup>[a]</sup>	9.5	3.6 $\pm$ 0.8
OCP-W288_BTA pink	<b>canthaxanthin</b>	<b>7.6</b>	<b>90.9 <math>\pm</math> 6.8</b>
	echinenone	9.7	8.5 $\pm$ 7.3

<sup>a</sup>The *ECN, prepeak* species was not further characterized. It may consist of *cis*-echinenone, which may have formed during the extraction procedure, the relative content depends on parameters such as sample concentration or residual water content—it is not considered to be present as *cis*-echinenone in the protein. Bold values denote the dominant carotenoid species.



wild-type OCP were subjected to electrospray LC-MS analysis (Figure 3C), which revealed 100% incorporation of BTA in the variant protein.

The carotenoid content of the protein samples was determined by HPLC (Table 1). While wild-type OCP almost exclusively contained ECN, the orange OCP-W288\_BTA(O) species had two peaks of approximately the same amplitude (54%/46%) corresponding to ECN and CAN. Different amounts of ECN/CAN in OCP samples obtained from xanthophyll-producing *E. coli* strains is a common phenomenon (de Carbon et al., 2015) and may depend on the employed  $\beta$ -carotene ketolase gene or on the expression level of OCP in the cells. However, the photoswitching ability is not affected by the presence of ECN or CAN (de Carbon et al., 2015) (see below). In contrast, the pink OCP-W288\_BTA(P) species contained more than 96% CAN.

When both the orange and pink OCP-W288\_BTA protein fractions were illuminated with blue light (3-W LED, 455 nm) at 15°C, the OCP-W288\_BTA(O) species was photoswitchable (Figure 4A) like wild-type OCP (Figure 4B), whereas no spectral changes were observed for the OCP-W288\_BTA(P) species (Figure 4C). Spectral decomposition showed that the spectrum of the dark-adapted OCP-W288\_BTA(O) species contained a much smaller contribution of a red non-photoswitchable species compared with the wild-type OCP (Figure 4A) than the dark-adapted wild-type OCP (5% vs. 12%, respectively) (Figure 4B). The absorption spectrum of the orange OCP-W288\_BTA(O) species exhibits a distinct vibrational fine structure, with even more pronounced vibrational bands than the wild-type OCP<sup>o</sup> (Figure 4A, inset). These features of the OCP-W288\_BTA(O) protein, a spectrally more homogenous dark-adapted state and a deeper vibrational substructure are typical attributes of so-called “super-orange” OCP species, as first described for the OCP-Y201W mutant (Yaroshevich et al., 2021).

In the OCP-Y201W mutant protein, the lack of one H-bond donor [according to the X-ray crystal structure solved in (Yaroshevich et al., 2021), the introduced Trp residue is too far away from the C=O group of the ketocarotenoid] leads to an increase in the importance of the remaining H-bond (from the C=O group of the carotenoid to the N-H group of Trp288), making the ground state less susceptible to spontaneous carotenoid fluctuations. In the OCP-W288\_BTA variant, the same effect could also occur in the opposite direction: Removal of the imino nitrogen by a sulphur atom also strengthens the importance of the remaining H-bond to Tyr201.

TABLE 2 OCP-W288\_BTA X-ray data collection and refinement statistics.

Data collection	
PDB ID	7ZXV <sup>1</sup>
Diffraction source	P14 beamline (EMBL Hamburg, Germany) at Petra III storage ring (DESY, Hamburg, Germany)
Wavelength (Å)	0.9763
Space group	P 3 <sub>2</sub> 2 1
Cell dimensions: a, b, c (Å)	83.195, 83.195, 88.648
α, β, γ (°)	90.00, 90.00, 120.00
Resolution range (Å) *	72.05-1.80 (1.84-1.80)
Unique reflections	32,925 (1940)
R <sub>merge</sub>	0.114 (4.998)
R <sub>pim</sub>	0.037 (1.628)
R <sub>meas</sub>	0.120 (5.574)
<I/σ>	13.7 (0.6)
CC <sub>1/2</sub>	0.999 (0.477)
Wilson B factor (Å <sup>2</sup> )	37.07
Completeness for range (%)	98.7 (100)
Redundancy	20.2 (20.1)
Refinement	
Resolution range (Å)	55.91-1.80 [55.91-4.12] (1.85-1.80)
No. of reflections: total	32,864 [2,789] (2,587)
“working” set	31,224 [2,627] (2,458)
“free” set	1,640 [162] (129)
R <sub>work</sub> (%)	17.30 [13.89] (44.76)
R <sub>free</sub> (%)	22.23 [18.94] (52.70)
Mean B value (overall Å <sup>2</sup> )	40.75
Completeness (working + test) (%)	98.58
No. of non-H atoms: protein/ligands/solvent	2,464/83/308
R.m.s.d. bond lengths (Å)/angles (°)	0.012/1.079
Ramachandran favored/outliers (%)	99.03/0.0
Molprobrity/All-atom clashscore	1.156/2.3

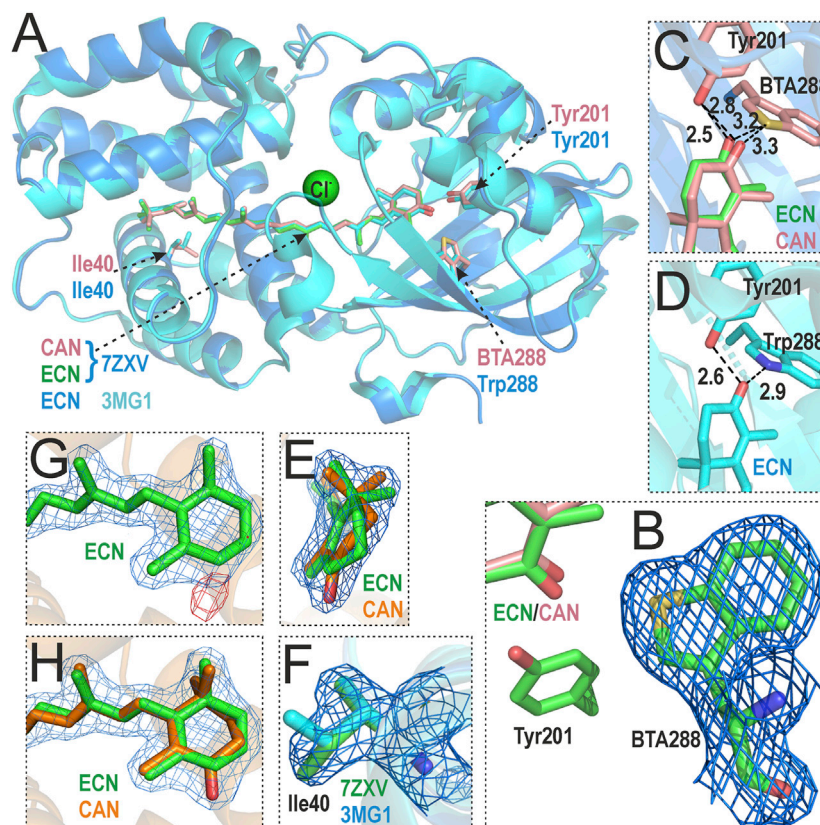
\* Statistics for the lowest and highest resolution shells are indicated in square brackets and parentheses, respectively.

The spectra of the OCP-W288\_BTA(P) species and the light-adapted OCP-W288\_BTA(O) species are very similar (Supplementary Material S1, Supplementary Figure S2). This suggests that the spectroscopic properties of the OCP-W288\_BTA(P) species are determined by features attributable to domain separation and complete translocation of the chromophore into the NTD, as present in the photoactivated form of OCP. The aforementioned spectra are also very similar to the absorbance spectrum of the Red Carotenoid Protein when it coordinates CAN (i.e., the OCP-NTD holoprotein produced by heterologous expression in *E. coli*) (Moldenhauer et al., 2017), but slightly blue-shifted and distinct

from the absorbance spectrum of the pink fraction of the OCP-W288A mutant protein previously characterized (Maksimov et al., 2016) (Supplementary Material S1, Supplementary Figure S2).

We were able to crystallize the orange species of OCP-W288\_BTA containing a mixture of ECN and CAN and determined its structure (PDB: 7ZXV) with a resolution of 1.8 Å (see Table 2 for crystallographic information) by molecular substitution with the

<sup>1</sup> to be released upon publication.



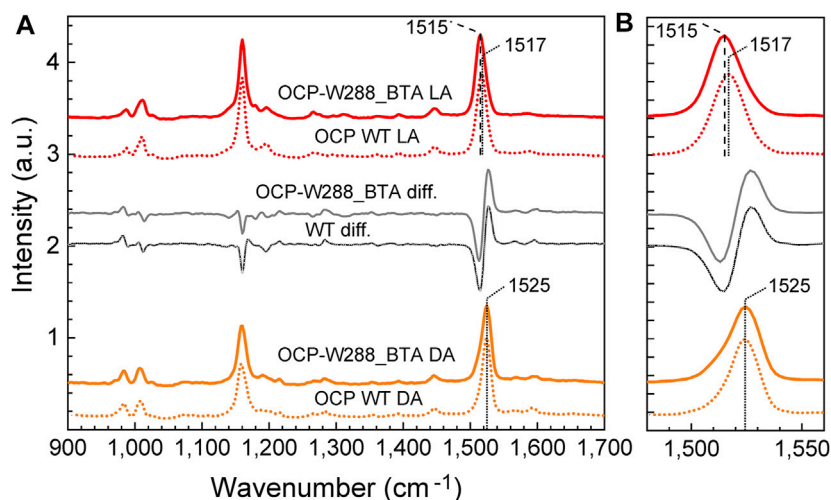
**FIGURE 5**

Crystal structure of the OCP-W288\_BTA variant (PDB: 7ZXV) and comparison with wild-type *Synechocystis* OCP (PDB: 3MG1). (A) Overlay of OCP-W288\_BTA (blue) and wild-type OCP (cyan) shows no global structural changes. Residues Tyr201, Trp/BTA288 and Ile40 are shown as ball-and-stick models (note the different colors of the C atoms: rose/orange/green in 7ZXV, and cyan in 3MG1). A  $\text{Cl}^-$  ion discovered in the 7ZXV structure is shown in green. (B) Sigma-A-weighted  $2m\text{Fo}-Df_c$  electron density map around BTA288, reflecting the distinct ring geometry of the BTA side chain. (C,D) Cofactor coordination in detail for OCP-W288\_BTA (C) and wild-type OCP (D). The distance between BTA288 and the xanthophyllic cofactors increases only slightly from 2.9 Å to 3.2/3.3 Å compared to coordination by Trp288 in wild-type OCP, the distance between Tyr201 and ECN is comparable to wild-type OCP and slightly increased for CAN. (E) The  $2m\text{Fo}-Df_c$  electron density map is shown around the ECN/CAN end rings in the NTD to illustrate the different ring geometries corresponding to the  $sp^2/sp^3$  hybridization state of the carbonyl carbon atom in CAN/ECN. (F) The  $2m\text{Fo}-Df_c$  electron density map around Ile40 in 7ZXV shows a different rotamer state of the side chain (green) than in the 3MG1 structure (cyan). (G) The  $2m\text{Fo}-Df_c$  (blue) and  $\text{Fo}-F_c$  (red) electron density maps around the xanthophyll cofactor in OCP-W288\_BTA, in which only ECN is fitted, clearly show the difference electron density attributable to a missing keto oxygen atom (red). (H) The  $2m\text{Fo}-Df_c$  electron density map in which both ECN and CAN were fitted showed occupancies of 45% and 55%, respectively, after refinement.

structure of wild-type OCP as a template (PDB: 5TUW) containing an ECN cofactor. As expected, the overall structure is very similar to the wild-type OCP, with RMSD values of 0.180 Å and 0.203 Å for monomer and dimer alignments at the Ca-positions, respectively. A close inspection of Figure 5A clearly shows that the BTA mutation has no significant effect on the global OCP structure, and the different geometry of the BTA side chain (Figure 5B; C-S-C angle of 92.5°) compared to Trp288 (C-N-C angle of 109°) is nicely reflected by the electron density (Figure 5B and Supplementary Material S1, Supplementary Figure S3). The introduced BTA residue matches well with its native Trp288 counterpart, resulting in an almost identical spatial arrangement (i.e., isomorphic replacement). Interestingly, the distance to the keto oxygen atoms of the ECN/CAN cofactors increases only slightly from 2.9 Å in the wild-type protein to 3.2 Å and 3.3 Å, respectively (Figures 5C, D), despite the potential repulsion of the lone electron pairs of sulphur, which, however, do not appear as a significant electron density in the

$2m\text{Fo}-Df_c$  map (Figure 5B, and Supplementary Material S1, Supplementary Figure S3). In addition, the distance between Tyr201 and the keto oxygen of ECN and is almost identical at 2.5 Å compared to 2.6 Å in wild-type OCP. However, the distance to CAN is slightly increased to 2.8 Å (Figures 5C, D). Here, the carbonyl carbon atom of CAN has an  $sp^2$  hybridization state, whereas the corresponding carbon atom of ECN is  $sp^3$ -hybridized, which affects the overall ring geometry in the NTD (Figure 5E) and is probably responsible for the small shift at the distal end of the xanthophyll in the CTD. Nevertheless, the protein environment of the xanthophylls seems to be unaffected by the presence of an additional keto function in the NTD, and the CAN configuration is almost identical to the one of the CAN cofactor in another *Synechocystis* wild-type OCP structure (PDB: 4XB5; Supplementary Material S1, Supplementary Figure S4). The only observable difference is that Cδ of Ile40 seems to be rotated 120° compared to the wild-type (Figure 5F), yet this does not result in a subsequent conformational





**FIGURE 6**

Resonance Raman spectra of wild-type OCP and the OCP-W288\_BTA variant in the dark-adapted and illuminated states. **(A)** Resonance Raman spectra of wild-type OCP (dotted curves) and the OCP-W288\_BTA variant (solid curves) in the dark-adapted (DA, orange curves) and light-adapted (LA, red curves) states. The peak wavenumbers of the  $\nu_1$  band are indicated. The grey curves show the difference spectra DA minus LA. The region of the  $\nu_1$ -band is shown in **(B)** at higher magnification. All dark- and light-matched spectra were intensity normalized based on the amplitude of the  $\nu_1$ -band and vertically offset for clarity.

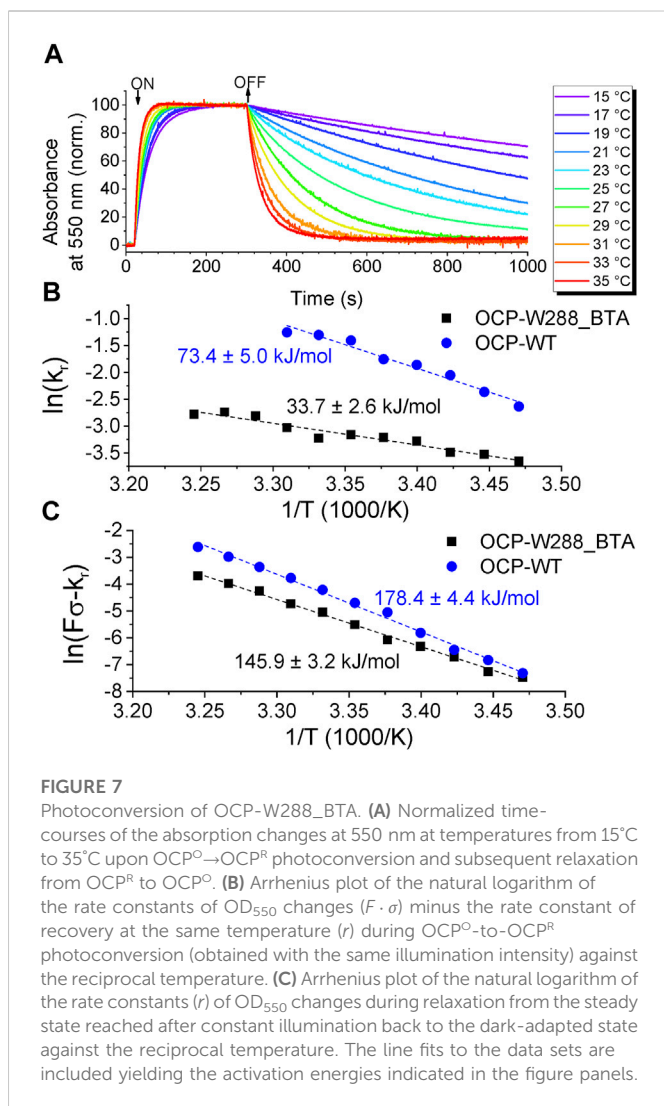
change. Overall, the differences due to the BTA mutation and the different cofactor notably result in very minor structural changes.

During refinement, the electron density of a  $\text{Cl}^-$  ion was delineated that was also present in another OCP mutant (Y201W, PDB: 6T6K) at the same position; which, however, is not present in the 3MG1 structure. The only potential OCP interaction partner in the proximity is Thr275. Interpretation of the xanthophyll cofactor electron density based on the wild-type structure with only one ECN molecule resulted in a difference density in the NTD (Figure 5G). Subsequently, the electron density was interpreted with an additional CAN molecule and occupancy refinement (Figure 5H), which led to the conclusion that both xanthophylls are present in the OCP-W288\_BTA mutant: The ratio of ECN/CAN ratio (45:55) corresponds to the ratio of xanthophylls extracted from the purified protein (Table 1).

To gain insight into the character of the interaction in the ECN-Tyr201-BTA288 complex, we performed a 10 ps quantum dynamics (QD) calculation. A comparison of the distances between the key atoms in the ECN-Tyr201-Trp288/BTA complexes (Supplementary Material S1, Supplementary Figure S5) clearly shows the absence of the energetically favourable interaction between the sulphur atom in BTA and the keto oxygen in ECN. The distance of 3.3 Å between the keto oxygen and the sulphur atom indicates the dispersion nature of the interaction in the ECN-BTA complex. Despite the high polarizability of the bulky sulphur atom, such an interaction does not significantly affect the OCP-W288\_BTA absorption spectrum in the dark-adapted state. Moreover, a slight trend toward shortening of the ECN-Tyr201 H-bond is observed in the presence of BTA. This effect is consistent with the hypothesis that the absence of a competing H-bond donor for the keto oxygen enhances the interaction with the remaining H-bond donor (Yaroshevich et al., 2021). In the case of OCP-W288\_BTA, the remaining H-bond is formed with Tyr201, making it similar to the previously reported OCP-Y201W mutant (Yaroshevich et al., 2021).

With the introduction of BTA into OCP, several new effects could be expected from the appearance of a bivalent sulphur atom in close proximity to the keto group of the bound xanthophyll. For example, the electrostatic interaction between the sulphur atom of a methionine and the carboxylate oxygen atom was suggested to play a major role in enzymatic catalysis (Taylor and Markham, 1999). Further analysis of structural data revealed that the bivalent sulphur atom forms specific non-bonded (sometimes called hypervalent) interactions with non-hydrogen atoms (O, N, C) within the protein backbone (Iwaoka and Isozumi, 2012). These relatively weak interactions can be classified into several types: Two-center-three-electron (Mishra et al., 2009), bifurcated (Zhang and Wang, 2009) and sulphur-carbene (Zhao et al., 2011) interactions. In addition, a theoretical analysis has shown that electron deficient, bivalent sulphur atoms can interact with electron donors by accepting their lone pair into low-lying  $\sigma^*$  orbitals (Beno et al., 2015). All these types of specific interactions are sensitive to the mutual orientations of the interacting counterparts and expected to be energetically significant at S...O distances less than 3 Å (Iwaoka and Isozumi, 2012; Beno et al., 2015). However, with respect to the interaction of low-lying  $\sigma^*$  orbitals of the C-S in the BTA bond with electron donors, we note that, first, the electron deficiency at the sulphur atom in BTA should be low [see Figure 8 in (Beno et al., 2015)], and, second, the sulphur (acceptor), oxygen (donor), and carbon (covalently bound to the sulphur in BTA) should be collinear ( $180^\circ$  C-S-O angle) for a favourable interaction to occur. Yet, the corresponding C-S-O angle from the 7ZXV crystal structure deviates substantially ( $160^\circ$ ), arguing against a significant contribution of the aforementioned non-bonded interactions.

Resonance Raman spectra of the photoswitchable OCP-W288\_BTA protein were recorded for dark- and light-adapted samples and compared with those of wild-type OCP (Figure 6), and those in the literature (Leverenz et al., 2014; Kish et al., 2015; Maksimov et al., 2016). In agreement with the established band assignment for carotenoids (Schlücker et al., 2003), the  $\nu_1$  band position around



1,525  $\text{cm}^{-1}$  is mainly due to C=C double bond stretching vibrations of the conjugated polyene chain, while the  $\nu_2$  band around 1,160–1,180  $\text{cm}^{-1}$  consists of contributions from stretching vibrations of C–C single bonds coupled with C–H in-plane bending modes. The  $\nu_3$  band at 1,000  $\text{cm}^{-1}$  arises from in-plane rocking vibrations of methyl groups on the conjugated polyene chain, coupled with in-plane bending modes of the neighboring C–H groups, and collectively considered as a fingerprint of the conjugated end-cycle configuration. Finally, the  $\nu_4$  band around 960  $\text{cm}^{-1}$  stems from C–H out-of-plane wagging motions coupled with C=C torsional modes (out-of-plane twists of the carbon backbone). When the conjugated system of the carotenoid is planar, these out-of-plane modes are not coupled with the electronic transition, and these bands are not resonance-enhanced. However, distortions around C–C single bonds increase the coupling of these modes with the electronic transition, leading to an increase in their intensity (Kish et al., 2015). The  $\nu_1$ ,  $\nu_2$ ,  $\nu_3$  and  $\nu_4$  bands of dark-adapted wild-type OCP and the OCP-W288\_BTA variant were at the same positions and had very similar amplitudes (Figure 6).

After photoconversion, there was a clear red shift of the  $\nu_1$  band from 1,525  $\text{cm}^{-1}$  to 1,517  $\text{cm}^{-1}$  for wild-type OCP, and even to 1,515  $\text{cm}^{-1}$  for OCP-W288\_BTA, which is consistent with the

correlation between the position of the  $\nu_1$  band and the position of the maximum of the chromophore's absorbance spectrum (Kish et al., 2015). Theoretical studies have shown that this correlation is due to a combination of two effects: (i) the rotation of the  $\beta$ -ionone rings with respect to the conjugated system (Mori, 2016) and (ii) the local electric field in the chromophore-binding pocket (Gamiz-Hernandez et al., 2015). The red form of OCP-W288\_BTA even appears to be slightly more red-shifted than wild-type OCP, which can be attributed to the effect of CAN embedment, which also causes a slight red-shift in the absorbance spectrum compared to ECN embedment (de Carbon et al., 2015). After photoconversion, both proteins show a decrease in the intensity of the  $\nu_4$  band since the carotenoid assumes a more planar configuration in the photoactivated state. The dark-minus-light difference spectra for the wild-type OCP and the BTA-bearing variant are virtually identical, leaving only the slight red shift in the  $\nu_1$  band of the photoactivated OCP-W288\_BTA by 2  $\text{cm}^{-1}$  as a difference (Figure 6B).

Next, we investigated the kinetics of transitions between the dark-adapted OCP<sup>O</sup> state and the light-adapted OCP<sup>R</sup> state and monitored the OCP<sup>O</sup>→OCP<sup>R</sup> photoconversion and subsequent recovery as the change in absorbance at 550 nm, always at the same illumination intensity (fluence  $F$ ) at different temperatures (Figure 7A). Fitting the rising and decaying phase of the curves with a monoexponential function yielded the rate constants for photoconversion and recovery [rate constants  $F \cdot \sigma$  and  $k_r$ , respectively, see the procedure described by (Shirshin et al., 2017)]. The resulting Arrhenius diagrams for the photoconversion and recovery rate constants for OCP-W288\_BTA variant and wild-type OCP are shown in Figures 7B, C, respectively. The data show that the OCP<sup>O</sup>→OCP<sup>R</sup> photoconversion of OCP-W288\_BTA is significantly slowed compared with wild-type OCP. Moreover, the activation energy ( $E_A$ ) for OCP-W288\_BTA determined from the Arrhenius plots is substantially reduced by ~40 kJ/mol compared with wild-type OCP. The same trend is observed for the rate constant for recovery to the ground state: The rate constant of OCP-W288\_BTA is slowed down compared to wild-type OCP, and  $E_A$  is lower by ~33 kJ/mol. Assuming that the differences in  $E_A$  are due to the deletion of a single H-bond, the determined energy differences classify the bond between Trp288 and ECN as a moderately strong intermolecular H-bond, since the dissociation energies of H-bonds in the condensed phase range from 0.8 kJ/mol to 162 kJ/mol [the latter being an exceptionally large value unique for HF<sub>2</sub><sup>-</sup> (Steiner, 2002)], which is much larger than the value of 9 kJ/mol for H-bonds in water at 298 K (van der Spoel et al., 2006).

## 4 Conclusion

Since the first high-resolution crystal structures of OCP were published (Kerfeld et al., 2003; Wilson et al., 2010), several studies on OCP have investigated the role of the absolutely conserved Tyr201 and Trp288 residues in the functional activity of this protein (Wilson et al., 2010; Wilson et al., 2011; Maksimov et al., 2016; Kerfeld et al., 2017; Yaroshevich et al., 2021). However, the limited possibilities to substitute aromatic amino acids within the standard amino acid repertoire led to ambiguous results, and overwhelmingly to the loss of the most important property of OCP—its photoactivity. To understand the role of H-bonds, one

must be able to control their presence or absence, the most difficult task being neither to disrupt the OCP<sup>O</sup> form of the protein, nor the ability to photoswitch. These requirements are met by the approach adopted in this work, in which the isosteric exchange of a nitrogen by a sulphur atom in the OCP-W288\_BTA variant precisely removes one of the two H-bond donors to the ketocarotenoid.

The X-ray crystal structure clearly reflects the minimal-invasive, hence isomorphous nature of this exchange, and the structural and spectroscopic properties of the OCP<sup>O</sup> state as well as photoactivity were preserved. Thus, we clearly demonstrate that Trp288 is not *required* for OCP photoactivity. The same is true for Tyr201, as shown by our previous work on the OCP-Y201W variant generated by classical site-directed mutagenesis (Yaroshevich et al., 2021). Thus, it is obvious that the presence of at least one H-bond donor at either of the two sites is *sufficient* to ensure proper formation of the OCP<sup>O</sup> state and photoactivity. Interestingly, removing one of the two H-bond donors competing for the keto oxygen leads to the intriguing consequence that the carotenoid forms a *more stable interaction* with the remaining H-bond donor. As a result, the protein lacks the spectral heterogeneity characteristic of the dark-adapted state of the wild-type OCP protein.

From the point of view of a protein engineer or a spectroscopist, the spectral homogeneity of the dark-adapted OCP<sup>O</sup> state and, consequently, the higher yield of the S\* photoproduct, which was proposed to be crucial for the productive photocycle (Yaroshevich et al., 2021), could be considered as advantageous evolutionary selection criteria. Therefore, one may speculate why nature solved the problem of designing the OCP photoswitch with two H-bond donors, because obviously the selection criterion was neither the spectral homogeneity of the OCP<sup>O</sup> ground state nor the quantum efficiency of the photoconversion. Rather, the incorporation of two H-bond donors competing for the ketocarotenoid was a compromise to ensure that the formation of the dark-adapted OCP<sup>O</sup> state from the highly flexible structural arrangement of the domain-separated OCP<sup>R</sup> state occurs with high reliability. Apparently, the two competing H-bond donors provide the necessary flexibility or functional redundancy for the compact dark-adapted OCP<sup>O</sup> state to be formed.

Replacement of Trp288 with BTA reduced the activation energies of photoconversion and recovery by approximately 40 kJ/mol and 33 kJ/mol, respectively, compared with wild-type OCP. Attributing these enthalpy differences to the deletion of a single H-bond, the determined values classify the N-H...O = C bond between Trp288 and ECN as a moderately strong intermolecular H-bond. Since all H-bonds can be considered as precursors to proton transfer reactions, which may be well advanced for strong H-bonds (Steiner, 2002), the determined enthalpy differences due to the deletion of an H-bond in OCP are important, especially in view of the recently proposed photoactivation pathway of OCP involving an oxocarbenium ion at the ECN cofactor (Yaroshevich et al., 2021).

Undoubtedly, we now have a much better understanding of the nature of the interactions between the carotenoid ligands and protein matrix of the OCP. Moreover, how the presence of a “surrogate” or “redundant” hydrogen bond may affect the photoactivation process and its biological role has not been studied before. The low quantum yield in the formation of the primary and all subsequent photoproducts in the OCP photocycle (Shirshin et al., 2017) is obviously important for maintaining the balance between photosynthetic activity and photoprotective processes in

cyanobacterial cells. We have demonstrated that the use of chemically controlled molecular operations based on the incorporation of non-canonical amino acids offers great potential for better elucidation of protein design by nature, but also for the engineering of such complex and biologically relevant chemical photoswitches in the future.

## Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/Supplementary Material.

## Author contributions

MM and H-WT, performed molecular biology, protein biochemistry, processed and interpreted data and prepared figures. MM, performed UV-Vis spectroscopy experiments. H-WT, performed mass spectrometric analysis. AK, performed Raman spectroscopic experiments, processed and interpreted data. NT, performed protein biochemistry experiments and purified proteins. IY, performed quantum chemical calculations and interpreted data. PH designed experiments and interpreted Raman spectroscopy data. NS, structured and interpreted structural data. GH, designed experiments and processed X-ray crystallography data. L-OE, designed experiments, processed X-ray crystallography data. LK: acquired, analysed and processed X-ray crystallography data and prepared figures. EM, designed experiments, analysed, structured and interpreted data. NB and TF, designed experiments, administered project, analysed data and wrote the paper. All authors contributed to correction and proof-reading of the manuscript.

## Funding

H-WT received funding under the Marie-Skłodowska-Curie grant agreement No. 764591 by the SynCrop ETN consortium of the European Union (Horizon 2020 research and innovation program), NB by the Canada Research Chairs Program grant agreement No. 950-23197, EM and NS by the Ministry of Science and Higher Education of the Russian Federation in the framework of the Agreement no. 075-15-2021-1354, and TF by the German Research Foundation (DFG, grant no. 1276/6-1).

## Acknowledgments

The authors are grateful for financial support by the SynCrop ETN consortium of the European Union (Horizon 2020 research and innovation program) under the Marie Skłodowska-Curie grant agreement No. 764591 (to H-WT), by the Canada Research Chairs Program grant agreement No. 950-23197 (to NB), the Ministry of Science and Higher Education of the Russian Federation in the framework of the Agreement no. 075-15-2021-1354 to EM and NS (07.10.2021), the German Research Foundation to TF (grant no. 1276/6-1) and the Einstein Foundation Berlin for granting an Azura FPLC machine. We also acknowledge Prof. Christian Hackenberger for

providing the Waters XEVO G2-XS QToF analyzer, and Philipp Ochtrop, Christian Stieger, and Mark de Geus for assistance in protein mass-spectrometry and analysis at Leibniz-Forschungsinstitut für Molekulare Pharmakologie (FMP) in Berlin. We thank the staff of the P14 beamline (EMBL Hamburg, Germany) at Petra III storage ring (DESY, Hamburg, Germany) and the staff of the Marburg crystallization facility (MarXtal) for technical support.

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## References

- Afonine, P. V., Grosse-Kunstleve, R. W., Echols, N., Headd, J. J., Moriarty, N. W., Mustyakimov, M., et al. (2012). Towards automated crystallographic structure refinement with phenix.refine. *Acta Cryst. D.* 68, 352–367. doi:10.1107/S0907444912001308
- Beno, B. R., Yeung, K.-S., Bartberger, M. D., Pennington, L. D., and Meanwell, N. A. (2015). A survey of the role of noncovalent sulfur interactions in drug design. *J. Med. Chem.* 58, 4383–4438. doi:10.1021/jm501853m
- Bussi, G., and Parrinello, M. (2007). Accurate sampling using Langevin dynamics. *Phys. Rev. E* 75, 056707. doi:10.1103/PhysRevE.75.056707
- de Carbon, C. B., Thurotte, A., Wilson, A., Perreau, F., and Kirilovsky, D. (2015). Biosynthesis of soluble carotenoid holoproteins in *Escherichia coli*. *Sci. Rep.* 5, 9085. doi:10.1038/srep09085
- Ditchfield, R., Hehre, W. J., and Pople, J. A. (1971). Self-consistent molecular-orbital methods. IX. An extended Gaussian-type basis for molecular-orbital studies of organic molecules. *J. Chem. Phys.* 54, 724–728. doi:10.1063/1.1674902
- Emsley, P., and Cowtan, K. (2004). Coot: Model-building tools for molecular graphics. *Acta Crystallogr. D. Biol. Crystallogr.* 60, 2126–2132. doi:10.1107/S0907444904019158
- Evans, P. R., and Murshudov, G. N. (2013). How good are my data and what is the resolution? *Acta Crystallogr. D. Biol. Crystallogr.* 69, 1204–1214. doi:10.1107/S0907444913000061
- Gamiz-Hernandez, A. P., Angelova, I. N., Send, R., Sundholm, D., and Kaila, V. R. I. (2015). Protein-induced color shift of carotenoids in  $\beta$ -crustacyanin. *Angew. Chem. Int. Ed.* 54, 11564–11566. doi:10.1002/anie.201501609
- Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M. R., Appel, R. D., et al. (2005). "Protein identification and analysis tools on the ExPASy server," in *The proteomics protocols handbook*. Editor J. M. Walker (Totowa, NJ: Humana Press), 571–607. doi:10.1385/1-59259-890-0:571
- Gupta, S., Guttman, M., Leverenz, R. L., Zhumadilova, K., Pawlowski, E. G., Petzold, C. J., et al. (2015). Local and global structural drivers for the photoactivation of the orange carotenoid protein. *Proc. Natl. Acad. Sci.* 112, E5567–E5574. doi:10.1073/pnas.1512240112
- Hauf, M., Richter, F., Schneider, T., Faidt, T., Martins, B. M., Baumann, T., et al. (2017). Photoactivatable mussel-based underwater adhesive proteins by an expanded genetic code. *ChemBioChem* 18, 1819–1823. doi:10.1002/cbic.201700327
- Holt, T. K., and Krogmann, D. W. (1981). A carotenoid-protein from cyanobacteria. *Biochimica Biophysica Acta (BBA) - Bioenergetics* 637, 408–414. doi:10.1016/0005-2728(81)90045-1
- Hoopes, J. T., Elbersson, M. A., Preston, R. J., Reddy, P. T., and Kelman, Z. (2015). Protein labeling in *Escherichia coli* with (2)H, (13)C, and (15)N. *Methods Enzymol.* 565, 27–44. doi:10.1016/bs.mie.2015.08.023
- Iwaoka, M., and Isozumi, N. (2012). Hypervalent nonbonded interactions of a divalent sulfur atom. Implications in protein architecture and the functions. *Molecules* 17, 7266–7283. doi:10.3390/molecules17067266
- Kabsch, W. (1993). Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. *J. Appl. Cryst.* 26, 795–800. doi:10.1107/S0021889893005588
- Kerfeld, C. A., Melnick, 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. doi:10.1111/nph.14670
- Kerfeld, C. A., Sawaya, M. R., Brahmamandam, V., Cascio, D., Ho, K. K., Trevithick-Sutton, C. C., et al. (2003). The crystal structure of a cyanobacterial water-soluble carotenoid binding protein. *Structure* 11, 55–65. doi:10.1016/s0969-2126(02)00936-x
- Kish, E., Pinto, M. M. M., Kirilovsky, D., Spezia, R., and Robert, B. (2015). Echinone vibrational properties: From solvents to the orange carotenoid protein. *Biochimica Biophysica Acta (BBA) - Bioenergetics* 1847, 1044–1054. doi:10.1016/j.bbabi.2015.05.010
- Ko, J., Wang, Y.-S., Nakamura, A., Guo, L.-T., Söll, D., and Umehara, T. (2013). Pyrrolysyl-tRNA synthetase variants reveal ancestral aminoacylation function. *FEBS Lett.* 587, 3243–3248. doi:10.1016/j.febslet.2013.08.018
- 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. *Plant Cell* 26, 426–437. doi:10.1105/tpc.113.118588
- Leverenz, R. L., Sutter, M., Wilson, A., Gupta, S., Thurotte, A., Bourcier de Carbon, C., et al. (2015). PHOTOSYNTHESIS. A 12 Å carotenoid translocation in a photoswitch associated with cyanobacterial photoprotection. *Science* 348, 1463–1466. doi:10.1126/science.aaa7234
- Maksimov, E. G., Moldenhauer, M., Shirshin, E. A., Parshina, E. A., Sluchanko, N. N., Klementiev, K. E., et al. (2016). A comparative study of three signaling forms of the orange carotenoid protein. *Photosynth. Res.* 130, 389–401. doi:10.1007/s11120-016-0272-8
- Maksimov, E. G., Shirshin, E. A., Sluchanko, N. N., Zlenko, D. V., Parshina, E. Y., Tsoraev, G. V., et al. (2015). The signaling state of orange carotenoid protein. *Biophys. J.* 109, 595–607. doi:10.1016/j.bpj.2015.06.052
- Mardirossian, N., and Head-Gordon, M. (2014).  $\omega$ B97X-V: A 10-parameter, range-separated hybrid, generalized gradient approximation density functional with nonlocal correlation, designed by a survival-of-the-fittest strategy. *Phys. Chem. Chem. Phys.* 16, 9904–9924. doi:10.1039/C3CP54374A
- McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007). Phaser crystallographic software. *J. Appl. Crystallogr.* 40, 658–674. doi:10.1107/S0021889807021206
- Minks, C., Huber, R., Moroder, L., and Budisa, N. (1999). Atomic mutations at the single tryptophan residue of human recombinant annexin V: Effects on structure, stability, and activity. *Biochemistry* 38, 10649–10659. doi:10.1021/bi990580g
- Misawa, N., Satomi, Y., Kondo, K., Yokoyama, A., Kajiwara, S., Saito, T., et al. (1995). Structure and functional analysis of a marine bacterial carotenoid biosynthesis gene cluster and astaxanthin biosynthetic pathway proposed at the gene level. *J. Bacteriol.* 177, 6575–6584. doi:10.1128/jb.177.22.6575-6584.1995
- Mishra, B., Sharma, A., Naumov, S., and Priyadarsini, K. I. (2009). Novel reactions of one-electron oxidized radicals of selenomethionine in comparison with methionine. *J. Phys. Chem. B* 113, 7709–7715. doi:10.1021/jp900322z
- Moldenhauer, M., Sluchanko, N. N., Bührke, D., Zlenko, D. V., Tavraz, N. N., Schmitt, F.-J., et al. (2017). Assembly of photoactive orange carotenoid protein from its domains unravels a carotenoid shuttle mechanism. *Photosynth. Res.* 133, 327–341. doi:10.1007/s11120-017-0353-3
- Mori, Y. (2016). Computational study on the color change of 3'-hydroxyechinenone in the orange carotenoid protein. *Chem. Phys. Lett.* 652, 184–189. doi:10.1016/j.cplett.2016.04.062
- Muzzopappa, F., and Kirilovsky, D. (2020). Changing color for photoprotection: The orange carotenoid protein. *Trends Plant Sci.* 25, 92–104. doi:10.1016/j.tplants.2019.09.013
- Neese, F., Wennmohs, F., Becker, U., and Riplinger, C. (2020). The ORCA quantum chemistry program package. *J. Chem. Phys.* 152, 224108. doi:10.1063/5.0004608
- Schlücker, S., Szeghalmi, A., Schmitt, M., Popp, J., and Kiefer, W. (2003). Density functional and vibrational spectroscopic analysis of  $\beta$ -carotene. *J. Raman Spectrosc.* 34, 413–419. doi:10.1002/jrs.1013
- Seritan, S., Bannwarth, C., Fales, B. S., Hohenstein, E. G., Kokkila-Schumacher, S. I. L., Luehr, N., et al. (2020). TeraChem: Accelerating electronic structure and *ab initio* molecular dynamics with graphical processing units. *J. Chem. Phys.* 152, 224110. doi:10.1063/5.0007615
- Shirshin, E. A., Nikonova, E. E., Kuzminov, F. I., Sluchanko, N. N., Elanskaya, I. V., Gorbunov, M. Y., et al. (2017). Biophysical modeling of *in vitro* and *in vivo* processes underlying regulated photoprotective mechanism in cyanobacteria. *Photosynth. Res.* 133, 261–271. doi:10.1007/s11120-017-0377-8

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmolb.2023.1072606/full#supplementary-material>

- Sluchanko, N. N., Klementiev, K. E., Shirshin, E. A., Tsoraev, G. V., Friedrich, T., and Maksimov, E. G. (2017a). The purple Trp288Ala mutant of *Synechocystis* OCP persistently quenches phycobilisome fluorescence and tightly interacts with FRP. *Biochim. Biophys. Acta Bioenerg.* 1858, 1–11. doi:10.1016/j.bbabi.2016.10.005
- Sluchanko, N. N., Slonimskiy, Y. B., Moldenhauer, M., Friedrich, T., and Maksimov, E. G. (2017b). Deletion of the short N-terminal extension in OCP reveals the main site for FRP binding. *FEBS Lett.* 591, 1667–1676. doi:10.1002/1873-3468.12680
- Steiner, T. (2002). The hydrogen bond in the solid state. *Angew. Chem. Int. Ed.* 41, 48–76. doi:10.1002/1521-3773(20021014)41:1<48::AID-ANIE48>3.0.CO;2-U
- Taylor, J. C., and Markham, G. D. (1999). The bifunctional active site of s-adenosylmethionine synthetase. Roles of the active site aspartates. *J. Biol. Chem.* 274, 32909–32914. doi:10.1074/jbc.274.46.32909
- Thi To, T. M., Kubyshkin, V., Schmitt, F.-J., Budisa, N., and Friedrich, T. (2022). Residue-specific exchange of proline by proline analogs in fluorescent proteins: How “molecular surgery” of the backbone affects folding and stability. *J. Vis. Exp.* 180, e63320. doi:10.3791/63320
- Tseng, H.-W., Baumann, T., Sun, H., Wang, Y.-S., Ignatova, Z., and Budisa, N. (2020). Expanding the scope of orthogonal translation with pyrrolysyl-tRNA synthetases dedicated to aromatic amino acids. *Molecules* 25, E4418. doi:10.3390/molecules25194418
- Tyler, R. C., Sreenath, H. K., Singh, S., Aceti, D. J., Bingman, C. A., Markley, J. L., et al. (2005). Auto-induction medium for the production of [U-15N]- and [U-13C, U-15N]-labeled proteins for NMR screening and structure determination. *Protein Expr. Purif.* 40, 268–278. doi:10.1016/j.pep.2004.12.024
- van der Spoel, D., van Maaren, P. J., Larsson, P., and Timneanu, N. (2006). Thermodynamics of hydrogen bonding in hydrophilic and hydrophobic media. *J. Phys. Chem. B* 110, 4393–4398. doi:10.1021/jp0572535
- Weigend, F., and Ahlrichs, R. (2005). Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: Design and assessment of accuracy. *Phys. Chem. Chem. Phys.* 7, 3297–3305. doi:10.1039/B508541A
- Wilson, A., Kinney, J. N., Zwart, P. H., Punginelli, C., D’Haene, S., Perreau, F., et al. (2010). Structural determinants underlying photoprotection in the photoactive orange carotenoid protein of cyanobacteria. *J. Biol. Chem.* 285, 18364–18375. doi:10.1074/jbc.M110.115709
- Wilson, A., Punginelli, C., Couturier, M., Perreau, F., and Kirilovsky, D. (2011). Essential role of two tyrosines and two tryptophans on the photoprotection activity of the Orange Carotenoid Protein. *Biochim. Biophys. Acta Bioenerg.* 1807, 293–301. doi:10.1016/j.bbabi.2010.12.009
- Yaroshevich, I. A., Maksimov, E. G., Sluchanko, N. N., Zlenko, D. V., Stepanov, A. V., Slutskaia, E. A., et al. (2021). Role of hydrogen bond alternation and charge transfer states in photoactivation of the Orange Carotenoid Protein. *Commun. Biol.* 4, 539. doi:10.1038/s42003-021-02022-3
- Zhang, Y., and Wang, W. (2009). The bifurcate chalcogen bond: Some theoretical observations. *J. Mol. Struct. THEOCHEM* 916, 135–138. doi:10.1016/j.theochem.2009.09.021
- Zhao, Q., Feng, D., Sun, Y., Hao, J., and Cai, Z. (2011). Theoretical investigations on the weak nonbonded C-S...CH<sub>2</sub> interactions: Chalcogen-bonded complexes with singlet carbene as an electron donor. *Int. J. Quantum Chem.* 111, 3881–3887. doi:10.1002/qua.22928