Supplementary Information for

On the Origin and Variation of Colors in Lobster Carapace

Authors: Shamima Begum[a], Michele Cianci[b], Bo Durbeej[c], Olle Falklöf[c], Alfons Hädener[a,¶], John R. Helliwell[a]*, Madeleine Helliwell[a], Andrew C. Regan[a], C. Ian F. Watt[a]

Affiliations:

[a]School of Chemistry, University of Manchester, Manchester, M13 9PL, UK.

[b]EMBL, PETRA III, DESY, Notkestrasse 85, 22603 Hamburg, Germany.

[c]Division of Computational Physics, IFM, Linköping University, SE-58183 Linköping, Sweden.

*Correspondence to: john.helliwell@manchester.ac.uk

¶Present address: Pfaffenlohweg 29, CH-4125 Riehen, Switzerland.

This PDF file includes:

Materials and Methods

1. Spectroscopic studies of the behavior of 3-hydroxy-4-oxo-β-ionone in aqueous medium

1.1. General methods

NMR spectra were recorded on Brüker 300, 400 or 500 AMX spectrometers. Chemical shifts are reported in parts per million (ppm) downfield from tetramethylsilane, using the solvent as an internal standard. Signal splittings are quoted in Hz. Signals are listed as: singlet (s), doublet (d), triplet (t), quartet (q) or multiplet (m) with width (W) to outside lines. Infrared spectra were recorded on a Perkin-Elmer Spectrum FTIR and were run as liquid films or solids in an ATR attachment. UV-vis spectra were recorded on a Cary 50 UV-visible spectrophotometer. Mass spectra were recorded on a Hewlett Packard HP5791A MSD. Modes of ionization are electron impact (EI), positive chemical ionization (+CI) using ammonia and fast atom bombardment (FAB), positive electrospray (+ES) and negative electrospray (-ES). Melting points were recorded on a Kofler heated stage microscope and are uncorrected. Thin layer chromatography (TLC) was carried out on Polygram Sil G/UV_{254} 0.25 mm silica gel plates with solvent systems as indicated. Analytical gas chromatography was carried out using Perkin-Elmer capillary gas chromatography model 8310 with flame ionization detection on a 30 m x 0.25 mm polydimethylsiloxane capillary column (Restek RTX-1). Analytical HPLC was carried out using a Bondclone C18 100 x 8 mm column with Perkin-Elmer LC480 Diode Array System (detection at 255 nm and 285 nm). Solvents were premixed and degassed (75:25 MeOH:H_2O) with a flow rate of 2 ml/min.

1.2. Preparation of compounds

3-hydroxy-4-oxo-β-ionone (systematically 6-hydroxy-2,4,4-trimethyl-3-[(1*E*)-3-oxo-1-buten-1-yl]-2-cyclohexen-1-one), **2**, was prepared by the published methods from β-ionone [1, 2], and obtained as off-white crystals from hexane whose spectroscopic and physical properties were in accord with literature reports; mp 48–49 ºC (literature [1]: 49.5–51 ºC); δH (300 MHz, $CDCl_3$) 7.07 (1H, d, *J* = 16.6), 6.08 (1H, d, *J* = 16.6), 4.23 (1H, dd, *J* = 5.8, 13.9), 2.08 (1H, dd, *J* = 5.9, 12.8), 2.02 (3H, s), 1.73 (1H, dd, *J* = 12.9, 13.8), 1.72 (3H, d, *J* = 1.1), 1.2 (3H, s) and 1.03 (3H, s).

2,3-dehydro-3-hydroxy-4-oxo-β-ionone (systematically 6-hydroxy-2,4,4-trimethyl-3-[(1E)-3-oxo-1-buten-1-yl]-2,4-cyclohexdien-1-one), **3**, was prepared by dissolving **2** (0.011 g) in dry methanol (5 ml) in a 50 ml wide-neck RB flask. Aqueous NaOH solution (20 ml of 0.2 M) was added to provide an immediate deep orange colored solution, which was shaken vigorously to ensure full aeration. The open flask was then allowed to stand at ambient temperature (21 ºC) with periodic vigorous shaking for 24 hours during which time the orange color faded noticeably. Glacial acetic acid was then added by microdropper until the residual color was discharged (p*H* < 8). The colorless solution was then evaporated to dryness under vacuum maintaining T < 30 ºC, the oily residue was taken up in chloroform, and filtered through a cotton wool wad. Examination by TLC (silica, eluting with 50:50 vol:vol ethyl acetate:hexane) showed that the solution contained traces of residual unchanged **2** and a single major, more mobile product, which was isolated by preparative TLC on silica and crystallized from ether-hexane as off-white needles; mp 111–112 ºC; δH (400 MHz, $CDCl_3$) 1.23 (6H, s), 1.97 (3H, d, *J* = 1.0), 2.32 (3H, s),

6.04 (1H, s), 6.25 (1H, d, J = 16.4), 6.28(1H, s), 7.24 (1H, dq, J = 16.4, 1.0 Hz); δC (75MHz, $CDCl_3$) 13.44, 27.71, 28.36, 39.13, 126.27, 130.61, 134.59, 138.76, 144.74, 157.86, 182.02, 197.31; ν_{max}/cm^{-1} 3328, 2964, 1685, 1636, 1617, 1326, 1249, 1065, 976; λ_{max} (water)/nm 267 (ε_{max} = 9500 $dm^3mol^{-1}cm^{-1}$); m/z (+ve EI) 220(40.5), 177(84.9), 159(37.1), 149, (100), 145(35.7), 91 (38.1); found 220.1094 (calculated 220.1099 for $C_{13}H_{16}O_3$).

1.3. Physical measurements

Solutions: Water was distilled in an all glass apparatus and deoxygenated with nitrogen. Methanol was taken from a freshly opened bottle. Stock aqueous solutions of NaOH in water were prepared by dilution of certified 2 M concentrates (Aldrich Chemical Company), and were degassed by freeze-pump-thaw cycles, releasing to nitrogen on each cycle, and then stored in vials fitted with 3-way taps until use. Solution p*H* was measured at 25 °C using an EDT p*H* meter and a FisherBrand glass electrode calibrated against standard buffers at p*H* 4 and p*H* 10 (Hydrion buffer capsules). For the investigations of **2**, a stock solution, 1.65×10^{-2} M, was prepared by dissolving **2** (0.076 g) in methanol (20.0 ml) and also stored under nitrogen.

1.4. UV-vis spectroscopy

UV-vis investigations were carried out on a Cary-50 BIO spectrometer fitted with a thermostatted block set to 30 °C. The spectra were run in quartz cells fitted with 3-way PTFE taps to permit flushing and transfers under nitrogen atmosphere. Typically, a cell was flushed with dry nitrogen by using a long syringe needle connected to the nitrogen supply. The solvent (2.5 ml) was then transferred by syringe from its storage vial under nitrogen to the cell, and the tap closed. The cell was then equilibrated to temperature for 30 minutes and a small plastic bag clipped over the tap. Nitrogen from the supply was piped by needle into the bag to provide a convenient temporary additional external nitrogen atmosphere. To initiate reactions, an aliquot (20 μl) of the stock solution of **2** in methanol was injected by microsyringe through the bag and stopcock, which was then immediately closed. The cell was then shaken for 10 seconds before timed scanning was initiated.

2. Quantum chemical calculations

2.1. General methods

Density functional theory (DFT) and Møller-Plesset second-order perturbation theory (MP2) calculations were performed with the GAUSSIAN 09 program [3], whereas *ab initio* approximate coupled-cluster singles and doubles (CC2 [4, 5]) and complete active space second-order perturbation theory (CASPT2 [6]) calculations were performed with the TURBOMOLE 6.3 [7, 8] and MOLCAS 7.6 [9] programs, respectively. Bulk solvation effects were accounted for using the CPCM polarizable conductor calculation model [10]. Geometry optimizations were throughout carried out in a water or methanol solvent using the B3LYP hybrid density functional in combination with a Pople style basis set, either 6-31G(d,p) that does not include diffuse functions, or 6-31+G(d,p) that includes diffuse functions on second-row atoms. Complementary geometry optimizations were also carried out using three other density functionals (BP86, BLYP and M06-2X), as well as MP2 and CC2. Some of these optimizations were done in the gas phase with the cc-pVDZ basis set.

2.2. UV-vis calculations

For each specific form of the molecules under investigation and for a number of relevant stereoisomers thereof, λ_{max} were obtained as vertical excitation energies for the strongly absorbing "B_u-like" state responsible for the color of astaxanthin. Based on optimized ground-state geometries of the stereoisomers in question, the excitation energies were calculated using time-dependent DFT (TD-DFT) [11-13] and CC2 (and in some cases also CASPT2). For the carotenoid system, the following stereoisomers were considered: *s-trans*, *s-cis*, *s-trans-s'-cis* and *s-cis-s'-trans*. For the 3-hydroxy-4-oxo-β-ionone model system and its 2,3-dehydro-3-hydroxy-4-oxo-β-ionone oxidation product, all unique local minima as obtained from mapping the corresponding potential energy surfaces along the C5-C6-C7-C8 and C7-C8-C9-O9 dihedral coordinates, were considered. Seven different density functionals were used for the TD-DFT calculations: M06-HF, LC-ωPBE, LC-BLYP, CAM-B3LYP, ωB97X-D, BHandH and BHandHLYP. Of these, the first five have been specifically designed for treating long-range electron-electron interactions, which make them particularly useful for excited-state calculations [14-16].

In order to account for the conformational flexibility of the molecules at the experimental temperature, the reported λ_{max} were throughout obtained as Boltzmann averages (at 30 °C) over all relevant stereoisomers. In each case, the Boltzmann factors were evaluated from the relative ground-state energies of the stereoisomers at the level of theory used for the geometry optimizations.

While all TD-DFT solution-phase excitation energies were calculated by directly combining TD-DFT with a CPCM description of the water or methanol solvent, as made possible by the GAUSSIAN 09 program, the CC2 solution-phase excitation energies were obtained *a posteriori* by adding to the corresponding CC2 gas-phase excitation energies a correction term accounting for the presence of the solvent in an indirect fashion. Specifically, using TD-DFT, this correction term was calculated as the difference between LC-ωPBE/gas-phase and LC-ωPBE/CPCM excitation energies.

2.3. pK_a calculations

Room-temperature pK_a values of the 3-hydroxy-4-oxo-β-ionone model system were calculated using a hybrid cluster-continuum approach [17], wherein a CPCM description of bulk solvation effects was combined with the inclusion of explicit water molecules to account for specific solute-solvent interactions. The calculations were performed considering only the most stable (in the absence of explicit water molecules) stereoisomers of the parent and deprotonated forms of the molecule in question. All calculations were carried out with the B3LYP hybrid density functional in combination with the 6-31+G(d,p) basis set.

Throughout, the calculations on the parent and deprotonated forms included the same number of explicit water molecules. With the aim to facilitate the formation of as many hydrogen bonds as possible, the water molecules were included in the following way: for systems with two water molecules, one molecule was incorporated in the vicinity of O9 and the other close to O3 and O4; for systems with three water molecules, one molecule was incorporated to interact with O9 and the other two to form a hydrogen-bonding network with O3 and O4.

Assuming ideal-gas behavior and using the 1 M standard state, gas-phase Gibbs free energies were calculated for gas-phase-optimized structures as the sum of electronic energies and (obtained from frequency calculations) thermal free energies. Free energies of solvation, in turn, were calculated for solution-phase-optimized

structures as the difference between electronic energies in aqueous solution and electronic energies in the gas phase. Deprotonation free energies were then obtained by adding the free energies of solvation to the gas-phase Gibbs free energies, from which the relevant pK_a values were readily obtained. For the proton, the values used for the gas-phase Gibbs free energy (–6.28 kcal mol^{-1}) and the free energy of solvation (–265.9 kcal mol^{-1}) were taken from the literature [18, 19].

3. Low-resolution refinement of astaxanthin in its enolic form in the structure of β-crustacyanin

The starting model and structure factors were downloaded from the PDB repository with codes 1GKA and r1gkasf [20]. CNS solve V1.3 [21] was used to refine the β-crustacyanin structure in its complex with the two astaxanthin molecules in the ground state (AXT, Fig. S9a) and the enolate state (AXE, Fig. S9b). Topology and parameter description files were generated using the PRODRG server [22] and inspected for correctness. For both forms, the 6-s-*trans*-conformation around the C6-C7 and C6′-C7′ bonds was retained. Bond angles and bond distances for the end rings and polyene chain were derived from the published models of astaxanthin [23, 24]. For CNS, $B_{individual}$ refinement was used as suggested in the CNS on-line manual (http://cns-online.org/v1.3/).

Supplementary Results and Discussion

4.1. Calculation on model system

The choice of density functionals for the calculations presented in the main text was based on analogous studies of the 3-hydroxy-4-oxo-β-ionone model system, for which we have been able to experimentally record UV-vis spectra of the hydroxyketone, mono-anion and di-anion forms (**2**, $\mathbf{2}^{-}$, and **2**-dianion, Fig. 2), but not of the ene-diol form (**2**ed, Fig. 2). These studies are summarized in Tables S1a and S1b.

Considering first the results in Table S1b, which were obtained with a basis set (aug-cc-pVDZ) including diffuse functions, the best performance for the hydroxyketone is shown by the M06-HF, LC-ωPBE and LC-BLYP functionals that were used also for the full carotenoid system. These methods reproduce the experimental λ_{max} with an accuracy of ~5 nm (~0.1 eV), which is better than that achieved by the other functionals tested (CAM-B3LYP, ωB97X-D, BHandH and BHandHLYP), and comparable to the performance of the CC2 method. Furthermore calculations performed on 2,3-dihydroxy-2-propenal [25] show again satisfactory agreement between theory and experiment (see section 4.2).

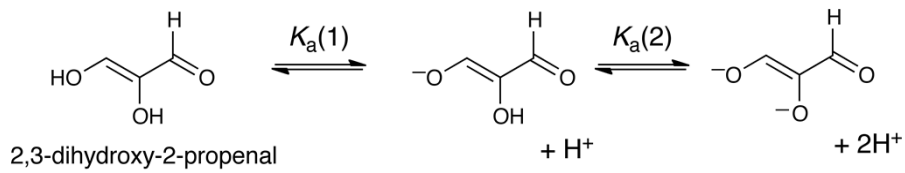
For the mono-anion, on the other hand, all the density functionals consistently yield λ_{max} at much longer wavelengths (454–539 nm) than that observed experimentally (356 nm). Despite further investigations (as further detailed below), we have not found a satisfactory explanation for this discrepancy or any evidence that factors typically having a negative impact on the accuracy of TD-DFT excitation energies, such as charge-transfer or near degeneracy effects, are at play. The calculations are, nonetheless, in qualitative if not exact quantitative accord with the experimental UV-vis spectra in that they predict a substantial red shift in the absorption upon formation of the mono-anion from the parent hydroxyketone, albeit with an exaggerated shift.

Turning to the calculations on the di-anion, we note that the M06-HF, LC-ωPBE and LC-BLYP functionals perform well, with results that for both peaks agree with the experimental values to within 10 nm (0.15 eV) or better, but also that the other functionals are of comparable accuracy.

Finally, as for the data in Table S1a obtained with basis sets (6-31G(d,p) and cc-pVDZ) lacking diffuse functions, these support the same overall conclusions as the data in Table S1b.

4.2. Calculations on 2,3-dihydroxy-2-propenal

As noted in the main text, 2,3-dihydroxy-2-propenal is the minimum carbon framework in which conjugation between a carbonyl group (here aldehydic) and an ene-diol may occur. For this molecule, dissociation to produce both anion and dianion forms has been observed and quantified experimentally [25].



To further test our computational approach, we have applied it to the calculation of the UV-vis λ_{max} of the 2,3-dihydroxy-2-propenal system. The results, given in Table S3, show satisfactory agreement between theory and experiment.

4.3. Calculations on the oxidation product

As discussed above, the quantum chemical calculations on the different forms of the 3-hydroxy-4-oxo-β-ionone model system predict an appreciable red shift in the absorption on conversion of the parent hydroxyketone to its mono-anion, but the quantitative agreement between theory and experiment as to the size of the shift is not very good. To further ascertain that theory gives results that are qualitatively consistent with experimental data, we have investigated how well our computational approach can reproduce the UV-vis spectra of the oxidation product of the model system hydroxyketone (**3**) and its anion. These calculations were carried out at exactly the same levels of theory as the previous calculations, and are presented in Tables S4a and S4b. Since the results obtained without diffuse functions (Table S4a) are similar to the results obtained with diffuse functions (Table S4b), we will in the interest of brevity only discuss the latter.

From Table S4b, it can be seen that all density functionals reproduce the experimental λ_{max} (at 270 nm) of the oxidation product with good accuracy, and fall within three different groups: those that give a peak at 10–17 nm (0.17–0.31 eV) shorter wavelengths (M06-HF, LC-ωPBE and LC-BLYP); those that give a peak at 10–11 nm (0.18 eV) longer wavelengths (CAM-B3LYP and ωB97X-D); and those that together with CC2 give a peak at essentially the same wavelength as the experiment (BHandH and BHandHLYP).

The experimental UV-vis spectrum of the anion of the oxidation product, in turn, has two peaks, both of which are red-shifted relative to the λ_{max} of the parent species. For the anion, all density functionals reproduce the stronger peak at 276 nm rather well, yielding λ_{max} at 9–20 nm (0.14–0.31 eV) longer wavelengths. The weaker peak at 371 nm, on the other hand, poses a somewhat greater challenge, although our principal methods (M06-HF, LC-ωPBE and LC-BLYP) perform markedly better than the other methods (CAM-B3LYP, ωB97X-D, BHandH and BHandHLYP). Specifically, the methods of the two groups give λ_{max} at 9–49 nm (0.08–0.39 eV) and 103–129 nm (0.72–0.86 eV) longer wavelengths than experiment, respectively.

Overall, then, the results in Table S4b seem to strengthen our belief that quantum chemical calculations using the principal density functionals of this work constitute a reliable tool for obtaining a qualitative assessment of how acid-base chemistry influences the UV-vis spectra of astaxanthin and models thereof. In this regard, it should also be mentioned that the pK_a calculations discussed in the main text (Table S3) also included the pK_a for deprotonation of the oxidation product, and that the computed values (15.5 $\rightarrow$ 11.1 $\rightarrow$ 9.2) were found to agree very well with the experimental estimate (9.9) as the number of explicit water molecules was increased.

4.4. Comparative *ab initio* calculations

While the size of astaxanthin – corresponding to 920 basis functions for the keto form at the 6-31G(d,p) basis-set level – precludes a quantum chemical treatment based on high-level *ab initio* methods and although we have not found any indication

that such methods are strictly needed for this study, it is nonetheless of interest to assess the difference in performance between TD-DFT and *ab initio* methods for the present molecules. To this end, we have compared the gas-phase λ_{max} of the different forms of the 3-hydroxy-4-oxo-β-ionone model system obtained with LC-ωPBE, one of the principal density functionals of this work, with those obtained with the *ab initio* CC2 and CASPT2 methods, the latter of which is widely regarded as the gold standard in the field. This comparison is presented in Table S5.

Interestingly, we find that the differences between LC-ωPBE and CASPT2 are relatively small for the parent hydroxyketone (0.09 eV), ene-diol (0.18 eV) and di-anion (0.26 and 0.24 eV for the two peaks, respectively) forms, with LC-ωPBE systematically yielding λ_{max} at shorter wavelengths. For the mono-anion, the difference is larger (0.42 eV), but not of such a magnitude that one would call into question the applicability of TD-DFT to the overall chemical problem at hand. Furthermore, it is notable that CASPT2 predicts an even larger red shift than LC-ωPBE for the gas-phase absorption of the mono-anion. Thus, it appears unlikely that the substantial red shift obtained in solution using TD-DFT is due to an artifact that could be rectified with a CASPT2 treatment.

4.5. Influence of molecular geometries on calculated λ_{max}

All excited-state calculations discussed up to this point have been carried out based on molecular geometries optimized using the B3LYP hybrid density functional and the 6-31G(d,p) basis set (in combination with a CPCM description of the water or methanol solvent), which is a standard approach in contemporary quantum chemistry. To assess how this procedure may influence the resulting λ_{max}, we have also subjected geometries optimized at other levels of theory to excited-state calculations. These investigations, which in large part are motivated by related studies showing that calculated excitation energies of conjugated systems can be very sensitive to the bond length alternation [26], are summarized in Tables S6–S9.

Starting with Table S6 and results on how the λ_{max} of the different forms of the 3-hydroxy-4-oxo-β-ionone model system change when the basis set used for geometry optimization is augmented with diffuse functions (6-31G(d,p) → 6-31+G(d,p)), it is clear that such functions can be omitted from this stage of the calculations. Indeed, for all forms of the model system, the geometric effect attributable to diffuse functions is of the order of a few nm only.

Table S7, in turn, compares M06-HF, LC-ωPBE and LC-BLYP excitation energies for the strongly absorbing "B_u-like" state of the different forms of the model system using geometries optimized with on the one hand B3LYP and on the other three other commonly used density functionals (BP86, BLYP and M06-2X), as well as MP2. From this comparison, we find that the calculated excitation energies are quite sensitive to the method used for geometry optimization, with maximum absolute deviations between them that amount to ~0.4–0.5 eV for the hydroxyketone, ene-diol and di-anion forms, and to ~0.2 eV for the mono-anion. However, given that the maximum absolute deviations for the corresponding *shifts* in excitation energies relative to the parent hydroxyketone are considerably smaller (Table S7), this effect is of no major consequence for the present investigation. The same exact conclusion can be drawn from analogous calculations on the full carotenoid system summarized in Table S8.

As a final test of the appropriateness of using B3LYP for the geometry optimizations, we have also performed excited-state calculations at B3LYP structures of the model and full carotenoid systems optimized in the gas phase, and evaluated how the results subsequently change when the calculations instead are performed at structures re-optimized with MP2 and CC2. Encouragingly, this investigation, presented in Table S9, shows that the gas-phase excitation energies obtained at the B3LYP geometries are similar to those obtained at the MP2 and CC2 geometries. Indeed, for each form of the two systems, the resulting energies agree to within ~0.1 (model system) and ~0.2 eV (full system).

5. Low-resolution refinement of astaxanthin in its enolic form in the structure of β-crustacyanin

At 3.1 Å resolution the Cruickshank Diffraction Precision Index ('DPI') of the protein model coordinate errors [27] calculated on R_{factor} and R_{free} are 0.31 Å and 0.25 Å, respectively. Thus, the magnitude of these estimated coordinate errors does not allow for the discrimination between double and single bonds in the case of C=O or C-O for instance. Hence, a confirmation of whether we in fact observe enolization of astaxanthin in the β-crustacyanin structure in its complex with the two astaxanthin molecules requires much higher X-ray crystallographic resolution. Some evidence of enolization may only come from changes in $B_{factors}$ due to different constraints in the geometry of the end rings and polyene chains of the two forms.

For convenience, we define astaxanthin (1) as the carotenoid with the C1-C6 end ring in the cavity of the A1 subunit and with the C21-C26 end ring close to His90 of the A2 subunit. We define astaxanthin (2) as the carotenoid with the C1-C6 end ring in the cavity of the A2 subunit and with the C21-C26 end ring close to His92 of the A1 subunit.

Whereas astaxanthin in its ground-state form (AXT, Fig. S1a) is symmetric, the enolate form is not since one of the end rings is enolized but not the other (AXE and AXI, Figs. S1b and S1c).

The two astaxanthin molecules were refined in the ground-state form (AXT, Figs. S2a and S2d) and in two possible orientations of the enolic form relative to the asymmetric binding sites of the protein subunits (AXE and AXI, Figs. S2b, S2c, S2e and S2f).

Fo-Fc omit electron density maps, with a 10 sigma cut-off contour level calculated omitting O4, O24 or O3, O23 from AXT and AXE on both carotenoids, show that both the AXT and AXE forms are perfectly compatible with the expected oxygen positions (Fig. S3). Fo-Fc omit electron density maps, with 10 sigma cut-off contour level calculated omitting the C18, C19, C20 or C38, C39, C40 from AXT and AXE on both carotenoids confirm the overall orientation of the molecules (data not shown).

The crystal structure refinement shows that, in any case, the difference in B_{factor} for all the atoms of AXT, AXE and AXI are generally < 1.5 $Å^2$ (Tables S10a and S10b). Given that the average rms-B for the full structure is > 1.5 $Å^2$, these differences are not significant. Hence, with the available data, the AXT and AXE forms are equally plausible in the β-crustacyanin model. In the case of astaxanthin (1), the $B_{factors}$ are partially improved when the AXE form was used. In the case of astaxanthin (2), all the $B_{factors}$ are improved when the AXI form was used.

The overall refinement statistics for the β-crustacyanin model slightly improve when the AXT form is used for astaxanthin (1) and the AXI form is used for

astaxanthin (2), compared to when the AXT form is used for both astaxanthin (1) and astaxanthin (2) (Table S11). This might indicate that the enolate form of astaxanthin represents a partially improved model for both astaxanthins, or that astaxanthin might coexist in both its ground state and its enolate state within β-crustacyanin.

Table S1a. Calculated λ_{max} (in nm) and the corresponding excitation energies (ΔE, in eV) and oscillator strengths (f) of different forms of the model system obtained without diffuse functions.[a]

Method		Hydroxyketone	Ene-diol	Mono-anion	Di-anion	
M06-HF	λ_{max}	260	343	477	273	429
	ΔE	4.78	3.61	2.60	4.54	2.89
	f	0.59	0.61	0.82	0.33	0.68
LC-ωPBE	λ_{max}	264	349	453	277	446
	ΔE	4.69	3.55	2.74	4.48	2.78
	f	0.60	0.66	0.84	0.47	0.58
LC-BLYP	λ_{max}	260	342	446	271	432
	ΔE	4.77	3.63	2.78	4.58	2.87
	f	0.62	0.67	0.86	0.45	0.64
CAM-B3LYP	λ_{max}	281	368	452	289	488
	ΔE	4.42	3.37	2.74	4.29	2.54
	f	0.56	0.66	0.79	0.33	0.50
ωB97X-D	λ_{max}	282	368	454	289	488
	ΔE	4.40	3.37	2.73	4.29	2.54
	f	0.55	0.66	0.78	0.38	0.49
BHandH	λ_{max}	275	361	435	276	464
	ΔE	4.51	3.44	2.85	4.49	2.67
	f	0.59	0.69	0.82	0.34	0.57
BHandHLYP	λ_{max}	274	358	434	276	463
	ΔE	4.52	3.46	2.86	4.49	2.68
	f	0.60	0.70	0.83	0.53	0.57
CC2	λ_{max}	260	364	506	301	512
	ΔE	4.77	3.41	2.45	4.12	2.42
	f	0.52	0.37	0.73	0.51	0.35
Experiment	λ_{max}	275	–	356	278	454
	ΔE	4.51	–	3.49	4.46	2.73
	ε_{max}[b]	5100	–	4000	3900	3300

[a]All calculations carried out with the 6-31G(d,p) basis set (cc-pVDZ for CC2) and geometries optimized using B3LYP/6-31G(d,p) in combination with a CPCM description of the water solvent.
[b]ε_{max} given in units of $dm^3mol^{-1}cm^{-1}$.

Table S1b. Calculated λ_{max} (in nm) and the corresponding excitation energies (ΔE, in eV) and oscillator strengths (f) of different forms of the model system obtained with diffuse functions.[a]

Method		Hydroxyketone	Ene-diol	Mono-anion	Di-anion	
M06-HF	λ_{max}	270	359	500	288	453
	ΔE	4.59	3.45	2.48	4.31	2.74
	f	0.61	0.54	0.84	0.19	0.47
LC-ωPBE	λ_{max}	273	361	474	288	461
	ΔE	4.54	3.43	2.61	4.31	2.69
	f	0.62	0.68	0.88	0.29	0.70
LC-BLYP	λ_{max}	270	355	468	285	449
	ΔE	4.60	3.49	2.65	4.36	2.76
	f	0.64	0.69	0.90	0.26	0.75
CAM-B3LYP	λ_{max}	291	383	475	284	501
	ΔE	4.26	3.24	2.61	4.37	2.48
	f	0.58	0.68	0.83	0.17	0.61
ωB97X-D	λ_{max}	291	381	475	303	498
	ΔE	4.27	3.25	2.61	4.09	2.49
	f	0.58	0.68	0.83	0.27	0.61
BHandH	λ_{max}	285	376	456	280	479
	ΔE	4.35	3.30	2.72	4.44	2.59
	f	0.58	0.71	0.86	0.16	0.65
BHandHLYP	λ_{max}	283	373	454	285	477
	ΔE	4.37	3.33	2.73	4.35	2.60
	f	0.61	0.71	0.87	0.19	0.66
CC2	λ_{max}	273	382	539	–	–
	ΔE	4.53	3.25	2.30	–	–
	f	0.54	0.45	0.61	–	–
Experiment	λ_{max}	275	–	356	278	454
	ΔE	4.51	–	3.49	4.46	2.73
	ε_{max}[b]	5100	–	4000	3900	3300

[a]All calculations carried out with the aug-cc-pVDZ basis set and geometries optimized using B3LYP/6-31G(d,p) in combination with a CPCM description of the water solvent.
[b]ε_{max} given in units of $dm^3mol^{-1}cm^{-1}$.

Table S2. Calculated pK_a values of different forms of the model system.[a]

	Hydroxyketone		Ketone[b]	Mono-anion	Ox. product[c]
# H_2O[d]	pK_a(C3-H)	pK_a(O3-H)	pK_a(C3-H)	pK_a(O4-H)	pK_a(O3-H)
0	10.6	24.3	25.6	23.6	15.5
2	7.7	17.5	22.6	17.9	11.1
3	9.1	19.1	19.9	14.5	9.2
Experiment	10.4	–	–	13.2	9.9

[a]All calculations carried out using B3LYP in combination with the 6-31+G(d,p) basis set and a CPCM description of the water solvent.

[b]4-oxo-β-ionone: the parent model compound without its C3 hydroxyl group.

[c]The oxidation product of the hydroxyketone form.

[d]Number of explicit water molecules included in the calculations.

Table S3. Calculated λ_{max} (in nm) and the corresponding excitation energies (ΔE, in eV) of different forms of 2,3-dihydroxy-2-propenal.[a]

	Ene-diol		Mono-anion				Di-anion			
Method	λ_{max}	ΔE	λ_{max}	$\Delta\lambda_{max}$[b]	ΔE	$\Delta\Delta E$[c]	λ_{max}	$\Delta\lambda_{max}$[b]	ΔE	$\Delta\Delta E$[c]
M06-HF	244	5.08	279	35	4.44	□0.64	324	80	3.82	–1.26
LC-ωPBE	252	4.92	271	19	4.58	□0.34	324	72	3.83	–1.09
LC-BLYP	247	5.02	268	21	4.62	□0.40	320	73	3.88	–1.14
Experiment	268	4.63	293	25	4.23	□0.40	323	55	3.84	–0.79

[a]All calculations carried out with the aug-cc-pVDZ basis set and geometries optimized using B3LYP/6-31G(d,p) in combination with a CPCM description of the water solvent.
[b]Change in λ_{max} relative to the ene-diol form.
[c]Change in ΔE relative to the ene-diol form.

Table S4a. Calculated λ_{max} (in nm) and the corresponding excitation energies (ΔE, in eV) and oscillator strengths (f) of the oxidation product of the model system and its anion obtained without diffuse functions.[a]

Method		Oxidation product	Anion of the oxidation product	
M06-HF	λ_{max}	252	274	394
	ΔE	4.92	4.52	3.14
	f	0.45	0.61	0.09
LC-ωPBE	λ_{max}	257	281	450
	ΔE	4.83	4.41	2.76
	f	0.53	0.54	0.07
LC-BLYP	λ_{max}	248	276	419
	ΔE	4.99	4.48	2.96
	f	0.53	0.55	0.07
CAM-B3LYP	λ_{max}	276	290	550
	ΔE	4.49	4.28	2.25
	f	0.45	0.68	0.08
ωB97X-D	λ_{max}	276	288	541
	ΔE	4.48	4.30	2.29
	f	0.46	0.63	0.08
BHandH	λ_{max}	268	280	516
	ΔE	4.62	4.42	2.40
	f	0.43	0.66	0.10
BHandHLYP	λ_{max}	267	279	512
	ΔE	4.65	4.44	2.43
	f	0.43	0.64	0.09
CC2	λ_{max}	262	286	–
	ΔE	4.74	4.34	–
	f	0.44	0.41	–
Experiment	λ_{max}	270	276	371
	ΔE	4.60	4.49	3.34
	ε_{max}[b]	9400	9500	1800

[a]All calculations carried out with the 6-31G(d,p) basis set (cc-pVDZ for CC2) and geometries optimized using B3LYP/6-31G(d,p) in combination with a CPCM description of the water solvent.
[b]ε_{max} given in units of $dm^3mol^{-1}cm^{-1}$.

Table S4b. Calculated λ_{max} (in nm) and the corresponding excitation energies (ΔE, in eV) and oscillator strengths (f) of the oxidation product of the model system and its anion obtained with diffuse functions.[a]

Method		Oxidation product	Anion of the oxidation product	
M06-HF	λ_{max}	259	285	380
	ΔE	4.78	4.35	3.26
	f	0.47	0.71	0.11
LC-ωPBE	λ_{max}	260	289	420
	ΔE	4.77	4.28	2.95
	f	0.54	0.67	0.09
LC-BLYP	λ_{max}	253	286	397
	ΔE	4.91	4.34	3.12
	f	0.55	0.71	0.09
CAM-B3LYP	λ_{max}	281	296	500
	ΔE	4.42	4.18	2.48
	f	0.46	0.62	0.10
ωB97X-D	λ_{max}	280	293	494
	ΔE	4.42	4.23	2.51
	f	0.47	0.48	0.10
BHandH	λ_{max}	272	287	478
	ΔE	4.56	4.32	2.59
	f	0.43	0.62	0.12
BHandHLYP	λ_{max}	270	286	474
	ΔE	4.59	4.34	2.62
	f	0.44	0.61	0.11
CC2	λ_{max}	272	–	–
	ΔE	4.56	–	–
	f	0.46	–	–
Experiment	λ_{max}	270	276	371
	ΔE	4.60	4.49	3.34
	ε_{max}[b]	9400	9500	1800

[a]All calculations carried out with the aug-cc-pVDZ basis set and geometries optimized using B3LYP/6-31G(d,p) in combination with a CPCM description of the water solvent.
[b]ε_{max} given in units of $dm^3mol^{-1}cm^{-1}$.

Table S5. Comparison of LC-ωPBE, CC2 and CASPT2 gas-phase λ_{max} (in nm) and the corresponding excitation energies (ΔE, in eV) and oscillator strengths (f) of different forms of the model system.[a]

Method[b]		Hydroxyketone	Ene-diol	Mono-anion	Di-anion	
LC-ωPBE	λ_{max}	257	333	425	284	450
	ΔE	4.82	3.72	2.92	4.36	2.75
	f	0.49	0.54	0.71	0.51	0.39
CC2	λ_{max}	252	344	472	306	521
	ΔE	4.91	3.60	2.63	4.05	2.38
	f	0.51	0.34	0.76	0.54	0.34
CASPT2[c]	λ_{max}	262	350	495	302	494
	ΔE	4.73	3.54	2.50	4.10	2.51
	f[d]	0.77	0.88	1.15	0.56	0.39

[a]All calculations based on geometries optimized using B3LYP/6-31G(d,p) in combination with a CPCM description of the water solvent (*i.e.*, excited-state gas-phase calculations on solution-phase geometries).

[b]LC-ωPBE and CC2 calculations with the cc-pVDZ basis set, CASPT2 calculations with the ANO-S-VDZP basis set.

[c]CASPT2(8,8) for the hydroxyketone and ene-diol forms, CASPT2(10,9) for the mono-anion form, and CASPT2(12,10) for the di-anion form, with the corresponding active spaces comprising the full π-system, the full π-system and one oxygen lone pair, and the full π-system and two oxygen lone pairs, respectively.

[d]Oscillator strengths computed from CASSCF wavefunctions.

Table S6. Calculated changes in λ_{max} (in nm) of different forms of the model system when molecular geometries are optimized with diffuse functions.

Method[a]	Hydroxy-ketone	Ene-diol	Mono-anion	Di-anion (Peak 1)	Di-anion (Peak 2)
M06-HF	$260^b \rightarrow 257^c$	$343^b \rightarrow 342^c$	$477^b \rightarrow 477^c$	$273^b \rightarrow 274^c$	$429^b \rightarrow 430^c$
LC-ωPBE	$264^b \rightarrow 262^c$	$349^b \rightarrow 349^c$	$453^b \rightarrow 453^c$	$277^b \rightarrow 278^c$	$446^b \rightarrow 449^c$
LC-BLYP	$260^b \rightarrow 257^c$	$342^b \rightarrow 341^c$	$446^b \rightarrow 446^c$	$271^b \rightarrow 272^c$	$432^b \rightarrow 434^c$
CAM-B3LYP	$281^b \rightarrow 278^c$	$368^b \rightarrow 369^c$	$452^b \rightarrow 452^c$	$289^b \rightarrow 286^c$	$488^b \rightarrow 490^c$
ωB97X-D	$282^b \rightarrow 279^c$	$368^b \rightarrow 368^c$	$454^b \rightarrow 454^c$	$289^b \rightarrow 282^c$	$488^b \rightarrow 490^c$
BHandH	$275^b \rightarrow 273^c$	$361^b \rightarrow 361^c$	$435^b \rightarrow 435^c$	$276^b \rightarrow 277^c$	$464^b \rightarrow 466^c$
BHandHLYP	$274^b \rightarrow 272^c$	$358^b \rightarrow 359^c$	$434^b \rightarrow 434^c$	$276^b \rightarrow 277^c$	$463^b \rightarrow 464^c$

[a]Excited-state singlepoint calculations carried out with the 6-31G(d,p) basis set.
[b]Calculations based on geometries optimized using B3LYP/6-31G(d,p) in combination with a CPCM description of the water solvent.
[c]Calculations based on geometries optimized using B3LYP/6-31+G(d,p) in combination with a CPCM description of the water solvent.

Table S7. Calculated excitation energies (ΔE, in eV) for the strongly absorbing "B_u-like" state of different forms of the model system at different molecular geometries.[a]

		M06-HF		LC-ωPBE		LC-BLYP	
Form	Geometry	ΔE	ΔΔE[b]	ΔE	ΔΔE[b]	ΔE	ΔΔE[b]
Hydroxyketone	BP86	4.57	–	4.49	–	4.57	–
	BLYP	4.60	–	4.52	–	4.60	–
	B3LYP	4.79	–	4.71	–	4.79	–
	M06-2X	4.96	–	4.87	–	4.96	–
	MP2	4.82	–	4.75	–	4.83	–
	MAD[c]	0.39	–	0.38	–	0.39	–
Ene-diol	BP86	3.38	–1.19	3.34	–1.15	3.41	–1.16
	BLYP	3.42	–1.18	3.38	–1.14	3.44	–1.16
	B3LYP	3.62	–1.17	3.56	–1.15	3.64	–1.15
	M06-2X	3.82	–1.14	3.74	–1.13	3.83	–1.13
	MP2	3.75	–1.07	3.67	–1.08	3.76	–1.07
	MAD[c]	0.44	0.12	0.40	0.07	0.42	0.09
Mono-anion	BP86	2.55	–2.02	2.69	–1.80	2.73	–1.84
	BLYP	2.55	–2.05	2.69	–1.83	2.73	–1.87
	B3LYP	2.64	–2.15	2.77	–1.94	2.81	–1.98
	M06-2X	2.74	–2.22	2.85	–2.02	2.90	–2.06
	MP2	2.67	–2.15	2.80	–1.95	2.84	–1.99
	MAD[c]	0.19	0.20	0.16	0.22	0.17	0.22
Di-anion	BP86	2.67	–1.90	2.59	–1.90	2.67	–1.90
	BLYP	2.68	–1.92	2.59	–1.93	2.67	–1.93
	B3LYP	2.90	–1.89	2.79	–1.92	2.88	–1.91
	M06-2X	3.15	–1.81	3.00	–1.87	3.11	–1.85
	MP2	2.97	–1.85	2.85	–1.90	2.95	–1.88
	MAD[c]	0.48	0.11	0.41	0.06	0.44	0.08

[a]All calculations based on the most stable stereoisomer of the respective form and carried out with the 6-31G(d,p) basis set in combination with a CPCM description of the water solvent.
[b]Shift relative to the corresponding excitation energy for the hydroxyketone form.
[c]Maximum absolute deviation.

Table S8. Calculated excitation energies (ΔE, in eV) for the strongly absorbing "B_u-like" state of different forms of astaxanthin at different molecular geometries.[a]

		M06-HF		LC-ωPBE		LC-BLYP	
Form	Geometry	ΔE	ΔΔE[b]	ΔE	ΔΔE[b]	ΔE	ΔΔE[b]
Keto	BP86	2.39	–	2.30	–	2.34	–
	BLYP	2.42	–	2.32	–	2.37	–
	B3LYP	2.74	–	2.64	–	2.70	–
	M06-2X	3.08	–	2.98	–	3.05	–
	MAD[c]	0.69	–	0.68	–	0.71	–
Mono-enol	BP86	2.24	–0.15	2.14	–0.16	2.19	–0.15
	BLYP	2.29	–0.13	2.19	–0.13	2.24	–0.13
	B3LYP	2.71	–0.03	2.61	–0.03	2.67	–0.03
	M06-2X	3.08	0.00	2.97	–0.01	3.04	–0.01
	MAD[c]	0.84	0.15	0.83	0.15	0.85	0.14
Di-enol	BP86	2.31	–0.08	2.21	–0.09	2.26	–0.08
	BLYP	2.36	–0.06	2.26	–0.06	2.31	–0.06
	B3LYP	2.73	–0.01	2.63	–0.01	2.69	–0.01
	M06-2X	3.08	0.00	2.97	–0.01	3.05	0.00
	MAD[c]	0.77	0.08	0.76	0.08	0.79	0.08
Mono-anion	BP86	1.60	–0.79	1.60	–0.70	1.63	–0.71
	BLYP	1.60	–0.82	1.60	–0.72	1.63	–0.74
	B3LYP	1.89	–0.85	1.89	–0.75	1.93	–0.77
	M06-2X	2.37	–0.71	2.37	–0.61	2.44	–0.61
	MAD[c]	0.77	0.14	0.77	0.14	0.81	0.16
Di-anion	BP86	1.40	–0.99	1.42	–0.88	1.47	–0.87
	BLYP	1.42	–1.00	1.43	–0.89	1.48	–0.89
	B3LYP	1.60	–1.14	1.62	–1.02	1.67	–1.03
	M06-2X	1.88	–1.20	1.90	–1.08	1.97	–1.08
	MAD[c]	0.48	0.21	0.48	0.20	0.50	0.21

[a]All calculations based on the most stable stereoisomer of the respective form and carried out with the 6-31G(d,p) basis set in combination with a CPCM description of the water solvent.
[b]Shift relative to the corresponding excitation energy for the keto form.
[c]Maximum absolute deviation.

Table S9. Calculated excitation energies (in eV) for the strongly absorbing "B_u-like" state of different forms of the model and full carotenoid systems at B3LYP, MP2 and CC2 gas-phase geometries.[a]

System	Form	Geometry	M06-HF	LC-ωPBE	LC-BLYP
Model	Hydroxyketone	B3LYP	4.87	4.79	4.87
		MP2	4.86	4.79	4.87
		CC2	4.79	4.72	4.80
Model	Ene-diol	B3LYP	3.75	3.72	3.80
		MP2	3.78	3.74	3.81
		CC2	3.68	3.65	3.72
Model	Mono-anion	B3LYP	2.79	2.91	2.96
		MP2	2.78	2.90	2.95
		CC2	2.73	2.86	2.90
Model	Di-anion	B3LYP	2.87	2.75	2.84
		MP2	2.86	2.75	2.84
		CC2	2.76	2.65	2.74
Full	Keto	B3LYP	2.85	2.78	2.83
		MP2	2.82	2.75	2.80
		CC2	2.72	2.64	2.69
Full	Mono-enol	B3LYP	2.79	2.71	2.76
		MP2	2.79	2.71	2.77
		CC2	2.67	2.59	2.64
Full	Di-enol	B3LYP	2.77	2.69	2.75
		MP2	2.78	2.70	2.75
		CC2	2.66	2.58	2.64
Full	Mono-anion	B3LYP	1.50	1.55	1.57
		MP2	1.53	1.58	1.60
		CC2	1.31	1.36	1.37
Full	Di-anion	B3LYP	1.61	1.66	1.70
		MP2	1.53	1.58	1.61
		CC2	1.39	1.44	1.46

[a]All calculations based on the most stable stereoisomer of the respective form and carried out with the cc-pVDZ basis set in the gas phase.

Table S10a. B-factor values (in Å²) for AXE and AXT (1) atoms after refinement and their respective difference values.

		B_{AXT}	B_{AXE}	B_{AXI}	$(B_{AXT})-(B_{AXE})$	$(B_{AXT})-(B_{AXI})$	
ATOM	C1	68.3	68.08	69.07	0.22	-0.77	Ring 1 – Tyr51
ATOM	C2	67.5	67.21	68.2	0.29	-0.7	
ATOM	C3	66.93	66.72	67.7	0.21	-0.77	
ATOM	C4	67.69	67.34	68.4	0.35	-0.71	
ATOM	C5	68.1	67.65	68.78	0.45	-0.68	
ATOM	C6	67.27	66.79	68.14	0.48	-0.87	
ATOM	C7	68.61	68.13	69.18	0.48	-0.57	Polyene chain
ATOM	C8	68.08	67.26	68.28	0.82	-0.2	
ATOM	C9	67.91	67.5	68.46	0.41	-0.55	
ATOM	C10	65.94	65.57	66.56	0.37	-0.62	
ATOM	C11	67.34	67.04	67.97	0.3	-0.63	
ATOM	C12	65.52	65.35	66.17	0.17	-0.65	
ATOM	C13	62.19	62.14	62.91	0.05	-0.72	
ATOM	C14	60.89	60.84	61.57	0.05	-0.68	
ATOM	C15	60.16	60.27	61	-0.11	-0.84	
ATOM	C16	61.33	61.29	61.99	0.04	-0.66	
ATOM	C17	61.83	61.71	62.35	0.12	-0.52	
ATOM	C18	61.69	61.53	62.17	0.16	-0.48	
ATOM	C19	62.33	62.15	62.77	0.18	-0.44	
ATOM	C20	62.89	62.63	63.25	0.26	-0.36	
ATOM	O3	62.31	62.15	62.75	0.16	-0.44	Ring 1 – Oxygens
ATOM	O4	61.74	61.62	62.18	0.12	-0.44	
ATOM	C21	61.75	61.64	62.18	0.11	-0.43	Ring 2 – His92
ATOM	C22	62.04	61.93	62.42	0.11	-0.38	
ATOM	C23	62.46	62.41	62.84	0.05	-0.38	
ATOM	C24	62.72	62.59	63.07	0.13	-0.35	
ATOM	C25	61.43	61.52	61.9	-0.09	-0.47	
ATOM	C26	60.41	60.56	60.9	-0.15	-0.49	
ATOM	C27	61.45	61.68	62.01	-0.23	-0.56	Polyene chain
ATOM	C28	61.65	62	62.33	-0.35	-0.68	
ATOM	C29	63.7	64	64.35	-0.3	-0.65	
ATOM	C30	60.73	61.27	61.59	-0.54	-0.86	
ATOM	C31	59.26	60.03	60.31	-0.77	-1.05	
ATOM	C32	58.26	59.24	59.46	-0.98	-1.2	
ATOM	C33	58.29	59.4	59.57	-1.11	-1.28	
ATOM	C34	59.13	60.35	60.48	-1.22	-1.35	
ATOM	C35	57.86	58.94	59.17	-1.08	-1.31	
ATOM	C36	58.3	59.43	59.53	-1.13	-1.23	
ATOM	C37	58.76	59.84	59.96	-1.08	-1.2	
ATOM	C38	59.34	60.65	60.51	-1.31	-1.17	
ATOM	C39	58.56	59.54	59.83	-0.98	-1.27	
ATOM	C40	58.88	54.28	52.46	-0.86	0.96	
ATOM	O24	57.79	53.63	52.35	-0.45	0.83	Ring 2 – Oxygens
ATOM	O23	55.47	52.79	51.42	-0.34	1.03	

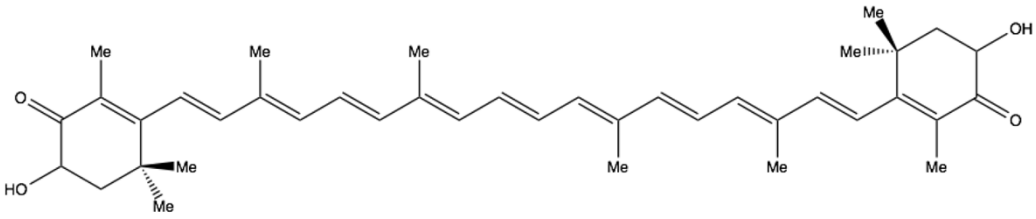
Table S10b. B-factor values (in Å²) for AXE and AXT (2) atoms after refinement and their respective difference values.

	B_{AXT}	B_{AXE}	B_{AXI}	$(B_{AXT})-(B_{AXE})$	$(B_{AXT})-(B_{AXI})$	
ATOM C1	40.72	40.09	40.4	0.63	0.32	Ring 1 – Tyr51
ATOM C2	40.18	39.59	39.92	0.59	0.26	
ATOM C3	38.68	38.19	38.46	0.49	0.22	
ATOM C4	40.75	39.93	40.49	0.82	0.26	
ATOM C5	41.78	40.6	41.25	1.18	0.53	
ATOM C6	43.85	42.75	43.42	1.1	0.43	
ATOM C7	42.47	41.34	41.76	1.13	0.71	Polyene chain
ATOM C8	42.36	40.51	40.89	1.85	1.47	
ATOM C9	40.67	39.74	40.13	0.93	0.54	
ATOM C10	40.99	40.21	40.63	0.78	0.36	
ATOM C11	39.92	39.24	39.62	0.68	0.3	
ATOM C12	38.82	38.17	38.52	0.65	0.3	
ATOM C13	36.62	35.94	36.22	0.68	0.4	
ATOM C14	36.87	36.22	36.44	0.65	0.43	
ATOM C15	36.62	36.01	36.21	0.61	0.41	
ATOM C16	37.89	37.42	37.63	0.47	0.26	
ATOM C17	35.78	35.49	35.66	0.29	0.12	
ATOM C18	34.35	33.87	34.02	0.48	0.33	
ATOM C19	34.79	34.25	34.33	0.54	0.46	
ATOM C20	34.11	33.54	33.65	0.57	0.46	
ATOM O3	36.06	35.64	35.65	0.42	0.41	Ring 1 – Oxygens
ATOM O4	36.48	36.2	36.09	0.28	0.39	
ATOM C21	38.29	38.04	37.93	0.25	0.36	Ring 2 – His92
ATOM C22	38.26	38.13	37.91	0.13	0.35	
ATOM C23	38.83	38.74	38.46	0.09	0.37	
ATOM C24	38.48	38.49	38.14	-0.01	0.34	
ATOM C25	41.95	41.9	41.53	0.05	0.42	
ATOM C26	44.14	44.17	43.63	-0.03	0.51	
ATOM C27	46.5	46.7	46.03	-0.2	0.47	Polyene chain
ATOM C28	48.9	49.19	48.42	-0.29	0.48	
ATOM C29	49.63	50.04	49.28	-0.41	0.35	
ATOM C30	49.48	49.96	49.06	-0.48	0.42	
ATOM C31	50.99	51.57	50.5	-0.58	0.49	
ATOM C32	54.03	54.53	53.36	-0.5	0.67	
ATOM C33	55.22	55.74	54.54	-0.52	0.68	
ATOM C34	56.6	57.04	55.94	-0.44	0.66	
ATOM C35	55.2	55.65	54.5	-0.45	0.7	
ATOM C36	55.59	56.07	54.76	-0.48	0.83	
ATOM C37	55.8	56.38	54.84	-0.58	0.96	
ATOM C38	57.89	58.54	57.05	-0.65	0.84	
ATOM C39	53.66	54.13	52.76	-0.47	0.9	
ATOM C40	53.42	54.28	52.46	-0.86	0.96	
ATOM O24	53.18	53.63	52.35	-0.45	0.83	Ring 2 – Oxygens
ATOM O23	52.45	52.79	51.42	-0.34	1.03	

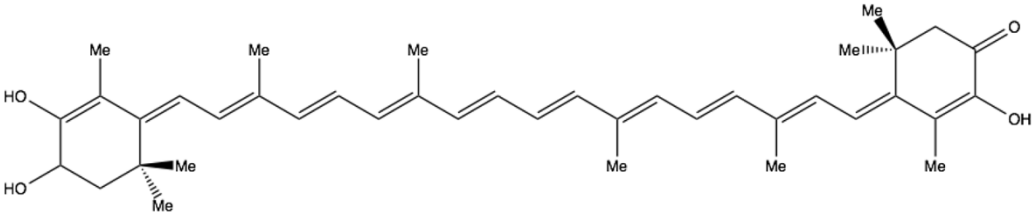
Table S11. Overall R-values for the working and test sets (Rfree) of the CNS refinement runs.

Astaxanthin (1) form	Astaxanthin (2) form	Overall R-value for the working set[a]	Overall R-value for the test set[a]
AXT	AXT	0.204	0.245
AXE	AXE	0.205	0.245
AXI	AXI	0.205	0.245
AXT	AXI	0.203	0.245
AXE	AXI	0.204	0.246
AXI	AXE	0.204	0.246

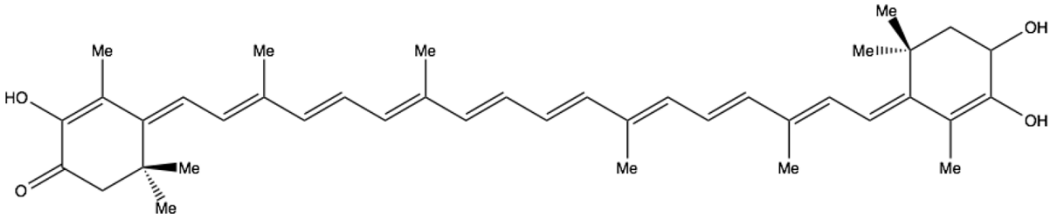
[a]Whilst the community convention is followed of quoting these R values to 3 decimal places the final decimal place is probably only just significant, if at all. Thus any differences seen in the 3rd decimal place are likely insignificant.



(a)



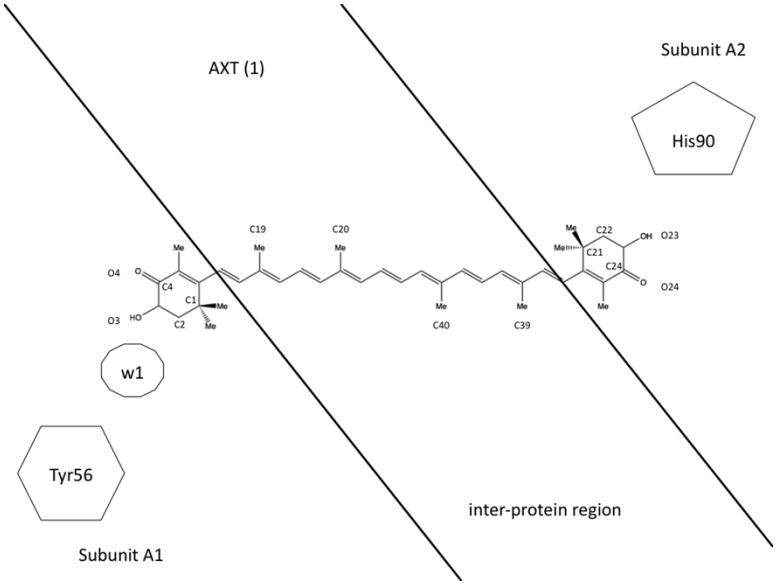
(b)



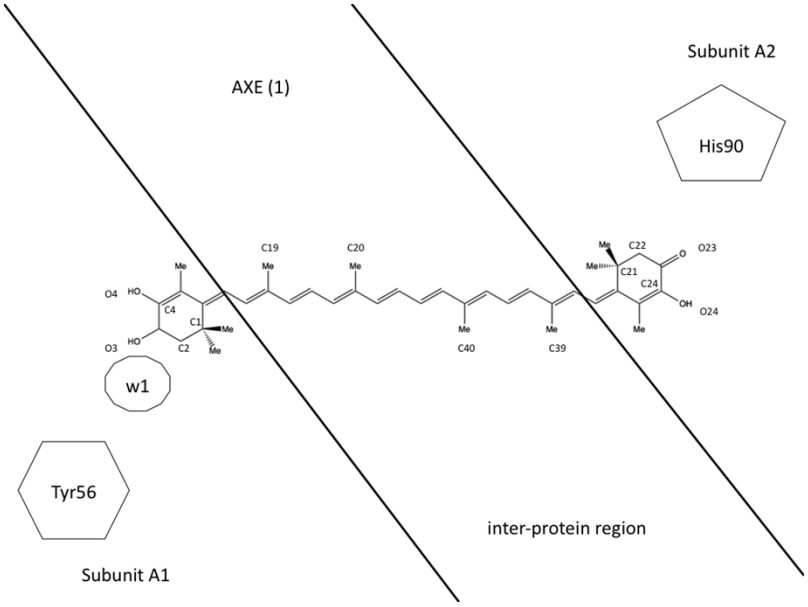
(c)

Fig. S1.

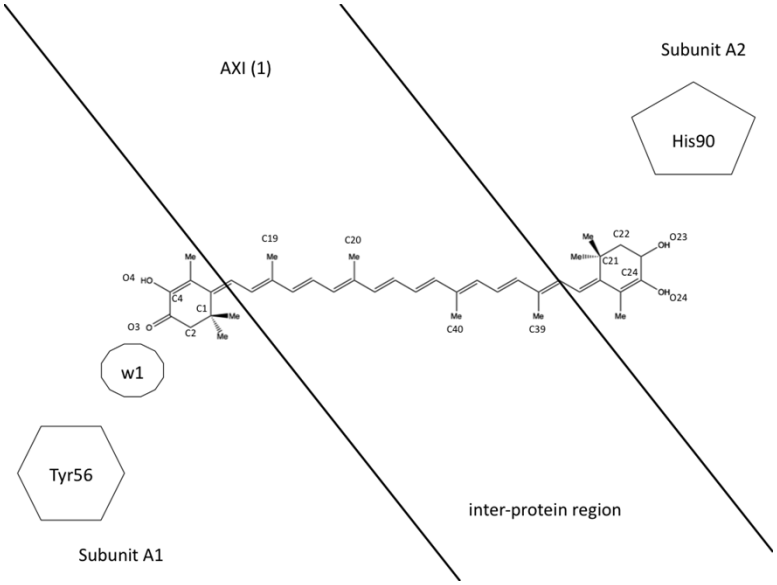
Astaxanthin in the s-*trans* conformation shown in its ground-state form (AXT, Fig. S1a), in an enolic form in which the right end ring is enolized (AXE, Fig. S1b), and an inverted enolic form in which the left end ring is enolized (AXI, Fig. S1c). Chemically, AXE and AXI are identical and correspond to the right resonance structure of the enolate $\mathbf{1}^-$ depicted in Scheme 1 of the main text, there shown deprotonated and in the s-*cis* conformation.


AXT (1)
Subunit A2
His90
w1
Tyr56
inter-protein region
Subunit A1

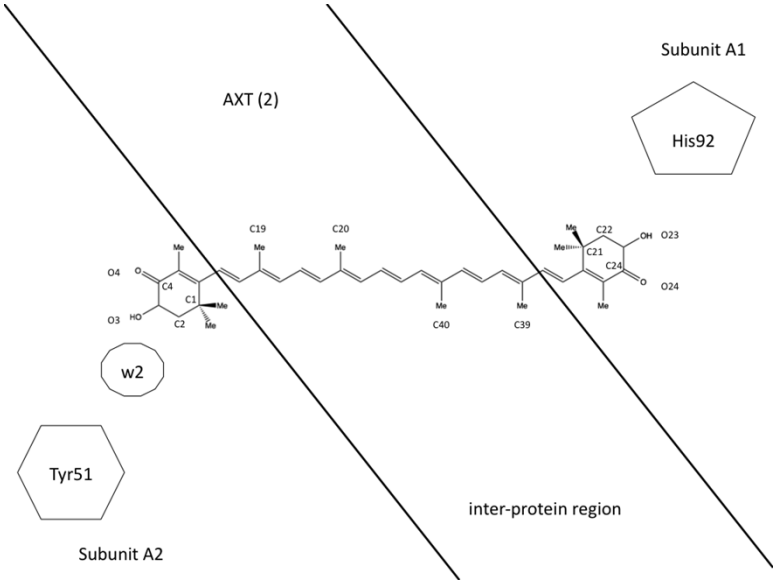
(a)


AXE (1)
Subunit A2
His90
w1
Tyr56
inter-protein region
Subunit A1

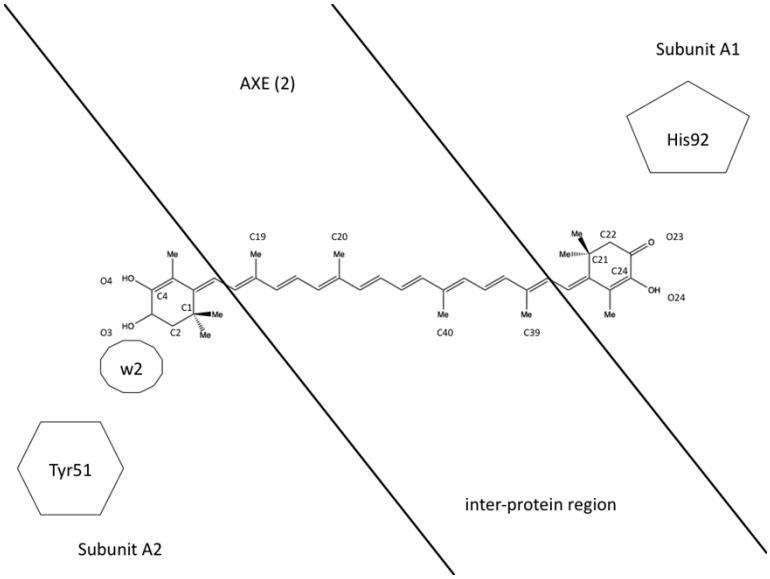
(b)


AXI (1)
Subunit A2
His90
w1
Tyr56
inter-protein region
Subunit A1

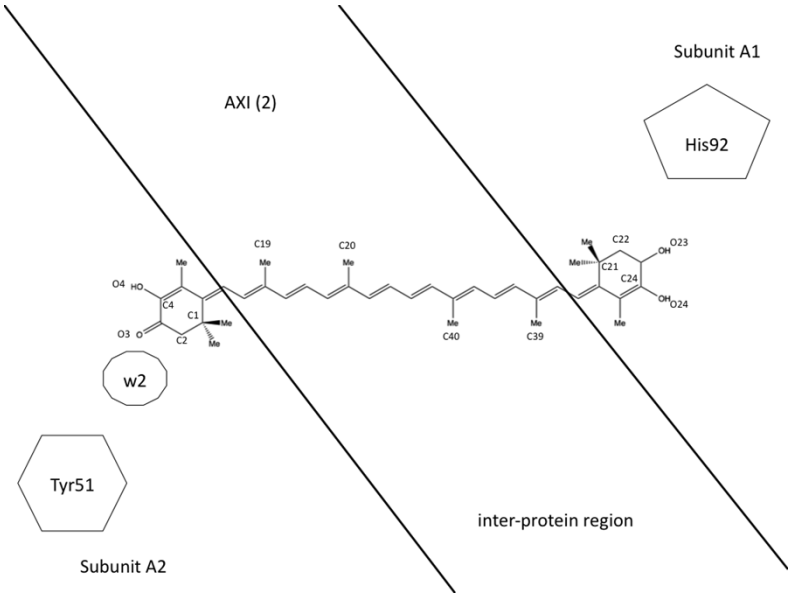
(c)


AXT (2)
Subunit A1
His92
w2
Tyr51
inter-protein region
Subunit A2

(d)



(e)



(f)

Fig. S2.

The two protein-bound astaxanthin molecules shown in the ground-state form (AXT1, Fig. S2a, and AXT2, Fig. S2d) and in two possible orientations of the enolic form. In one orientation the enolized end ring of one molecule (AXE1) faces His90 of subunit A2 (Fig. S2b) and that of the other molecule (AXE2) faces His92 of subunit A1 (Fig. S2e). In the other orientation the enolized end ring of one molecule (AXI1) faces Tyr56 of subunit A1 (Fig. S2c) and that of the other molecule (AXI2) faces Tyr51 of subunit A2 (Fig. S2f). The transversal black lines mark the three different regions clamping each carotenoid.

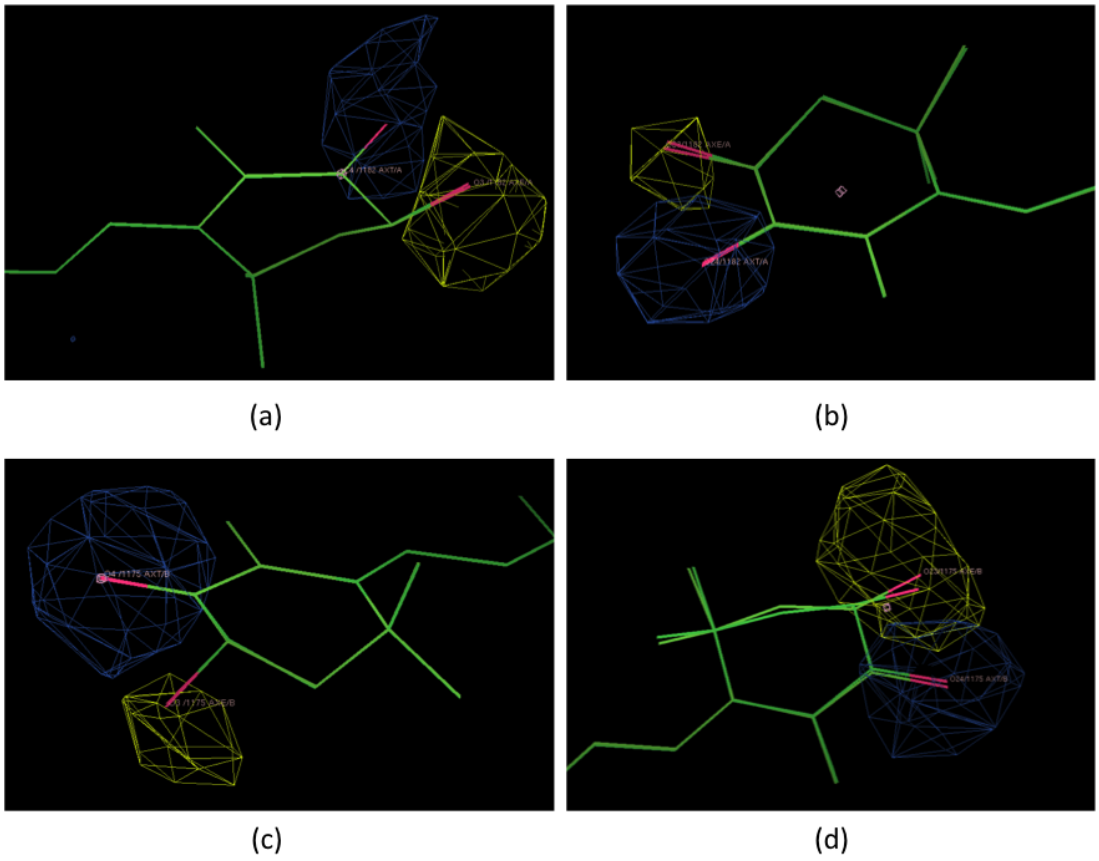


Fig. S3.

Fo-Fc omit electron density maps (blue color for atoms O4 and O24; yellow color for atoms O3/O23), with a 10 sigma cut-off contour level calculated omitting O3/O4 (Fig. S3a) and O23/O24 (Fig. S3b) from AXT/AXE 1; and omitting O3/O4 (Fig. S3c) and O23/O24 (Fig. S3d) from AXT/AXE 2.

References

1. E. Becher, R. Albrecht, K. Bernhard, H. G. W. Leuenberger, H. Mayer, R. K. Müller, W. Schüep and H. P. Wagner, *Helv. Chim. Acta*, 1981, **64**, 2419–2435.
2. F. D. Negra, G. Santoni, M. Stivanello, M. Soukup and M. Fachini, *PCT Int. Appl.*, 2007, WO2007072529 A2 20070628.
3. M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery Jr., J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, N. J. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, Ö. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski and D. J. Fox, *Gaussian 09, Revision D.01* (2009) Gaussian, Inc., Wallingford, CT, USA.
4. O. Christiansen, H. Koch and P. Jørgensen, *Chem. Phys. Lett.*, 1995, **243**, 409-418.
5. C. Hättig, *J. Chem. Phys.*, 2003, **118**, 7751-7761.
6. K. Andersson, P.-Å. Malmqvist and B. O. Roos, *J. Chem. Phys.*, 1992, **96**, 1218-1226.
7. R. Ahlrichs, M. Bär, M. Häser, H. Horn and C. Kölmel, *Chem. Phys. Lett.*, 1989, **162**, 165-169.
8. *TURBOMOLE V6.3 2011, a development of University of Karlsruhe and Forschungszentrum Karlsruhe GmbH, 1989-2007, TURBOMOLE GmbH, since 2007; available at:* http://www.turbomole.com *(accessed July 25, 2013)*.
9. F. Aquilante, L. D. Vico, N. Ferré, G. Ghigo, P.-Å. Malmqvist, P. Neogrády, T. B. Pedersen, M. Pitoňák, M. Reiher, B. O. Roos, L. Serrano-Andrés, M. Urban, V. Veryazov and R. Lindh, *J. Comput. Chem.*, 2010, **31**, 224–247.
10. M. Cossi, N. Rega, G. Scalmani and V. Barone, *J. Comput. Chem.*, 2003, **24**, 669-681.
11. M. A. L. Marques and E. K. U. Gross, *Annu. Rev. Phys. Chem.*, 2004, **55**, 427-455.
12. A. Dreuw and M. Head-Gordon, *Chem. Rev.*, 2005, **105**, 4009-4037.
13. M. E. Casida, *J. Mol. Struct. (Theochem)*, 2009, **914**, 3-18.
14. T. Yanai, D. P. Tew and N. C. Handy, *Chem. Phys. Lett.*, 2004, **393**, 51-57.
15. O. A. Vydrov and G. E. Scuseria, *J. Chem. Phys.*, 2006, **125**, 234109.
16. Y. Zhao and D. G. Truhlar, *J. Phys. Chem. A*, 2006, **110**, 13126-13130.
17. J. R. Pliego Jr and J. M. Riveros, *J. Phys. Chem. A*, 2002, **106**, 7434–7439
18. M. D. Liptak and G. C. Shields, *J. Am. Chem. Soc.*, 2001, **123**, 7314–7319.
19. M. D. Tissandier, K. A. Cowen, W. Y. Feng, E. Gundlach, M. H. Cohen, A. D. Earhart and J. V. Coe, *J. Phys. Chem. A*, 1998, **102**, 7787–7794.
20. M. Cianci, P. J. Rizkallah, A. Olczak, J. Raftery, N. E. Chayen, P. F. Zagalsky and J. R. Helliwell, *Proc. Natl. Acad. Sci. U. S. A.*, 2002, **99**, 9795-9800.
21. A. T. Brünger, *Nat. Protoc.*, 2007, **2**, 2728–2733.
22. A. W. Schuettelkopf and D. M. F. v. Aalten, *Acta Crystallogr. Sect. D. Biol. Crystallogr.*, 2004, **60**, 1355-1363.
23. G. Bartalucci, S. Fisher, J. R. Helliwell, M. Helliwell, S. Liaaen-Jensen, J. E. Warren and J. Wilkinson, *Acta Cryst.* , 2009, **B65**, 238-247.
24. G. Bartalucci, J. Coppin, S. Fisher, G. Hall, J. R. Helliwell, M. Helliwell and S. Liaaen-Jensen, *Acta Cryst.*, 2007, **B63**, 328-337.
25. Y. Abe, H. Horii, S. Taniguchi, K. Kamai and M. Takagi, *Bull. Chem. Soc. Jpn.*, 1983, **56**, 467-470.
26. M. Wanko, M. Hoffmann, P. Strodel, A. Koslowski, W. Thiel, F. Neese, T. Frauenheim and M. Elstner, *J. Phys. Chem. B*, 2005, **109**, 3606-3615.
27. D. W. J. Cruickshank, *Acta Crystallogr. Sect. D. Biol. Crystallogr.*, 1999, **55**, 583-601.