Computational Study of the Structure and Electronic Circular Dichroism Spectroscopy of Blue Copper Proteins

Hainam Do[†], Robert J. Deeth[‡] and Nicholas A. Besley[†*]

[†]*School of Chemistry, University of Nottingham, University Park, Nottingham, NG7 2RD, UK and*

[‡] *Inorganic Computational Chemistry Group, Department of Chemistry, University of Warwick, Coventry, CV4 7AL, UK.*

E-mail: nick.besley@nottingham.ac.uk

[*]To whom correspondence should be addressed

Abstract

The calculation of the electronic circular dichroism (CD) spectra of the oxidised form of the blue copper proteins plastocyanin and cucumber basic protein and the relationship between the observed spectral features and the structure of the active site of the protein is investigated. Excitation energies and transition strengths are computed using multi reference configuration interaction, and it is shown that computed spectra based on coordinates from the crystal structure or a single structure optimised in quantum mechanics/molecular mechanics (QM/MM) or ligand field molecular mechanics (LFMM) are qualitatively incorrect. In particular, the rotational strength of the ligand to metal charge transfer band is predicted to be too small or have the incorrect sign. By considering calculations on active site models with modified structures it is shown that the intensity of this band is sensitive to the non-planarity of the histidine and cysteine ligands coordinated to copper. Calculation of the ultraviolet absorption and CD spectra based upon averaging over many structures drawn from a LFMM molecular dynamics simulation are in good agreement with experiment, and superior to analogous calculations based upon structures from a classical molecular dynamics simulation. This provides evidence that the LFMM force field provides an accurate description of the molecular dynamics of these proteins.

keywords: plastocyanin, cucumber basic protein, ligand-field molecular mechanics, molecular dynamics

Introduction

Blue copper proteins play an important role in a number of biological processes such as photosynthesis and nitrogen fixation, where the primary role of the proteins is to facilitate electron transfer.[1] In these proteins, copper is bound in an approximate trigonal plane to the sulphur atom of cysteine and the nitrogen atoms of two histidine ligands. In most blue copper proteins, there is also a methionine ligand in the axial position. The precise nature of this coordination and its relationship to the function of the protein has fascinated inorganic and biological chemists. The distorted trigonal geometry of the proteins is largely determined by the strong Cu-thiolate bond[2] and is intermediate between the stable geometries preferred by Cu(I) and Cu(II) complexes. This leads to a small geometry change on reduction from Cu(II) to Cu(I),[3,4] resulting in a small reorganization energy and high rate of electron transfer.[5]

The oxidised forms of blue copper proteins have a d^9 electronic configuration that results in distinctive absorption spectra,[6–9] with an intense band at about 2.1 eV (16 700 cm^{-1}). This band is often referred to as arising from a ligand to metal charge transfer (LMCT) transition, which corresponds to an excitation from the Cys_π orbital to the singly occupied orbital (SOMO). These orbitals are formed from the bonding and antibonding combinations of the Cu $3d_{x^2-y^2}$ and $S_{cys}3p_\pi$ orbitals, respectively (see Figure 1). Ligand field d←d excitations lie at approximately 1.6 eV (12 800 cm^{-1}) and a further band is observed at 2.7 eV (21 400 cm^{-1}), which is assigned to the excitation from the Cys_σ orbital to the SOMO. The absorption spectra are sensitive to small variations in the underlying structure of the copper active site. On going from axial type proteins, such as plastocyanin, to more rhombic type proteins, such as cucumber basic protein and nitrite reductase, there is an increase in the intensity of the higher energy feature at 2.7 eV and a decrease in intensity of the band near 2.1 eV, with the combined intensities of the bands remaining approximately constant.[10]

Blue copper proteins have also been studied with electronic circular dichroism (CD) spec-

troscopy.[6,7] There are more distinct bands in the CD spectra compared to the absorption spectra. For plastocyanin, a small positive band at 1.3 eV (10 800 cm^{-1}) and larger negative band at about 1.6 eV are observed. The LMCT transition leads to an intense band at 2.1 eV and at higher energy there is a small negative band at 2.6 eV (21 000 cm^{-1}) followed by a further positive band at about 3.0 eV (24 000 cm^{-1}).[6] For the perturbed blue copper proteins, cucumber basic protein and nitrite reductase, the pattern of the bands is similar, but the relative intensities of the bands change. In particular, the intensity of the band arising from the LMCT excitation is reduced relative to the ligand field bands.[6,7]

In addition to experimental work, a number of theoretical studies have investigated the structure or spectroscopy of the oxidised form of blue copper proteins.[11–20,22,24–26] Calculations confirm that the intense band observed in the absorption spectrum arises from an excitation from the orbital comprising a bonding mixture of the Cu $3d_{x^2-y^2}$ and $S_{cys}3p_{\pi}$ orbitals to the SOMO.[11,12] Complete active space self-consistent field (CASSCF) with multiconfigurational second-order perturbation theory (CASPT2) calculations have been reported by Roos and co-workers[14,16] and Vancoillie and Pierloot.[25] In these calculations, the active site is treated explicitly within the quantum chemical calculations with the remainder of the protein included as point charges. Overall, good agreement with the experiment was obtained, with the excitation energies of the six lowest excitations predicted within 0.25 eV of their experimental values. Time dependent density functional theory (TDDFT) has also been applied to compute the absorption,[20,27] electronic circular dichroism (CD)[27,28] and magnetic circular dichroism (MCD)[27] of blue copper proteins. While the qualitative shape of the absorption spectra was reproduced, the excitation energies were found to be too high and the computed CD spectra show a relatively poor agreement with experiment. These studies generally use the structure of the active site reported from X-ray diffraction. In a study on plastocyanin combining classical molecular dynamics simulations with multireference configuration interaction (MRCI) and TDDFT calculations,[29] it was shown that there is a very strong correlation between the energy of the intense absorption band and the Cu-S_{cys} bond length, and it was argued

that the bond length reported in the crystal structure of 2.07 Å should be longer, and a value of 2.15 Å is more accurate. The computed CD spectrum of plastocyanin from averaging over many structural snapshots taken from the classical molecular dynamics simulation was in better agreement with experiment than the computed spectrum based upon the crystal structure. However, significant deviations from experiment remained.

Fewer spectroscopic studies have considered the reduced form of the proteins, primarily since the d^{10} electronic configuration does not have the rich absorption spectroscopy of the d^9 electronic configuration of the ground state. The electronic structure of the reduced form has been studied in detail using photoelectron spectroscopy combined with Xα molecular orbitals calculations.[30] The spectroscopy of small molecules that incorporate the copper-cysteine bond has been studied through experiment and calculations[31,32] providing some insight into the electronic structure of the reduced form.

A number of authors have addressed the problem of determining the structure of the active sites of blue copper proteins directly through calculations. Quantum mechanics/molecular mechanics (QM/MM) approaches wherein the active site is described using density functional theory (DFT) while the remainder of the protein is treated using molecular mechanics have proved successful for treating metalloproteins.[33] Sinnecker and Neese reported a QM/MM geometry optimization of plastocyanin using the BP86 functional and split-valence SV(P) basis set, and found a bond length of 2.22 Å for Cu-S_{cys}.[20] Ryde and Olsson studied the structure of the oxidised and reduced forms of a variety of blue copper proteins within a QM/MM framework where the protein active site was treated using DFT with the B3LYP functional and found a Cu-S_{cys} bond length of 2.14 Å.[17] A limitation of QM/MM approaches is that the computational cost of the calculations makes structural optimization or the study of the molecular dynamics of the proteins challenging. To address this problem, the ligand-field molecular mechanics approach was developed by Deeth for oxidised Type I centers.[21,22] This method has been shown to provide a comparable accuracy to QM/MM

calculations but at a computational cost reduced by several orders of magnitude.

A precise knowledge of the structure of the active site of blue copper proteins is key to modelling their function. Structures determined from X-ray diffraction have an associated uncertainty which can be significant. Calculations have the potential to provide more accurate structures and have been used to refine crystal structures for systems including blue copper proteins.[23] Furthermore, it has been demonstrated that conformational averaging is necessary to reproduce properties such as the redox potential.[18] However, without accurate experimental structures, it is difficult to assess the accuracy of structures determined through calculations. Spectroscopic methods, such as CD spectroscopy, that show a great sensitivity to the underlying structure of the active site provide one mechanism through which the accuracy of the structure can be tested. In this paper, the calculated CD spectroscopy of the oxidised forms of plastocyanin and cucumber basic protein based upon single structures optimised using QM/MM or LFMM, compared to conformationally averaged data derived from LFMM molecular dynamics simulation is assessed, and the relationship between the structure of the active site and the different features observed in the CD spectra is explored.

Computational Details

QM/MM geometry optimizations were performed for the oxidised forms of plastocyanin (PDB code 1PLC) and cucumber basic protein (2CBP). The active site $[Cu(Im)_2(SCH_3)(S(CH_3)_2)]^+$ where Im denotes imidazole was treated using DFT with the B97-1 functional[34] and the remainder of the protein was described at the MM level using the CHARMM force field.[35] In the geometry optimisations, the 28 residues closest to the active site were allowed to relax while the remaining resides were kept fixed in their positions as given by the crystal structure. The 6-311G* basis set[36,37] was used for all atom types except copper. Two different basis sets were used for copper, the all electron 6-31G* basis set and effective core potential Stuttgart relativistic small core (SRSC)

basis sets.[38] These calculations exploit the interface between CHARMM and Q-Chem.[39,40]

The ligand field molecular dynamics (LFMD) simulations started from the optimised structures reported previously[22] which comprise the respective proteins surrounded by a layer of TIP3P water molecules approximately 10 Å thick (728 and 1536 water molecules for plastocytanin and cucumber basic protein respectively). While this protocol introduces interactions between the outer, 'surface' water molecules which tends to constrain the tertiary structure, we have shown[41] that the local geometry in the vicinity of the transition metal centre is able to relax fully. The canonical ensemble was employed using the Nose-Poincare-Anderson equations of motion at a temperature of 300 K, a timestep of 2 fs and a temperture relaxation time of 0.1 ps as implemented in DommiMOE,[42] our ligand field extended version of the Molecular Operating Environment (MOE), version 2011.10.[43] All bond lengths to hydrogen atoms were constrained. The simulation was run for 1.5 ns with a sample taken every 0.5 ps. The final 200 snapshots of each simulation was used for the spectral computations. Classical MD simulations have also been performed following a protocol detailed elsewhere.[29]

Excitation energies and intensities were computed using MRCI with the MOLPRO suite of programs.[44] Reference orbitals for the MRCI calculations were obtained from state averaged multiconfigurational self-consistent field calculations, and the subsequent MRCI calculations used the projection procedure introduced by Knowles and Werner.[45] These calculations were performed on the active site alone, and the remainder of the protein was not included. We find that inclusion of the remainder of the protein as points charges with the MRCI calculations has little affect on the computed CD (shown later). An active space comprising the four doubly occupied d orbitals of Cu with the Cys_{π}, Cys_{σ} orbitals in addition to the SOMO was chosen. A further two occupied orbitals were also included in the active space, resulting in an active space of 17 electrons in 9 orbitals. CD spectra were computed at the MRCI level using the origin independent (velocity) form, where the

rotational strength for a electronic transition A←0 is given by

$$R_{0A} = \frac{-i}{2\omega}\rho_{0A}.\mathbf{L}_{A0} \tag{1}$$

where ω, ρ and $\mathbf{L}$ represent the excitation energy, momentum and angular momentum, respectively.[46] Graphical representations of the spectra were generated by representing each electronic transition with its associated oscillator or rotary strength as a Gaussian function with a full width at half maximum of 0.1 eV.

Results and Discussion

Tables 1 and 2 show the values for structural parameters describing the structure of the active site resulting from the geometry optimisations for plastocyanin and cucumber basic protein. For plastocyanin, all of the calculations predict a Cu-S_{cys} bond length that is longer than the value given by the crystal structure, with the QM/MM calculations with the SRSC basis set for copper giving a value of 2.15 Å. If the QM is region is increased to include the complete residues of the coordinating ligands the value for the Cu-S_{cys} bond length increases to 2.19 Å. This range of values is consistent with previous work on plastocyanin.[17,20] For cucumber basic protein, the predicted Cu-S_{cys} bond lengths are close to the crystal structure value, with a slightly higher value from LFMM. The Cu-S_{met} bond length is much longer, and the calculations give a range of values with LFMM predicting the largest value for plastocyanin and the shortest value for cucumber basic protein, and the calculated Cu-N_{his} bond lengths are quite close to the crystal structure values. The predicted bond angles from the calculations and the crystal structure are largely consistent with each other. One exception to this are the the S_{cys}-Cu-N_{his} angles in plastocyanin, where the angles from LFMM deviate from the QM/MM calculations and the crystal structure. The dihedral and improper torsion angles are also in good agreement with each other, the largest variation is found for the C_β(cys)-S(cys)-Cu-S(met) dihedral angle. This is perhaps not surprising due to the long, and presumable relatively weak, Cu-S(met) bond.

Figure 2 shows the computed CD spectra based upon the structures of the active site from the crystal structure and geometry optimisation calculations. For plastocyanin, it is striking that none of the computed spectra provides even a qualitative match to the experiment. All of the structures result in positive and negative bands at low energy arising from the ligand field transitions. However, for the optimised structures the relative intensities of these bands is incorrect, with the negative band predicted to be less intense than the positive band. The relative intensity of these bands is given most accurately from the crystal structure. One of the most prominent features in experiment is the intense positive band arising from the SOMO$\leftarrow$Cys$_\pi$ (LMCT) transition. However, all of the calculations predict a negative band for this transition, which is shifted to higher energy for the crystal structure. The optimised structures lead correctly to a negative band for the SOMO$\leftarrow$Cys$_\sigma$ transition, while a small positive band is given by the crystal structure. The positive band at about 3 eV in experiment is not reproduced by any of the calculations and may arise from transitions not associated with the active site.

For cucumber basic protein, there is a better agreement of the computed spectra with experiment. The bands arising from the ligand field transitions are reproduced well by the optimised structures from the QM/MM calculations, particularly from the calculations with the all electron basis set. For the crystal structure and LFMM structure, the positive ligand field band is not predicted. The experimental spectrum for cucumber basic protein has a less intense positive SOMO$\leftarrow$Cys$_\pi$ band, and most of the calculations do predict a weak positive feature for this transition, with the exception of the QM/MM with SRSC basis set where this transition is predicted to have a very small rotary strength. All of the calculations lead to a negative band for the SOMO$\leftarrow$Cys$_\sigma$ transition, although for the calculated structures this band is too intense compared to the ligand field bands.

The calculated CD spectra for the different optimised structures demonstrates that for even

structures that are quite similar, significant differences in the resulting CD spectra are evident. To identify what features in the computed CD are sensitive to which structural parameter, Figure 3 shows the computed CD spectra as different structural parameters are modified. For these calculations the bond lengths and angles are increased by 0.03 Å and 10-15°, respectively in the structure given by the QM/MM with SRSC basis set calculation. The results show that variation in the copper-ligand bond lengths has little affect on the computed rotary strengths and the only change evident in the spectra is a decrease in the excitation energies for the SOMO←Cys_{π} and SOMO←Cys_{σ} transitions. Variation in the bond angles of the coordinating ligands does not change the excitation energies but does have some affect on the computed rotary strengths. In particular, increasing the $N_{his}CuN_{his}$ bond angle results a significant decrease in the intensity of the Cys_{σ} band while the intensity of the Cys_{π} band is sensitive to the $S_{cys}CuS_{meth}$ bond angle. Furthermore, changes to the bond angles involving the histidine ligands also affects the intensity of the negative band arising from ligand field excitations. Even for these variations in structure, a positive Cys_{π} band is not predicted. However, variation of the dihedral angles does result in a positive LMCT band. In particular, decreasing the magnitude of the N(his37)-N(his87)-Cu-S(cys) dihedral angle (i.e. a smaller negative value) leads to a positive LMCT band. This corresponds to a greater deviation from planarity of the histidine and cysteine ligands and suggests that the value for this angle in the crystal structure of plastocyanin is not accurate and its value should be closer to that given by the crystal structure of cucumber basic protein.

It is unlikely that calculations on a single structure will be sufficient to capture all the features of the CD spectra of these proteins, and averaging over many structures is necessary. This has been demonstrated previously in the literature for much smaller systems.[47,48] Molecular dynamics simulations within a QM/MM framework are computationally demanding and to perform simulations of a sufficiently long timescale is challenging. Consequently, it is of interest to examine the computed spectra resulting from averaging over structures drawn from computationally less expensive molecular dynamics simulations. These simulations sample a single Gibbs energy minimum.

Figures 4 and 5 show computed UV absorption and CD spectra based upon LFMM and classical molecular dynamics simulations, respectively. The spectra were based upon 200 structures for the LFMD simulations and 115 structures for the classical molecular dynamics simulations. The position of the bands in the UV absorption spectra for the LFMM simulations are in excellent agreement with experiment. For both plastocyanin and cucumber basic protein the intensity of the ligand field band is underestimated. However, the increased intensity of the higher energy band arising from the SOMO$\leftarrow$Cys$_\sigma$ transition in cucumber basic protein compared to plastocyanin in described well. The computed CD spectra also reproduce the overall shape of the experimental spectra well with a positive LMCT band. Furthermore, the reduced intensity of the SOMO$\leftarrow$Cys$_\pi$ band in cucumber basic protein is predicted. Closer inspection shows that in plastocyanin the intensity of the SOMO$\leftarrow$Cys$_\pi$ band remains underestimated and the positive ligand field band in cucumber basic protein is not reproduced. The spectra computed based upon the classical MD simulations show a significantly poorer agreement with experiment. In particular, in the increased intensity of the SOMO$\leftarrow$Cys$_\pi$ band in the UV absorption of cucumber basic protein is not evident and the characteristic shape of the CD spectra is not reproduced well, providing evidence that the LFMM force field gives an accurate description of dynamics which provides a cost effective alternative to DFT based QM/MM simulations.

Figure 6 illustrates the evolution of the computed spectrum for plastocyanin as a different number of structures is included in the conformational averaging. These structures have been extracted from the initial part of the trajectory, but also from the complete trajectory with a larger time step in between. The fewest number of structures shown is 5, and in both cases there the resulting spectra do not agree qualitatively with experiment. However for 20 structures, the characteristic shape of the experimental spectrum begins to emerge. With the inclusion of additional structures, the computed spectrum becomes in better agreement with experiment. The primary discrepancy with experiment is at 3 eV, where the positive feature observed in experiment is not reproduced by the calculations. One possible reason for this discrepancy could be the finite active site model used

in the calculations of the transition energies and rotational strengths. Considering the remainder of the protein, the lowest energy transitions are associated with the aromatic side chains, but these lie significantly higher in energy at approximately 5 eV.[49] Despite this difference in energy, the work of Hsu and Woody[50] has suggested that the rotational strength of the Soret band, which arises from $\pi\pi^*$ transitions on the heme group, in myoglobin and haemoglobin is affected by coupling to the aromatic side chain transitions on neighbouring aromatic side chain residues. While it may be expected that the heme $\pi\pi^*$ transition would be coupled more strongly to the aromatic side chain $\pi\pi^*$ transitions than the type of metal-ligand transitions relevant here, it cannot be discounted. However, to include such transitions within the MRCI calculations is not possible. Furthermore, the extended protein environment can also directly influence the computed transition properties. This can be incorporated into the MRCI calculations through the inclusion of point charges, which can be assigned from the CHARMM22 force field. Figure 7 compares the spectra computed with and without point charges to model the extended protein environment based upon the classical MD simulation. These spectra are very similar, particularly for the ligand field and ligand to metal charge transfer band, indicating that this is not a critical effect for these transitions. However, for the higher energy band there is more of an effect.

Conclusions

The structure and CD spectra of plastocyanin and cucumber basic protein using DFT within a QM/MM framework and the LFMM approach, with MRCI used to determine the excitation energies and transition strength has been investigated. Structural optimisation using the different methods led to active site structures that are largely consistent with the crystal structure except for the length of the Cu-S_{cys} bond which is predicted to be longer than the crystal structure value for all of the methods. The computed CD spectra using the crystal structure and single, optimised structures for the active site are in poor agreement with experiment. In particular, the intensity of the positive LMCT band is predicted to be negative or with very small intensity. It is shown that the intensity

of this band is sensitive to the N(his37)-N(his87)-Cu-S(cys) dihedral angle, and a greater deviation from planarity of the histidine and cysteine ligands results in an increasingly positive band. In contrast, UV absorption and CD spectra computed by averaging over many structures drawn from a LFMD simulation are in good agreement with experiment. The increased intensity of the SOMO$\leftarrow$Cys$_\sigma$ transition in the UV absorption spectrum of cucumber basic protein and the major features of the CD spectra are reproduced. These spectra are closer to experiment than equivalent calculations using structures taken from a classical molecular dynamics simulations providing evidence that LFMM provides a more accurate description of the molecular dynamics and represents a cost effective alternative to DFT-based QM/MM approaches. More generally, if quantum chemical calculations of the transition properties can be performed with sufficient accuracy, comparing CD spectra computed in this way with experiment provides a mechanism for assessing the accuracy of force fields.

Acknowledgments

We thank the University of Nottingham High Performance Computing facility for providing computer recourses, the Engineering and Physical Sciences Research Council (EPSRC) for funding (Grant No. EP/I012303) and the Chemical Computing Group for software support.

Supporting Information Available

Full citations for references 35, 39 and 44 are provided in the Supporting Information. This material is available free of charge via the Internet at http://pubs.acs.org

References

(1) Gray, H. B.; Solomon, E. I. *Copper Proteins*; Ed: Spiro, T. G.; Wiley, New York, **1981**.

(2) Ryde, U; Olsson, M. H. M.; Pierloot, K.; Roos, B. O. The Cupric Geometry of Blue Copper Proteins Is Not Strained. *J. Mol. Biol.*, **1996**, *261*, 586-596.

(3) Shepard, W. E. B.; Anderson, B. F.; Lewandoski, D. A.; Norris, G. E.; Baker, E. N. Copper Coordination Geometry in Azurin Undergoes Minimal Change on Reduction of Copper(II) to Copper(I). *J. Am. Chem. Soc.*, **1990**, *112*, 7817-7819.

(4) Adman, T. Copper Protein Structures. Adv. Prot. Chem., **1991**, *42*, 145-197.

(5) Marcus, R. A.; Sutin, N. Electron Transfers in Chemistry and Biology. *Biochim. Biophys. Acta*, **1985**; *811*; 265-322.

(6) LaCroix, L. B.; Shadle, S. E.; Wang, Y.; Averill, B. A.; Hedman, B.; Hodgson, K. O.; Solomon, E. I. Electronic Structure of the Perturbed Blue Copper Site in Nitrite Reductase: Spectroscopic Properties, Bonding, and Implications for the Entatic/Rack State. *J. Am. Chem. Soc.*, **1996**, *118*, 7755-7768.

(7) LaCroix, L. B.; Randall, D. W.; Nerissian, A. M.; Hoitnik, C. W. G.; Canters, G. W.; Valentine, J. S.; Solomon, E. I. Spectroscopic and Geometric Variations in Perturbed Blue Copper Centers: Electronic Structures of Stellacyanin and Cucumber Basic Protein. *J. Am. Chem. Soc.*, **1998**, *120*, 9621-9631.

(8) Solomon, E. I.; Szilagyi, R. K.; George, S. D.; Basumallick, L. Electronic Structures of Metal Sites in Proteins and Models: Contributions to Function in Blue Copper Proteins. *Chem. Rev.*, **2004**, *104*, 419-458.

(9) Solomon, E. I. Spectroscopic Methods in Bioinorganic Chemistry: Blue to Green to Red Copper Sites. *Inorg. Chem.*, **2006**, *45*, 8012-8025.

(10) Han, J.; Loehr, T. M.; Lu, Y.; Valentine, J. S.; Averill, B. A.; Sanders-Loehr, J. Resonance Raman Excitation Profiles Indicate Multiple Cys→Cu Charge Transfer Transitions in Type 1 Copper Proteins. *J. Am. Chem. Soc.*, **1993**, *115*, 4256-4263.

(11) Penfield, K. W.; Gewirth; A. A.; Solomon, E. I. Electronic Structure and Bonding of The Blue Copper Site in Plastocyanin. *J. Am. Chem. Soc.*, **1985**, *107*, 4519-4529.

(12) Gewirth, A. A.; Solomon, E. I. Electronic Structure of Plastocyanin: Excited State Spectral Features. *J. Am. Chem. Soc.*, **1988**, *110*, 3811-3819.

(13) Larsson, S.; Broo, A.; Sjoelin, L. Connection between Structure, Electronic Spectrum, and Electron-Transfer Properties of Blue Copper Proteins. *J. Phys. Chem.*, **1995**, *99*, 4860-4865.

(14) Pierloot, K.; De Kerpel, J. O. A.; Ryde, U.; Roos, B. O. Theoretical Study of the Electronic Spectrum of Plastocyanin. *J. Am. Chem. Soc.*, **1997**, *119*, 218-226.

(15) Ungar, L. W.; Scherer, N. F.; Voth, G. A. Classical Molecular Dynamics Simulation of the Photoinduced Electron Transfer Dynamics of Plastocyanin. *Biophys. J.*, **1997**, *72*, 5-17.

(16) Pierloot, K.; De Kerpel, J. O. A.; Ryde, U.; Olsson, M. H. M.; Roos, B. O. Relation between the Structure and Spectroscopic Properties of Blue Copper Proteins. *J. Am. Chem. Soc.*, **1998**, *120*, 13156-13166.

(17) Ryde, U.; Olsson, M. H. M. Structure, Strain, and Reorganization Energy of Blue Copper Models in the Protein. *Int. J. Quantum Chem.*, **2001**, *81*, 335-347.

(18) Olsson, M. H. M.; Hong, G.; Warshel, A. Frozen Density Functional Free Energy Simulations of Redox Proteins: Computational Studies of the Reduction Potential of Plastocyanin and Rusticyanin *J. Am. Chem. Soc.*, **2003**, *125*, 5025-5039.

(19) Ando, K. Excited State Potentials and Ligand Force Field of a Blue Copper Protein Plastocyanin. *J. Phys. Chem. B*, **2004**, *108*, 3940-3946.

(20) Sinnecker, S.; Neese, F. QM/MM Calculations With DFT for Taking Into Account Protein Effects on the EPR and Optical Spectra of Metalloproteins. Plastocyanin as a Case Study *J. Comput. Chem.*, **2006**, *27*, 1463-1475.

(21) Deeth, R. J. A Test of Ligand Field Molecular Mechanics as an Efficient Alternative to QM/MM for Modelling Metalloproteins: the Structures of Oxidised Type I Copper Centres. *Chem. Comm.*, **2006**, 2551-2553.

(22) Deeth, R. J. Comprehensive Molecular Mechanics Model for Oxidized Type I Copper Proteins: Active Site Structures, Strain Energies, and Entatic Bulging. *Inorg. Chem.*, **2007**, *46*, 4492-4503.

(23) Ryde, U. ; Nilsson, K. Quantum Refinement - a Combination of Quantum Chemistry and Protein Crystallography. *J. Mol. Struc. - Theochem*, **2003**, *632*, 259-275.

(24) Cascella, M.; Cuendet, M. A.; Tavernelli, I.; Rothlisberger, U. Optical Spectra of Cu(II)-Azurin by Hybrid TDDFT-Molecular Dynamics Simulations. *J. Phys. Chem. B*, **2007**, *111*, 10248-10252.

(25) Vancoillie, S.; Pierloot, K. Multiconfigurational g Tensor Calculations as a Probe for the Covalency of the Copper-Ligand Bonds in Copper(II) Complexes: $[CuCl_4]^{2-}$, $[Cu(NH_3)_4]^{2+}$, and Plastocyanin *J. Phys. Chem. A*, **2008**, *112*, 4011-4019.

(26) Ando, K. Ligand-to-Metal Charge-Transfer Dynamics in a Blue Copper Protein Plastocyanin: A Molecular Dynamics Study. *J. Phys. Chem. B*, **2008**, *112*, 250-256.

(27) Zhekova, H. R.; Seth, M. ; Ziegler, T. A Magnetic and Electronic Circular Dichroism Study of Azurin, Plastocyanin, Cucumber Basic Protein, and Nitrite Reductase Based on Time-Dependent Density Functional Theory Calculations. *J. Phys. Chem. A*, **2010**, *114*, 6308-6321.

(28) Monari, A.; Very, T.; Rivail, J.-L; Assfeld, X. Effects of Mutations on the Absorption Spectra of Copper Proteins: a QM/MM Study. *Theor. Chem. Acc.*, **2012**, *131*, 1221.

(29) Robinson, D.; Besley, N. A. Modelling the Spectroscopy and Dynamics of Plastocyanin. *Phys. Chem. Chem. Phys.*, **2010**, *12*, 9667-9676.

(30) Guckert, J. A.; Lowery, M. D.; Solomon, E. I. Electronic-Structure of the Reduced Blue Copper Active-Site - Contributions to Reduction Potentials and Geometry. *J. Am. Chem. Soc.*, **1995**, *117*, 2817-2844.

(31) Sunahori, F. X.; Zhang, X.; Clouthier, D. J. The Electronic Spectrum of Jet-Cooled Copper Hydrosulfide (CuSH). *J. Chem. Phys.*, **2006**, *125*, 084310.

(32) Do, H.; Besley, N. A. Theoretical Study of the Electronic Spectra of Small Molecules That Incorporate Analogues of the Copper-Cysteine Bond. *J. Phys. Chem. A*, **2012**, *116*, 8507-8514.

(33) Ryde, U. Combined Quantum and Molecular Mechanics Calculations on Metalloproteins. *Curr. Opin. Chem. Biol.*, **2003**, *7*, 136-142.

(34) Hamprecht, F. A.; Cohen, A. J.; Tozer, D. J.; Handy, N. C. Development and Assessment of New Exchange-Correlation Functionals. *J. Chem. Phys.*, **1998**, *109*, 6264-6271.

(35) Brooks, B. R.; Brooks, C. L. III; Mackerell, A. D.; Nilsson, L. ; Petrella, R. J.; Roux, B.; Won, Y.; G. Archontis, G.; Bartels, C.; Boresch, S. et al. CHARMM: The Biomolecular Simulation Program. *J. Comp. Chem.*, **2009**, *30*, 1545-1614.

(36) McLean, A. D.; Chandler, G. S. Contracted Gaussian-Basis sets for Molecular Calculations .1. 2nd row atoms, Z=11-18. *J. Chem. Phys.*, **1980**, *72*, 5639-5648.

(37) Krishnan, R.; Binkley, J. S.; Seeger, R.; Pople, J. A. Contribution of Triple Substitutions to the Electron Correlation-Energy in 4^{th} Order Perturbation-Theory. *J. Chem. Phys.*, **1980**, *72*, 4244-4245.

(38) Kaupp, M., Schleyer, P. V. R., Stoll, H.; Preuss, H. Pseudopotential Approaches to Ca, Sr, and Ba Hydrides - Why are Some Alkaline-Earth MX2 Compounds Bent? *J. Chem. Phys.*, **1991**, *94*, 1360-1366.

(39) Shao, Y.; Molnar, L. F.; Jung, Y.; Kussmann, J.; Ochsenfeld, C.; Brown, S. T.; Gilbert, A. T. B.; Slipchenko, L. V.; Levchenko, S. V.; O'Neill, D. P. et al. Advances in Methods and Algorithms in a Modern Quantum Chemistry Program Package. *Phys. Chem. Chem. Phys.*, **2006**, *8*, 3172-3191.

(40) Woodcock, H. L. III; Hodoscek, M.; Gilbert, A. T. B.; Gill, P. M. W.; Schaefer, H. E. III; Brooks, B. R. Interfacing Q-chem and CHARMM to Perform QM/MM Reaction Path Calculations. *J. Comput. Chem.*, **2007**, *28*, 1485-1502.

(41) Tai, H.-C.; Brodbeck, R.; Kasparkova, J.; Farrer, N. J.; Brabec, V.; Sadler, P. J.; Deeth, R. J. Combined Theoretical and Computational Study of Interstrand DNA Guanine-Guanine Cross-Linking by *trans*-[Pt(pyridine)$_2$] Derived from the Photoactivated Prodrug *trans*,*trans*,*trans*-[Pt(N$_3$)$_2$(OH)$_2$(pyridine)$_2$]. *Inorg. Chem.*, **2012**, *51*, 6830-6841.

(42) Deeth, R. J.; Fey, N.; Williams-Hubbard, B. J. DommiMOE: An Implementation of Ligand Field Molecular Mechanics in the Molecular Operating Environment. *J. Comput. Chem.*, **2005**, *26*, 123-130.

(43) MOE version 2011.10, Chemical Computing Group: Montreal, 2011.

(44) MOLPRO, version 2006.1, a package of ab initio programs, Werner, H.-J. P.; Knowles, P. J.; Lindh, R.; Manby, F. R.; Schütz, M.; Celani, P.; Korona, T.; Rauhut, G.; Amos, R.D.; Bernhardsson, A. et al.

(45) Knowles, P. J.; Werner, H. -J. Internally Contracted Multiconfiguration-Reference Configuration-Interaction Calculations for Excited-States. *Theor. Chim. Acta*, **1992**, *84*, 95-103.

(46) Besley, N. A; Brienne, M. -J.; Hirst, J. D. Electronic Structure of a Rigid Cyclic Diamide. *J. Phys. Chem. B*, **2000**, *104*, 12371-12377.

(47) Frelek, J.; Kowalska, P.; Masnyk, M.; Kazimierski, A.; Korda, A.; Woznica, M.; Chmielewski, M.; Furche, F. Circular Dichroism and Conformational Dynamics of Cephams and Their Carba and Oxa Analogues. *Chem. Eur. J.*, **2007**, *13*, 6732-6744.

(48) Gaigeot, M.-P.; Besley, N. A; Hirst, J. D. Modelling the Infrared and Circular Dichroism Spectroscopy of a Bridged Cyclic Diamide. *J. Phys. Chem. B*, **2011**, *115*, 5526-5535.

(49) Rogers, D. M.; Hirst, J. D. Ab Initio Study of Aromatic Side Chains of Amino Acids in Gas Phase and Solution. *J. Phys. Chem. A*, **2003**, *107*, 11191-11200.

(50) Hsu, M. C.; Woody, R. W. The Origin of the Heme Cotton Effects in Myoglobin and Hemoglobib. *J. Am. Chem. Soc.*, **1971**, *93*, 3515-3525.

Table 1: Comparison of the structural parameters for the active site of plastocyanin.

	Cystal 1PLC	QM/MM 6-31G* for Cu	QM/MM SRSC for Cu	LFMM
Distances (Å)				
Cu-S(cys)	2.067	2.131	2.149	2.140
Cu-N(his87)	2.058	1.934	1.972	2.020
Cu-N(his37)	1.907	1.963	1.980	2.010
Cu-S(met)	2.822	2.904	2.723	3.050
Angles (°)				
S(cys)-Cu-N(his87)	121.0	125.8	124.3	141.0
S(cys)-Cu-N(his37)	131.7	127.2	125.5	119.4
N(his87)-Cu-N(his37)	97.2	98.2	92.2	97.0
S(cys)-Cu-S(met)	109.9	105.2	113.0	101.8
Dihedral angles (°)				
C_β(cys)-S(cys)-Cu-S(met)	-3.2	-24.0	-23.9	-30.3
C_γ(his87)-N(his87)-Cu-S(met)	77.0	71.7	74.5	56.6
C_γ(his37)-N(his37)-Cu-S(met)	-111.8	-108.3	-111.8	-113.6
C_γ(his87)-N(his87)-Cu-S(cys)	-43.8	-53.1	-54.7	-68.5
N(his37)-N(his87)-Cu-S(cys)	-143.8	-148.4	-136.1	-165.6
Improper torsions (°)				
Cu-S(cys)-N(his87)-N(his37)	-21.7	-30.9	-20.9	-15.2
Cu-S(cys)-N(his37)-N(his87)	22.0	31.5	26.6	10.0
Cu-N(his87)-N(his37)-S(cys)	15.5	21.1	16.1	7.8

Table 2: Comparison of the structural parameters for the active site of cucumber basis protein.

	Cystal 2CBP	QM/MM 6-31G* for Cu	QM/MM SRSC for Cu	LFMM
Distances (Å)				
Cu-S(cys)	2.155	2.153	2.153	2.187
Cu-N(his87)	1.951	1.955	1.938	2.027
Cu-N(his37)	1.931	1.955	1.943	2.029
Cu-S(met)	2.606	2.672	2.673	2.489
Angles (°)				
S(cys)-Cu-N(his87)	110.4	114.3	117.6	106.1
S(cys)-Cu-N(his37)	138.1	132.5	130.6	137.1
N(his87)-Cu-N(his37)	99.3	99.0	95.2	96.2
S(cys)-Cu-S(met)	110.6	111.2	111.2	120.5
Dihedral angles (°)				
C_β(cys)-S(cys)-Cu-S(met)	-6.3	-15.5	-15.5	-13.4
C_γ(his87)-N(his87)-Cu-S(met)	78.0	68.4	71.2	50.6
C_γ(his37)-N(his37)-Cu-S(met)	-114.9	-117.4	-111.3	-116.5
C_γ(his87)-N(his87)-Cu-S(cys)	45.9	57.1	54.4	50.6
N(his37)-N(his87)-Cu-S(cys)	-135.4	-135.2	-133.7	-121.8
Improper torsions (°)				
Cu-S(cys)-N(his87)-N(his37)	-19.6	-23.0	-26.3	-24.5
Cu-S(cys)-N(his37)-N(his87)	32.6	31.8	33.3	43.0
Cu-N(his87)-N(his37)-S(cys)	18.2	20.0	21.0	22.8

Figure Captions

Figure 1: Molecular orbitals involved in the most intense electronic transitions in plastocyanin and cucumber basic protein.

Figure 2: Calculated CD spectra of plastocyanin and cucumber basic protein based upon the optimised structures. (a) - crystal structure, (b) - SRSC basis set for Cu, (c) - 6-31G* basis set for Cu, (d) - LFMM. Experimental data from references [5,6].

Figure 3: Variation of the computed CD spectra from arising from modification of the structure of the active site. Black line - unmodified structure, red line - modified structure.

Figure 4: Computed UV absorption and CD spectra for plastocyanin and cucumber basic protein using structures from a LFMM molecular dynamics simulation. Black line - experiment, red line - calculation. Experimental data from references [5,6].

Figure 5: Computed UV absorption and CD spectra for plastocyanin and cucumber basic protein using structures from a classical molecular dynamics simulation. Black line - experiment, red line - calculation. Experimental data from references [5,6].

Figure 6: Convergence of the computed CD spectra for plastocyanin using structures from the LFMD simulation. Black line - full 200 structures, red line - (a) first 5 structures, (b) first 20 structures, (c) first 40 structures, (d) first 100 structures, (e) every 40^{th} structure, (f) every 10^{th} structure, (g) every 5^{th} structure, (h) every 2^{nd} structure.

Figure 7: Computed CD spectra for plastocyanin using structures from a classical molecular dynamics simulation. Black line - no points charges, red line - including point charges.

Figure 1: Molecular orbitals involved in the most intense electronic transitions in plastocyanin and cucumber basic protein.

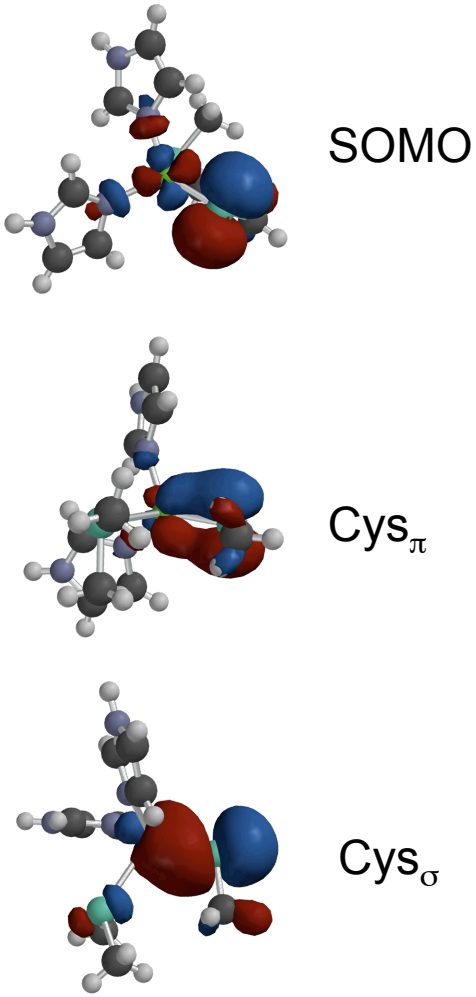

Figure 2: Calculated CD spectra of plastocyanin and cucumber basic protein based upon the optimised structures. (a) - crystal structure, (b) - SRSC basis set for Cu, (c) - 6-31G* basis set for Cu, (d) - LFMM. Experimental data from references [5,6].

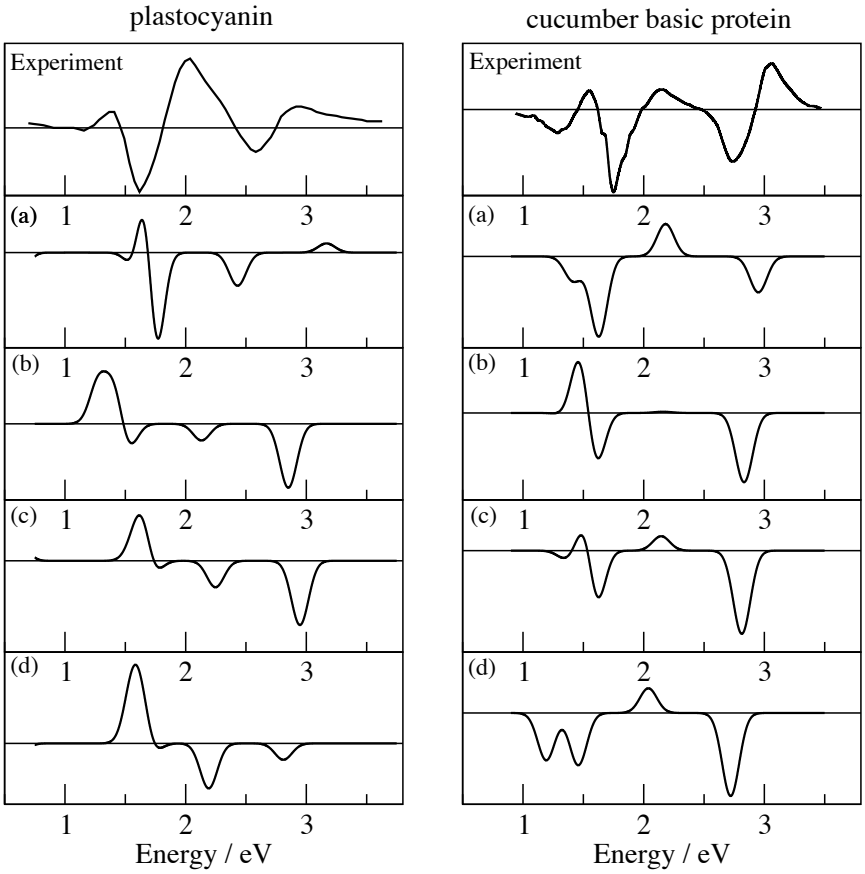

Figure 3: Variation of the computed CD spectra from arising from modification of the structure of the active site. Black line - unmodified structure, red line - modified structure.

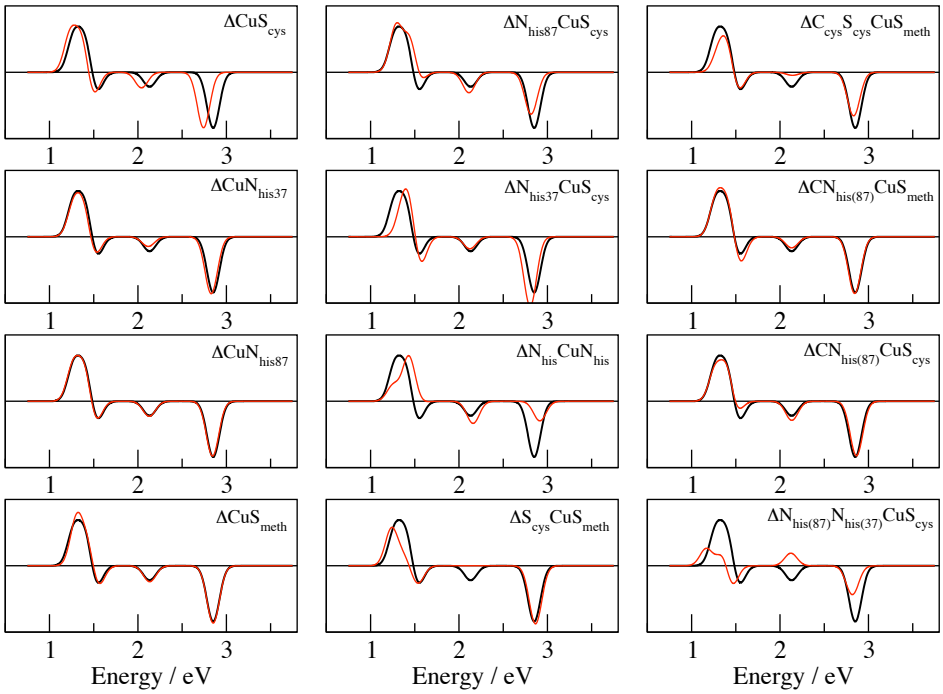

Figure 4: Computed UV absorption and CD spectra for plastocyanin and cucumber basic protein using structures from a LFMM molecular dynamics simulation. Black line - experiment, red line - calculation. Experimental data from references [5,6].

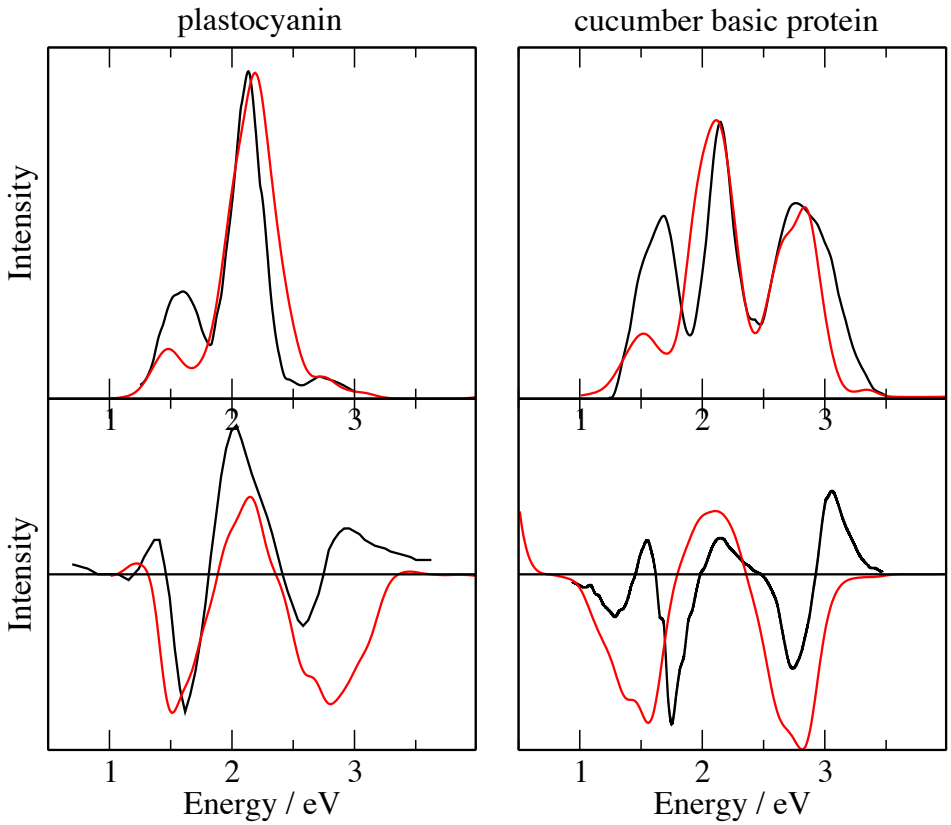

Figure 5: Computed UV absorption and CD spectra for plastocyanin and cucumber basic protein using structures from a classical molecular dynamics simulation. Black line - experiment, red line - calculation. Experimental data from references [5,6].

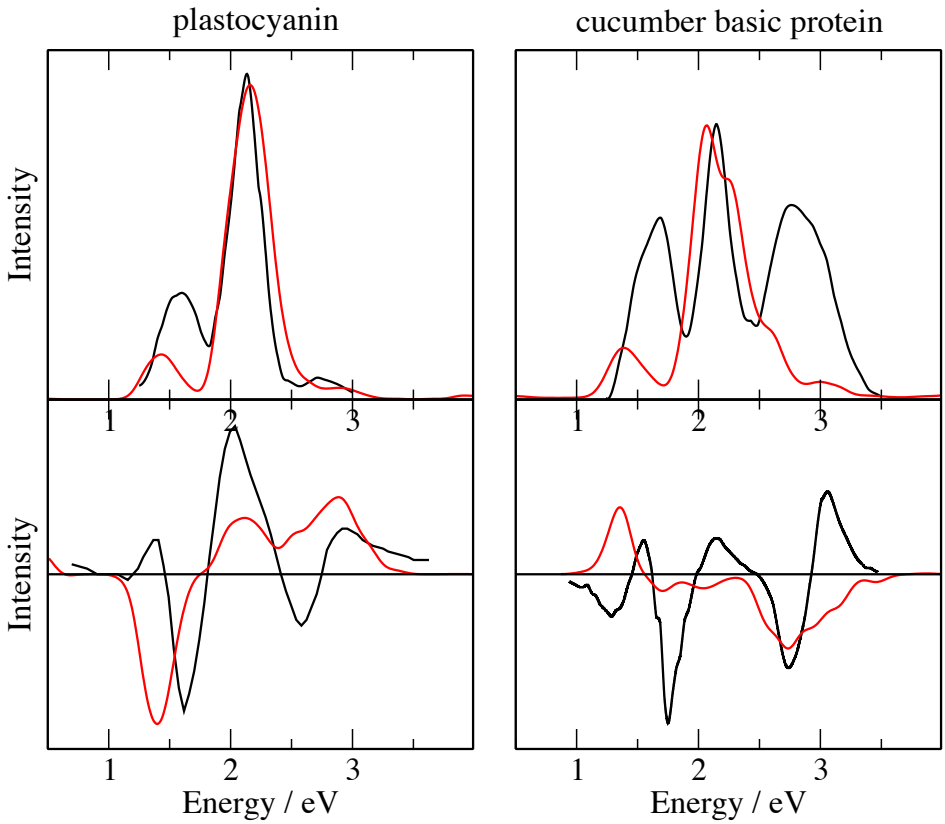

Figure 6: Convergence of the computed CD spectra for plastocyanin using structures from the LFMD simulation. Black line - full 200 structures, red line - (a) first 5 structures, (b) first 20 structures, (c) first 40 structures, (d) first 100 structures, (e) every 40th structure, (f) every 10th structure, (g) every 5th structure, (h) every 2nd structure.

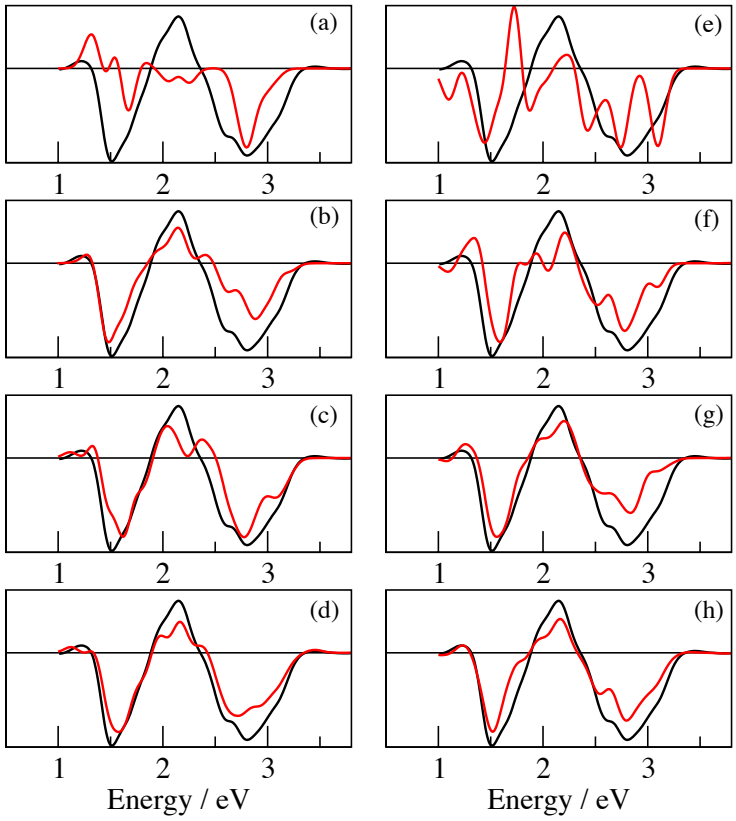

Figure 7: Computed CD spectra for plastocyanin using structures from a classical molecular dynamics simulation. Black line - no points charges, red line - including point charges.

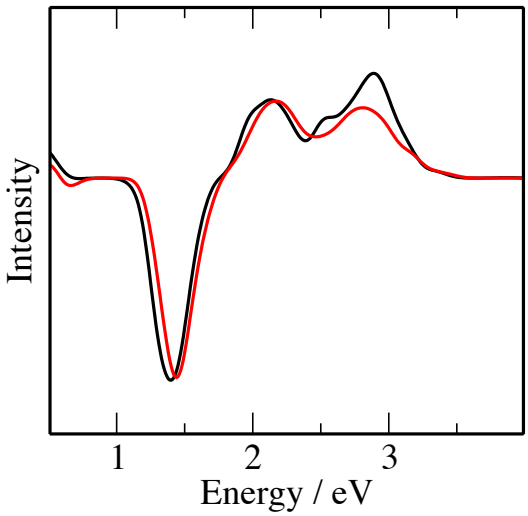

Table of Contents Graphic