Relation between the Structure and Spectroscopic Properties of Blue Copper Proteins

Kristine Pierloot,[†] Jan O. A. De Kerpel,[†] Ulf Ryde,*[,‡] Mats H. M. Olsson,[‡] 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, Chemical Center, Lund University, P.O.B. 124, S-221 00 Lund, Sweden

Received July 7, 1998. Revised Manuscript Received September 28, 1998

Abstract: The electronic spectra of three rhombic type 1 blue copper proteins, nitrite reductase, pseudoazurin, and cucumber basic protein, have been studied by ab initio multiconfigurational second-order perturbation theory (the CASPT2 method). The six lowest excitations have been calculated and assigned with an error of less than 1800 cm^{-1}. The singly occupied orbital in the ground-state forms a strongly covalent antibond between the copper ion and the thiolate group of the cysteine ligand with a mixture of σ and π character. This is in contrast to the axial type 1 copper protein plastocyanin which has an almost pure $Cu-S_{Cys}$ π interaction. The two brightest lines in the absorption spectrum originate from transitions to the corresponding σ (∼460 nm) and π (∼600 nm) bonding orbitals. The relative intensity of these two lines is determined by the character of the ground-state orbital. It is possible to obtain a structure closely similar to the one found in nitrite reductase by geometry optimizations with the hybrid density functional B3LYP method in vacuum. It is a tetragonal structure with bonds of mainly σ character to the four ligands like normal square-planar Cu(II) complexes, but the cysteine thiolate group donates much charge to the copper ion and thereby makes the structure strongly distorted toward a tetrahedron. Both this structure and a trigonal π-bonded structure, which also can be obtained for all complexes and is an excellent model of plastocyanin, are equilibrium structures (although usually not with the same ligand models). They have virtually the same energy (within ∼7 kJ/mol), and the barrier between them is low. Therefore, small differences in the structure and electrostatics of different proteins may lead to stabilization of one or the other of the structures. The results indicate that axial type 1 proteins have a trigonal structure with an almost pure $Cu-S_{Cys}$ π bond, whereas rhombic type 1 proteins have tetragonal structures with a significant σ character in this bond. Type 1.5 and 2 copper−cysteinate proteins arise when the tetragonal structure becomes more flattened than in nitrite reductase, probably by the inclusion of stronger (type 1.5) and more (type 2) ligands.

Introduction

Blue copper proteins, in their oxidized form, exhibit a number of specific macroscopic properties that distinguish them from small-molecule cupric complexes,[1,2] e.g., an intense blue color, distinctive electron spin resonance spectra, and unusually high reduction potentials. These extraordinary properties are related to the specific coordination geometry of the copper ligand sphere in this class of proteins. A typical example is plastocyanin, a small water-soluble protein involved in photosynthesis. In this protein, the copper ion is surrounded by a distorted trigonal plane, formed by two histidine residues and a cysteine thiolate group, the latter being found at an unusually short distance (∼210 pm). A fourth ligand, a methionine thioether group, is bound in an axial position at a large distance (∼290 pm).[3] Such a trigonal copper surrounding is widely different from the tetragonal surrounding normally found in small inorganic Cu(II) complexes.[4] Instead, it is closer to the (distorted) tetrahedral surrounding found for the reduced Cu(I) form of the same protein.[5] The close resemblance between the two geometries accounts for a small reorganization energy accompanying reduction and, hence, for the high rate of electron transfer in this class of proteins.[6] This fact has led to the suggestion that the unusual trigonal surrounding of the oxidized site is not the one preferred by Cu(II) itself but is instead imposed upon it by the tertiary protein structure, in order to facilitate electron-transfer, the induced-rack[7,8] and entatic-state[9,10] theories.

However, we have previously shown[11] that the role of the protein in determining the Cu(II) coordination geometry is considerably less important than suggested by these theories. This was achieved by performing quantum chemical geometry optimizations on realistic models for plastocyanin, containing

[†] University of Leuven.
[‡] Lund University.

(1) Sykes, A. G. *Adv. Inorg. Chem.* **1990**, *36*, 377−408.
(2) Adman, E. T. *Adv. Protein Chem.* **1991**, *42*, 145−197.
(3) Guss, J. M.; Bartunik, H. D.; Freeman, H. C. *Acta Crystallogr.* **1992**, *B48*, 790−807.
(4) Lever, A. B. P. *Inorganic Electronic Spectroscopy*; Elsevier: Amsterdam, 1984.
(5) Guss, J. M.; Harrowell, P. R.; Murata, M.; Norris, V. A.; Freeman, H. C. *J. Mol. Biol.* **1986**, *192*, 361−387.
(6) Marcus, R. A.; Sutin, N. *Biochim. Biophys. Acta* **1985**, *811*, 265−322.
(7) Gray, H. B.; Malmström, B. G. *Comments Inorg. Chem.* **1983**, *2*, 203−209.
(8) Malmström, B. G. *Eur. J. Biochem.* **1994**, *223*, 207−216.
(9) Vallee, B. L.; Williams, R. J. P. *Proc. Natl. Acad. Sci. U.S.A.* **1968**, *59*, 498−505.
(10) Williams, R. J. P. *Eur. J. Biochem.* **1995**, *234*, 363−381.
(11) Ryde, U.; Olsson, M. H. M.; Pierloot, K.; Roos, B. O. *J. Mol. Biol.* **1996**, *261*, 586−596.

10.1021/ja982385f CCC: $15.00
Published on Web 12/02/1998

the appropriate first coordination sphere around Cu(II) (e.g., $Cu(imidazole)_2(SCH_3)(S(CH_3)_2)^+$) but not the rest of the protein. With these models, we obtained a structure that is virtually identical to the experimental plastocyanin structure, indicating that protein strain is insignificant for the Cu(II) structure. The origin of the diverging Cu(II) structure in these proteins (i.e., trigonal instead of tetragonal) must therefore be traced back to the nature of the ligands themselves, and we showed in a recent study[12] on a series of four-coordinate Cu(II) model compounds, in which the thiolate group was systematically replaced by other ligands, that the softness and high polarizability of the thiolate group are crucial for stabilizing a trigonal Cu(II) surrounding, in which a strongly covalent $Cu-S_{Cys}$ π bond is formed. Furthermore, the covalency of the Cu−thiolate π bond is also at the origin of the extraordinary spectroscopic characteristics of blue copper proteins, in particular the appearance of the intense band at 600 nm (the blue band). This has been shown in several theoretical studies of the electronic spectrum of plastocyanin, where the blue band was found to originate from a charge-transfer excitation from the bonding to the antibonding Cu−thiolate π orbital, gaining its high intensity from the large overlap between the two orbitals involved.[13−16]

But if the geometric structure around Cu(II) in these proteins is solely determined by its ligands, would one not then expect different proteins containing the same ligand set to have almost the same structure? Examples of proteins with the same ligand set as plastocyanin are cucumber basic protein, pseudoazurin, and nitrite reductase. To a first approximation, the structure of the copper surrounding in these proteins is indeed close to the one in plastocyanin, but there are some differences, the most prominent being a longer bond to cysteine and a shorter bond to methionine. Nitrite reductase from *Achromobacter cycloclastes* is the extreme case, with an average $Cu-S_{Cys}$ bond length of 217 pm (as compared to 211 pm in plastocyanin) and a $Cu-S_{Met}$ bond length of 256 pm (288 pm in plastocyanin).[11,17] Even if these geometric differences are subtle, they give rise to quite strongly differing spectroscopic characteristics. In the ESR spectra, a varying degree of rhombic distortion is found among the various proteins. This has actually led to a further classification of blue copper proteins into axial type 1 proteins, e.g., plastocyanin and azurin, with no rhombic distortion, and rhombic type 1 proteins such as pseudoazurin, cucumber basic protein, nitrite reductase, and also stellacyanin. In the electronic spectra, an additional feature at 460 nm, barely visible in the plastocyanin spectrum, becomes considerably more prominent in the spectra of the rhombic type 1 proteins. At the same time, the intensity of the 600-nm band decreases, so that the sum of ϵ_{460} and ϵ_{600} remains virtually constant.[18] Nitrite reductase is again a limiting case, with the band at 460 nm more intense than the 600-nm band, so that the color of this protein is green rather than of blue. It has been suggested that the varying rhombicity of the ESR spectra and relative intensity of the 460- versus 600-nm bands are related to the structural variations found between the different proteins.[19] This conjecture will be explored in this work, where we calculate the electronic spectra of several rhombic type 1 proteins.

Methods

(1) Geometry Optimizations. Geometry optimizations were performed on model complexes of the type CuX_2YZ, where X represents either NH_3 or imidazole as a model for histidine, Y stands for the cysteine thiolate group, modeled by either SH^- or SCH_3^-, and Z represents the methionine thioether group, modeled by either SH_2 or $S(CH_3)_2$. The geometries were optimized with the hybrid density functional B3LYP method, as implemented in either the Gaussian-94[20] or the Mulliken-2.31h[21] software (denoted MB3LYP). The two implementations differ slightly in that they use two different solutions to the uniform electron gas in the Vosko−Wilk−Nusair correlation functional. However, very similar structures are obtained with the two methods in all cases (within 2 pm for bond lengths and 2° for bond angles). Harmonic frequencies were calculated for some of the complexes using the Gaussian-94 software (on structures optimized with the same code). For copper, we used the double-ζ basis (62111111/3111/311) of Schäfer et al.,[22] enhanced with diffuse p, d, and f functions with exponents 0.174, 0.132, and 0.390, respectively, and for the other atoms, the 6-31G* basis sets were used.[23] Only the pure five d-type and seven f-type functions were used and the standard fine grid was employed.

A set of constrained geometry optimizations were also performed with the $Cu(NH_3)_2(SH)(SH_2)^+$ model. Here, the dihedral angle between the N−Cu−N and $S_{Cys}-Cu-S_{Met}$ planes, denoted as ϕ, was kept fixed at several angles between 0° (strictly tetragonal) and 90° (strictly trigonal), while all other geometrical parameters were optimized using the B3LYP method. In a second step, the various B3LYP structures were used as a starting point for a constrained CASPT2 optimization, in which only the $Cu-S_{Cys}$ and $Cu-S_{Met}$ distances were reoptimized, while all other parameters were kept fixed. These CASPT2 calculations were performed with an active space of 11 electrons in 11 orbitals (Cu 3d, 4d and the highest S_{Cys} 3π orbital), and all valence electrons (but not Cu 3s and 3p) were correlated in the CASPT2 step, in order to keep basis set superposition error to a minimum.

(2) Calculation of the Electronic Spectra. The electronic spectra were calculated by the CASSCF/CASPT2 method (second-order many-body perturbation theory with a multiconfigurational reference wave function).[24,25] The calculations were performed with moderate-sized basis sets of the generally contracted atomic natural orbital (ANO)

(12) Olsson, M. H. M.; Ryde, U.; Roos, B. O.; Pierloot, K. *J. Biol. Inorg. Chem.* **1998**, *3*, 109−125.
(13) Penfield, K. W.; Gewirth, A. A.; Solomon, E. I. *J. Am. Chem. Soc.* **1985**, *107*, 4519−4519.
(14) Gewirth, A. A.; Solomon, E. I. *J. Am. Chem. Soc.* **1988**, *110*, 3811−3819.
(15) Larsson, S.; Broo, A.; Sjölin, L. *J. Phys. Chem.* **1995**, *99*, 4860−4865.
(16) Pierloot, K.; De Kerpel, J. O. A.; Ryde, U.; Roos, B. O. *J. Am. Chem. Soc.* **1997**, *119*, 218−226.
(17) Adman, E. T.; Godden, J. W.; Turley, S. *J. Biol. Chem.* **1995**, *270*, 27458−27474.
(18) Han, J.; Loehr, T. M.; Lu, Y.; Valentine, J. S.; Averill, B. A.; Sanders-Loehr, J. *J. Am. Chem. Soc.* **1993**, *115*, 4256−4263.
(19) Lu, Y.; Roe, J. A.; Gralla, E. B.; Valentine, J. S. In *Bioinorganic Chemistry of Copper*; Karlin, K. D., Tyeklár, Z., Eds.; Chapman & Hall: New York, 1993; p 64.
(20) 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.
(21) 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.; Linch, 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™ Version 2.31h, internal release*; IBM Corp.: Almaden, CA, 1995.
(22) Schäfer, A.; Horn, H.; Ahlrichs, R. *J. Chem. Phys.* **1992**, *97*, 2571−2577.
(23) Hehre, W. J.; Radom, L.; v. R. Schleyer, P.; Pople, J. A. *Ab initio molecular orbital theory*; Wiley−Interscience: New York, 1986.
(24) Andersson, K.; Malmqvist, P.-Å.; Roos, B. O.; Sadlej, A. J.; Wolinski, K. *J. Phys. Chem.* **1990**, *94*, 5483−5488.
(25) 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: 219−331.

Table 1. Metal Coordination Geometry of the Different (Oxidized) Blue Copper Proteins Considered in This Work[a]

protein		distance to Cu (pm)			angle subtended at Cu (deg)						
		S_{Cys}	N_1, N_2	S_{Met}	N_1-N_2	$S_{Cys}-S_{Met}$	$S_{Cys}-N_1$	$S_{Cys}-N_2$	$S_{Met}-N_1$	$S_{Met}-N_2$	ϕ
plastocyanin	range (8)	202−221	189−222	278−310	96−104	102−110	125−145	112−123	85−95	91−108	77−89
	1plc	207	191, 206	282	97	110	132	121	89	101	82
pseudoazurin	range (4)	214−216	188−227	266−276	94−100	107−109	136−167	110−114	86−88	109−112	70−75
	1paz	214	220, 227	267	100	107	138	112	86	109	74
cucumber basic protein	2cbp	216	193, 195	261	99	111	138	110	83	112	70
nitrite reductase	range (15)	202−236	189−222	246−270	96−102	103−109	129−139	98−109	84−91	122−137	56−65
	1nic	218	202, 203	265	100	106	133	103	88	131	61

[a] Data compiled from all entries in the June 1997 version of the Brookhaven Protein Data Bank.

type,[26] [6s4p3d1f] on Cu, [4s3p1d] on S, and [3s2p] on C, N, whereas the size of the basis set used for H depends on the ligand in which it occurs: [2s] for NH_3, SH^-, and SH_2, but [1s] for imidazole, SCH_3, and $S(CH_3)_2$.

The CASSCF wave function was obtained by distributing 13 electrons in an active space of 12 orbitals, consisting of the Cu 3d orbitals, a correlating 4d shell, and the S_{Cys} 3pσ and π lone-pair orbitals. This active space allows for the description of the ground state and the six lowest excited states in the spectra, i.e., four states that can be formally described as ligand-field states, and two ligand-to-metal charge-transfer states originating from the thiolate group. All calculations were performed without symmetry, and the CASPT2 calculations were based on one state-averaged CASSCF calculation including all seven states. In the CASPT2 calculations the core orbitals, consisting of C, N 1s, and S, Cu 1s−2p were kept frozen. The Cu 3s, 3p semicore electrons were included in the calculations. 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. Both the effect of adding relativistic corrections and of correlating the semicore Cu 3s, 3p electrons on the calculated excitation energies were investigated in detail in our previous study of the plastocyanin spectrum,[16] where it was shown that even if both effects are small (a few hundred reciprocal centimeters), 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.

Based on the results of a series of test calculations on $Cu(imidazole)_2(SH)(SH_2)^+$,[27] a level shift of 0.3 eV was used in all CASPT2 calculations. The effect of the level shift on the second-order energy was removed by a back-correction technique.[28] Finally, oscillator strengths were calculated using the CAS state-interaction method.[29] The transition moments were obtained at the CASSCF level, while CASPT2 excitation energies were used in the expression for the oscillator strength. All CASSCF/CASPT2 calculations were performed with the MOLCAS-3.1 quantum chemistry software,[30] and all calculations were run on IBM RS6000 workstations. Throughout the discussion, a coordinate system is defined with the copper ion at the origin, the *z* axis along the $Cu-S_{Met}$ bond, and the $Cu-S_{Cys}$ bond situated in the *xz* plane.

(3) Models for the Surrounding Protein and Solvent. Calculations of the spectra were also performed on $Cu(imidazole)_2(SH)(SH_2)^+$ without symmetry, using the experimental geometry of the type 1 copper ions in the cucumber basic protein, pseudoazurin, and nitrite reductase.[17,31−33] The model was truncated by hydrogen atoms placed at standard bond lengths in the direction of the removed carbon atom. The copper coordination geometries in the various crystal structures are compared in Table 1.

In order to estimate the effect of the surrounding protein, a second set of calculations was performed on these models, surrounded by a large number of point charges, representing the remaining protein atoms and solvent water molecules. 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 proteins were also charged. The protonation status of the histidine residues were decided from the hydrogen bond pattern: In nitrite reductase, His-26, -60, -231, -255, -306, and -319 were found to have a proton on the $N^{\delta 1}$ atom, His-76 was protonated on the $N^{\epsilon 2}$ atom, whereas His-28, -217, -245, and -260 had hydrogens on both nitrogen atoms. For pseudoazurin, His-6 was found to be protonated on both nitrogen atoms. All other His residues (e.g., the only two residues in cucumber basic protein) are either copper ligands or were not included in the calculations.

Each atom was assigned a point charge, taken from the Amber 4.1 force field.[34] 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^β in Cys and His and C^γ 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).

For cucumber basic protein, the coordinates in the PDB file 2cbp were used.[31] All atoms in the 97 residues of the protein were treated, as well as the 120 crystal water molecules. For several residues, two alternative configurations are given, each with an occupancy of 0.5. In every case, we (quite arbitrarily) chose the A conformation. All half-occupied water molecules were included in the calculations, except HOH-121 which makes close contacts with two other water molecules. In addition, a cap of solvent water molecules with a radius of 2.4 nm was included, centered on the mass center of the molecule. In total, the system contained 5263 atoms.

For pseudoazurin, the crystal structure in the PDB file 2paz (pseudoazurin from Alcaligenes faecalis) was used.[32,33] All 123 residues were included in the simulations, as well as 89 crystal water molecules. A 2.6-nm sphere of solvent water molecules was added to the system as for cucumber basic protein, leading to a total of 6933 atoms.

The calculations of nitrite reductase were based on the coordinates in the PDB file 1nic.[17] This protein is much larger than the other two proteins; it is composed of three identical subunits, each with 340 residues and two copper ions (one type 1 blue copper ion and one normal type 2 copper ion, constituting the active site of the enzyme). Therefore, the protein was truncated: all residues in the trimer with no atoms closer than 3.0 nm of the type 1 copper ion in subunit A were deleted. This resulted in 390 amino acids, 2 copper ions, and 173 crystal water molecules from all three subunits of the protein. A 2.4-nm water cap (720 water molecules) centered on the type 1 copper ion (in subunit A) was included in the calculations, leading to a total

(26) Pierloot, K.; Dumez, B.; Widmark, P.-O.; Roos, B. O. *Theor. Chim. Acta* **1995**, *90*, 87−114.

(27) Roos, B. O.; Andersson, K.; Fülscher, M. P.; Serrano-Andrés, L.; Pierloot, K.; Merchán, M.; Molina, V. *J. Mol. Struct.* (*THEOCHEM*) **1996**, *388*, 257−331.

(28) Roos, B. O.; Andersson, K. *Chem. Phys. Lett.* **1995**, *245*, 215−223.

(29) Malmqvist, P. Å.; Roos, B. O. *Chem. Phys. Lett.* **1989**, *155*, 189−194.

(30) Andersson, K.; Fülscher, M. P.; Karlström, G.; Lindh, R.; Malmqvist, P.-Å.; 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, Lund University, Lund, 1994.

(31) Guss, J. M.; Merritt, E. A.; Phizackerley, R. P.; Freeman, H. C. *J. Mol. Biol.* **1996**, *262*, 686−705.

(32) Inoue, T.; Kai, Y.; Harada, S.; Kasai, N.; Ohshiro, Y.; Suzuki, S.; Kohzuma, T.; Tobari, J. *Acta Crystallogr.* **1994**, *D50*, 317.

(33) Adman, E. T.; Turley, S.; Bramson, R.; Petranos, K.; Banner, D.; Tsernoglou, D.; Beppu, T.; Watanbe, H. *J. Biol. Chem.* **1989**, *264*, 87.

(34) 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.

Table 2. Comparison of the Electronic Spectra of Four Type 1 Copper Proteins

	plastocyanin		cucumber basic protein	
	no crystal	with crystal	no crystal	with crystal
(Cu−S)σ^* → (Cu−S)π^*	5 411(0.0000)	4 836(0.0000)	4 861(0.0001)	5 540(0.0000)
Cu3d$_{z^2}$ → (Cu−S)π^* LF	14 475(0.0000)	12 537(0.0000)	12 878(0.0003)	11 291(0.0001)
Cu3d$_{yz}$ → (Cu−S)π^* LF	15 356(0.0028)	13 556(0.0028)	14 699(0.0022)	13 252(0.0012)
Cu3d$_{xz}$ → (Cu−S)π^* LF	15 084(0.0004)	13 235(0.0003)	14 586(0.0001)	13 002(0.0001)
(Cu−S)π → (Cu−S)π^* CT	15 948(0.1189)	16 970(0.1032)	13 535(0.0674)	15 860(0.0662)
(Cu−S)σ → (Cu−S)π^* CT	22 093(0.0010)	23 156(0.0014)	20 505(0.0387)	23 001(0.0272)

pseudoazurin		nitrite reductase		
no crystal	with crystal	no crystal	with crystal	
4 937(0.0001)	5 660(0.0000)	3 487(0.0000)	4 391(0.0000)	(Cu−S)π^* → (Cu−S)σ^*
12 991(0.0015)	11 680(0.0006)	12 870(0.0009)	11 618(0.0006)	Cu3d$_{z^2}$ → (Cu−S)σ^* LF
14 230(0.0001)	12 673(0.0002)	13 344(0.0002)	11 595(0.0000)	Cu3d$_{yz}$ → (Cu−S)σ^* LF
14 904(0.0008)	13 419(0.0005)	13 860(0.0025)	12 657(0.0014)	Cu3d$_{xz}$ → (Cu−S)σ^* LF
13 967(0.0487)	16 486(0.0491)	12 331(0.0342)	13 863(0.0323)	(Cu−S)π → (Cu−S)σ^* CT
20 235(0.0629)	22 861(0.0494)	19 049(0.0865)	20 772(0.0749)	(Cu−S)σ → (Cu−S)σ^* CT

of 8687 atoms. A few residues in the crystal structure are disordered; we chose the conformation with the highest occupancy, i.e., conformation B for Arg-54, and conformation A for Glu-197, Arg-250, and Arg-271. Disordered crystal water molecules were removed, except DIS-1250, -1275, and -1282. A sulfate ion was also removed.

The positions of the solvent water molecules and the protein hydrogen atoms were determined by a 22.5-ps simulated annealing molecular dynamics simulation. 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 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 30 fs), the neighbor list was updated every 20 time steps (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 the SHAKE algorithm. All heavy atoms present in the PDB files were kept fixed (except for the half-occupied residues and water molecules in cucumber basic protein), 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.[11] The software Amber version 4.1[35] with the Cornell et al. 1995 force field[34] was used. Further computational details can be found from the analogous equilibration of plastocyanin.[16]

Results and Discussion

(1) Electronic Spectra of the Four Proteins. Electronic spectra have been calculated with the CASPT2 method for four blue copper proteins: plastocyanin, cucumber basic protein, pseudoazurin, and nitrite reductase. In Table 2 we present excitation energies and oscillator strengths calculated on the Cu(imidazole)$_2$(SH)(SH$_2$)$^+$ model at the geometry found in the crystal structures.[3,17,31,32] Two sets of results are given for each protein, one calculated with the surrounding protein modeled by an array of point charges, and one without these charges. The results for plastocyanin have been presented before,[16] and they are included to facilitate comparisons.

The various excited states can be labeled by their characteristic singly occupied molecular orbital, i.e., the orbital involved in the excitation. The main atomic components of these orbitals are listed in Table 3. All orbitals are strongly dominated (91−100%) by contributions from Cu 3d and S$_{Cys}$, so only these are given. They are further divided according to their symmetry around the Cu−S$_{Cys}$ bond, i.e., as σ or π. Interestingly, the ground-state orbital (X^{2}A) differs between the various proteins. Although the orbital in all proteins shows a strong covalent interaction between copper and S$_{Cys}$, in plastocyanin it is almost purely of π character (96%), whereas for the other three proteins it is a mixture of both π and σ interactions, with π interactions dominating for cucumber basic protein, but σ interactions dominating for pseudoazurin and even more for nitrite reductase. This can also be seen from the ground-state singly occupied orbitals shown in Figure 1.

Table 3. Composition (%) of the Singly Occupied Orbital in the Ground State and the Different Excited States of the Blue Copper Proteins Considered in This Work[a]

protein	atomic orbital	state						
		X^{2}A	b^{2}A	c^{2}A	d^{2}A	e^{2}A	f^{2}A	g^{2}A
plastocyanin	Cu 3d$_\sigma$	2	73	89	87	21	0	25
	Cu 3d$_\pi$	79	3	9	12	77	19	0
	S$_{Cys}$ 3p$_\sigma$	0	18	0	0	0	0	67
	S$_{Cys}$ 3p$_\pi$	17	0	0	0	0	78	0
cucumber basic protein	Cu 3d$_\sigma$	39	47	75	86	37	1	21
	Cu 3d$_\pi$	53	34	23	13	61	14	1
	S$_{Cys}$ 3p$_\sigma$	5	11	0	0	0	2	70
	S$_{Cys}$ 3p$_\pi$	9	4	0	0	0	82	1
pseudoazurin	Cu 3d$_\sigma$	43	31	71	96	34	1	20
	Cu 3d$_\pi$	29	51	27	4	65	14	0
	S$_{Cys}$ 3p$_\sigma$	8	7	0	0	0	1	70
	S$_{Cys}$ 3p$_\pi$	6	7	0	0	0	83	1
nitrite reductase	Cu 3d$_\sigma$	47	35	66	71	51	3	24
	Cu 3d$_\pi$	32	46	32	27	48	15	0
	S$_{Cys}$ 3p$_\sigma$	10	10	0	0	0	13	66
	S$_{Cys}$ 3p$_\pi$	6	7	0	0	0	69	1

[a] With our definition of the coordinate axes, the sum of the contributions from 3d$_{xy}$ and 3d$_{yz}$ whereas 3d$_\sigma$ is the sum of 3d$_{x^2-y^2}$, 3d$_{z^2}$, and 3d$_{xz}$.

The singly occupied orbital of the lowest excited state (b^{2}A) also shows a strong covalent interaction between copper and S$_{Cys}$. In plastocyanin, it is almost a pure σ antibonding orbital, whereas for the other proteins it is a mixture with approximately the opposite proportion between σ and π interactions compared to the ground-state orbital. Thus, this excitation only involves a redistribution of charge between σ and π orbitals and no charge is transferred between the copper ion and its ligands. The excitation energy is calculated around 5000 cm^{-1} for all proteins.

The singly occupied orbitals of the next three excitations (c^{2}A, d^{2}A, e^{2}A) are almost pure copper orbitals for all four proteins. With our definition of the coordinate system, they can be assigned as excitations to the Cu 3d$_{z^2}$, 3d$_{yz}$, and 3d$_{xz}$ orbitals. Although these excitations, together with the lowest one,

(35) 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.

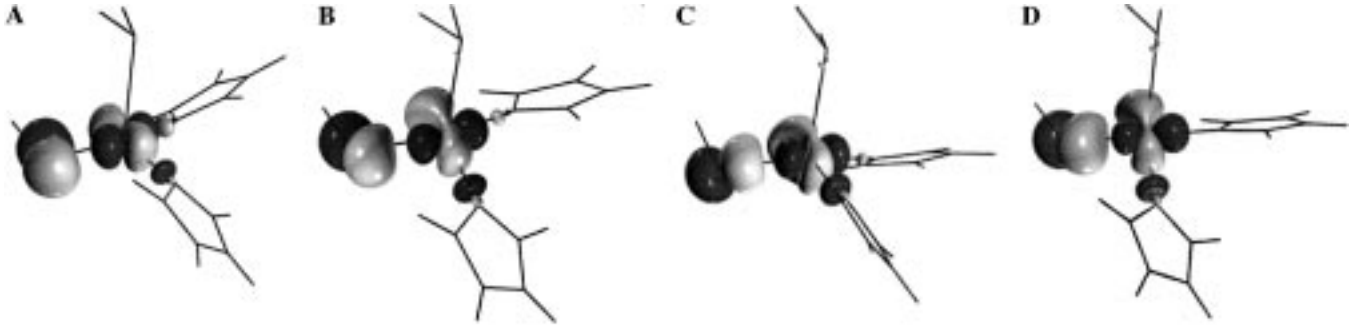

Figure 1. Ground-state singly occupied orbitals of (A) plastocyanin, (B) cucumber basic protein, (C) pseudoazurin, and (D) nitrite reductase, using the $Cu(imidazole)_2(SH)(SH_2)^+$ model at the crystal geometries.

formally should be classified as ligand-field excitations, they involve a significant movement of charge from the copper ion to S_{Cys} (since the singly occupied orbital in the ground state is strongly delocalized onto S_{Cys}). These three excitations occur between 11 000 and 15 000 cm^{-1}.

The last two excitations are charge-transfer transitions. The singly occupied orbital of the f^2A state is the bonding π interaction corresponding to the π antibond in the ground state of plastocyanin (it is an almost pure π state, except in nitrite reductase), whereas the singly occupied orbital of the g^2A state contains a pure $Cu-S_{Cys}$ σ bond in all proteins. As has been noted before for plastocyanin,[16] not very much charge is transferred in these transitions (typically 0.2 *e*). This classification of the excitations is included in Table 3, showing the assignment valid for plastocyanin with a π ground state on the left-hand side and the assignment for a σ ground state on the right-hand side.

As was discussed above, the blue copper proteins show absorption peaks around 600 and 460 nm (17 000 and 22 000 cm^{-1}). These two peaks correspond to the two charge-transfer excitations (f^2A and g^2A in Table 2). Our results indicate that, for plastocyanin, only the 600-nm excitation would be visible, whereas for the other three proteins, both excitations have significant intensities. The reason for this behavior is, of course, the change in the ground-state orbital. The intensity of an excitation is determined mainly by the differential overlap between the ground-state and excited-state orbitals. The blue line in plastocyanin gains its large intensity from the large overlap between bonding and antibonding $Cu-S_{Cys}$ π orbitals in the X^2A and f^2A states. However, in the other proteins, the ground-state orbital is a mixture of π and σ interactions. Therefore, both the f^2A state (with a π bond) and the g^2A state (with a σ bond) gain intensity. Moreover, since the character of the f^2A and g^2A states is fairly constant among the four proteins, the relative intensity of the two excitations is determined mainly by the character of the ground state. Thus, the 600-nm excitation is predicted to be stronger than the 460-nm excitation for cucumber basic protein, whereas for pseudoazurin they are equally bright, and for nitrite reductase, the 460-nm excitation is more intense; the calculated ratios of the oscillator strengths are (with the protein model) 0.01, 0.42, 1.01, and 2.32, which can be compared to the ratio between the experimental absorption coefficients of the 460- and 600-nm excitations, $\epsilon_{460}/\epsilon_{600}$: 0.06, 0.60, 0.41, and 1.22.[36] Apparently, the trends are reasonably reproduced even if the actual values are rather uncertain. However, the results do not reproduce the quantitative trends observed experimentally in the excitation energies of plastocyanin and nitrite reductase. Instead of the increase found for all transitions between nitrite reductase and plastocyanin in Solomon's analysis of the two spectra,[37] a decrease is found, amounting to more than 3000 cm^{-1} for the blue band. The latter is calculated at 16 970 cm^{-1} in plastocyanin, close to the experimental excitation energy of 16 700 cm^{-1}, but in nitrite reductase the calculated result is 13 863 cm^{-1}, almost 4000 cm^{-1} lower than the experimental band at 17 550 cm^{-1}. This discrepancy will be explained below.

The results in Table 2 also clearly show that the effect of the surroundings cannot be neglected: the excitation energy of the two $S_{Cys} \rightarrow Cu$ charge-transfer states increases considerably when the effect of the surroundings is included (with up to 2500 cm^{-1}), whereas the ligand-field excitations, which involve an appreciable charge flow from Cu to S_{Cys}, decrease by almost 2000 cm^{-1}. A considerably smaller effect is found for the lowest transition, which is essentially a transition within the $Cu-S_{Cys}$ bond. It is also notable that the surrounding crystal reduces all oscillator strengths by a factor of up to 1.75.

(2) Geometries. The most important result in the previous section is that the rhombic type 1 copper proteins (cucumber basic protein, pseudoazurin, nitrite reductase) have a ground-state electronic structure that is different from that of the axial type 1 proteins (plastocyanin). The latter proteins have a pure π interaction between copper and S_{Cys}, whereas the former group has a mixture of σ and π interactions. Interestingly, this reminds one of the difference between trigonal and tetragonal copper complexes, which we have studied before.[12] In that investigation, we showed that normal tetragonal (square-planar or elongated octahedral) Cu(II) complexes arise when the singly occupied Cu 3d orbital forms σ bonds with four ligands. This is the effect of a D_{2d} distortion that lifts the orbital degeneracy in a hypothetical tetrahedral structure, which is Jahn−Teller unstable. However, the degeneracy can also be lifted by other distortions. For example, a C_{2v} distortion leads instead to a trigonal structure. With σ donors, such a structure is appreciably less stable than the corresponding tetragonal structures, but with large and polarizable ligands that may form favorable π bonds to copper, the trigonal structures may actually become more stable than the tetragonal ones. In such a structure, the π-bonding ligand formally occupies two ligand positions in a square coordination, whereas the other two positions are occupied by the other two, σ-bonding, ligands in the trigonal plane. Thus, the trigonal structure encountered in plastocyanin and azurin is caused by the strong π bond between copper and S_{Cys} in which considerable amounts of charge is transferred from the thiolate group to copper.

(36) Lu, Y.; LaCroix, L. B.; Lowery, M. D.; Solomon, E. I.; Bender, C. J.; Peisach, J.; Roe, J. A.; Gralla, E. B.; Valentine, J. S. *J. Am. Chem. Soc.* **1993**, *115*, 5907.

(37) LaCroix, L. B.; Shadle, S. E.; Wang, Y.; Averill, B. A.; Hedman, B.; Hodgson, K. O.; Solomon, E. I. *J. Am. Chem. Soc.* **1996**, *118*, 7755−7768.

Table 4. MB3LYP Structures Obtained for a Number of Model Complexes[a]

	distance to Cu (pm)			angle subtended at Cu					
model	S_{Cys}	N	S_{Met}	N−N	S_{Cys}−N	S_{Cys}−S_{Met}	N−S_{Met}	ϕ	ΔE (kJ/mol)
$Cu(NH_3)_2(SH)(SH_2)^+$ trig-cr	217	207	284	103	124	110	93	90	3.4
$Cu(NH_3)_2(SH)(SH_2)^+$ trig-opt	217	207	282	103	126	101	94	90	0.0
$Cu(NH_3)_2(SH)(SH_2)^+$ tetr	226	205−206	257	143	93−100	134	94−101	34−55	12.6
$Cu(NH_3)_2(SH)(S(CH_3)_2)^+$ trig-cr[b]	218	209	258	102	121	114	98	90	6.5
$Cu(NH_3)_2(SH)(S(CH_3)_2)^+$ trig-opt[b]	218	208	259	103	124	104	99	90	2.8
$Cu(NH_3)_2(SH)(S(CH_3)_2)^+$ tetr	221	209−214	240	99	94−142	105	96−123	63	0.0
$Cu(NH_3)_2(SCH_3)(S(CH_3)_2)^+$ trig-cr	218	211	258	103	119	121	96	90	7.4
$Cu(NH_3)_2(SCH_3)(S(CH_3)_2)^+$ trig-opt	218	210	259	103	124	102	98	90	0.0
$Cu(NH_3)_2(SCH_3)(S(CH_3)_2)^+$ tetr[c]	219	210−221	240	101	99−131	118	99−106	80	2.4
$Cu(imidazole)_2(SH)(S(CH_3)_2)^+$ trig-cr[b]	219	203	262	104	121	116	97	90	9.3
$Cu(imidazole)_2(SH)(S(CH_3)_2)^+$ trig-opt[b]	219	203	263	103	124	101	99	90	6.5
$Cu(imidazole)_2(SH)(S(CH_3)_2)^+$ tetr	223	205−206	242	100	97−141	103	95−126	62	0.0
$Cu(imidazole)_2(SCH_3)(S(CH_3)_2)^+$ trig-cr	218	204	267	103	120−122	116	94−95	90	0.0
$Cu(imidazole)_2(SCH_3)(S(CH_3)_2)^+$ tetr[c]	220	204−213	242	102	102−132	115	99−104	81	2.6
plastocyanin	211	201−208	288	100	119−132	107	89−102	82	
nitrite reductase	217	200−202	258	99	104−134	106	88−131	61	

[a] For most models, a tetragonal and two trigonal structures are listed. The two trigonal structures differ in the $H/C_{Cys}-S_{Cys}-Cu-S_{Met}$ dihedral angle, which is 180° in the optimal structure (trig-opt), but 0° in the crystal structure (trig-cr; cf. the text). [b] A transition state obtained using C_s symmetry. [c] These structures are actually also trigonal but with one of the imidazole groups rather than methionine as the axial ligand, cf. the text.

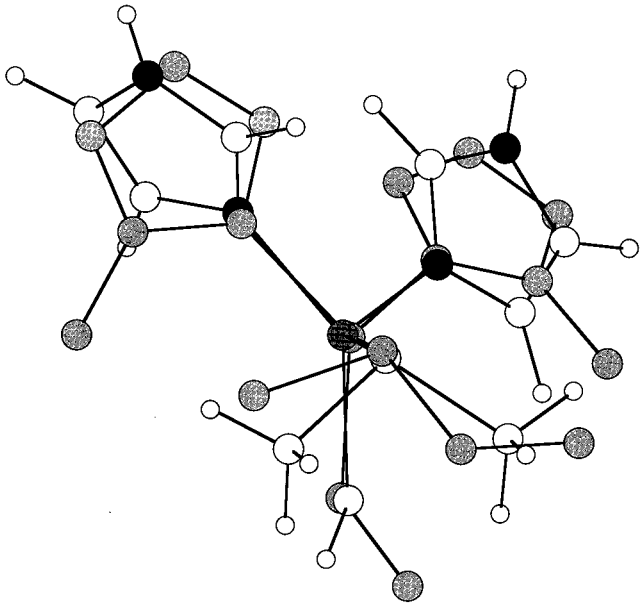

Figure 2. Comparison between the optimized tetragonal structure of $Cu(imidazole)_2(SH)(S(CH_3)_2)^+$ and the crystal structure of oxidized nitrite reductase (shaded).

However, also in a tetragonal structure involving Cu(II) and thiolate, much charge is transferred from thiolate to copper. This causes the geometry of such complexes to become rather similar to the structure observed for Cu(I) complexes, i.e., quite close to a tetrahedron. In fact, it turns out that the tetragonal structure of models of the ligand sphere in the blue copper proteins ($CuHis_2CysMet$) is not at all square-planar, but instead very similar to the structure encountered in nitrite reductase. This is illustrated in Figure 2, where the optimized structure of our largest model, $Cu(imidazole)_2(SH)(S(CH_3)_2)^+$, is compared to the crystal structure of nitrite reductase. Except for a too short $Cu-S_{Met}$ bond in the optimized structure (it is well documented that B3LYP gives a too short distance for this bond[11]), the two structures are closely similar. This interpretation is also supported by calculations of the electron spin resonance spectra of model systems as well as protein models;[38] the spectra of tetragonal models invariably show a pronounced rhombic distortion, whereas those of trigonal models are close to axial.

In Table 4, trigonal and tetragonal structures of a number of models of the blue copper proteins (from $Cu(NH_3)_2(SH)(SH_2)^+$ to $Cu(imidazole)_2(SCH_3)(S(CH_3)_2)^+$) are compared to each other and to the crystal structures of plastocyanin and nitrite reductase (average values of the 8 and 15 structures present in the June 1997 version of Brookhaven protein data bank). The tetragonal structures always give a longer $Cu-S_{Cys}$ bond and a shorter $Cu-S_{Met}$ bond than the trigonal structures. This is in good agreement with the structural difference between plastocyanin and nitrite reductase.

Moreover, whereas the trigonal structures have very similar S−Cu−N angles for the two nitrogen ligands (as an effect of being close to C_s symmetry), the tetragonal structures always have distinctly different S−Cu−N angles for the two nitrogen ligands. Again, this is in agreement with the crystal structures, although plastocyanin shows a larger variation in the angles than the optimized structures. This difference between the two types of structures is a simple effect of the geometry of the complexes. In an ideal trigonal complex, all angles between the axial ligand and the other ligands are 90°, as is the angle between the two σ-bonded trigonal ligands, whereas the other two angles in the trigonal plane should be 135°. Thus, the two $S_{Cys}-Cu-N$ angles in the trigonal complexes should be large and equal (∼120° in the optimized complexes) and the two $S_{Met}-Cu-N$ angles should be close to 90° (93−99°). In the ideal square-planar complex where the two nitrogen ligands occupy neighboring positions, on the other hand, one of the $S_{Cys}-Cu-N$ angles is 90° and the other is 180°, and the same holds true for the two $S_{Met}-Cu-N$ angles. Due to the strong tetrahedral distortion of the trigonal complexes, the angles in the optimized complexes are more displaced from their ideal values, so the $S_{Cys}-Cu-N$ angles are around 95° and 140° and the $S_{Met}-Cu-N$ angles are around 95° and 123° in the optimized complexes.

Examining the angles of the tetragonal complexes, a variation in the structures obtained with different ligand models is observed. The tetragonal $Cu(NH_3)_2(SH)(SH_2)^+$ complex falls out by having a very small difference between the S−Cu−N angles of the two ammonia ligands. The reason is that this complex has the ammonia ligands in opposite (trans) positions.

(38) Eriksson, L.; Ryde, U. Calculations of ESR parameters of rhombic and axial blue copper proteins. Manuscript in preparation, 1998.

We have tried to optimize a tetragonal $Cu(NH_3)_2(SH)(SH_2)^+$ complex with the ammonia ligands in cis positions (as in the protein), but such a complex reorganizes to the trigonal structure. Second, the $Cu(NH_3)_2(SCH_3)(S(CH_3)_2^+$ and $Cu(imidazole)_2(SCH_3)(S(CH_3)_2)^+$ complexes deviate from the other complexes by having rather large ϕ angles (80° compared to 62°; ϕ is the angle between the N−Cu−N and S_{Cys}−Cu−S_{Met} planes), smaller differences between the S−Cu−N angles, and one long Cu−N bond (but still with a longer Cu−S_{Cys} and a shorter Cu−S_{Met} bond than in the trigonal structures). An examination of the electronic structure of these complexes shows that they actually have a Cu−S_{Cys} bond of mainly π character. Thus, the structure is best described as (chiefly) trigonal with one of the imidazoles as an axial ligand, whereas the methionine model is the third ligand in the (approximate) trigonal plane. Again, all attempts to optimize a true tetragonal structure have been fruitless.

For the models with the same set of ligands, we can calculate the relative energy of the trigonal and tetragonal complexes. This is done in the last column in Table 4. For each trigonal structure, two entries are provided. One is the conformation encountered in the crystal structures (trig-cr). However, this conformation is not optimal in vacuum. Instead, a structure with a H/C_{Cys}−S_{Cys}−Cu−S_{Met} dihedral angle of 180° (instead of 0°) is more stable (by 3−7 kJ/mol; trig-opt). As has been discussed before,[11] this conformation is stabilized by hydrogen bond interactions between the methionine model and S_{Cys}. Yet, in the protein, better hydrogen bond donors are present, making the crystal conformation more favorable. From the result in Table 4, it can be seen that all three conformations are very close in energy. Except for the $Cu(NH_3)_2(SH)(SH_2)^+$ model, where the tetragonal structure has the nitrogen ligands in trans positions, the energy difference between the trigonal and tetragonal structures is less than 7 kJ/mol. Interestingly, the tetragonal structure is most stable if methionine is modeled by $S(CH_3)_2$ and cysteine by SH^-, whereas the trigonal structure is most stable otherwise.

However, it should be noted that no true equilibrium states of the trigonal structures could be obtained with the $Cu(NH_3)_2(SH)(S(CH_3)_2)^+$ and $Cu(imidazole)_2(SH)(S(CH_3)_2)^+$ models. The structures presented in Table 4 are optimized with C_s symmetry; if the symmetry restriction is removed and the structures are reoptimized, the geometry reorganizes to the tetragonal structure. In fact, the only system for which both structures are equilibrium states is $Cu(NH_3)_2(SH)(SH_2)^+$ (for $Cu(NH_3)_2(SH)(S(CH_3)_2)^+$, the trigonal structure is a transition state). Thus, the relative energies must be examined with care. Still, it is clear that the two types of structures have very similar energies and that the potential surface is extremely sensitive to the models used for the copper ligands. This means that the structure should also be very sensitive to the protein environment, which explains why some proteins stabilize the trigonal structure, whereas other stabilize the tetragonal structure, although the copper ligands are the same. In fact, free energy perturbations indicate that plastocyanin stabilizes the trigonal structure by favoring its bond lengths and electrostatics, whereas nitrite reductase prefers the angles of the tetragonal structure.[39]

(3) Relation between Geometry and Spectroscopy. Our results up to now indicate that there is a strong correlation between the structures and spectra of the blue copper proteins: Trigonal structures give rise to spectra with a low $\epsilon_{460}/\epsilon_{600}$ ratio, whereas tetragonal structures give rise to spectra with high ratios. In order to study this relation and its applicability to other types of copper−cysteinate proteins more closely, we performed a series of geometry optimizations of $Cu(NH_3)_2(SH)(SH_2)^+$ at different values of the ϕ angle between the S_{Cys}−Cu−S_{Met} and N−Cu−N planes. First, we optimized the structures with the B3LYP method, constraining the ϕ angle and optimizing all other geometrical parameters (for three structures, ϕ = 75°−85°, the methionine model tended to dissociate; therefore, the Cu−S_{Met} distance was also constrained to values obtained by extrapolating the results for ϕ = 70° and 90°).

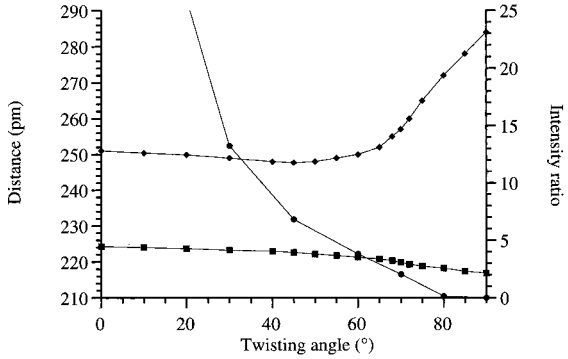

Figure 3. Cu−S_{Cys} (squares) and Cu−S_{Met} (diamonds) distances (optimized by the CASPT2 method), and the calculated quotient of the oscillator strengths of the lines around 460 and 600 nm (bullets) as a function of the ϕ angle, using the $Cu(NH_3)_2(SH)(SH_2)^+$ model.

Second, the Cu−S distances were reoptimized with the more accurate CASPT2 method, keeping the rest of the geometry fixed, since we know that the B3LYP method gives significant errors for these distances.[11] This led to a decrease in the Cu−S_{Cys} bond distance (by 3−8 pm) as well as in the Cu−S_{Met} distance (by 6−10 pm; smaller differences were obtained for the structures where this distance was constrained in the B3LYP optimizations, but at the CASPT2 level, the methionine model showed no tendency to dissociate). The resulting distances in Figure 3 clearly show that the Cu−S_{Met} bond length decreases and the Cu−S_{Cys} bond length increases when going from the trigonal structure (ϕ = 90°) to the tetragonal structures ($\phi \approx$ 60°). However, only small changes in the bond distances are seen when the ϕ angle is further decreased to the ideal square-planar structure (ϕ = 0°) (Figure 3).

Finally, electronic spectra were calculated on seven representative structures using the CASPT2 method. The results are shown in Table 5. The assignment of the excitations are the same as in Table 2. However, both the trigonal and square-planar structures possess C_s symmetry, and the different excited states are labeled according to their symmetry representation. Looking at the spectrum of the trigonal structure first (ϕ = 90°), we recognize the main characteristics of the plastocyanin spectrum. The c^2A'' state, which gives rise to the dominant blue band, is calculated at 15 042 cm^{-1}, while the second S_{Cys} → Cu charge-transfer band is found at 22 036 cm^{-1} with little intensity. As for $Cu(imidazole)_2(SH)(SH_2)^+$, the only ligand-field state giving rise to a transition with significant intensity is b^2A''.

In the square-planar structure (ϕ = 0°), the situation is reversed. There, all ligands are located in the *xz* plane, and the ground-state singly occupied orbital is a σ-antibonding orbital between the Cu $3d_{z^2-x^2}$ orbital and the four ligands, containing a considerable amount (21%) of S_{Cys} $3p_\sigma$ character and a much smaller amount of S_{Met} and NH_3 character, 3% and 2%, respectively. Thus, the ground state is X^2A' and the Cu−S_{Cys} $\sigma \rightarrow \sigma^*$ excitation to the d^2A' state has by far acquired the largest

(39) De Kerpel, J. O. A.; Ryde, U. Protein strain in blue copper proteins studied by free energy perturbations. *Proteins, Struct., Funct., Genet.*, submitted.

Table 5. Calculated Spectrum of $Cu(NH_3)_2(SH)(SH_2)^+$ as a Function of the Angle ϕ between the $S_{Cys}-Cu-S_{Met}$ and N−Cu−N Planes[a]

	$\phi = 90°$	$\phi = 80°$	$\phi = 70°$	$\phi = 60°$	$\phi = 45°$	$\phi = 30°$	$\phi = 0°$	
a^2A' (Cu−S)σ^* → (Cu−S)π^*	4 313(0.0000)	3 655(0.0001)	3 833(0.0001)	5 416(0.0000)	7 456(0.0001)	9 131(0.0000)	10 583(0.0000)	a^2A'' (Cu−S)π^* → (Cu−S)σ^*
b^2A' Cu3d_{z^2} → (Cu−S)π^* LF	12 842(0.0000)	12 334(0.0000)	12 287(0.0005)	13 785(0.0009)	16 097(0.0016)	17 256(0.0014)	18 164(0.0013)	b^2A' Cu3d_{xz} → (Cu−S)σ^* LF
c^2A' Cu3d_{xz} → (Cu−S)π^* LF	14 774(0.0000)	14 252(0.0001)	14 423(0.0006)	15 243(0.0006)	16 661(0.0001)	18 757(0.0006)	20 005(0.0007)	c^2A' Cu3d_{z^2} → (Cu−S)σ^* LF
b^2A'' Cu3d_{yz} → (Cu−S)π^* LF	14 861(0.0019)	14 206(0.0023)	14 766(0.0013)	16 023(0.0011)	17 519(0.0012)	18 868(0.0007)	20 824(0.0000)	b^2A'' Cu3d_{yz} → (Cu−S)σ^* LF
c^2A'' (Cu−S)π → (Cu−S)π^* CT	15 042(0.1225)	14 575(0.1053)	14 981(0.0441)	15 453(0.0296)	16 432(0.0191)	17 670(0.0108)	18 866(0.0007)	c^2A'' Cu−S)π → (Cu−S)σ^* CT
d^2A' (Cu−S)σ → (Cu−S)π^* CT	22 036(0.0008)	21 427(0.0158)	21 412(0.0902)	21 638(0.1128)	22 053(0.1305)	22 440(0.1432)	22 527(0.1560)	d^2A' Cu−S)σ → (Cu−S)σ^* CT

[a] The axes are chosen such that S_{Met} remains on the *z* axis and S_{Cys} in the *xz* plane along the entire rotation path, while the plane containing the two NH_3 groups is rotated.

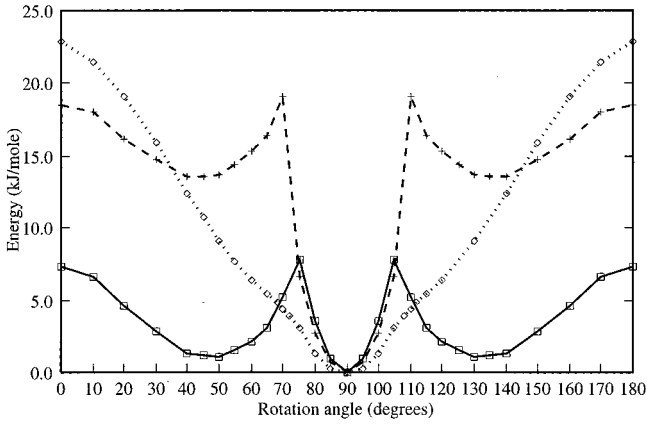

Figure 4. B3LYP (dotted line) and the CASPT2 energies with (solid line) and without (dashed line) relativistic corrections calculated for the $Cu(NH_3)_2(SH)(SH_2)^+$ model at the geometry optimized by the B3LYP method and (for the CASPT2 energies) with the Cu−S distances reoptimized by the CASPT2 method.

intensity, whereas the excitation to the c^2A'' state has almost completely vanished. The intensity of the ligand-field states also reflects the change in ground state: the intensity of the b^2A'' state has dropped to zero, while the b,c^2A' states now gain intensity from the presence of a small amount of S_{Cys} $3p_\sigma$ character in the corresponding singly occupied orbitals. All excitation energies have increased in the square-planar structure as compared to the trigonal structure. The strongest increase, around 6000 cm^{-1}, is found for the first excited state and the ligand-field states, reflecting the stronger ligand field exerted in the square-planar structure, with four, rather than three, strongly bound ligands. The excitation energies of the two charge-transfer states also have higher excitation energies, but to a smaller extent. For the c^2A'' state (π), an increase of 3800 cm^{-1} is found, whereas the d^2A' state (σ) is raised by less than 500 cm^{-1}.

Even if the calculations were performed on a simple model, the calculated $\epsilon_{460}/\epsilon_{600}$ ratio presented in Figure 3 nicely reflects the structure−electronic spectroscopy relationship between the different types of copper cysteinate proteins. The copper coordination geometry of axial type 1 proteins is close to trigonal, and their spectroscopic characteristics are reflected by the results obtained for $\phi = 90°$. Rhombic type 1 proteins such as pseudoazurin and cucumber blue protein on the other hand typically have ϕ angles between 70° and 80°. As can be seen from Table 5, even at such a small deviation from orthogonality, the transition to the d^2A' state has already gained significant intensity due to mixing of σ character into the ground-state singly occupied orbital. The largest deviation from orthogonality within the type 1 copper proteins is found for nitrite reductase from A. cycloclastes, which has $\phi = 56°-65°$.[17] Accordingly, the most important spectroscopic characteristics of this protein are reflected in the results obtained for $\phi = 60°$. Interestingly, all calculated excitation energies are higher than for $\phi = 90°$, in agreement with the shifts found in the experimental nitrite reductase spectrum compared to plastocyanin.[37] The redistribution of the intensities of the two charge-transfer bands between $\phi = 90°$ and $\phi = 60°$ is also notable, with the second transition being the most intense at $\phi = 60°$. This is in accord with the green color of nitrite reductase.

Figure 4 shows the energy of the $Cu(NH_3)_2(SH)(SH_2)^+$ models as a function of the ϕ angle, obtained with three different methods. The dotted line shows the energies obtained by the B3LYP method on the B3LYP-optimized structures. The solid and the dashed lines were obtained by the CASPT2 method on

Table 6. Comparison between the Calculated and Experimental Spectra of Plastocyanin and Nitrite Reductase[a]

	plastocyanin		nitrite reductase		
	calc	exp	exp	calc	
(Cu−S)σ^* → (Cu−S)π^*	4 363(0.0001)	5 000(0.0000)	5 600(0.0000)	4 408(0.0000)	(Cu−S)π^* → (Cu−S)σ^*
Cu3d$_{z^2}$ → (Cu−S)π^* LF	11 645(0.0000)	10 800(0.0031)	11 900(0.0026)	12 329(0.0003)	Cu3d$_{z^2}$ → (Cu−S)σ^* LF
Cu3d$_{yz}$ → (Cu−S)π^* LF	12 981(0.0025)	12 800(0.0114)	13 500(0.0086)	12 872(0.0004)	Cu3d$_{yz}$ → (Cu−S)σ^* LF
Cu3d$_{xz}$ → (Cu−S)π^* LF	12 671(0.0002)	13 950(0.0043)	14 900(0.0101)	13 873(0.0028)	Cu3d$_{xz}$ → (Cu−S)σ^* LF
(Cu−S)π → (Cu−S)π^* CT	15 654(0.1162)	16 700(0.0496)	17 550(0.0198)	15 789(0.0325)	(Cu−S)π → (Cu−S)σ^* CT
(Cu−S)σ → (Cu−S)π^* CT	21 974(0.0012)	21 390(0.0035)	21 900(0.0299)	22 461(0.1192)	(Cu−S)σ → (Cu−S)σ^* CT

[a] The calculations were performed on the Cu(imidazole)$_2$(SH)(SH$_2$)$^+$ model at the crystal geometry but with the Cu−S$_{Cys}$ and Cu−S$_{Met}$ distances taken from the CASPT2 optimizations of Cu(NH$_3$)$_2$(SH)(SH$_2$)$^+$.

the structures optimized by the B3LYP method with the Cu−S distances reoptimized with the CASPT2 method. They differ in whether relativistic effects are included (solid line) or not (dashed line). The three curves differ quite appreciably. The B3LYP energy shows only one minimum, the trigonal structure, whereas the tetragonal structure is not even a transition state (actually, there is a local minimum representing a tetragonal structure, but as has been discussed before, this structure has the two ammonia ligands in opposite positions, and it is therefore not obtained by the twist described by the ϕ angle). On the other hand, both CASPT2 energy curves show a local minimum for the tetragonal state around $\phi = 50°$. Yet, the relative energies of this minimum and the transition state between this structure and the trigonal structure are very different. With relativistic effects, the tetragonal structure is only 1 kJ/mol less stable than the trigonal one and the transition state is 8 kJ/mol above the trigonal structure. Without relativistic effects, the two energies are ∼12 kJ/mol higher.

As has been discussed earlier,[12] the sizeable effect of the relativistic corrections is due to the large difference in the Cu 3d population between the trigonal and tetragonal structures in the CASSCF calculations. The trigonal structures have a Cu 3d population close to 10, whereas in the tetragonal structures, the population is close to 9. Relativistic effects always favor the structure with the lowest Cu 3d population and therefore lead to a strong stabilization of the tetragonal structure. The cusps, which appear in the two CASPT2 curves in the plot, are also caused by the difference in Cu 3d population. The CASSCF calculations always converge to either a d^9 or a d^{10} state. At small ϕ angles, the d^{10} state is lower in energy, but for ϕ angles between 70° and 80°, the d^9 state becomes lower and this change in the electronic state gives rise to the cusps. In the B3LYP calculations, the Cu 3d population is ∼9.5 and approximately the same for the trigonal and tetragonal structures. Therefore, only small effects are expected if relativistic corrections were included in these calculations (which unfortunately is not possible with the programs used).

Thus, these calculations confirm that the energy difference between the trigonal and tetragonal state of models of the blue copper is very small and they also show that the barrier between the two structures is low. If the surrounding protein matrix does not change the energies significantly, this means that both structures would have significant populations at normal temperatures and that the geometries as well as the spectra observed experimentally should be a Boltzmann-averaged mixture of the geometries and spectra obtained for the trigonal and tetragonal states (and also all intermediate states). However, the distribution is very sensitive to the relative energies. If the energy difference is 10 kJ/mol higher, the population of the less stable state would only be ∼1% of the one in the more stable state and would hardly be detected. Moreover, it is also feasible that the proteins may stabilize states with geometries intermediate between the trigonal and tetragonal structures. According to Table 5, such structures would also give rise to an intermediate spectrum. Thus, it is hard to tell from experiments or from theoretical calculations, with the accuracy achievable today, whether the intermediate structures and spectra observed for cucumber basic protein and pseudoazurin are the result of a Boltzmann average of two structures or whether they are obtained from a distinct intermediate structure stabilized by the protein surrounding.

(4) Comparison of the Experimental and Calculated Spectra for Plastocyanin and Nitrite Reductase. It is notable that the spectra calculated on the small Cu(NH$_3$)$_2$(SH)(SH$_2$)$^+$ model at various ϕ angles (Table 5) reproduced the experimentally observed increase in all excitation energies between plastocyanin and nitrite reductase, whereas the spectra obtained on the crystal structures failed to do so (Table 2). This indicates that the large error found in the latter data should not be traced back to deficiencies in the theoretical approach. Instead, it seems likely that it is at least partly caused by the uncertainty in the crystal structures.

In order to check this, a second set of calculations was performed on the spectra of plastocyanin and nitrite reductase, using the same crystal structures, but replacing the Cu−S$_{Met}$ and Cu−S$_{Cys}$ bond distances by their CASPT2 value as optimized for the Cu(NH$_3$)$_2$(SH)(SH$_2$)$^+$ model (see the previous section) at ϕ angles 82° and 61° (the angles found in the crystal structures), respectively. For plastocyanin, this leads to a lengthening of the Cu−S$_{Cys}$ bond, from 207 to 210 pm, while in nitrite reductase the corresponding bond is shortened, from 218 to 214 pm. The differences for the Cu−S$_{Met}$ bonds are larger and in both cases this bond is shortened: from 282 to 271 pm in plastocyanin, and from 265 to 241 pm in nitrite reductase. Table 6 contains the results of this second set of calculations, together with the experimental band positions found for the two proteins. The results shown were obtained for the naked Cu(imidazole)$_2$(SH)(SH$_2$)$^+$ cluster and were corrected afterward for the effect of the crystal surrounding using the results of Table 1 (no correction was applied to the oscillator strengths, however).

The effect of the change in geometry is profound. For plastocyanin, the excitation energies of the two charge-transfer states decrease by 1316 and 1766 cm^{-1}, respectively, whereas in nitrite reductase, the charge-transfer excitation energies increase by 1926 and 1689 cm^{-1}. Similar, but smaller effects are found for the four lowest states in the spectra: all excitation energies decrease for plastocyanin and increase for nitrite reductase when the Cu−S bond lengths are changed. As a result, the experimentally observed trends in the excitation energies between nitrite reductase and plastocyanin are well reproduced. Except for the second ligand-field state, all calculated excitation energies are higher for nitrite reductase than for plastocyanin, with differences up to 1200 cm^{-1} (the experimental differences range between 500 and 1100 cm^{-1}). The overall agreement between the calculated and experimental

excitation energies is also satisfactory. The largest error is still found for the blue band in nitrite reductase, but it is reduced to less than 1800 cm^{-1}, which is within the uncertainty of the CASPT2 method.

In fact, these changes in the spectra are mainly an effect of the change in the Cu−S_{Cys} bond length. The change in this bond length (an increase by 3 pm in plastocyanin and a decrease by 4 pm in nitrite reductase) is within the uncertainty of the crystal data (typically ∼7 pm[40]). However, the change in the Cu−S_{Met} distances, 11−24 pm, is larger than the experimental uncertainty. The reason for this is probably that a too small model was used in the CASPT2 optimizations. It is well documented that the Cu−S_{Met} distance is very sensitive to the ligand models, and a partial CASPT2 optimization of this bound length with the slightly large model $Cu(NH_3)_2(SH)(S(CH_3)_2)^+$ gave actually an *increase* by 6.5 pm for plastocyanin.[11] Therefore, we have also calculated the spectra for plastocyanin and nitrite reductase with only the Cu−S_{Cys} bond length from the CASPT2 calculation (the rest of the structure was taken from the crystal data). This gave only small changes in the excitation energies and oscillator strengths compared to the results in Table 6 (−20 to +330 cm^{-1} for plastocyanin, and −16 to −610 cm^{-1} for nitrite reductase).

In summary, it is clear that changes in geometry as well as in the electrostatic properties of the surrounding protein may affect the calculated transition energies substantially and that the uncertainties connected with the former as well as the approximate model used for the latter may seriously affect the theoretical results obtained for the spectra. All CASPT2 results obtained in this and the previous sections are therefore afflicted with some uncertainties. However, it has been demonstrated that the quality of these results is high enough to provide an understanding of the spectroscopic characteristics of the different types of cysteine-containing copper proteins, in relation to their structure.

(5) Other Types of Copper Proteins. The relevance of the results in Table 5 and Figure 3 is not restricted to the type 1 copper proteins. By site-directed mutagenesis, copper sites involving a cysteine ligand have been produced, which have properties that diverge from the rhombic type 1 copper proteins by even higher $\epsilon_{460}/\epsilon_{600}$ ratios and different resonance Raman spectra. These copper sites have been termed type 1.5 and type 2 copper sites. The results in Figure 3 show that when the copper complex becomes more flattened tetragonal than in the rhombic type 1 proteins, the excitation around 460 nm gains even more in intensity at the expense of the 600-nm band. Moreover, it can also be seen that the Cu−S_{Cys} bond length increases slightly, which would lead to changes in the resonance Raman spectrum. Thus, it seems most likely that the type 1.5 and 2 copper−cysteinate proteins have tetragonal structures that are more flattened than the rhombic type 1 proteins.

In order to investigate this suggestion more closely, we have calculated the structures and spectra of some more realistic models of type 1.5 and 2 mutants. Table 7 shows the spectra obtained for two such models: $Cu(NH_3)_3(SH)^+$, serving as a model for the Met121His azurin mutant, and $Cu(NH_3)(OH_2)_2(SH)^+$, as a tentative model for His117Gly(H_2O) azurin, a mutant obtained by replacing one histidine ligand by two water molecules[41] (the structure was obtained by first optimizing $Cu(NH_3)(OH_2)_2(SH)(SH_2)(H_2CO)^+$ and then removing the axial ligands). MB3LYP-optimized structures were used, with only the Cu−S_{Cys}, distance reoptimized at the CASPT2 level. The optimum $Cu(NH_3)_3(SH)^+$ structure has an angle $\phi = 53°-54°$ (depending on which of the two S_{Cys}−Cu−N planes is considered), while $Cu(NH_3)(OH_2)_2(SH)^+$ is considerably more flattened, with $\phi = 28°-29°$. This is in accord with the classification of Met121His azurin as type 1.5 and His117Gly(H_2O) azurin as type 2. The difference between the two types is also reflected in the calculated spectra. The excitation from the (Cu−S)σ bonding orbital is by far the most intense in both cases, but more in $Cu(NH_3)(OH_2)_2(SH)^+$ than in $Cu(NH_3)_3(SH)^+$, while the intensity of the blue band decreases in the same order. Finally, it is also notable that the calculated Cu−S_{Cys} $\sigma \rightarrow \sigma^*$ excitation energy is considerably (more than 1000 cm^{-1}) higher for $Cu(NH_3)(OH_2)_2(SH)^+$ than for $Cu(NH_3)_3(SH)^+$ or $Cu(NH_3)_2(SH)(SH_2)^+$. This is in agreement with the experimentally observed blue-shift of the second $S_{Cys} \rightarrow$ Cu charge-transfer band when going from type 1 to type 2 copper−cysteinate proteins, from 460 (21 800 cm^{-1}) to 410 nm (24 400 cm^{-1}).[40]

Table 7. Calculated Spectrum of $Cu(NH_3)_3(SH)^+$ and $Cu(NH_3)(OH_2)_2(SH)^+$ as Models of Type 1.5 and Type 2 Cu−Cysteinate Proteins, Respectively[a]

$Cu(NH_3)_3(SH)^+$	$Cu(NH_3)(OH_2)_2(SH)^+$	
7 400(0.0000)	9 501(0.0000)	(Cu−S)$\pi^* \rightarrow$ (Cu−S)σ^*
14 053(0.0001)	17 292(0.0010)	Cu3d $\rightarrow$ (Cu−S)σ^* LF
14 617(0.0000)	18 187(0.0001)	Cu3d $\rightarrow$ (Cu−S)σ^* LF
15 288(0.0020)	18 414(0.0019)	Cu3d $\rightarrow$ (Cu−S)σ^* LF
16 086(0.0157)	16 212(0.0029)	(Cu−S)$\pi \rightarrow$ (Cu−S)σ^* CT
22 559(0.1340)	23 634(0.1728)	(Cu−S)$\sigma \rightarrow$ (Cu−S)σ^* CT

[a] The calculations were performed on MB3LYP-optimized models with the Cu−S_{Cys} distance reoptimized with the CASPT2 method.

In conclusion, the theoretical studies show that the most pronounced difference among the various types of copper proteins is between the axial and the rhombic type 1 proteins, i.e., between a tetragonal σ-bonded structure and a trigonal π-bonded structure. For the other groups, there is no distinct difference; the three tetragonal types differ mainly by the extent of tetrahedral distortion and the number and properties of the ligands.

Conclusions

We have presented theoretical calculations of the structure and the electronic spectra of several blue copper protein models. The geometries were optimized using the B3LYP density functional method and the spectra were computed using multiconfigurational wave functions with dynamic correlation effects estimated by second-order perturbation theory (CASPT2). The effect of the protein surrounding on the spectra have been simulated by a simple point-charge model. The calculated spectra reproduce the experimental spectra with an error of less than 1800 cm^{-1}. Interestingly, the spectra depend strongly on the geometry of the copper complex, especially on the Cu−S_{Cys} bond length, and it has turned out to be necessary to reoptimize this bond length with the CASPT2 method, in order to obtain accurate spectra.

We have shown that the spectral differences that exist between the axial and rhombic type 1 copper proteins originate from differences in the ground-state electronic structure of the copper complex; the axial proteins exhibit a π bond between the S_{Cys} atom and the copper ion, whereas in the rhombic proteins, this bond exhibits a mixture of σ and π character. In fact, these two electronic structures represent two local minimums on the potential surface of the copper coordination sphere. The

(40) 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.

(41) den Blaauwen, T.; Hoitinink, C. W. G.; Canters, G. W.; Han, J.; Loehr, T. M.; Sanders-Loehr, J. *Biochemistry* **1993**, *32*, 12455−12464.

π-bonded structure gives rise to a trigonal geometry, whereas the σ-bonded structure forms a tetragonal geometry. For normal hard ligands, the tetragonal structure is close to square-planar (typically with one or two additional axial ligands), as is observed for most small inorganic complexes. However, with a large and soft (polarizable) ligand, such as the cysteine thiolate group, much charge is transferred from the thiolate group to copper, yielding an effective charge on copper close to +1 and therefore a structure that is strongly distorted toward a tetrahedron. Therefore, both the trigonal and tetragonal structures of models of the blue copper proteins are rather similar and quite tetrahedral, which explains their low reorganization energies (Cu(I) is typically tetrahedral). Interestingly, the energy difference between the two structures is low (less than 7 kJ/mol) and the barrier between them is low. Therefore, some proteins stabilize the trigonal structure (e.g., plastocyanin), whereas others stabilize the tetragonal structure (e.g., nitrite reductase) although they have exactly the same copper ligands.

The difference between the trigonal and tetragonal structures can be described by the angle between the N−Cu−N and S_{Cys}−Cu−S_{Met} planes (ϕ); trigonal structures have ϕ close to 90°, whereas tetragonal structures have $\phi < 80°$. We have shown that most of the geometric and spectral differences between the axial and rhombic type 1 copper proteins can be reproduced by model complexes optimized at $\phi = 90°$ and $\phi = 60°$ (cf. Figure 3). However, this comparison is not restricted to the type 1 copper proteins. Complexes obtained with ϕ around 50° and 30° are good models for the so-called type 1.5 and type 2 copper proteins, respectively. Our modeling studies indicate that type 1 copper sites form when there are three strong and one or two weak ligands (thioether or carbonyl), whereas type 1.5 structures arise in complexes with four strong ligands, and additional axial ligands seem to be needed for type 2 structures. Thus, together with our study of the copper site in stellacyanin[42] (which has a trigonal structure but a rhombic spectrum due to a rather strong axial ligand), we now have gained a fairly complete understanding of the relation between the spectra and structure of the mononuclear copper−cysteinate proteins.

Acknowledgment. This investigation has been supported by grants from the Flemish Science Foundation (FWO), the Concerted Action of the Flemish Government, the Swedish Natural Science Research Council (NFR), and the European Commission through the TMR program (Grant ERBFMRXCT960079). The work was also supported with computer resources by the Swedish Council for Planning and Coordination of Research (FRN) and Parallelldatorcentrum (PDC) at the Royal Institute of Technology, Stockholm.

JA982385F

(42) De Kerpel, J. O. A.; Pierloot, K.; Ryde, U.; Roos, B. O. *J. Phys. Chem. B* **1998**, *102*, 4638−4647.