

Article

# Spectral Features of Canthaxanthin in HCP2. A QM/MM Approach

Kevin Clark <sup>1</sup>, Natalia B. Pigni <sup>1,2</sup> , Kithmini Wijesiri <sup>1</sup> and José A. Gascón <sup>1,\*</sup> 

<sup>1</sup> Department of Chemistry, University of Connecticut, Storrs, CT 06269-3060, USA; kevin.l.clark@uconn.edu (K.C.); natalia.pigni@uconn.edu (N.B.P.); kithmini.wijesiri@uconn.edu (K.W.)

<sup>2</sup> Instituto de Ciencia y Tecnología de Alimentos Córdoba (ICYTAC-CONICET), Ciudad Universitaria, Córdoba X5000HUA, Argentina; npigni@unc.edu.ar

\* Correspondence: jose.gascon@uconn.edu; Tel.: +1-860-486-0591

**Abstract:** The increased interest in sequencing cyanobacterial genomes has allowed the identification of new homologs to both the N-terminal domain (NTD) and C-terminal domain (CTD) of the Orange Carotenoid Protein (OCP). The N-terminal domain homologs are known as Helical Carotenoid Proteins (HCPs). Although some of these paralogs have been reported to act as singlet oxygen quenchers, their distinct functional roles remain unclear. One of these paralogs (HCP2) exclusively binds canthaxanthin (CAN) and its crystal structure has been recently characterized. Its absorption spectrum is significantly red-shifted, in comparison to the protein in solution, due to a dimerization where the two carotenoids are closely placed, favoring an electronic coupling interaction. Both the crystal and solution spectra are red-shifted by more than 50 nm when compared to canthaxanthin in solution. Using molecular dynamics (MD) and quantum mechanical/molecular mechanical (QM/MM) studies of HCP2, we aim to simulate these shifts as well as obtain insight into the environmental and coupling effects of carotenoid–protein interactions.

**Keywords:** HCP2; QM/MM; canthaxanthin



**Citation:** Clark, K.; Pigni, N.B.; Wijesiri, K.; Gascón, J.A. Spectral Features of Canthaxanthin in HCP2. A QM/MM Approach. *Molecules* **2021**, *26*, 2441. <https://doi.org/10.3390/molecules26092441>

Academic Editor: Stacey Wetmore

Received: 29 March 2021

Accepted: 18 April 2021

Published: 22 April 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Photosynthetic organisms produce a specific type of pigment known as carotenoids. Although carotenoids are highly hydrophobic, their solubility is increased by binding to proteins. Once bound to photosynthetic related proteins, carotenoids can adopt specific orientations, functioning as accessory pigments by absorbing excess photons not captured by other chromophores [1]. In addition, carotenoid-binding proteins play protective roles such as stopping the formation of reactive oxygen species (ROS) by dissipating excess light energy or directly quenching singlet oxygen [2,3]. Many functions and structures of carotenoid-binding proteins require further study and are still being identified to this day [4].

Orange Carotenoid Protein (OCP) is part of an exceptional photoprotective mechanism in cyanobacteria. It has two well-defined domains with a ketocarotenoid spanning both domains. The N-terminal domain (NTD), unique to cyanobacteria, consists of an all  $\alpha$ -helix structure and acts as the “effector” domain [5]. Recently, there has been an increased interest in sequencing cyanobacterial genomes, which has resulted in the identification of new OCP families as well as homologs of both domains [6–8]. To date, at least nine different NTD homologs have been discovered with diverse functionalities [4]. These homologs are known as Helical Carotenoid Proteins (HCPs). They present similar helical folds to those described for the NTD of OCP and also have the ability to bind carotenoids. Homologs to the C-terminal domain (CTD) of OCP, known as CTDHs, have also been found in almost every genome encoding HCPs [4,7]. Protein evolution studies, combining gene fusion with taxonomic species distribution, have suggested that OCP was likely derived from a domain fusion event between an HCP and a CTDH [5,6,8,9].

In this work, we focus our investigation on HCP2; one of the most widespread subtypes of the HCP family found in cyanobacteria. Recently, it has been isolated from *Tolypothrix* sp.

and its crystal structure was reported with a 1.7 Å resolution (PDB: 6MCJ) [9]. HCP2 exclusively binds canthaxanthin (CAN) and is an effective singlet oxygen quencher. However, unlike the active form of OCP, it is unable to bind to the phycobilisome (PBS) to quench excess energy [9]. In solution, HCP2 is found as a monomer, while the crystal structure reveals a dimeric structure with a 5 Å separation between the stacked  $\beta$ 1-rings of the two CAN molecules. The UV–visible absorption maximum of the crystal structure of HCP2 is red-shifted by 18 nm when compared to HCP2 in solution (548 nm and 530 nm, respectively), and both are significantly red-shifted (>50 nm) with respect to CAN in solution [4,9]. The shape of the carotenoid absorption bands is directly related to the conjugation of the polyene chain and the conformation of the terminal rings [10–12].

Experimental studies of CAN in solution have been performed using different solvents, including benzene, methanol and *n*-hexane [9,13]. The observed absorption maxima vary slightly depending on the solvent. Besides the shift in the absorption maximum, the absorption lineshape of HCP2 in solution is very similar to that of the isolated CAN in solution. This suggests that the carotenoid ending rings have a comparable freedom to rotate inside the protein [14]. In this work, we aim to explain the spectral features observed in experimental studies by means of quantum mechanical/molecular mechanical (QM/MM) and excitonic coupling calculations. Additionally, we use molecular dynamics (MD) simulations in order to reveal the interactions and conformations associated with the UV–vis spectral features of the complex in solution.

## 2. Results and Discussion

### 2.1. QM/MM Calculations of HCP2 Models from the Crystal Structure

The crystal structure of HCP2 (PDB: 6MCJ) was used as the starting point for our QM/MM calculations. The protein complexes were prepared with Maestro (Schrödinger 2019-1), using the entire PDB structure for calculations of the dimer, and the two individual chains separately (A and B) for calculations of the monomeric structures. We performed a local minimization of the carotenoid in the QM region, leaving the protein frozen. On average, this local minimization produced only small changes of the polyene chain in the protein cavity ( $\sim 0.20$  Å).

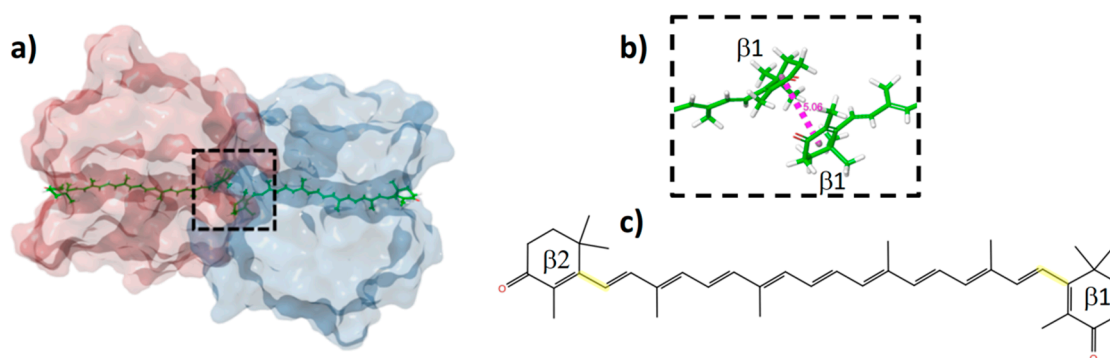
Figure 1a shows the HCP2 model from the crystal structure. The dimer is composed of two monomers, each non-covalently bound to a CAN molecule. The  $\beta$ 1 rings of both carotenoids are placed in close proximity ( $\sim 5$  Å) showing a stacked alignment. Due to this interaction, the solvent-exposed area of the carotenoids is reduced when compared to the isolated monomers. It has been demonstrated that the main oligomeric state of HCP2 in solution is as a monomer. Thus, in solution, it is expected that the chromophore experiences an enhanced conformational freedom, in contrast to the dimeric structure where the protein environment constrains the  $\beta$ 1 terminal ring rotation [9].

During the system preparation process of the dimer (previous to the QM/MM calculations), we found one particular glutamic acid (Glu94), located in the proximity of  $\beta$ 1 ring ( $< 5$  Å), which became protonated (neutrally charged) after hydrogen-bonding optimization at pH 7. In contrast, the same procedure applied to each individual monomer assigns a negatively charged state for Glu94. It is already known that the pKa values of ionizable groups of acidic and basic residues can be affected by the protein microenvironment, differing from the normal values observed in water [15]. Interestingly, this particular residue is the closest charged amino acid to the  $\beta$ 1 terminal ring in the monomer. As previously discussed, the  $\beta$ 1 terminal rings have significant environmental differences between monomeric and dimeric structures.

Post-optimization single-point QM/MM excitation energy calculations were performed on both oligomeric states of HCP2, and the results are shown in Table 1. Note that the monomers extracted from the crystal structure show a 14 nm difference between the calculated  $\lambda_{\max}$  values, which is probably explained by the existence of different conformations in  $\beta$ 1. The end-ring torsions in  $\beta$ -carotene-like molecules are known to have a marked influence on the electronic properties and excitation energies [16]. In the case of CAN in

solution, despite the freedom to rotate, the lowest energy conformations are predominant with torsions around  $45^\circ$ , restricting the conjugation of the system to  $\sim 465\text{--}480$  nm absorption maximum values, depending on the particular solvent [9,13]. Red-shifted states occur for a completely planar conjugated system, *s-cis* ( $0^\circ$ ) or *s-trans* ( $180^\circ$ ), but this is prevented in solution by intrinsic steric interactions between the rings and the methyl groups of the isoprenoid chain. In the crystal structure of HCP2, the  $\beta 2$  rings of both monomers are nominally *s-trans*, while the  $\beta 1$  rings show a nominally *s-cis* configuration. After performing a QM optimization of CAN inside the protein environment, both dihedrals maintain the same nominal conformation with dihedrals of  $\sim 179^\circ$  for  $\beta 2$ . However, the *s-cis*  $\beta 1$  ring torsions reveal different values for each monomer favoring a larger conjugation for ligand B, which is reflected in a red-shifted absorption maximum (Table 1), in much closer agreement to the experimental absorption  $\lambda_{\max}$  (530 nm). Thus, we conclude that the best crystal structure to construct solvated models of the of HCP2 in its natural form is monomer B.

Regarding the experimental 18 nm shift observed between the  $\lambda_{\max}$  of the dimer (548 nm) and monomer in solution (530 nm) [9], the results from our calculations on the crystal structures show a calculated shift of 23 nm (comparing the dimer with monomer B), which is in very good agreement with the experimental result.



**Figure 1.** HCP2 dimer model (PDB: 6MCJ). (a) Protein surface representation with each monomer colored differently and canthaxanthin (CAN) molecules shown as green sticks. (b) Enlarged view of the dashed square in panel (a) to highlight the close interaction between the two  $\beta 1$  rings. (c) CAN with labeled  $\beta 1$  and  $\beta 2$  rings. Highlighted in yellow are the bonds for which dihedral angles will be discussed throughout this work.

**Table 1.** Quantum mechanical/molecular mechanical (QM/MM) calculations on optimized structure of HCP2. The  $\beta 1$  and  $\beta 2$  dihedral angles correspond to the QM/MM minimized structure.

Model	QM/MM $\lambda_{\max}$ (nm)	$\beta 1$ Dihedrals	$\beta 2$ Dihedrals	Experimental $\lambda_{\max}$ (nm)
Dimer	515			548 [9]
Monomer Ligand A	478	$74.0^\circ$	$179.3^\circ$	530 [9]
Monomer Ligand B	492	$37.2^\circ$	$178.3^\circ$	
Shift (Dimer–Monomer)	23 *			18
CAN (in <i>n</i> -hexane)	465	$40.7^\circ$	$143.3^\circ$	465 [13]
CAN (in water)	470	$41.4^\circ$	$140.3^\circ$	N/A

\* Reported shift corresponds to the difference between the dimer and the monomer with highest  $\lambda_{\max}$ .

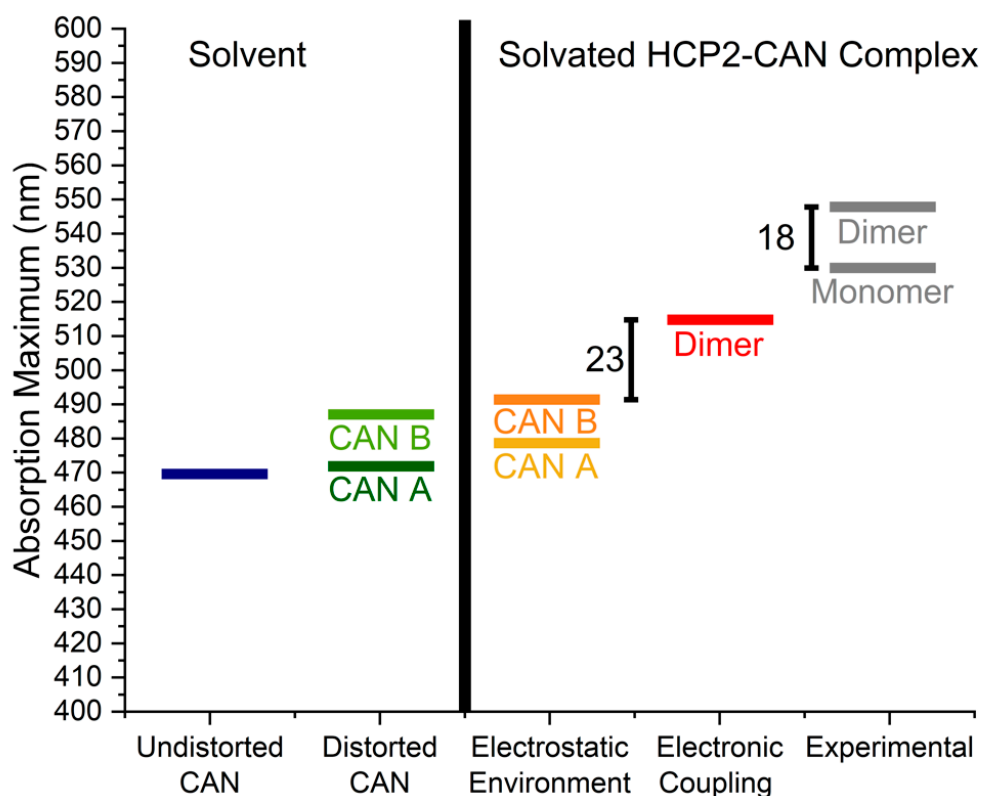
## 2.2. Distortions, Environmental Influence, and Excitonic Coupling

There are three effects on the absorption spectrum that were evaluated: (1) the mechanical distortion as a result of change in conformation from solvent to protein, (2) the combined electrostatic and screening effects due to the protein and solvent, and (3) the effect of electronic coupling between the two chromophores. Figure 2 illustrates how these effects progressively affect the excitation energy. While there are no experimental values of the excitation energy of CAN in water, we still considered CAN in water theoretically to be able to see the effects as CAN becomes surrounded by the protein cavity in an aqueous

environment. For monomer A, the mechanical distortion exerted by the protein produces a minimal change in absorption. This is in part due to the large out of plane value of  $\beta_1$  ( $74^\circ$ ) that breaks conjugation. On the other hand, the mechanical distortion of CAN in monomer B produces a substantial red shift due to a more nominally *s-cis* conformation in  $\beta_1$  ( $37^\circ$ ), which favors conjugation.

The electrostatic environment of the protein plus solvent causes a very small red shift of 3–5 nm in the calculated  $\lambda_{\max}$  for both monomers. This observation also applies to other related carotenoid-containing proteins (e.g., OCP, RCP) [17]. It is interesting to note the contrast with other light-driven proteins, such as type I and II rhodopsins, for which the tuning mechanism of the protein chromophores is strongly influenced by the electrostatic environment [18].

The last effect to consider is the strength of excitonic coupling (i.e., the electronic coupling between chromophores). While this is not relevant for the natural function of HCP2 (which occurs as a monomer in solution), it is worth asking what level of energy transfer we would expect between the two CANs in the crystalline state. To determine this, we computed the electronic coupling between the  $S_0 \rightarrow S_2$  excited states of the two CAN ligands in the presence of the protein and solvent environment using the electronic energy transfer methodology [19,20]. This last effect adds a substantial red shift (23 nm from monomer B), which explains the origin of the observed experimental red shift from solution to crystal. Thus, the 18 nm shift reported experimentally for HCP2 is closely recreated by our QM/MM calculations.



**Figure 2.** Absorption maxima calculated for canthaxanthin (CAN) after QM optimization in different environments using implicit solvent: CAN alone (undistorted), CAN isolated from the protein (distorted), and CAN inside HCP2 in two different oligomeric states. Experimental values are from reference [9].

Table 2 shows the site energies (i.e., localized excitons), the electronic energy transfer (EET) coupling between the two localized excitons, and the corresponding eigenvalues of the  $2 \times 2$  exciton Hamiltonian. The accuracy of the coupling calculations (0.061 eV) is evident in the close comparison between the eigenvalues and the two lowest excitation

energies of the full dimer calculation at the QM/MM level. It is interesting to compare the magnitude of the excitonic coupling ( $0.061 \text{ eV} = 492 \text{ cm}^{-1}$ ) with recent calculations by our group on the antenna peridinin chlorophyll *a* protein (PCP) [21]. In PCP, there is a 4:1 peridinin (Per) to Chl *a* ratio. The biological role of PCP is to funnel excitation energy through paths connecting the four peridinins, ending in a localized exciton on Chl *a* [22,23]. Per–Per–Per and Per–Chl couplings range between  $\sim 200 \text{ cm}^{-1}$  and  $\sim 600 \text{ cm}^{-1}$ . Thus, the CAN–CAN coupling found in the crystal structure of the HCP2 dimer is as large as the largest couplings in PCP.

**Table 2.** Comparison of QM/MM and EET calculations of absorption maxima of HCP2.

Model	Site Energy $\lambda_{\text{max}}$ nm (eV)	EET Coupling (eV)	Eigenvalues (nm)	QM/MM $\lambda_{\text{max}}$ (nm)	Exp. $\lambda_{\text{max}}$ (nm)
Monomer A	486 (2.551)	0.061	481	484	548
Monomer B	511 (2.427)		514	515	

### 2.3. Molecular Dynamics of HCP2 Monomers

HCP2 in solution is predominantly found as a monomer in a dynamical equilibrium between multiple conformations, in contrast with the crystal structure that reflects spectral properties of a more restrained conformation. We were interested in evaluating the effect of temperature on the computation of the relevant excitation energies. In order to obtain a sampling of representative states of the complex in solution, we performed 1000 ns long classical MD simulations for the individual monomers of HCP2. In our approach, we kept the general parametrization of the OPLS3e force field, applying a local minimization of the carotenoid for each snapshot under a QM/MM treatment. This procedure has been demonstrated to yield very accurate results for OCP and the Red Carotenoid Protein (RCP), within the range of the reported experimental values [17]. In that previous work, we found that the in-place root mean square deviation (RMSD), before and after QM/MM minimization, was, on average,  $0.28 \text{ \AA}$ . Thus, the local QM/MM minimization, while producing very small displacements within the protein cavity, creates accurate structures that are suitable for the evaluation of spectroscopic properties at the QM level.

The MD simulations for both monomers show stable RMSD values for the protein ( $\sim 1.8 \text{ \AA}$ ) and CAN ( $\sim 1.3 \text{ \AA}$ ) during the entire trajectories (Figure S1, Supplementary Materials). Besides the predominant hydrophobic contacts, one of the most frequent protein–ligand interactions is a water bridge with Glu94 (Figures S2 and S3, Supplementary Materials). Similar contacts have been described for RCP, where the carotenoid interacts mostly through hydrophobic contacts between residues of the cavity and the isoprenoid chain, while both rings show an increased solvent exposure when compared to the OCP cavity [17].

From the resulting MD trajectories, 51 snapshots evenly distributed along the simulation time (every 20 ns) were extracted for each monomer. After deleting the solvent and ions, the structures were locally optimized using QSite, including the carotenoid in the QM part, and freezing the entire protein. Single points were run on each optimized snapshot in the presence of implicit solvent to obtain excitation energies. The results are shown in Table 3. The MD simulations allow both monomers to reach more similar average  $\lambda_{\text{max}}$  values between them, thus decreasing the shift observed in the crystal structure. However, the absolute absorption maxima calculated show a  $\sim 50 \text{ nm}$  discrepancy with respect to the experimental value reported for HCP2 in solution. We were puzzled by the fact that this same protocol and same level of theory produced very accurate results in OCP and RCP [17]. Thus, we explored what aspects in the modeling of HCP2 could be deficient.

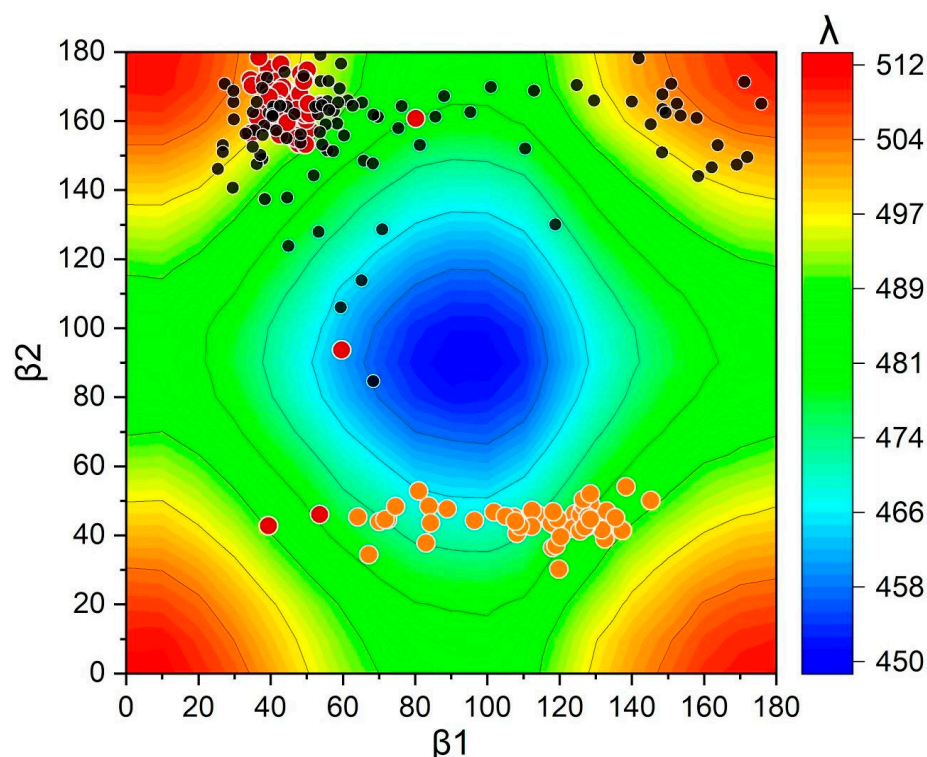
**Table 3.** QM/MM calculations on optimized structures from MD of HCP2 monomers. Results are expressed as the mean of 51 snapshots. One-sample *t* test was used to obtain the confidence intervals.

Model	$\lambda_{\max}$ (nm)	Confidence Interval (95%)		SD	Experimental $\lambda_{\max}$ (nm)
Monomer Lig. A	481.9	478.3	485.5	12.7	530
Monomer Lig. B	477.0	472.4	481.5	16.1	

As described before, the torsions of the end-terminal rings are among the main factors affecting the conjugation of a carotenoid, which are reflected on its spectral properties [12,16]. Being a homolog of the NTD of OCP, HCP2 is structurally similar to RCP. In solution, the terminal rings of the carotenoids bound to these two proteins are exposed to solvent. However, as we show below, this exposure to solvent does not necessarily result in enhanced conformational freedom for the  $\beta$ 1 ring to rotate. Figure 3 shows the conformational surface explored by the end-ring dihedrals of the optimized snapshots for OCP<sup>O</sup>, RCP and HCP2 monomers. The conformations are plotted over an underlying contour graph showing the calculated absorption maxima for a full scan of CAN dihedrals in vacuum (see Methods section for details). The torsions explored by OCP<sup>O</sup> and RCP are much more restricted (at least in one of the angles) than those observed for HCP2. Particularly, the environment of RCP influences the selection of a subset of end-ring torsions, mostly restricted to a nominal *s*-cis  $\beta$ 1 (~45°) combined with a nominal *s*-trans  $\beta$ 2 (~160°). This predominant configuration leads to red-shifted absorption maxima in RCP, in agreement with experiments. In contrast, although the conformational surface explored by CAN in HCP2 is restricted to absolute values over ~90° for  $\beta$ 2 (nominal *s*-trans), the values are widely distributed within that range. In addition, the variability is even higher for  $\beta$ 1, which explores values covering almost the whole range of torsions from 20° to 180°. Only in a small number of snapshots do  $\beta$ 1-torsions in HCP2 assume values over ~160°, which combined with similar values of  $\beta$ 2-torsions, allows higher absorption maxima. Thus, differently than in our previously reported work, in which we matched the experimental values of RCP and OCP through a similar computational approach [17], here we found a discrepancy between the HCP2-calculated values and those reported from experimental measurements.

Besides the homology between RCP and HCP2, the two crystal structures differ in the relative orientation of the monomers in the dimeric structure. Specifically, in HCP2 the orientation of monomers is such that both CAN molecules are positioned head-to-head, while in RCP (PDB code: 4XB4), both CANs are located in a perpendicular orientation without showing any close contacts between them. In addition, RCP and HCP2 monomeric structures present a different set of missing residues, which are positioned to potentially interact with the  $\beta$ 1 ring and may be affecting the conformational freedom in a different manner in both proteins. Therefore, even though we are carrying out MD simulations, it is possible that our starting structure of HCP2 does not fully evolve into its most stable conformation in solution.

Within the context of end-rings torsions analyses in these proteins, we also have included data from an MD simulation of CAN in water to compare the behavior of CAN alone in solution [17]. The distribution of  $\beta$ 1 and  $\beta$ 2 dihedrals of 1000 snapshots of CAN in water is plotted in Figure S4 (Supplementary Materials) over the same graph shown in Figure 3. As expected, the values for CAN in solution show a wider distribution when compared to CAN inside OCP, RCP, or HCP2. However, besides the expected unlimited freedom to rotate of CAN in solution, there is a marked trend to select values between 60° and 80°. This causes the prevalence of blue-shifted absorption maxima when compared to CAN in any of these proteins. Despite the difficulties in accurately reproducing HCP2 experimental values, our computational approach demonstrates the ability to achieve simulations reflecting the general trends and the influence of the protein environment on the carotenoid spectral features.



**Figure 3.** End-ring dihedrals distribution of the optimized snapshots from MD simulations for: OCP<sup>0</sup> (orange circles), RCP (red circles) [17], and HCP2 monomers (black circles). The underlying contour plot shows the calculated absorption maxima values for a full scan of  $\beta_1/\beta_2$  dihedrals of CAN in vacuum.

### 3. Materials and Methods

#### 3.1. Model Preparation

The crystal structure of HCP2 (PDB: 6MCJ) was processed using Maestro (Schrödinger Release 2019-1: Maestro, Schrödinger, LLC, New York, NY, USA). For the monomer calculations of HCP2, each alternative chain was taken from the dimeric structure. Water, pentaethylene glycol, and iodide ions were deleted before using the protein preparation module, which was used for assigning bond orders and adding hydrogens. Optimization of hydrogen bond assignment was performed with default parameters (pH 7.0) and a restrained minimization was applied for hydrogens only.

#### 3.2. Structure Optimizations and QM/MM Calculations

QM/MM calculations were run on the processed structures. CAN was defined as the QM region while the rest of the protein was treated at the MM level using the OPLS3e force field [24]. A local optimization was done on CAN to correct the bond length alternation (BLA) of the carotenoid, which is overestimated by the force field. This optimization was performed using QSite (Schrödinger Release 2019-1: Maestro, Schrödinger, LLC, New York, NY, USA) at the DFT B3LYP/LACVP\* level with CAN in the QM region and with the protein frozen. Excited state energies were then calculated from the optimized structures using the Tamm–Dancoff approximation (TDA) at the CAM-B3LYP/LACVP\* level. This particular combination of functionals for optimization and single point calculations have produced accurate results by us [17,25] and others before [26]. Solvent effects were treated using the Poisson–Boltzmann (PB) implicit solvent in QSite.

For the dihedrals scan of CAN in vacuum, a local geometry optimization was performed at the DFT B3LYP/LACVP\* level using Jaguar (Schrödinger Release 2019-3: Maestro, Schrödinger, LLC, New York, NY, USA). A relaxed scan was performed on the optimized structure, fixing one end-ring dihedral and changing the other dihedral from 0° to

180°, sequentially. A single point energy calculation was performed on the configurations of all scan coordinates. Excited state energies for each pair of  $\beta$ 1-dihedral and  $\beta$ 2-dihedral were calculated applying TDA at the CAM-B3LYP/LACVP\* level. The absorption maxima obtained from the scan were red-shifted by a constant value of 25 nm so that the range of absorptions in OCP and RCP lie roughly on top of the green/blue regions and yellow/red regions, respectively.

### 3.3. Electronic Coupling Calculations

Electronic energy transfer (EET) calculations were implemented with the Gaussian 16 program using TDA at the CAM-B3LYP/6-31g(d) level. EET was used to compute coupling constants between the two carotenoids of the HCP2 dimer. Each CAN molecule was defined as a fragment for the EET analysis. To perform these calculations in the presence of the HCP2 protein complex, partial atomic charges from the protein as well as “solvent” charges from the Poisson–Boltzmann Qsite calculation were added as the background charge distribution.

### 3.4. Molecular Dynamics Simulations

The prepared HCP2 monomers were solvated in an orthorhombic box (10 Å buffer size) using the TIP3P water model [27]. The relative orientation of the box with respect to the protein was optimized to produce a minimum number of waters. After this volume minimization procedure, the resulting number of water molecules was ~5800. Na<sup>+</sup> ions were added to neutralize the negative charge of the system. All minimizations and simulations used the OPLS3e force field for the protein and CAN. MD simulations were run for solvated systems on Nvidia GPU hardware with Desmond (Schrödinger Release 2019-4: Maestro, Schrödinger, LLC, New York, NY, USA). Prior to MD production, the default relaxation protocol was performed as detailed in our previous publication [17]. The total time for each MD simulation was 1000 ns using the NPT ensemble (300 K, 1 atm). Snapshots were extracted every 20 ns and sent to QSite for QM/MM calculations. RMSD and protein–ligand interactions analyses are described in the Supplementary Materials.

## 4. Conclusions

Using the recently reported dimeric crystal structure of the cyanobacterial HCP2 carotenoid-binding protein, we have applied a computational approach, including QM/MM calculations and MD simulations, to study the electronic and molecular properties behind the spectral features of HCP2. Through QM/MM calculations we were able to reproduce the experimental shift observed between the HCP2 crystal structure (dimer) and the protein in solution (monomer). We have analyzed the influence of different environments on the carotenoid spectral features, demonstrating the progressive effects caused by changing from a solvent to a protein–solvent environment, and the effects of the electronic coupling of the two chromophores in the dimer.

The combination of MD and QM/MM methodologies used here has previously demonstrated the ability to reproduce the experimental absorption maxima of reference proteins (OCP and RCP). Despite the limitations in reproducing HCP2 experimental values, our computational approach shows the ability to simulate the general spectral trends of the studied system, as well as the influence of protein environment on the carotenoid spectral features.

**Supplementary Materials:** The following are available online. Figure S1: RMSD HCP2 Monomers MD. (A) Ligand A. (B) Ligand B, Figure S2: Protein–ligand interactions in HCP2 Monomer–Ligand A, Figure S3 Protein–ligand interactions in HCP2 Monomer–Ligand B, Figure S4: CAN  $\beta$ 1/ $\beta$ 2 dihedrals scan over calculated absorption maxima.

**Author Contributions:** Conceptualization, N.B.P. and J.A.G.; investigation, K.C., N.B.P. and K.W.; supervision, J.A.G.; writing—original draft, K.C. and N.B.P.; writing—review and editing, N.B.P. and J.A.G. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by NSF Grant CHE-1904700 to J.A.G.



**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The data presented in this study are available on request from the corresponding author.

**Conflicts of Interest:** The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

**Sample Availability:** Not applicable.

## References

1. Kerfeld, C.A.; Kirilovsky, D. Chapter One—Structural, Mechanistic and Genomic Insights into OCP-Mediated Photoprotection. In *Advances in Botanical Research*; Chauvat, F., Cassier-Chauvat, C., Eds.; Academic Press: Cambridge, MA, USA, 2013; Volume 65, pp. 1–26.
2. Steiger, S.; Schäfer, L.; Sandmann, G. High-light-dependent upregulation of carotenoids and their antioxidative properties in the cyanobacterium *Synechocystis* PCC 6803. *J. Photochem. Photobiol. B* **1999**, *52*, 14–18. [[CrossRef](#)]
3. Zhu, Y.; Graham, J.E.; Ludwig, M.; Xiong, W.; Alvey, R.M.; Shen, G.; Bryant, D.A. Roles of xanthophyll carotenoids in protection against photoinhibition and oxidative stress in the cyanobacterium *Synechococcus* sp. strain PCC 7002. *Arch. Biochem. Biophys.* **2010**, *504*, 86–99. [[CrossRef](#)]
4. Kerfeld, C.A.; Melnicki, M.R.; Sutter, M.; Dominguez-Martin, M.A. Structure, function and evolution of the cyanobacterial orange carotenoid protein and its homologs. *New Phytol.* **2017**, *215*, 937–951. [[CrossRef](#)]
5. Kerfeld, C.A.; Sawaya, M.R.; Brahmamdam, V.; Cascio, D.; Ho, K.K.; Trevithick-Sutton, C.C.; Krogmann, D.W.; Yeates, T.O. The Crystal Structure of a Cyanobacterial Water-Soluble Carotenoid Binding Protein. *Structure* **2003**, *11*, 55–65. [[CrossRef](#)]
6. Melnicki, M.R.; Leverenz, R.L.; Sutter, M.; López-Igual, R.; Wilson, A.; Pawlowski, E.G.; Perreau, F.; Kirilovsky, D.; Kerfeld, C.A. Structure, Diversity, and Evolution of a New Family of Soluble Carotenoid-Binding Proteins in Cyanobacteria. *Mol. Plant* **2016**, *9*, 1379–1394. [[CrossRef](#)]
7. Muzzopappa, F.; Wilson, A.; Yogarajah, V.; Cot, S.; Perreau, F.; Montigny, C.; Bourcier de Carbon, C.; Kirilovsky, D. The paralogs to the C-terminal domain of the cyanobacterial OCP are carotenoid donors to HCPs. *Plant Physiol.* **2017**. [[CrossRef](#)]
8. Muzzopappa, F.; Wilson, A.; Kirilovsky, D. Interdomain interactions reveal the molecular evolution of the orange carotenoid protein. *Nat. Plants* **2019**, *5*, 1076–1086. [[CrossRef](#)]
9. Dominguez-Martin, M.A.; Polívka, T.; Sutter, M.; Ferlez, B.; Lechno-Yossef, S.; Montgomery, B.L.; Kerfeld, C.A. Structural and spectroscopic characterization of HCP2. *Biochim. Biophys. Acta Bioenerg.* **2019**, *1860*, 414–424. [[CrossRef](#)]
10. Wei, T.; Balevičius, V.; Polívka, T.; Ruban, A.V.; Duffy, C.D.P. How carotenoid distortions may determine optical properties: Lessons from the Orange Carotenoid Protein. *Phys. Chem. Chem. Phys.* **2019**, *21*, 23187–23197. [[CrossRef](#)] [[PubMed](#)]
11. Maksimov, E.G.; Shirshin, E.A.; Sluchanko, N.N.; Zlenko, D.V.; Parshina, E.Y.; Tsoraev, G.V.; Klementiev, K.E.; Budylin, G.S.; Schmitt, F.-J.; Friedrich, T.; et al. The Signaling State of Orange Carotenoid Protein. *Biophys. J.* **2015**, *109*, 595–607. [[CrossRef](#)] [[PubMed](#)]
12. Bondanza, M.; Cupellini, L.; Lipparini, F.; Mennucci, B. The Multiple Roles of the Protein in the Photoactivation of Orange Carotenoid Protein. *Chem* **2020**, *6*, 187–203. [[CrossRef](#)]
13. Khan, T.; Dominguez-Martin, M.A.; Šimová, I.; Fuciman, M.; Kerfeld, C.A.; Polívka, T. Excited-State Properties of Canthaxanthin in Cyanobacterial Carotenoid-Binding Proteins HCP2 and HCP3. *J. Phys. Chem. B* **2020**, *124*, 4896–4905. [[CrossRef](#)] [[PubMed](#)]
14. Young, A.J.; Phillip, D.M.; Hashimoto, H. 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.* **2002**, *642*, 137–145. [[CrossRef](#)]
15. Harms, M.J.; Castañeda, C.A.; Schlessman, J.L.; Sue, G.R.; Isom, D.G.; Cannon, B.R.; García-Moreno, E.B. The pKa Values of Acidic and Basic Residues Buried at the Same Internal Location in a Protein Are Governed by Different Factors. *J. Mol. Biol.* **2009**, *389*, 34–47. [[CrossRef](#)]
16. Šlouf, V.; Kuznetsova, V.; Fuciman, M.; de Carbon, C.B.; Wilson, A.; Kirilovsky, D.; Polívka, T. Ultrafast spectroscopy tracks carotenoid configurations in the orange and red carotenoid proteins from cyanobacteria. *Photosynth. Res.* **2017**, *131*, 105–117. [[CrossRef](#)]
17. Pigni, N.B.; Clark, K.L.; Beck, W.F.; Gascón, J.A. Spectral Signatures of Canthaxanthin Translocation in the Orange Carotenoid Protein. *J. Phys. Chem. B* **2020**, *124*, 11387–11395. [[CrossRef](#)]
18. Gascón, J.A.; Sproviero, E.M.; Batista, V.S. Computational Studies of the Primary Phototransduction Event in Visual Rhodopsin. *Acc. Chem. Res.* **2006**, *39*, 184–193. [[CrossRef](#)] [[PubMed](#)]
19. Curutchet, C.; Mennucci, B. Toward a Molecular Scale Interpretation of Excitation Energy Transfer in Solvated Bichromophoric Systems. *J. Am. Chem. Soc.* **2005**, *127*, 16733–16744. [[CrossRef](#)]
20. Russo, V.; Curutchet, C.; Mennucci, B. Towards a Molecular Scale Interpretation of Excitation Energy Transfer in Solvated Bichromophoric Systems. II. The Through-Bond Contribution. *J. Phys. Chem. B* **2007**, *111*, 853–863. [[CrossRef](#)]

21. Guberman-Pfeffer, M.J.; Gascón, J.A. Carotenoid-Chlorophyll Interactions in a Photosynthetic Antenna Protein: A Supramolecular QM/MM Approach. *Molecules* **2018**, *23*, 2589. [[CrossRef](#)] [[PubMed](#)]
22. Guberman-Pfeffer, M.J.; Greco, J.A.; Birge, R.R.; Frank, H.A.; Gascón, J.A. Light Harvesting by Equally Contributing Mechanisms in a Photosynthetic Antenna Protein. *J. Phys. Chem. Lett.* **2018**, *9*, 563–568. [[CrossRef](#)] [[PubMed](#)]
23. Tilluck, R.W.; Ghosh, S.; Guberman-Pfeffer, M.J.; Roscioli, J.D.; Gurchiek, J.K.; LaFountain, A.M.; Frank, H.A.; Gascón, J.A.; Beck, W.F. Interexciton nonradiative relaxation pathways in the peridinin-chlorophyll protein. *Cell Rep. Phys. Sci.* **2021**, *2*, 100380. [[CrossRef](#)]
24. Roos, K.; Wu, C.; Damm, W.; Reboul, M.; Stevenson, J.M.; Lu, C.; Dahlgren, M.K.; Mondal, S.; Chen, W.; Wang, L.; et al. OPLS3e: Extending Force Field Coverage for Drug-Like Small Molecules. *J. Chem. Theory Comput.* **2019**, *15*, 1863–1874. [[CrossRef](#)] [[PubMed](#)]
25. Gurchiek, J.K.; Rose, J.B.; Guberman-Pfeffer, M.J.; Tilluck, R.W.; Ghosh, S.; Gascón, J.A.; Beck, W.F. Fluorescence Anisotropy Detection of Barrier Crossing and Ultrafast Conformational Dynamics in the S2 State of  $\beta$ -Carotene. *J. Phys. Chem. B* **2020**, *124*, 9029–9046. [[CrossRef](#)] [[PubMed](#)]
26. Andreussi, O.; Knecht, S.; Marian, C.M.; Kongsted, J.; Mennucci, B. Carotenoids and Light-Harvesting: From DFT/MRCI to the Tamm–Dancoff Approximation. *J. Chem. Theory Comput.* **2015**, *11*, 655–666. [[CrossRef](#)]
27. Jorgensen, W.L.; Chandrasekhar, J.; Madura, J.D.; Impey, R.W.; Klein, M.L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, *79*, 926–935. [[CrossRef](#)]