Theoretical Study of the Electronic Spectrum of Plastocyanin

Kristine Pierloot,*,† Jan O. A. De Kerpel,† Ulf Ryde,‡ and Björn O. Roos‡

Contribution from the Department of Chemistry, University of Leuven, Celestijnenlaan 200F, B-3001 Heverlee-Leuven, Belgium, and Department of Theoretical Chemistry, University of Lund, Chemical Center, P.O.B. 124, S-221 00 Lund, Sweden

Abstract: The electronic spectrum of the blue copper protein plastocyanin has been studied by ab initio multiconfigurational second-order perturbation theory (the CASPT2 method). The six lowest electronic transitions have been calculated and assigned with an error of less than 2000 cm^{-1}. The singly occupied orbital in the ground state is Cu 3d–S$_{Cys}$ 3pπ antibonding with some N$_{His}$ 2pσ character. The bright blue color originates from an electron transfer to this orbital from the corresponding Cu 3d–S$_{Cys}$ 3pπ bonding orbital. The influence of different ligand models on the spectrum has been thoroughly studied; Cu(imidazole)$_2$(SCH$_3$)(S(CH$_3$)$_2$)$^+$ as a model of CuHis$_2$CysMet is the smallest system that gives converged results. The spectrum is surprisingly sensitive to changes in the geometry, especially in the Cu–S bond distances; a 5 pm change in the Cu–S$_{Cys}$ bond length may change the excitation energies by as much as 2000 cm^{-1}. The effect of the surrounding protein and solvent on the transition energies has been modeled by point charges and is found to be significant for some of the transitions (up to 2000 cm^{-1}).

1. Introduction

Blue or type 1 copper proteins have long attracted chemists' interest by their intense blue color, distinctive electron paramagnetic resonance spectra, and unusually high reduction potentials.[1,2] A typical example is plastocyanin, a protein that plays a key role in photosynthesis, being a mobile electron shuttle between reduced cytochrome *f* in the cytochrome *b$_6f$* complex and pigment *P*700$^+$ of photosystem I in the photosynthetic electron transport chain of higher plants and algae. High-resolution crystal structures are known for plastocyanin[3–5] and several other proteins of this group, native as well as engineered.[2] Although Cu^{2+} normally forms tetragonal complexes, the crystal structure of blue copper proteins invariably shows a trigonal copper ion bound in a distorted plane formed by a cysteine thiolate group at an unusually short distance (207–220 pm) and two histidine nitrogen atoms at normal distances (190–210 pm). In addition, a methionine sulfur atom, and in some proteins also a backbone amide oxygen, binds in an axial position at a large distance (260–330 pm).[2–5]

The electronic spectra of several blue copper proteins are well-characterized and have been recorded by visual and near-infrared absorption spectroscopy, circular dichroism, and magnetic circular dichroism.[1,6–8] The spectra show a prominent peak around 16 700 cm^{-1} (600 nm) that gives rise to the blue color, and a weaker band around 12 000 cm^{-1} (830 nm). In addition, some proteins, e.g. stellacyanin and nitrite reductase, have a third peak near 22 000 cm^{-1} (450 nm).[6,9] By a Gaussian resolution of the spectrum, additional bands can be discerned; for plastocyanin, in total nine different absorption bands have been reported.[6,8] The dominant blue band in the protein spectrum is considerably more intense than absorption bands of normal inorganic copper complexes; the molar extinction coefficient, ϵ, amounts to 1000–6000 M^{-1} cm^{-1} for blue copper proteins, compared to $\epsilon < 100$ M^{-1} cm^{-1} for normal tetragonal copper complexes.[8]

Solomon and co-workers have studied the electronic spectrum of plastocyanin using the density functional Xα scattered-wave method.[8,10] They obtain a ground state with a singly occupied orbital that is an antibonding mixture of a Cu 3d orbital (42%) and a S$_{Cys}$ 3pπ orbital (36%).[8] The four bands with the lowest excitation energies are attributed to copper ligand-field transitions and the dominant blue band is assigned to a charge-transfer transition from the bonding S$_{Cys}$ pπ orbital (at 16 940 cm^{-1}). In addition, a prominent band at 18 700 cm^{-1}, hidden in the dominant blue band, is assigned to a transition from a so-called S$_{Cys}$ pseudo-σ orbital. Although the error in the predicted transitions is small for the four central bands, including the dominant blue band (less than 2000 cm^{-1}), the calculations are more approximate for the low- and high-energy bands, with errors in some cases exceeding 10 000 cm^{-1}.

Larsson *et al.* have recently published a calculation of the spectra of azurin and some blue copper protein mutants, using the semiempirical CNDO/S method.[11] They also obtain a ground state with a singly occupied antibonding Cu 3d–S$_{Cys}$ 3pπ orbital, now with 32% Cu and 50% S$_{Cys}$ character. The most intense transition is calculated around 13 800 cm^{-1} and corresponds to an excitation from the bonding Cu–S$_{Cys}$ pπ orbital. In variance with Solomon, they predict that three of the ligand-field transitions should have higher excitation energies than the dominant blue band.

Both CNDO/S and Xα are semiempirical methods that depend on a number of adjustable parameters, and they do not give a consistent description of the electronic structure and the

† University of Leuven.
‡ University of Lund.
⊗ Abstract published in *Advance ACS Abstracts,* December 1, 1996.
(1) Sykes, A. G. *Adv. Inorg. Chem.* **1990,** *36,* 377–408.
(2) Adman, E. T. *Adv. Protein Chem.* **1991,** *42,* 145–197.
(3) Colman, P. M.; Freeman, H. C.; Guss, J. M.; Murata, M.; Norris, V. A.; Ramshaw, J. A. M.; Venkatappa, M. P. *Nature* **1978,** *272,* 319–324.
(4) Guss, J. M.; Freeman, H. C. *J. Mol. Biol.* **1983,** *169,* 521–563.
(5) Guss, J. M.; Bartunik, H. D.; Freeman, H. C. *Acta Crystallogr.* **1992,** *B48,* 790–807.
(6) Solomon, E. I.; Hare, J. W.; Dooley, D. M.; Dawson, J. H.; Stephens, P. J.; Gray, H. B. *J. Am. Chem. Soc.* **1980,** *102,* 168.
(7) Penfield, K. W.; Gay, R. R.; Himmelwright, R. S.; Eickman, N. C.; Norris, V. A.; Freeman, H. C.; Solomon, E. I. *J. Am. Chem. Soc.* **1981,** *103,* 4382.
(8) Gewirth, A. A.; Solomon, E. I. *J. Am. Chem. Soc.* **1988,** *110,* 3811.

(9) Liu, M.-Y.; Liu, M.-C.; Payne, W. J.; Legall, J. *J. Bacteriol.* **1986,** *166,* 604–608.
(10) Penfield, K. W.; Gewirth, A. A.; Solomon, E. I. *J. Am. Chem. Soc.* **1985,** *107,* 4519.
(11) Larsson, S.; Broo, A.; Sjölin, L. *J. Phys. Chem.* **1995,** *99,* 4860.

Table 1. Comparison of the Optimized Structure of the Different Model Complexes with the Crystal Structures of Blue Copper Proteins

complex	distance to Cu (pm)			angle subtended at Cu (deg)			
	S_{Met}	S_{Cys}	N	$S_{Met}-S_{Cys}$	$S_{Met}-N$	$S_{Cys}-N$	N−N
$Cu(NH_3)_2(SH)(SH_2)^+$	284	217	207	110	93	124	103
$Cu(NH_3)_2(SCH_3)(SH_2)^+$	288	217	208	113	92	123	104
$Cu(NH_3)_2(SH)[S(CH_3)_2]^+$	258	218	209	114	98	121	102
$Cu(NH_3)_2(SCH_3)[S(CH_3)_2]^+$	258	218	211	121	96	119	103
$Cu(NH_3)_2(SC_2H_5)[S(CH_3)_2]^+$	258	218	211	121	96	119	103
$Cu(imidazole)_2(SH)(SH_2)^+$	291	217	202	108	93	124	106
$Cu(imidazole)_2(SCH_3)(SH_2)^+$	297	217	203	112	91	123	108
$Cu(imidazole)_2(SH)[S(CH_3)_2]^+$	262	219	203	112	97	121	104
$Cu(imidazole)_2(SCH_3)[S(CH_3)_2]^+$	262	218	205	119	95	119	105
$Cu(imidazole)_2(SC_2H_5)[S(CH_3)_2]^+$	262	218	205	119	95	119	105
plastocyanin[a]	282	207	191−206	110	89−101	121−132	97
experimental average[b]	283	212	206	111	98	120	103

[a] The experimental structure of poplar plastocyanin.[5] [b] The average geometry of 18 different blue copper proteins.[22]

spectrum. It would, therefore, be satisfying if a more stringent method could be used. The recently developed ab initio CASPT2 method (multiconfigurational second-order perturbation theory) has shown to be very promising for the interpretation of electronic spectra;[12,13] a large number of chromophores have been investigated with this method, and the error in the transition energies is less than 2000 cm^{-1} with few exceptions. The method has been used in a number of successful studies of ligand-field and charge-transfer spectra of transition-metal complexes.[13] These results have shown that accurate excitation energies can be obtained also for complex systems. We therefore expect to obtain results of the same quality.

This paper presents results from CASPT2 calculations of the electronic spectrum of several realistic models of the active-site copper ion in the blue protein plastocyanin. The sensitivity of the results to the choice to the model system, the geometry, and the basis sets have been investigated, as well as the influence of the surrounding protein on the spectrum. The results are encouraging: the six lowest transitions in the spectrum are predicted with an error of less than 2000 cm^{-1}. A full assignment of the spectrum is provided in the energy range up to 20 000 cm^{-1}.

2. Methods and Details of the Calculations

2.1. Geometry Optimizations. The copper coordination sphere in plastocyanin consists of two histidines, one cysteine, and a more distant methionine ligand. These ligands were modeled in several different ways: histidine was modeled either as ammonia or imidazole, cysteine as SH$^-$, SCH$_3^-$, or SC$_2$H$_5^-$, and methionine as SH$_2$ or S(CH$_3$)$_2$. The structures of the different model complexes were optimized with the hybrid density functional method MB3LYP as implemented in the Mulliken-2.31h software (unrestricted formalism).[14] This method

differs slightly from B3LYP defined in the Gaussian software[15] in that it uses the Ceperley−Alder solution to the uniform electron gas in the Vosko−Wilk−Nusair correlation functional (this is the fit recommended by the authors[16]). Hybrid density functional methods have been shown to give as good or better geometries as correlated ab initio methods for first-row transition-metal complexes.[17−19] The full geometry of the models was optimized until the maximum and root-mean-square gradients were below 22 and 15 J mol^{-1} pm^{-1}, and the maximum and root-mean-square changes in geometry were below 0.079 and 0.053 pm, respectively, using the standard fine grid.[14]

The basis set for copper was the double-ζ basis (62111111/3111/311) of Schäfer et al.,[20] enhanced with diffuse p, d, and f functions with exponents 0.174, 0.132, and 0.390, and 6-31G* for the other atoms.[21] Only the pure 5d and 7f functions were used. Both the basis sets and the grid have been thoroughly tested and shown to give accurate results for similar copper complexes.[22] In order to minimize the computational effort in the calculations of the spectra, C_s symmetry was imposed during the geometry optimizations. However, some calculations using the experimental geometry without symmetry were also performed.

Some geometry parameters of all optimized model systems used in the spectrum calculations are listed in Table 1. The optimized structure of $Cu(imidazole)_2(SCH_3)(S(CH_3)_2)^+$ is shown in Figure 1. All systems are in the conformation most similar to the crystal structure, i.e. with the dihedral angle $S_{Met}-Cu-S_{Cys}-C_{Cys} \approx 0°$. As we have discussed before, the structures optimized in vacuum are very similar to the experimental structure. This indicates that hypotheses suggesting that the protein strains the cupric geometry are not correct.[22]

Throughout the discussion, a coordinate system is defined in the following way: the copper ion is at the origin, the z axis is along the Cu−S$_{Met}$ bond, and the Cu−S$_{Cys}$ bond is situated in the xz plane. Thus, the coordinate system is defined in the same way as made by Larsson et al.,[11] but the x axis is turned 45° around the z axis compared to the coordinate system used by Solomon et al.[8]

2.2. Calculations of the Spectrum. The calculations of the electronic spectra of the model complexes were performed by the CASSCF/CASPT2 method (second-order perturbation theory with a

(12) Andersson, K.; Malmqvist, P.-Å.; Roos, B. O. *J. Chem. Phys.* **1992**, *96*, 1218.

(13) Roos, B. O.; Andersson, K.; Fülscher, M. P.; Malmqvist, P.-Å.; Serrano-Andrés, L.; Pierloot, K.; Merchán, M. In *Advances in Chemical Physics: New Methods in Computational Quantum Mechanics*; Prigogine, I., Rice, S. A., Eds.; John Wiley & Sons, New York, 1996; Vol. XCIII: 219.

(14) Rice, J. E.; Horn, H.; Lengsfiels, B. H.; McLean, A. D.; Carter, J. T.; Replogle, E. S.; Barnes, L. A.; Maluendes, S. A.; Lie, G. C.; Gutwski, M.; Rude, W. E.; Sauer, S. P. A.; Lindh, R.; Andersson, K.; Chevalier, T. S.; Widmark, P.-O.; Bouzida, D.; Pacansky, G.; Singh, K.; Gillan, C. J.; Carnevali, P.; Swope, W. C.; Liu, B. *Mulliken*^{TM} *Version 2.25b, internal release*; IBM Corporation: Almaden, CA, 1995.

(15) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Gill, P. M. W.; Johnson, B. G.; Robb, M. A.; Cheeseman, J. R.; Keith, T.; Petersson, G. A.; Montgomery, J. A.; Raghavachari, K.; Al-Laham, M. A.; Zakrzewski, V. G.; Ortiz, J. V.; Foresman, J. B.; Cioslowski, J.; Stefanov, B. B.; Nanayakkara, A.; Challacombe, M.; Peng, C. Y.; Ayala, P. Y.; Chen, W.; Wong, M. W.; Andres, J. L.; Replogle, E. S.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Binkley, J. S.; Defrees, D. J.; Baker, J.; Stewart, J. P.; Head-Gordon, M.; Gonzalez, C.; Pople, J. A. *Gaussian 94, Revision D.1*; Gaussian, Inc.: Pittsburgh, PA, 1995.

(16) Vosko, S. H.; Wilk, L.; Nusair, M. *J. Chem. Phys.* **1980**, *58*, 1200−1211.

(17) Ricca, A.; Bauschlicher, C. W. *J. Phys. Chem.* **1994**, *98*, 12899−12903.

(18) Ricca, A.; Bauschlicher, C. W. *Theor. Chim. Acta* **1995**, *92*, 123−131.

(19) Holthausen, M. C.; Mohr, M.; Koch, W. *Chem. Phys. Lett.* **1995**, *240*, 245−252.

(20) Schäfer, A.; Horn, H.; Ahlrichs, R. *J. Chem. Phys.* **1992**, *97*, 2571−2577.

(21) Hehre, W. J.; Radom, L.; Schleyer, P. v. R.; Pople, J. A. *Ab initio molecular orbital theory*; Wiley-Interscience: New York, 1986.

(22) Ryde, U.; Olsson, M. H.; Pierloot, K.; Roos, B. O. *J. Mol. Biol.* **1996**, *261*, 586−596.

(23) Andersson, K.; Malmqvist, P.-Å.; Roos, B. O.; Sadlej, A. J.; Wolinski, K. *J. Phys. Chem.* **1990**, *94*, 5483.

(24) Andersson, K.; Roos, B. O. In *Modern electron structure theory*; Yarkony, R., Ed.; Advanced Series in Physical Chemistry; World Scientific Publishing Co. Pte. Ltd.: Singapore, 1995; Vol. 2, Part I:55.
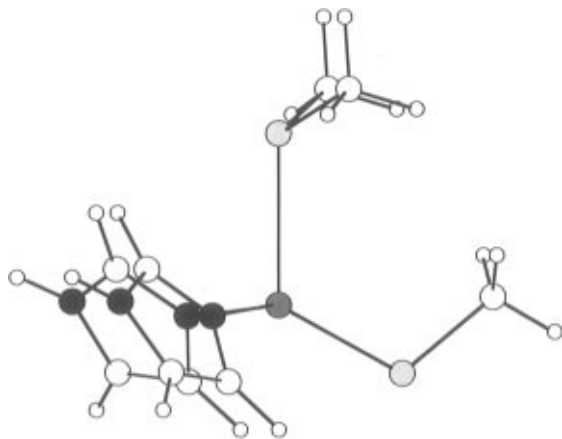
Figure 1. The optimized structure of Cu(imidazole)$_2$(SCH$_3$)(S(CH$_3$)$_2$)$^+$ in C_s symmetry.

multiconfigurational reference wave function).[13,23,24] For Cu(imidazole)$_2$-(SH)(SH$_2$)$^+$ without symmetry, the CASSCF wave function was obtained by distributing 13 electrons in an active space of 12 orbitals, consisting of the five Cu 3d orbitals, five correlating Cu 4d orbitals, and the S$_{Cys}$ 3pσ and π lone-pair orbitals. This active space can be used to describe the ground state and the six lowest excited states in the spectrum. The need for a second d shell in the active space has been discussed in earlier applications to transition-metal spectroscopy.[13]

A larger number of excited states can be studied for complexes in C_s symmetry, by allowing different active spaces for states of different symmetry. For the complexes containing imidazole, 13 electrons were correlated in 12 active orbitals, consisting of Cu 3d and 4d and two ligand orbitals belonging to the symmetry representation of the calculated state. For the complexes containing NH$_3$, only 11 active orbitals were used for the calculation of the A″ states, since the highest d^{2}A″ state is an imidazole π to Cu 3d charge-transfer state that is absent in the NH$_3$ systems. In all cases, only one state-averaged CASSCF calculation was performed for all states belonging to the same symmetry representation (A′ or A″ in C_s, A in C_1) with equal weights for all states included.

In the CASPT2 calculations all electrons originating from S 3s,3p, C 2s,2p, N 2s,2p, H 1s, and Cu 3s,3p,3d were correlated (113 in the largest system). The effect of correlating the semicore Cu 3s,3p electrons was investigated in more detail for Cu(NH$_3$)$_2$(SH)(SH$_2$)$^+$, by comparing the results obtained for the CASPT2 excitation energies both with and without the inclusion of the 3s,3p electrons in the correlation treatment with a large basis set. Relativistic corrections (Darwin and mass–velocity) were added to all CASPT2 excitation energies. They were obtained using first-order perturbation theory at the CASSCF level.

The CASPT2 calculations were performed using a level-shift technique in order to remove intruder states.[25,26] The presence of intruder states is manifested by a sudden drop of the CASSCF reference weight in the final first-order CASPT2 wave function. On the basis of the results of a series of test calculations on Cu(imidazole)$_2$(SH)-(SH$_2$)$^+$,[26] a level shift of 0.3 eV was used in all calculations (in the calculations for the e^{2}A′ state in the two largest models it had to be increased to 0.35). The effect of the level shift on the second-order energy was removed by a back-correction technique (the LS correction).[25]

The CASSCF/CASPT2 calculations were performed using generally contracted atomic natural orbital (ANO) type basis sets.[27] Four basis sets of different sizes were used. They are described in Table 2. Due to the size of the models studied (up to 36 atoms), moderate-size contracted basis sets were employed, including an f-type polarization function on Cu and a d function on S, but no polarization functions on C, N, and H. For the models involving NH$_3$, a [2s] contracted set was used for all hydrogen atoms (basis A), while for the imidazole

Table 2. ANO Basis Sets Used for the Calculation of the Spectra

basis	Cu	S	N	C	H$_a$[a]	H$_b$[a]
primitive	17s12p9d4f	13s10p4d	10s6p3d	10s6p3d	7s3p	7s3p
			contracted to			
A	6s4p3d1f	4s3p1d	3s2p	3s2p	2s	2s
B	6s4p3d1f	4s3p1d	3s2p	3s2p	2s	1s
C	6s5p4d2f	5s4p2d	4s3p1d	4s3p1d	2s1p	2s1p
D[b]	8s7p6d1f	4s3p1d	3s2p	3s2p	2s	2s

[a] H$_a$ = hydrogen bound to S; H$_b$ = hydrogen bound to N,C. [b] With uncontracted s,p,d functions in the Cu core-valence region.

model systems, the size of the hydrogen basis set was further reduced to [1s] for all hydrogens except those directly bound to sulfur (basis B). Basis C consists of considerably larger contracted sets on all atoms and was used to estimate the basis set truncation error. Finally, basis D was constructed with the purpose of providing a more accurate description of the Cu 3s,3p correlation effects, by uncontracting the appropriate primitives in the 3s,3p,3d region.

For the six lowest excited states in the spectra, oscillator strengths were calculated using the CAS State–Interaction method.[28] The transition moments were obtained at the CASSCF level, while CASPT2 excitation energies were used in the expression for the oscillator strength. The CASSCF/CASPT2 calculations were performed with the MOLCAS-3.1 quantum chemistry software,[29] and all calculations were run on IBM RS6000 workstations.

2.3. A Model of the Surrounding Protein and Solvent. A calculation of the spectrum was also performed on Cu(imidazole)$_2$-(SH)(SH$_2$)$^+$ without symmetry, using the experimental geometry of poplar plastocyanin at 133 pm resolution (PDB file 1PLC).[5] The system was truncated by hydrogen atoms placed at standard bond lengths in the direction of the removed carbon atom.

In order to estimate the effect of the surrounding protein, a second set of calculations was performed on this model surrounded by 5232 point charges, representing the remaining protein atoms, all crystal water molecules, and a sphere of solvent water molecules with a radius of 2.4 nm. Each atom was assigned a point charge taken from the Amber 4.1 force field.[30] All Asp and Glu residues were treated as anions, and all Lys and Arg residues as cations. The first and last amino acids of the protein were also charged. The only two His residues in the protein are copper ligands, and it is clear from the crystal structure that the protons are bound to the N$^{\epsilon 2}$ atoms. The only Cys residue is also a copper ligand and was treated as an anion. The charge of atoms directly connected to the junction carbon atoms (the atoms that are converted to hydrogen atoms in the quantum chemical calculations, i.e. C$^\beta$ in Cys and His and C$^\gamma$ in Met) were set to 0, and the charges of the rest of the atoms in the amino acid were uniformly scaled, so that the total residue charge vanishes (a change of less than 0.006 e/atom).

The positions of the solvent water molecules and the protein hydrogen atoms were determined by a 22.5 ps simulated annealing by molecular dynamics (hydrogen atoms are present in the PDB file, but they make several unphysically short bonds; therefore all hydrogen positions were recalculated). The temperature was kept at 300 K during the first 7.5 ps and was then decreased linearly with time to 0 K at the end of the simulation. The time constant for the temperature coupling to a heat bath was kept at 0.2 ps during the first 7.5 ps, it was then increased to 1.0 ps for 7.5 ps, decreased to 0.5 ps for 4.5 ps, and was set to 0.05 ps, during the last 3 ps of the simulation. The simulated annealing was concluded by a 10 000 step molecular mechanics minimization. In all simulations, a nonbounded cutoff of 2.0 nm was applied (nonbonded interactions between 1.5 and 2.0 nm were calculated only once every 25 time steps), the neighbor list was updated every 25 time step (the time step was 1.5 fs), a dielectric constant $\epsilon = 1$ was used, and the bond lengths were kept at their equilibrium values with

(25) Roos, B. O.; Andersson, K. *Chem. Phys. Lett.* **1995**, *245*, 215.

(26) Roos, B. O.; Andersson, K.; Fülscher, M. P.; Serrano-Andrés, L.; Pierloot, K.; Merchán, M.; Molina, V. *J. Mol. Struct. (Theochem)* **1996**, in press.

(27) Pierloot, K.; Dumez, B.; Widmark, P.-O.; Roos, B. O. *Theor. Chim. Acta* **1995**, *90*, 87.

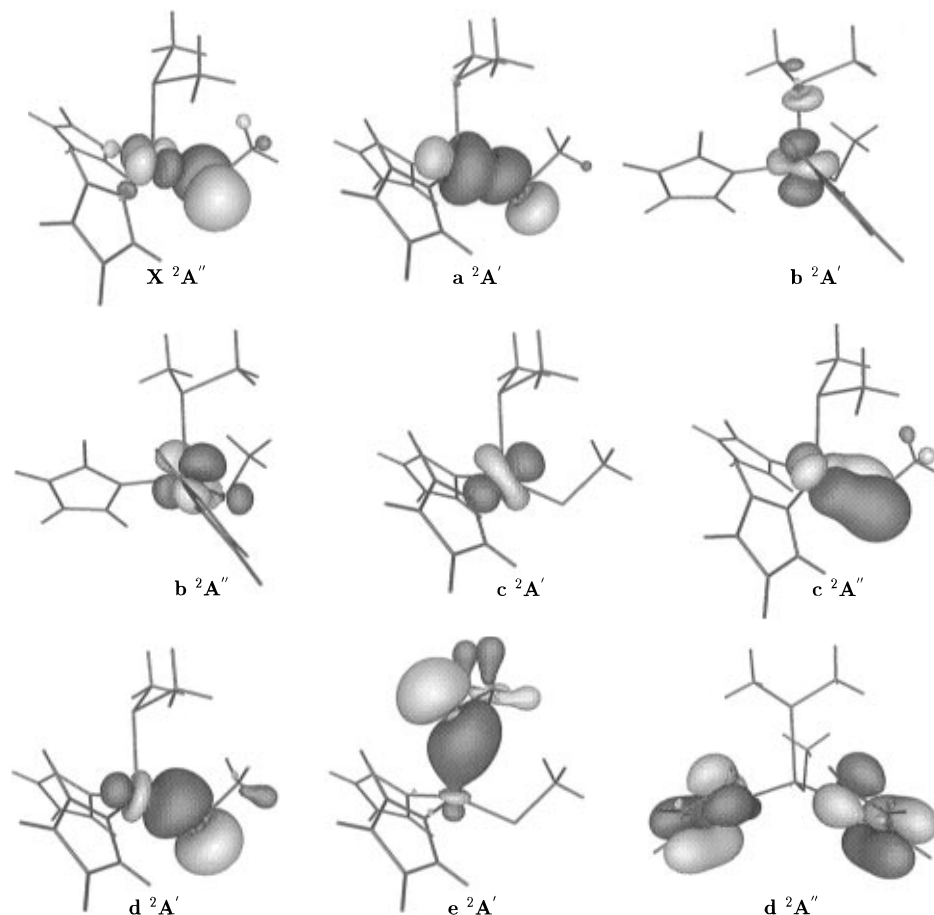(28) Malmqvist, P. Å.; Roos, B. O. *Chem. Phys. Lett.* **1989**, *155*, 189.

(29) Andersson, K.; Fülscher, M. P.; Karlström, G.; Lindh, R.; Malmqvist, P.-A.; Olsen, J.; Roos, B. O.; Sadlej, A. J.; Blomberg, M. R. A.; Siegbahn, P. E. M.; Kellö, V.; Noga, J.; Urban, M.; Widmark, P.-O. *MOLCAS Version 3*: Dept. of Theor. Chem., Chem. Center, Univ. of Lund, P.O.B. 124, S-221 00 Lund, Sweden, Lund, 1994.

(30) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179–5197.

Figure 2. The singly occupied natural orbitals for the ground state (X^2A'') and the eight lowest excited states of $Cu(imidazole)_2(SCH_3)(S(CH_3)_2)^+$.

the SHAKE algorithm. All heavy atoms present in the PDB file were kept fixed and no bonds between copper and the ligands were defined. The charges on the copper ion and its ligands were taken from a quantum chemical calculation on $Cu(imidazole)_2(SCH_3)(S(CH_3)_2)^+$ without symmetry.[22] The software Amber version 4.1[31] with the Cornell *et al.* 1995 force field[30] was used.

3. Results and Discussion

3.1. The Electronic Structure of Plastocyanin. The electronic structure of plastocyanin in the ground state and in the lowest excited states can best be characterized by analyzing the singly occupied orbital in each state. These orbitals for Cu-$(imidazole)_2(SCH_3)(S(CH_3)_2)^+$ are shown in Figure 2 and their main atomic orbital components are listed in Table 3. The singly occupied orbital in the ground state, X^2A'', is a π-antibonding orbital involving the Cu $3d_{xy}$ and S_{Cys} $3p_y$ atomic components. It also forms weak σ-antibonding interactions between copper and the two nitrogen ligands. The orbital is strongly delocalized over the Cu−S_{Cys} bond (54% Cu 3d and 45% S_{Cys} character). It does not, however, contain any contribution from S_{Met}, which indicates that the bond between the methionine residue and copper is purely electrostatic in nature. The c^2A'' state has the corresponding *bonding* orbital between Cu $3d_{xy}$ and S_{Cys} $3p\pi$ singly occupied. This orbital is also strongly delocalized over the bond, and has the same percentage S_{Cys} and copper character as the ground-state orbital.

Similarly, the lowest state in the other symmetry, a^2A', and the d^2A' state have a single electron in the σ combination of Cu $3d_{x^2-y^2}$ and S_{Cys} $3p_x$; the singly occupied orbital in a^2A' is

(31) Pearlman, D. A.; Case, D. A.; Caldwell, J. W.; Ross, W. S.; Cheatham, T. E.; Ferguson, D. M.; Seibel, G. L.; Singh, U. C.; Weiner, P. K.; Kollman, P. A. *Amber Version 4.1*; University of California: San Francisco, CA, 1995.

Table 3. Main Atomic Orbital Components of the Singly Occupied Orbital for the Different Excited States in $Cu(imidazole)_2(SCH_3)(S(CH_3)_2)^+$

state	molecular orbital[a]
X^2A''	51% $Cu3d_{xy}$ + 3% $Cu3d_{yz}$ + 45% S_{Cys} + 1% Im
a^2A'	3% $Cu3d_{z^2}$ + 19% $Cu3d_{xz}$ + 44% $Cu3d_{x^2-y^2}$ + 5% $Cu4s$ + 3% $Cu4p$ + 1% S_{Met} + 25% S_{Cys}
b^2A'	85% $Cu3d_{z^2}$ + 9% $Cu3d_{x^2-y^2}$ + 5% S_{Met} + 1% Im
b^2A''	19% $Cu3d_{xy}$ + 74% $Cu3d_{yz}$ + 6% S_{Cys}
c^2A'	3% $Cu3d_{z^2}$ + 69% $Cu3d_{xz}$ + 27% $Cu3d_{x^2-y^2}$
c^2A''	29% $Cu3d_{xy}$ + 24% $Cu3d_{yz}$ + 46% S_{Cys} + 1% Im
d^2A'	4% $C3d_{z^2}$ + 11% $Cu3d_{xz}$ + 20% $Cu3d_{x^2-y^2}$ + 4% $Cu4s$ + 3% $Cu4p$ + 58% S_{Cys}
e^2A'	5% $Cu3d_{z^2}$ + 2% $Cu4s$ + 2% $Cu4p$ + 91% S_{Met}
d^2A''	100% Im

[a] S_{Cys} = cysteine, S_{Met} = methionine, Im = imidazole.

the antibonding combination and in d^2A' the corresponding bonding combination. These two orbitals are also strongly delocalized along the Cu−S_{Cys} bond.

In the b^2A'', b^2A', and c^2A' states, the singly occupied orbitals are almost pure Cu 3d: d_{yz}, d_{z^2}, and d_{xz}, respectively. In b^2A' it contains a small amount of a σ antibond to S_{Met} $3p_z$. The corresponding bonding interaction is found in the e^2A' state, which is mainly localized on S_{Met} (except in $Cu(imidazole)_2$-$(SH)(SH_2)^+$, where this state represents a charge transfer from imidazole π; the $SH_2 \rightarrow Cu$ transition appears at even higher energy). The d^2A'' state, finally, has an almost pure imidazole π orbital singly occupied.

Formally, the c,d^2A'' and the d,e^2A' states represent charge-transfer transitions and the other four states represent ligand-field transitions. This interpretation is confirmed by the Cu 3d occupation numbers for the different states listed in Table 4. The ground state and the a^2A' states have a population around

Table 4.   Occupation Numbers of the Cu 3d Orbitals in the Ground State and the Different Excited States of the Model Complexes

	Cu(NH₃)₂SXSY₂ (C_s)					Cu(Im)₂SXSY₂ (C_s)					Cu(Im)₂SHSH₂ (C_1)	
state	X = H Y = H	X = CH₃ Y = H	X = H Y = CH₃	X = CH₃ Y = CH₃	X = C₂H₅ Y = CH₃	X = H Y = H	X = CH₃ Y = H	X = H Y = CH₃	X = CH₃ Y = CH₃	X = C₂H₅ Y = CH₃	opt[a]	exp[a]
X²A″	9.38	9.47	9.34	9.43	9.44	9.39	9.49	9.36	9.47	9.47	9.24	9.29
a²A′	9.38	9.54	9.36	9.43	9.43	9.33	9.46	9.31	9.37	9.36	9.30	9.34
b²A′	9.09	9.08	9.11	9.10	9.10	9.04	9.05	9.09	9.08	9.08	9.06	9.05
b²A″	9.07	9.08	9.08	9.09	9.09	9.04	9.06	9.06	9.08	9.08	9.05	9.05
c²A′	9.04	9.05	9.04	9.04	9.04	9.03	9.04	9.03	9.03	9.03	9.04	9.05
c²A″	9.69	9.58	9.71	9.60	9.60	9.58	9.45	9.58	9.46	9.51	9.81	9.74
d²A′	9.68	9.52	9.71	9.64	9.64	9.69	9.56	9.72	9.66	9.66	9.75	9.70
e²A′	9.97	9.95	9.94	9.95	9.95	9.99	9.94	9.93	9.94	9.94		
d²A″						9.97	9.97	9.97	9.97	9.97		

[a] opt = at the optimized C_s geometry, but calculated without symmetry; exp = using the experimental structure.

Table 5.   The Spectrum of Cu(NH₃)₂(SH)(SH₂)⁺ (in cm⁻¹), Calculated at the CASPT2 Level with Different Basis sets: $\Delta E = \Delta E_{val} + \Delta E_{rel} + \Delta E_{3s,3p}$[a]

	ΔE_{val}			ΔE_{rel}			$\Delta E_{3s,3p}$			ΔE		
state	A	C	D	A	C	D	A	C	D	A	C	D
a²A′	4 353	4 158	4 330	+144	+127	+126	−70	−73	−58	4 427	4 212	4 398
b²A′	13 447	13 161	13 187	−236	−239	−230	−555	−509	−553	12 656	12 413	12 404
c²A′	15 685	15 398	15 404	−290	−291	−282	−594	−547	−598	14 801	14 560	14 524
b²A″	15 452	15 306	15 217	−263	−257	−262	−353	−296	−369	14 836	14 753	14 586
c²A″	13 214	12 805	13 084	+277	+258	+263	+622	+613	+689	14 113	13 676	14 036
d²A′	19 137	18 600	19 047	+430	+404	+417	+536	+539	+613	20 103	19 543	20 077
e²A′	31 229	29 292	30 684	+546	+505	+523	+722	+395	+759	32 497	30 192	31 966

[a] ΔE_{val} includes no relativistic effects and only correlation of valence electrons; the effects of adding relativistic corrections and Cu 3s,3p correlation on the excitation energies are given in ΔE_{rel} and $\Delta E_{3s,3p}$, respectively.

9.4, indicating a delocalized orbital with mainly Cu character. The states with the corresponding bonding orbitals singly occupied (c²A″ and d²A′) have, in general, larger S_{Cys} character and thus a higher Cu 3d population, around 9.6. However, only a small amount of charge (0.2 electrons) is transferred in these excitations, so the term charge transfer is a bit misleading. They are better characterized as transitions from a bonding to an antibonding orbital (with only a small shift in the polarity), which also explains the large intensity. On the other hand, the almost pure ligand-field states (b,c²A′ and b²A″) have Cu 3d populations around 9.1 and the almost pure charge-transfer transitions (e²A′ and d²A″) have around 9.95 Cu 3d electrons. For these states, the classification as ligand-field and charge-transfer transitions is more appropriate.

3.2. The Copper−Cysteine Bond. The Cu−Cys bond seems to be the crucial factor for the spectral properties of the blue copper proteins and probably also for several other unusual characteristics of these proteins. It turns out that the polarity of this bond in plastocyanin strongly depends on the method used in the calculation. The problem is illustrated by the data in Table 4. Column 7 gives the Cu 3d populations for Cu-(imidazole)₂(SH)(SH₂)⁺ in C_s symmetry. The orbitals were obtained from two state-average CASSCF calculations including all states of ²A′ and ²A″ symmetry, respectively. This gives a 3d population in the ground state of 9.39 electrons and 9.58 electrons in the c²A″ state. Column 12 shows the results obtained in C_1 symmetry where the averaging included the lowest seven electronic states. The 3d population of the ground state has now decreased (a more ionic bond) to 9.24 electrons and that of the charge-transfer state has increased to 9.81. On the other hand, using fully optimized orbitals (single-state CASSCF calculations) leads to two different stationary CASSCF solutions for the X²A″ state with very similar energies, corresponding to d⁹ and d¹⁰ respectively, while the c²A″ state now obtains a 3d population close to 9.5. Unfortunately, this approach cannot be used to compute the electronic spectrum, since it is impossible to converge single-state CASSCF wave functions for all the excited states.

The sensitivity of the charge distribution of the Cu−Cys bond makes it very hard to arrive at a conclusive value for the polarity. Such a calculation would require a much larger MCSCF wave function, which includes the majority of the dynamic electron correlation effects. It might be worth mentioning that the MB3LYP calculations gives a 3d population close to 9.5 for the ground state.

3.3. Influence of the Basis Sets, Correlation of the Cu 3s and 3p Electrons, and Relativistic Effects on the Spectrum. The CASPT2 method has mainly been used on small chromophores where the size is a moderate problem.[13] However, the size of the chromophore in the blue copper proteins is so large that all reasonable approximations to reduce the size of the system are necessary. In order to verify that these approximations do not affect the result, several test calculations on the smallest model were performed.

First, the influence of the basis sets was examined. Table 5 shows the energies of the seven lowest transitions for Cu(NH₃)₂-(SH)(SH₂)⁺. It can be seen in the last three columns that when the large basis set C is truncated to the smaller basis set A (which is used for the other systems) the transition energies increase for all transitions. However, the change is small for all states (100−250 cm⁻¹ for ligand-field transitions, and about 500 cm⁻¹ for charge-transfer transitions), except for the highest state (e²A′) where it is 2300 cm⁻¹. Thus, the small basis sets can safely be used for the six lowest transitions.

In calculations on the model systems with imidazole ligands, the slightly smaller basis set B was used. It has a 1s contraction for all hydrogens, except those directly bound to sulfur. Test calculations on Cu(imidazole)₂(SH)(SH₂)⁺ showed that the effect on the computed spectrum of reducing the basis set for the imidazole hydrogens was small (less than 200 cm⁻¹) for all transitions except e²A′ and d²A″ (charge-transfer transitions from the imidazole rings), where it was 700−800 cm⁻¹. Reduction of the basis set on the methyl hydrogens gave a somewhat larger effect: calculations on Cu(NH₃)₂(SCH₃)-

Table 6. The Spectrum of Different Blue Copper Models Cu(NH3)$_2$SXSY$_2$, Calculated by the CASPT2 Method[a]

final state	excitation energy (oscillator strength)				
	X = H; Y = H	X = CH$_3$; Y = H	X = H; Y = CH$_3$	X = CH$_3$; Y = CH$_3$	X = C$_2$H$_5$; Y = CH$_3$
a^{2}A′	4 427	4 782	2 716	3 182	3 322
b^{2}A′	12 656	15 504	9 002	11 558	11 598
b^{2}A″	14 836 (0.0046)	16 924 (0.0118)	12 495 (0.0049)	14 218 (0.0111)	14 264 (0.0113)
c^{2}A′	14 801	17 340	12 768	15 148	15 192
c^{2}A″	14 113 (0.1370)	15 276 (0.1600)	13 308 (0.1211)	14 092 (0.1445)	14 135 (0.1456)
d^{2}A′	20 103 (0.0005)	20 144 (0.0004)	18 998 (0.0005)	17 851 (0.0004)	17 840 (0.0004)
e^{2}A′	32 497	36 346	26 200	29 177	29 185

[a] Excitation energies are given in cm^{-1}.

Table 7. The Spectrum of Different Blue Copper Models Cu(imidazole)$_2$SXSY$_2$, Calculated by the CASPT2 Method[a]

final state	excitation energy (oscillator strength)				
	X = H; Y = H	X = CH$_3$; Y = H	X = H; Y = CH$_3$	X = CH$_3$; Y = CH$_3$	X = C$_2$H$_5$; Y = CH$_3$
a^{2}A′	5 237	5 561	3 762	4 302	4 520
b^{2}A′	13 330	15 835	9 162	11 647	11 767
b^{2}A″	14 776 (0.0034)	16 650 (0.0082)	12 751 (0.0042)	14 243 (0.0091)	14 337 (0.0094)
c^{2}A′	14 636	16 815	12 630	14 890	14 894
c^{2}A″	15 403 (0.1345)	16 650 (0.1651)	14 908 (0.1202)	15 933 (0.1500)	16 004 (0.1517)
d^{2}A′	21 506 (0.0007)	20 890 (0.0006)	20 143 (0.0007)	18 861 (0.0005)	18 949 (0.0006)
e^{2}A′	33 037	38 283	27 899	31 264	31 272
d^{2}A″	33 127	35 802	32 611	34 992	35 083

[a] Excitation energies are given in cm^{-1}.

[S(CH$_3$)$_2$]$^+$ with basis set B resulted in excitation energies that were between 140 and 560 cm^{-1} lower than those obtained with basis A.

Second, the size of the relativistic corrections (ΔE_{rel}) was examined. They are shown in the second section of Table 5. The corrections are rather small for all states, positive for the charge-transfer transitions (260–550 cm^{-1}), and negative for the ligand-field transitions (−230 to −290 cm^{-1}), except for the lowest transition (+130 cm^{-1}). The basis sets effect is minimal; the maximum difference is 40 cm^{-1}.

Third, the Cu 3s,3p correlation was studied by including these electrons in the CASPT2 treatment of dynamic correlation effects. Since all pair excitations are included, the calculations include both core–core and core–valence effects. In general, normal basis sets are too small to give an appropriate description of the correlation of the electrons in the Cu 3s and 3p orbitals.[13] In order to estimate the size of this correlation, calculations with larger basis sets that are decontracted in the Cu 3s,3p,3d region have to be used. Basis set D in Table 2 was constructed with this aim. The size of the 3s,3p-correlation correction is shown in the third section in Table 5. Again, the effect is small: −70 to +600 cm^{-1} for ligand-field transitions and +610 to −760 cm^{-1} for charge-transfer transitions. Moreover, the basis set effect is small, maximally 80 cm^{-1}, indicating that for our blue copper models, the small effect of Cu 3s,3p correlation can be reasonably estimated already with the small basis set A.

Even if the effect of relativistic corrections and 3s,3p correlation is small, they act in the same direction and their effects differ between the ligand-field and charge-transfer transitions (both effects favor the states with the lowest Cu 3d population). Thus, their combined effect may significantly alter the relative energy of the excited states. Therefore, relativistic corrections, calculated at the CASSCF level are included in all energies and the effect of the Cu 3s and 3p orbitals are included in the correlation treatment.

3.4. Influence of the Choice of Model System and the Geometry on the Spectrum. In the blue copper proteins, the copper ion is coordinated to two histidines, one cysteine, and one methionine. These ligands were modeled in several different ways; histidine was modeled as NH$_3$ or imidazole; cysteine as SH$^-$, SCH$_3^-$, or SC$_2$H$_5^-$; and methionine as SH$_2$ or

S(CH$_3$)$_2$. In order to examine the influence of the choice of the models on the spectrum, all combinations were tested. The results are gathered in Tables 6 and 7.

When ammonia is replaced by imidazole the excitation energy of all states increase by 800–1900 cm^{-1}, except for the ligand-field transitions b^{2}A′, b^{2}A″, and c^{2}A′, which are barely affected (<700 cm^{-1}). It is notable that only complexes including imidazole give the correct order of the two most intense transitions, b^{2}A″ and c^{2}A″.

When SH$^-$ is replaced by SCH$_3^-$ all excitation energies increase by 300–5200 cm^{-1}, except for d^{2}A′. On the other hand, substituting S(CH$_3$)$_2$ for SH$_2$ *decreases* all excitation (by 500–6800 cm^{-1}; consequently, the results obtained with Cu-(imidazole)$_2$(SH)(SH$_2$)$^+$ and Cu(imidazole)$_2$(SCH$_3$)(S(CH$_3$)$_2$)$^+$ are quite similar, except for d^{2}A′ (the difference is less than 1900 cm^{-1}). Further replacing SCH$_3^-$ by SC$_2$H$_5^-$ has a very small effect on all excitation energies (less than 220 cm^{-1}). This is a bit surprising, since Zerner *et al.* report appreciable changes in the spectrum (2100 cm^{-1} down shifts) when the chromophore in rubredoxin is modeled by Fe(SC$_2$H$_5$)$_4$ instead of Fe(SCH$_3$)$_4$.[32]

In conclusion, Cu(imidazole)$_2$(SCH$_3$)(S(CH$_3$)$_2$)$^+$ seems to the smallest system that gives converged excitation energies. Unfortunately, such a large system can be treated by the CASPT2 method only if it is forced to have a plane of symmetry (as was done for the systems in Table 6 and 7). This is not the optimal structure, however; a more stable structure is obtained if the two imidazole rings are tilted with respect to each other, and this is also the conformation found in the proteins. Therefore, the spectrum was also calculated for Cu(imidazole)$_2$-(SH)(SH$_2$)$^+$ at the experimental geometry without symmetry. The results were corrected for the ligand truncations using the data in Table 7.

As can be seen in Table 1, the quantum chemically optimized structures differ from the experimental structure in several details (it must, although, be remembered that there is appreciable uncertainty in the experimental structure also). For example, the Cu−S$_{Cys}$ distance seems to be a few picometers too long and the Cu−S$_{Met}$ distance is about 20 pm too short in the larger

(32) Stavrev, K. K.; Zerner, M. C. *Int. J. Quantum Chem.: Quantum Biol., Symp.* **1995**, *22*, 155–159.

Table 8. The Spectrum of $Cu(NH_3)_2(SH)(S(CH_3)_2)^+$ at Different (Partially Optimized) Geometries, Calculated using CASPT2[a]

	excitation energy (oscillator strength)			
final state	$r(Cu-S_{Cys}) = 218$ pm $r(Cu-S_{Met}) = 258$ pm	$r(Cu-S_{Cys}) = 213$ pm $r(Cu-S_{Met}) = 259$ pm	$r(Cu-S_{Cys}) = 208$ pm $r(Cu-S_{Met}) = 263$ pm	$r(Cu-S_{Cys}) = 203$ pm $r(Cu-S_{Met}) = 265$ pm
a^2A'	2 716	3 261	3 744	4 092
b^2A'	9 002	10 210	11 208	12 201
b^2A''	12 495 (0.0049)	13 516 (0.0047)	14 449 (0.0043)	15 400 (0.0042)
c^2A'	12 768	14 002	14 864	15 780
c^2A''	13 308 (0.1211)	14 243 (0.1261)	15 950 (0.1378)	17 753 (0.1490)
d^2A'	18 998 (0.0005)	19 871 (0.0007)	21 408 (0.0007)	22 956 (0.0009)
e^2A'	26 200	25 418	25 593	25 907

	excitation energy (oscillator strength)			
	$r(Cu-S_{Met}) = 258$ pm $r(Cu-S_{Cys}) = 218$ pm	$r(Cu-S_{Met}) = 268$ pm $r(Cu-S_{Cys}) = 218$ pm	$r(Cu-S_{Met}) = 278$ pm $r(Cu-S_{Cys}) = 217$ pm	$r(Cu-S_{Met}) = 288$ pm $r(Cu-S_{Cys}) = 217$ pm
a^2A'	2 716	3 261	3 744	4 109
b^2A'	9 002	10 267	11 091	11 898
b^2A''	12 495 (0.0049)	13 303 (0.0045)	13 876 (0.0041)	14 398 (0.0034)
c^2A'	12 768	13 698	14 141	14 543
c^2A''	13 308 (0.1211)	12 825 (0.1173)	13 065 (0.1202)	13 084 (0.1220)
d^2A'	18 998 (0.0005)	18 709 (0.0005)	19 052 (0.0005)	19 346 (0.0005)
e^2A'	26 200	24 138	23 353	22 492

[a] Excitation energies are given in cm^{-1}.

systems. Furthermore, these two distances vary considerably between different types of blue copper proteins.[2] Therefore, we decided to examine the effect on the excitation energies when these distances were changed systematically.

The effect of the Cu−S bond lengths on the excitation energies can be seen in Table 8. The structures were obtained with the MB3LYP method by a constrained optimization of $Cu(NH_3)_2(SH)(S(CH_3)_2)^+$ with either the Cu−S_{Cys} distance fixed to 213, 208, and 203 pm, or the Cu−S_{Met} distance fixed to 268, 278, and 288 pm. The rest of the geometry was optimized. Of course, a change in one Cu−S distance affects the other distance, but the effect is not very large; when the Cu−S_{Cys} distance is decreased, the Cu−S_{Met} distance increases slightly, but the Cu−S_{Cys} distance hardly changes at all when the Cu−S_{Met} distance is changed.

When the Cu−S distances are changed from the optimized values (left column) to the near-experimental values (right column), the a^2A' transition energy increases about 1300 cm^{-1} both by the change in the Cu−S_{Cys} and Cu−S_{Met} distances. The ligand-field states b^2A', b^2A'', and c^2A' increase about 3000 cm^{-1} when the Cu−S_{Cys} distance is decreased and 1700−2900 cm^{-1} when the Cu−S_{Met} distance increases. For the two cysteine to copper charge-transfer states d^2A' and c^2A'', naturally, the effect of changing the Cu−S_{Met} distance is limited (<500 cm^{-1}), while a very strong effect of the Cu−S_{Cys} distance is found: 4000−4500 cm^{-1}. Similarly, the e^2A' state (charge transfer to S_{Met}) is indifferent to the Cu−S_{Cys} distance, but very sensitive to the Cu−S_{Met} bond length (it decreases by 3700 cm^{-1} when the distance increases with 30 pm). The d^2A'' state is not present in the ammonia complexes, but judging from the results for the other charge-transfer transitions, the effect of a change in the Cu−S bond lengths can be expected to be small, since the state involves a imidazole to copper charge transfer.

Overall, the excitation energies depend strongly on the geometry. For most states the effect of the geometry is more important than the effects due to the model substitutions shown in Tables 6 and 7. It is notable that the different optimized geometries have very similar Cu−S_{Cys} distances, so that the effect of the Cu−S_{Cys} distance and the ligand substitution are separable. This is not the case for the Cu−S_{Met} distance, however. It strongly depends on whether the methionine ligand is modeled as SH_2 or $S(CH_3)_2$. Oddly enough, the effect of the ligand substitution on the excitation energy for the e^2A' state is

opposite to what could be expected from the concomitant changes in the Cu−S_{Met} distance.

The strong dependence of the excitation energies on the Cu−S_{Cys} bond length is notable. Measured excitation energies of the dominant blue band for plastocyanin and other normal (not rhombic) type 1 blue copper proteins (e.g. azurin and amicyanin) from different sources range between 16 000 and 16 800 cm^{-1}.[33] According to the results in Table 8, such a range corresponds to a variation in the Cu−S_{Cys} bond length of less than 4.3 pm. This is much less than the variation found in crystal structures, which is probably due to the relatively low accuracy of measured distances (about 12 pm[34]). The small variation is a natural consequence of the strong Cu−S_{Cys} bond if the copper complex is in a relaxed equilibrium state as we have suggested.[22] However, it is hard to explain if the cupric geometry is imposed by the protein, as frequently has been assumed before.[35,36]

3.5. Comparison with the Experimental Spectrum. We are now prepared to collect the results and compare them to the experimental spectrum. This is done in Table 9 for the six lowest transitions. The most reliable results are obtained with $Cu(imidazole)_2(SH)(SH_2)^+$ at the crystal geometry (column C), including the effect of the surrounding protein and solvent by point charges (column D), and correcting for the truncation of the cysteine and methionine models (using the results in Table 7, column E). The resulting excitation energies are in agreement with the experimental energies (column F);[8] in no instance is the difference in the energies greater than 1800 cm^{-1}. The largest error occurs for the charge-transfer transition d^2A', while the ligand-field transition energies are almost exactly reproduced by the calculations.

This excellent agreement is probably somewhat fortuitous, however. The computed CASPT2 energies depend somewhat on the choice of orbitals. The excitation energies were obtained with orbitals optimized in state average CASSCF calculations, as discussed above. The first-order wave function includes single excitations and the method should therefore to some extent correct for the use of nonoptimal orbitals. Still, the

(33) Andrew, C. R.; Yeom, H.; Valentine, J. S.; Karlsson, B. G.; Bonander, N.; van Pouderoyen, G.; Canters, G. W.; Loehr, T. M.; Sanders-Loehr, J. *J. Am. Chem. Soc.* **1994**, *116*, 11489−11498.

(34) Fields, B. A.; Bartsch, H. H.; Bartunik, H. D.; Cordes, F.; Guss, J. M.; Freeman, H. C. *Acta Crystallogr.* **1994**, *D50*, 709−730.

(35) Malmström, B. G. *Eur. J. Biochem.* **1994**, *223*, 207−216.

(36) Williams, R. J. P. *Eur. J. Biochem.* **1995**, *234*, 363−381.

Table 9. The Spectrum of Cu(imidazole)$_2$(SH)(SH$_2$)$^+$ Calculated by the CASPT2 Method[a]

final state	excitation energy (oscillator strength)					
	A	B	C	D	E	F
a^{2}A′	5 237	3 880	5 441	4 836	4 119	5 000
b^{2}A′	13 330	11 577	14 475	12 537	10 974	10 800 (0.0031)
b^{2}A″	14 776 (0.0034)	12 996 (0.0014)	15 356 (0.0028)	13 556 (0.0015)	13 117 (0.0015)	12 800 (0.0114)
c^{2}A′	14 636	12 821	15 084 (0.0004)	13 235 (0.0003)	13 493 (0.0003)	13 950 (0.0043)
c^{2}A″	15 403 (0.1345)	12 393 (0.0886)	15 948 (0.1189)	16 970 (0.1032)	17 571 (0.1032)	16 700 (0.0496)
d^{2}A′	21 506 (0.0007)	20 243 (0.0008)	22 093 (0.0010)	23 156 (0.0014)	20 599 (0.0014)	18 700 (0.0048)

[a] (A) optimized geometry (C_s); (B) optimized geometry (C_1); (C) experimental geometry (C_1); (D) experimental geometry with surrounding point charges; (E) same as D but corrected for the effects of replacing SH by SC$_2$H$_5$ and SH$_2$ by S(CH$_3$)$_2$ (corrections taken from Table 7); (F) experimental excitation energies and oscillator strengths [8]. All excitation energies are given in cm^{-1}.

energies are affected. Thus, increasing the number of states in the average from four to seven, the ground state CASPT2 energy increases by 2200 cm^{-1}, while that of the c^{2}A″ state decreases by 1100 cm^{-1}. The other states are much less affected. Consequently, the increasing ground state energy decreases all excitation energies by between 1000 and 2000 cm^{-1}. The relative energy of the c^{2}A″ state is even more severely affected, by about 3000 cm^{-1} (compare columns A and B in Table 9). As a coincidence, the effect on the excitation energies due to the increase in the number of states in the average CASSCF calculations at the optimum geometry is of the same size but with reversed sign compared to the difference in excitation energies between the experimental and the optimum geometry. As a result, the C_s spectrum obtained for Cu(imidazole)$_2$(SH)-(SH$_2$)$^+$ at its optimum geometry (column A) and the C_1 spectrum at the experimental structure (column C) are only slightly different.

The difference in the excitation energies between the optimal and the experimental structure of Cu(imidazole)$_2$(SH)(SH$_2$)$^+$ (columns B and C) is mainly due to the change in the Cu$-$S$_{Cys}$ distance. The change in the Cu$-$S$_{Met}$ distance is rather small. Consequently, all energies increase. The influence of the surrounding protein on the spectrum should be noted. It varies between -1938 cm^{-1} for b^{2}A′ to $+1063$ cm^{-1} for d^{2}A′. The excitation energies of the ligand-field states decrease while they increase for the two charge-transfer states. Thus, the effect is considerable and cannot be neglected.

The calculated oscillator strengths are reasonable, although far from perfect. Still, they are accurate enough to confirm the assignment of the experimental bands based on the calculated excitation energies. The strongest band in the spectrum undoubtedly corresponds to the c^{2}A″ state, i.e. the transition from the π-bonding Cu 3d$-$S$_{Cys}$ 3p orbital. This excitation gains its high intensity from the extremely large overlap between the singly occupied orbitals in the X^{2}A″ and c^{2}A″ states (*c.f.* Figure 2). This interpretation is in accord with earlier assignments.[8,11]

The second most intense band corresponds to the ligand-field transition to the Cu 3d$_{yz}$ orbital (b^{2}A″) which gains intensity by mixing in a small amount of S$_{Cys}$ 3pπ (they are of the same symmetry). The band at 18 700 cm^{-1} corresponds to the d^{2}A′ state, the excitation from the S$_{Cys}$ σ-bonding orbital. This transition becomes slightly more intense in the C_1 structure, probably as the result of some mixing between the Cu$-$S$_{Cys}$ σ and Cu$-$S$_{Cys}$ π bonds. The remaining three bands are weak ligand-field transitions, namely Cu 3d$_{x^2-y^2}$ ($+$ S$_{Cys}$ 3pσ), 3d$_{z^2}$, and 3d$_{xz}$, ordered by increasing energy. This order of the d$_{x^2-y^2}$ and the d$_{z^2}$ states differs from the one suggested by Solomon *et al.*,[8] but their excitation energies support our assignment.

There are three weak bands in the experimental spectrum, at 21 390, 23 440, and 32 500 cm^{-1}, that are outside the range shown in Table 9. Solomon *et al.*[8] assign the 21 390 and 32 500 cm^{-1} bands to charge-transfer transitions from the imidazole π

system to copper and the 23 440 cm^{-1} band to a S$_{Met}$ to copper charge-transfer excitation. Our calculations include both the S$_{Met}$ $\rightarrow$ Cu and the lowest imidazole $\rightarrow$ Cu transitions (e^{2}A′ and d^{2}A″, respectively), but only for the models with a symmetry plane. The e^{2}A′ excitation is predicted at 31 300 cm^{-1} for the best model system, but the excitation energy depends strongly on the Cu$-$S$_{Met}$ bond length; if this distance was corrected to the experimental value, the excitation energy would decrease by almost 4000 cm^{-1} (Table 8). Furthermore, this transition suffers quite a bit from the basis set incompleteness; as shown in Table 5, a better basis set would also decrease the energy of this transition. Therefore, Solomon may be right in attributing the S$_{Met}$ to Cu charge-transfer transition to the 23 440 cm^{-1} band.

The d^{2}A″ transition, on the other hand, is calculated at 35 083 cm^{-1} for the best model system, and it seems to be much more stable to changes in the geometry and basis sets. Therefore, this transition most likely corresponds to the band observed around 32 500 cm^{-1}.

Thus, the 21 390 cm^{-1} band remains unassigned. Our calculations indicate that Solomon's interpretation of this band as an imidazole π to copper charge-transfer transition is incorrect, but give no clue to its actual character. Furthermore, there is strong experimental evidence that this transition involves the cysteine sulfur rather than imidazole: the resonance Raman spectrum of several blue copper proteins resulting from excitations around 460 nm (21 700 cm^{-1}) is very similar to those obtained from excitation of the bright blue band around 600 nm (16 700 cm^{-1}), which undoubtedly originate from a S$_{Cys}$ $\rightarrow$ Cu charge transfer.[37]

It is notable that the 21 390 cm^{-1} band directly corresponds to the bright line observed around 450 nm in the spectra of some blue copper proteins, e.g. nitrite reductase and stellacyanin. This leads us to suggest that the 21 390 cm^{-1} line is due to some sort of intrinsic contaminant, i.e. a copper center in plastocyanin with a modified structure. This modification may, for example, be a water molecule replacing one of the normal ligands. However, some evidence points to an even more suggestive explanation.

Nitrite reductase is an interesting protein, having exactly the same copper ligands, but a different coordination geometry, spectrum, and EPR characteristics.[9,38] Our preliminary results[39] indicate that this is because nitrite reductase has a different ground state, viz. a strongly distorted square-planar structure arising from the a^{2}A′ state with a Cu$-$S$_{Cys}$ σ bond. The structure is distinctly unsymmetric so that the a^{2}A′ and X^{2}A″

(37) Han, J.; Loehr, T. M.; Lu, Y.; Valentine, J. S.; Averill, B. A.; Sanders-Loehr, J. *J. Am. Chem. Soc.* **1993**, *115*, 4256$-$4263.

(38) Adman, E. T.; Godden, J. W.; Turley, S. *J. Biol. Chem.* **1995**, *270*, 27458$-$27474.

(39) Pierloot, K.; De Kerpel, J. O. A.; Olsson, M. H. M.; Ryde, U.; Eriksson, L.; Roos, B. O. Degeneracy in blue copper proteins $-$ a theoretical investigation of the structure, spectrum and EPR parameters of nitrite reductase. To be published, 1997.

states can mix. As a result, at least four transitions gain significant intensities.

Furthermore, the two structures are almost degenerate; the energy difference seems to be less than 5 kJ/mol. This means that both structures may be occupied at ambient temperatures, and therefore the spectrum could be a mixture of the two spectra arising from different geometries. Since the band around 21 700 cm^{-1} is so bright in the nitrite reductase spectrum, it might become visible also in the plastocyanin spectrum. Because it originates from a S$_{Cys}$ to copper transition in the nitrite reductase spectrum, it will have cysteine character also in the plastocyanin spectrum. Furthermore, this would explain why more lines have significant intensity in the experimental spectrum of plastocyanin than in the theoretical spectrum. These points will be thoroughly discussed in a forthcoming publication.[39]

4. Conclusions

We have presented results from a theoretical study of the electronic spectrum of the blue copper protein plastocyanin. Geometries were optimized using the MB3LYP density functional method and the spectrum was computed using multiconfigurational wave functions with dynamic correlation effects estimated by second-order perturbation theory (CASPT2). The calculated spectrum is in agreement with experiment. The comparison between theory and experiment is not without problems, however. The excitation energies have been found to depend strongly on the geometry, and uncertainties in the experimental determination of the structure as well as possible errors in the computed geometries affect the result.

It should also be emphasized that the theoretical methods used are afflicted with uncertainties. The size of the system has made it necessary to limit the basis sets; the effect from the part of the protein not included in the quantum chemical treatment and the solvent are only approximately taken into account; the flexibility of the structure introduces some uncertainty in the theoretical geometry optimizations, and the CASPT2 method has an inherent error of about 2000 cm^{-1}. However, the most serious problem encountered during this study is the difficulty in describing the polarity of the Cu$-$S$_{Cys}$ bond. The results strongly depend on the theoretical method used and this ambiguity will also affect the spectral properties. Problems of this type occur in systems where the character of the wave function strongly depends on dynamic electron correlation. Normally, such effects do not have a large influence on the electronic density. The CASSCF/CASPT2 approach is devised to deal with situations where the CASSCF wave function gives

a good representation of the true wave function. The situation here is more difficult, since the polarity of the Cu$-$S$_{Cys}$ bond depends strongly on electron correlation. The problem can be resolved by allowing the reference function to relax under the influence of the electron correlation terms. Such approaches have been devised, but are not yet at the stage where they can be applied to systems of the present size.

Most of the approximations introduced in the calculations have been tested by careful studies of the smaller models. This includes studies of effects due to the basis sets, geometry changes, the influence of the surrounding protein and solvent, etc. Some of these effects were found to be small but others influence the spectrum strongly. On the basis of these model studies and our experience from earlier studies, we expect the present results to give an adequate description of the electronic spectrum of plastocyanin and enable an assignment of the experimental spectrum in the energy range 0$-$30000 cm^{-1}.

Plastocyanin has been found to have a ground state of ^{2}A$''$ symmetry (assuming an approximate C_s symmetry for the structure), where the cysteine sulfur is π bonded to the metal center. This property determines the structure of the coordination sphere (distorted trigonal pyramid) and also the main features of the electronic spectrum. In a forthcoming publication we shall discuss the structural and spectral properties of a similar protein (nitrite reductase), where S$_{Cys}$ is instead σ bonded to Cu.[39] This results in a structure that is distorted square planar, and it also affects the spectral properties. Thus, the Cu$-$S$_{Cys}$ bond, which has a polarity that can be easily modified by the surrounding, and which can give rise to two different ground state symmetries with closely similar energies, is the key property for the understanding of the electronic and structural characteristics of the blue copper proteins.

Note Added in Proof: More recent results obtained from a comparative study of the spectrum of different type 1 blue copper proteins with the same copper ligands indicate that the S$_{Cys}$ 3pσ $\rightarrow$ Cu excitation (state d^{2}A$'$ in Table 9) should be assigned to the band at 21 390 cm^{-1} rather than to the one at 18 700 cm^{-1} in the plastocyanin spectrum.

JA962381F