

Chromophore-in-Protein Modeling of the Structures and Resonance Raman Spectra for Type 1 Copper Proteins

Di Qiu,[†] Siddharth Dasgupta,[‡] Pawel M. Kozlowski,[†] William A. Goddard III,[‡] and Thomas G. Spiro^{*,†}

Department of Chemistry, Princeton University, Princeton, New Jersey 08544, and Materials and Process Simulation Center, Beckman Institute (139-74), Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, California 91125

Received December 31, 1996. Revised Manuscript Received September 25, 1998

Abstract: Geometries and resonance Raman (RR) spectra have been modeled for the type 1 Cu active sites of plastocyanin and azurin, as well as two azurin site-directed mutants, M121G and H46D. Using force constants for the Cu coordination group chosen to fit the RR spectra in conjunction with the AMBER force field, we calculated geometries and vibrational spectra. The fitting procedure utilized a chromophore-in-protein approximation, in which a large fraction of the protein was included in the calculation, but only the atoms within a certain distance of the Cu were allowed to vibrate. This procedure reduces the size and complexity of the calculation while retaining all the protein forces. The calculation was tested against experimental RR frequencies, isotope shifts (⁶⁵Cu, ³⁴S, ¹⁵N, Cys-¹⁵N, Cys-C_βD₂) and relative intensities. We find that including six or more heavy atoms (C, N, O) along each Cu-ligating residue (along with the attached H atoms) leads to results essentially independent of the size of the vibrating unit. The calculated spectral features reproduced most observed features, including isotope shifts and the redistribution of RR intensity upon ³⁴S substitution. The spectral changes in the azurin mutants result mainly from a decreased Cu–S(Cys) force constant. The spectra of plastocyanin and azurin are markedly different, despite identical Cu–S force constants. The complexity of the RR spectra near 400 cm⁻¹ results from coordinate mixing among Cu–S stretching and several angle bending coordinates of the cysteine side chain (and small amounts of neighboring side chains).

1. Introduction

Type 1 copper proteins have been a continuing focus of biophysical studies because of their importance in biological electron-transfer processes and their unusual structural and spectroscopic properties.¹ Although Cu(II) complexes are usually tetragonal, this class of proteins leads to a highly distorted tetrahedron for the coordination group. The Cu atom is close to a trigonal plane formed by two histidine and one cysteinate ligands; it is coordinated weakly to a distant axial ligand, usually methionine. The short Cu–S(Cys) bond (~2.1 Å) and the trigonal coordination plane are responsible for unusual spectroscopic properties: reduced hyperfine splitting in the EPR spectrum¹ and intense optical absorption near 600 nm due to (Cys)S → Cu charge-transfer transitions.²

Unusual features are also seen in resonance Raman (RR) spectra excited in the ~600-nm absorption band.³ Understanding these features may yield insight into the function of the type 1 Cu proteins, since the resonant (Cys)S → Cu charge-

transfer transition may carry information about the electron-transfer process. The RR spectra contain multiple bands in the 400-cm⁻¹ region, whereas only a single Cu–ligand stretching vibration (associated with the short Cu–S(Cys) bond) is expected in this region. The large effective mass of the imidazole ring leads to the two Cu–N(His) bond stretches at much lower frequencies,^{3c} near 250 cm⁻¹. Likewise, the Cu–S(Met) stretch is expected at a low frequency (because of the long bond distance).

The origin of the multiple ~400-cm⁻¹ bands has been controversial, but accumulating evidence from isotope labeling studies^{3c,5–9} implicates mixing of the Cu–S(Cys) stretching coordinate with angle bending coordinates of the cysteine side chain as the principal source of the spectral complexity. This mixing is expected to be maximized by the coplanarity of the side chain with the Cu–S bond, a constant feature of all the extant type 1 Cu protein structures.^{1a} In a coplanar structure, the Cu–S stretch is maximally reinforced by angle bending coordinates of the side chain.

A similar anomaly is encountered for nitrile hydratase, which contains mononuclear Fe³⁺ bound to cysteine and histidine ligands. Nitrile hydratase likewise displays a very rich RR spectrum in the 400-cm⁻¹ region⁴ which, as established by isotope studies, involve Fe–S, cysteine side chain deformation, and participation of N donor ligands. These bands shift to higher energy at higher pH, suggesting a stronger Fe–S bond.

Insight into the nature of the coordinate mixing of the ~400-cm⁻¹ band has been garnered from model complex studies with

[†] Princeton University.

[‡] California Institute of Technology.

(1) (a) Adman, E. T. *Adv. Protein Chem.* **1991**, *42*, 145–197. (b) Solomon, E. I.; Baldwin, M. J.; Lowery, M. D. *Chem Rev.* **1992**, *92*, 521–542.

(2) (a) Gray, H. B.; Solomon, E. I. In *Copper Proteins*; Spiro, T. G., Ed.; J. Wiley & Sons: New York, 1981; pp 1–39. (b) Gewirth, A. A.; Solomon, E. I. *J. Am. Chem. Soc.* **1988**, *110*, 32811–32819.

(3) (a) Woodruff, W. H.; Dyer, R. B.; Schoonover, R. B. In *Biological Applications of Raman Spectroscopy*, Vol. 3; Spiro, T. G., Ed.; Wiley: New York, 1988; pp 413–438. (b) Sanders-Loehr, J. In *Bioinorganic Chemistry of Copper*; Karlin, K. D., Tyeklar, Z., Eds.; Chapman Hall: New York, 1993, pp 51–63. (c) Nestor, L.; Larrabee, J. A.; Woolery, G.; Reinhammar, B.; Spiro, T. G., *Biochemistry* **1984**, *23*, 1084–1093.

(4) Brennan, B. A.; Cummings, J. G.; Chase, D. B.; Turner, I. M.; Nelson, M. A. *Biochemistry* **1996**, *35*, 10068–10077.

thiolate adducts of a tris-pyrazolyl borate chelate, which closely mimics the type 1 Cu site.¹⁰ When the adduct was formed with *sec*-butylthiolate, the RR spectrum contained three strong bands near 400 cm⁻¹, whereas the *tert*-butylthiolate analogue complex gave only one band. The difference was traced to inequivalent C–C–C angle bending coordinates at the carbon atom next to the ligating sulfur atom and to the involvement of H–C–C–H torsion coordinates in the *sec*-butyl group. These structural features are absent in *tert*-butylthiolate but are present in *sec*-butylthiolate and in the cysteine side chain of proteins. However, there must be additional sources of complexity in type 1 Cu proteins because their RR spectra contains more than three prominent bands near 400 cm⁻¹.

Understanding this complexity requires calculation of the normal modes of the protein chromophore. However, identifying chromophore modes is not a straightforward task. The size of the calculation and the spectral density create formidable problems if the goal is to characterize specific modes. Alternatively, one can model the active site with a limited set of atoms radiating out from the Cu atom. When we tried this approach, using protein coordinates and a force field (FF) developed for the analogue complexes,¹⁰ we encountered unreasonably large contributions to the calculated modes from bending coordinates of whatever atoms terminate the model, an artifact of neglecting the bonds between these atoms and the rest of the protein. This problem was reported as well by Urushiyama and Tobari,¹¹ who were able to suppress the effect by increasing the size of the model to 169 non-hydrogen atoms, a significant fraction (~25%) of the non-hydrogen atoms in the protein. They found that the Cu-isotope-sensitive modes were not localized to the ligating residues but were widely distributed in the coordinate space of the model. The validity of this counterintuitive result is difficult to assess. Because of the large size of the calculation, the FF was necessarily quite simple. Also, the hydrogen atoms were neglected, whereas the analogue complex results cited above indicate that H–C–C–H torsions may be important.

To study this problem, we adopted a “chromophore-in-protein” approach (Figure 1) to modeling the experimental data. A large part of the protein is included in the calculation, using a molecular mechanics FF, but the vibrating unit is restricted to a limited number of atoms surrounding the copper. This number is increased until there is no further influence on the calculated quantities of interest, namely the frequencies and eigenvectors of the chromophore modes. The size of the calculation remains manageable, even with inclusion of hydrogen atoms, while important protein influences on the chromophore, including nonbonded interactions and electrostatic effects, are retained in the FF.

The chromophore modes are tracked via their isotope shifts. Many experimental isotope shifts are available for two key

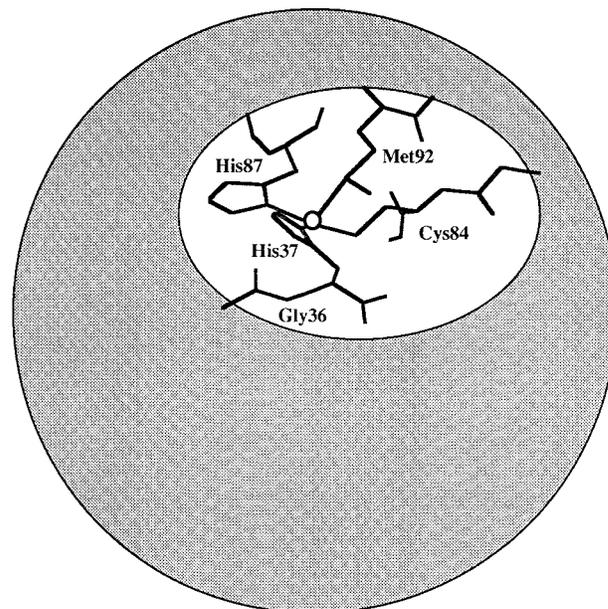


Figure 1. Schematic view of the “chromophore-within-protein” modeling of RR spectra for the case of Pcy. The entire protein is included in the calculation, but only the atoms of the chromophore are allowed to vibrate. The size of the chromophoric fragment was systematically increased until convergence was found for the modes having significant Cu–S contributions.

representatives of the type 1 Cu proteins, *Pseudomonas aeruginosa* azurin (Azu) and poplar plastocyanin (Pcy), thanks to the availability of bacterial expression systems. The proteins have been substituted with ⁶⁵Cu, ³⁴S, and ¹⁵N,^{5–7} and cysteine residues have been incorporated with C_βD₂ and ¹⁵N labels.^{8,9} ³⁴S data are also available for two site-directed mutants of Azu,⁵ one in which the axial methionine ligand is replaced with glycine (M121G) and the other in which one of the histidine ligands is replaced by aspartate (H46D). These multiple isotope shifts strongly constrain the calculation. A further check is provided by calculating the relative intensities of the chromophore RR bands. Although agreement with the data is not quantitative, we are able to reproduce all the observed spectral trends, including significant differences between azurin and plastocyanin and between azurin and the site-directed mutants. The outcome is a clearer picture of the type 1 Cu protein chromophore modes and their dependence on the details of protein structures.

2. Methodology

The BIOGRAF molecular modeling package³⁶ was used to predict the structure and normal modes using the procedure outlined in Scheme 1. In the vibrational calculation, 306 atoms were included in the calculations, but most were effectively held fixed by assigning them a masses of 999 amu. A set of atoms radiating out from the Cu atom were assigned their actual masses and allowed to vibrate at their natural frequencies. The size of this set was increased systematically until there was negligible change in the frequencies of the normal modes displaying significant ⁶⁵Cu or ³⁴S frequency shifts. This point was reached when the chain of atoms along each ligating side chain was extended to about six atoms (Figure 2); because of the branching of the chains, the total number of movable atoms (including H) was 81. The remaining 225 atoms provided a background of forces for the moving atoms.

(5) (a) Dave, B. C.; Germanas, J. P.; Czernuszewicz, R. S. *J. Am. Chem. Soc.* **1993**, *115*, 12175–12176. (b) Czernuszewicz, R. S.; Dave, B. C.; Germanas, J. P. In *Spectroscopic Methods in Bioinorganic Chemistry*; Solomon, E. I., Hodgson, K. O., Eds.; ACS Symposium Series 692; American Chemical Society: Washington, DC, 1998; pp 220–240.

(6) Qiu, D.; Dong, S.; Ybe, J.; Hecht, M.; Spiro, T. G. *J. Am. Chem. Soc.* **1995**, *117*, 6443–6446.

(7) Andrew, C. R.; Han, J.; den Blaauwen, T.; van Pouderooyen, G.; Vijgenboom, E.; Canters, G. W.; Loehr, T. M.; Sanders-Loehr, J. *J. Biol. Inorg. Chem.* **1997**, *2*, 98–107.

(8) Czernuszewicz, R. S., personal communication, 1995.

(9) Dong, S.; Spiro, T. G. *J. Am. Chem. Soc.* **1998**, *120*, 10434–10440.

(10) Qiu, D.; Kilpatrick, L.; Kitajima, N.; Spiro, T. G. *J. Am. Chem. Soc.* **1994**, *116*, 2585–2590.

(11) Urushiyama, A.; Tobari, J. *Bull. Chem. Soc. Jpn.* **1990**, *63*, 1563–1571.

Scheme 1

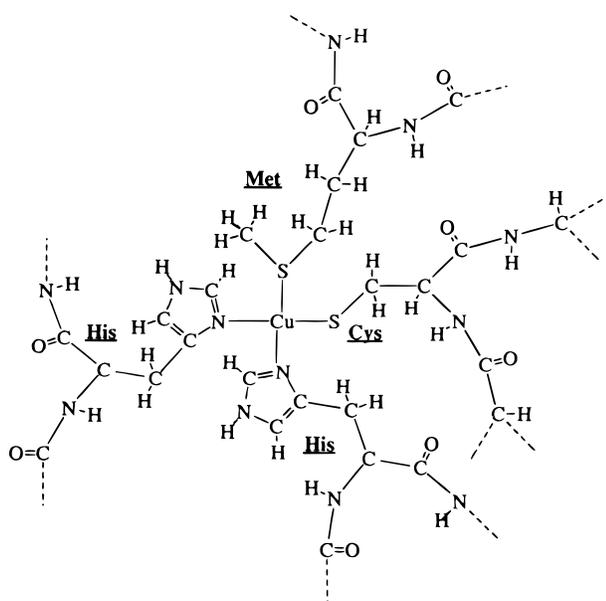
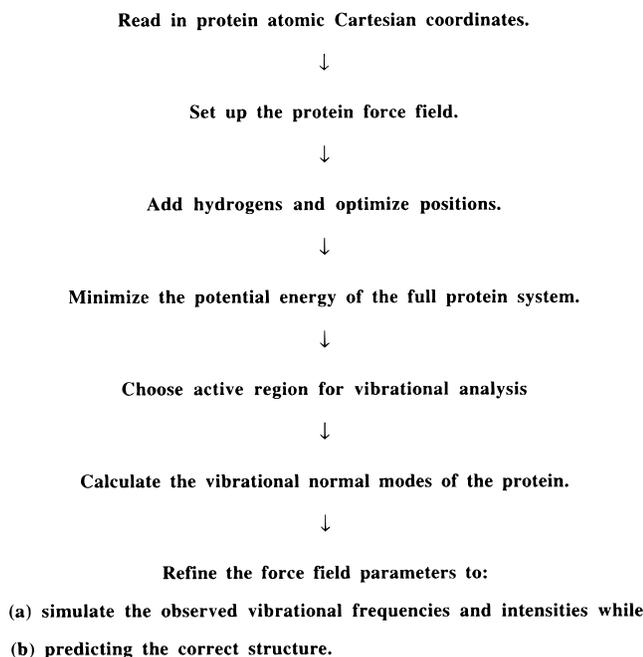


Figure 2. Chromophoric fragment for calculating the vibrational modes of Pcy.

Atomic coordinates were obtained from the Brookhaven Protein Databank for *poplar* plastocyanin¹² [1plc] and *Pseudomonas aeruginosa* azurin¹³ [4azu]. The azurin mutant structures were generated artificially by replacing methionine 121 with glycine or histidine 46 with aspartic acid in the wild-type structure, followed by energy minimization.

The FF parameters describing the copper–ligand bond lengths, angles, and partial charges were adjusted as described in the Results section, and the copper van der Waals parameters were adopted from universal force field (UFF) developed by Rappé et al.¹⁶ Otherwise we used parameters from the

AMBER¹⁴ FF (atom charges, hydrogen bonding and van der Waals parameters, and equilibrium bond lengths and angles), which has been used in many molecular dynamics and mechanics studies of proteins. AMBER incorporates defined charges and van der Waals parameters for hydrogen atoms (explicitly included in our study). The AMBER FF employs harmonic bond stretching and angle bending, a Lennard-Jones 6-12 function for the nonbonded van der Waals potential, and a 10–12 potential for hydrogen bonding:

$$U = \sum_{\text{bonds}} K_r (r - r_{\text{eq}})^2 + \sum_{\text{angles}} K_\theta (\theta - \theta_{\text{eq}})^2 + \sum_{\text{dihedrals}} \frac{K}{2} [1 + \cos(n\phi - \nu)] + \sum_{i < j} \left(\frac{B_{ij}}{r_{ij}^{12}} - \frac{A_{ij}}{r_{ij}^6} + \frac{q_i q_j}{\epsilon r_{ij}} \right) + \sum_{\text{Hbonds}} \left(\frac{B'_{ij}}{r_{ij}^{12}} - \frac{A'_{ij}}{r_{ij}^{10}} \right) \quad (1)$$

The dielectric constant was chosen to be $\epsilon = 4R_{ij}$, in accordance with the study by Whitlow and Teeter.¹⁵ The distance dependence represents the effect of solvent in attenuating the electrostatic interactions of charged groups in an aqueous environment (dampening long-range polarization effects). Non-bonding atom-based spline cutoffs were applied to both the electrostatic and the van der Waals interactions ($R_{\text{in}} = 8.0 \text{ \AA}$ and $R_{\text{out}} = 8.5 \text{ \AA}$).

The potential energy of the full protein was minimized using conjugate gradient methods to arrive at a final rms force of less than 0.1 (kcal/mol)/Å. The structure was reoptimized after each FF modification so that the rms force was less than the cutoff. (Geometry optimization is ignored in traditional normal coordinate analysis, which usually considers only local valence terms.) We included long-range Coulomb and van der Waals terms so that the FF can be used in the full protein to determine how mutations affect the modes. This required that the structure be optimized to obtain the minimum of the potential energy surface before calculating the normal modes. We then removed the translations and rotations from the genuine vibrational modes. Optimization of the FF parameters was carried out using the Hessian biased approach.¹⁷

The potential energy of the molecule for small distances from equilibrium is expanded as

$$E(X_1 \dots Z_N) = E_0 + \sum_{i=1}^{3N} \left(\frac{\partial E}{\partial R_i} \right) (\partial R_i) + \frac{1}{2} \sum_{i,j=1}^{3N} \left(\frac{\partial^2 E}{\partial R_i \partial R_j} \right) (\partial R_i) (\partial R_j) + \dots \quad (2)$$

where $F_i = -(\partial E / \partial R_i)_0$ is the force on atomic coordinate R_i , and $H_{ij} = (\partial^2 E / \partial R_i \partial R_j)_0$ is the Hessian. The vibrational states are calculated by transforming to the mass-weighted Hessian and diagonalizing to obtain the eigenvalues λ . The vibrational frequency (in cm^{-1}) is given by $\nu = 108.591 (\lambda^{1/2})$ (assuming that masses are in atomic mass units, energies are in kilocalories per mole, and distances are in angstroms).

Relative RR intensities were calculated as

(12) Guss, J. M.; Bartunik, H. D.; Freeman, H. C. *Acta Crystallogr.* **1992**, *B48*, 790.

(13) Nar, H.; Messerschmidt, A.; Huber, R.; van de Kamp, M.; Canters, G. W. *J. Mol. Biol.* **1991**, *221*, 765–772.

(14) Weiner, S. J.; Kollman, P. A.; Nguyen, D. T.; Case, D. A. *J. Comput. Chem.* **1986**, *7*, 230–252.

(15) Whitlow, M.; Teeter, M. M. *J. Am. Chem. Soc.* **1986**, *108*, 7163–7172.

(16) Rappé, A. K.; Casewit, C. J.; Colwell, K. S.; Goddard, W. A., III; Skiff, W. M. *J. Am. Chem. Soc.* **1992**, *114*, 10024–10035.

(17) Dasgupta, S.; Goddard, W. A., III. *J. Chem. Phys.* **1989**, *90*, 7207–7215. Dasgupta, S.; Yamasaki, T.; Goddard, W. A., III. *J. Chem. Phys.* **1996**, *104*, 2898.

$$I_j = K(\Delta_j^e)^2 \nu_j^2 \quad (3)$$

where K is a constant for a given incident exciting frequency, Δ^e is the distortion along the normal coordinate in the excited state from the ground-state geometry, and ν is the frequency of the normal mode.¹⁸ This expression assumes that the excited state leads to only small displacements without significant frequency shifts or coordinate rotation. It is valid in the “short-time” limit, which applies to preresonant excitation or fast electronic relaxation.¹⁹ Type 1 Cu proteins show a broad, unstructured absorption band with very weak fluorescence, suggesting fast electronic relaxation. We further assumed that Δ^e can be approximated by the displacement along the Cu–S_{Cys} bond in the ground state, $\Delta_{\text{Cu-S}}$. The rationale for this approximation is that the absorption intensity arises from cysteine-to-copper charge-transfer transitions. This suggests that the main distortion coordinate in the excited state would be a lengthened Cu–S bond. The excited-state displacement for a given mode should then be proportional to $\Delta_{\text{Cu-S}}$, the contribution of Cu–S stretching to the eigenvector. A similar assumption was made by Woodruff and co-workers²⁰ to evaluate variations in Cu–S bonding among type 1 Cu proteins.

3. Results and Discussion

3.1. Copper–Ligand Distances and Angles. To determine the reference values for the bond distances and angles involving Cu, we averaged the Cu–ligand distances and angles over the 15 crystal structures now available for type 1 Cu proteins.^{12,13,21–31} [It is not satisfactory to use crystallographically determined values for any single X-ray structure, because the Cu–ligand bond lengths are uncertain to at least 0.1 Å, even for quite high resolution (1.33 Å) data, as discussed by Guss et al.¹²] These values are given in Table 1. We allowed a longer Cu–S_{Met} bond, 3.18 Å, to be used for the azurins, because this parameter is systematically longer in the azurin structures, presumably because of the relatively close approach (3.1–3.5 Å) of a backbone carbonyl oxygen atom to the Cu on the side opposite the methionine ligand. The remaining averaged bond lengths and angles are all quite reasonable, although it can be seen that the range of crystallographically determined values is broad,

(18) (a) Warshel, A.; Karplus, M. *J. Am. Chem. Soc.* **1974**, *96*, 5677–5689. (b) Blazej, D. C.; Peticolas, W. L. *Proc. Natl. Acad. Sci. U.S.A.* **1977**, *74*, 2639–2643.

(19) Myers, A. B.; Mathies, R. A. In *Biological Applications of Raman Spectroscopy*, Vol. 2; Spiro, T. G., Ed.; Wiley: New York, 1987; pp 1–59.

(20) Blair, D. F.; Campbell, G. W.; Schoonover, J. R.; Chan, S. I.; Gray, H. B.; Malmström, B. G.; Pecht, I.; Swanson, B. I.; Woodruff, W. H.; Cho, W. K.; English, A. M.; Fry, H. A.; Lum, B.; Norton, K. A. *J. Am. Chem. Soc.* **1985**, *107*, 5755–5766.

(21) Fields, B. A.; Guss, J. M.; Freeman, H. C. *J. Mol. Biol.* **1991**, *222*, 1053–1065.

(22) Lowery, M. D.; Solomon, E. I. *Inorg. Chim. Acta* **1992**, *198–200*, 233–243.

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

(24) Baker, E. N. *J. Mol. Biol.* **1988**, *203*, 1071–1095.

(25) Durley, R. C. E.; Chen, L.; Lim, L. W.; Mathews, F. S.; Davidson, V. *Protein Sci.* **1993**, *2*, 739–752.

(26) Collyer, C. A.; Guss, J. M.; Sugimura, Y.; Yoshizaki, F.; Freeman, H. C. *J. Mol. Biol.* **1990**, *211*, 617–632.

(27) Messerschmidt, A.; Ladenstein, R.; Huber, R. *J. Mol. Biol.* **1992**, *224*, 179–205.

(28) Petratos, K.; Dauter, Z.; Wilson, K. S. *Acta Crystallogr.* **1988**, *B44*, 628–636.

(29) Romero, A.; Hoiink, C. W. G.; Nar, H.; Huber, R.; Messerschmidt, A.; Canters, G. W. *J. Mol. Biol.* **1993**, *229*, 1007–1021.

(30) Nar, H.; Messerschmidt, A.; Huber, R. *J. Mol. Biol.* **1991**, *218*, 427–447.

(31) Tsai, L.-C.; Sjölin, L.; Langer, V.; Pascher, T.; Nar, H. *Acta Crystallogr.* **1995**, *D51*, 168–176.

about 0.3 Å for distances and 10–40° for angles, consistent with the uncertainty analysis of Guss et al.¹² The average values $N_1\text{–Cu–S}_c = 133^\circ$, $N_2\text{–Cu–S}_c = 118^\circ$, and $N_1\text{–Cu–N}_2 = 102^\circ$ indicate a nearly planar trigonal system, while the $N_1\text{–Cu–S}_m = 84^\circ$ and $N_2\text{–Cu–S}_m = 97^\circ$ average values indicate a long Cu–S_m bond nearly perpendicular to this plane.

The effect of the Cu partial charge on the calculated geometries and vibrational frequencies was tested over the range from +0.5 to +2.0. The most consistent results were obtained for +0.75, which is consistent with values from quantum mechanical calculations. It is somewhat larger than the $q_{\text{Cu}} = +0.5$ used in molecular dynamics simulations of plastocyanin²¹ or the $q_{\text{Cu}} = +0.58$ used in $X\alpha$ -scattered wave calculations of the electronic structure.²²

Force constants for the Cu coordination group were adopted from literature values for model compounds with similar bonds¹⁰ (Table 2). Only diagonal valence force constants were used, except for Pcy, where a single coupling force constant was introduced, as discussed below. The critical Cu–S_{Cys} force constant was varied over a range of values, but the remaining valence force constants were held fixed in all the calculations. All the angle bending force constants were lumped into two values, one for L–Cu–L bends and one for Cu–L–X bends, where X is the atom next to the ligating atom. The precise values in this class of force constants are poorly known and have relatively little influence on the modes in the 400-cm⁻¹ region, which are of principal interest.

With force constants and average structural parameters in hand, we next calculated equilibrium geometries for the individual proteins under study (Table 3). For Pcy and Azu, the crystallographically determined coordinates were used as the starting values. For the H46D and M121G mutants of Azu, we substituted the appropriate residue into the native structure and then minimized the energy. In addition, a water molecule was inserted as the axial ligand in the M121G mutants.⁵ The reference Cu–O bond length was set to 2.2 Å for the water ligand in M121G and to 1.91 Å for the carboxylate ligand in H46D. Because the best fits to the RR spectra (see below) were obtained with the same Cu–S_{Cys} stretching force constant, Pcy and Azu have identical Cu–S_{Cys} distances in the minimized structures. The coplanarity of Cu atom and Cys ligand¹⁰ is retained upon minimization. In addition, the Cu atom lies very close (within 0.06 Å) to the $N_{\text{His}}\text{–}N_{\text{His}}\text{–S}_{\text{Cys}}$ coordination plane for both proteins, although for Pcy the starting value was significantly out of the plane (0.36 Å toward the methionine ligand). The root-mean-square coordinate disagreement between the calculated structure and the X-ray crystal structures is 0.51 and 0.38 Å for Pcy and Azu, respectively, indicating that the FF provides a consistent description of the structure. For both Azu mutants, the RR spectra were fit with a lower Cu–S_{Cys} force constant, and the Cu–S_{Cys} bond is calculated to lengthen by 0.03 Å relative to that of the native protein. The calculations indicate that the H46D substitution causes the Cu atom to be pulled 0.46 Å out of the coordination plane. This is consistent with the inference of enhanced tetrahedral character of the Cu,²³ as judged by increased 460-nm absorption and rhombicity of the EPR signal.⁵

3.2. Summary of Vibrational Results. Observed and calculated frequencies and isotope shifts in the 400-cm⁻¹ region are compared in Table 4 for Azu and Pcy. (Additional modes are calculated between 370 and 440 cm⁻¹, three for Azu and four for Pcy, but these have negligible contributions from the Cu coordination group.) The criterion for selecting modes for

Table 1. Average Values and Ranges of Crystallographic Parameters of the Type 1 Copper Site^a

coordinate ^b	range ^c	average	coordinate	range	average
Cu–S _c	2.03–2.27	2.14	Cu–N2–C ^ε	121.2–127.1	125.0
Cu–S _m	2.67–3.52	2.92	Cu–N2–C ^γ	122.2–131.0	127.1
Cu–N1	1.90–2.20	2.04	S–Cu–S	100.3–143.1	109.4
Cu–N2	1.96–2.27	2.10	N1–Cu–N2	95.0–109.1	101.9
Cu–S _c –C	82.6–122.2	108.5	N1–Cu–S _c	125.3–143.8	133.3
Cu–S _m –C ^ε	92.5–129.3	102.7	N2–Cu–S _c	109.0–125.6	117.8
Cu–S _m –C ^γ	121.1–144.8	134.3	N1–Cu–S _m	72.0–95.4	83.9
Cu–N1–C ^ε	111.6–132.4	123.0	N2–Cu–S _m	83.2–111.8	97.0
Cu–N1–C ^γ	115.6–139.2	128.5			
τ(Cu–S _c –C _b –Ca)	164.2–178.7	172.2	τ(S _c –C _b –C _a –N)	162.4–179.8	172.4

^a From 15 protein structures: azurin from *A. denitrificans*;²⁴ azurin from *P. aeruginosa*;¹³ amicyanin from *P. denitrificans*;²⁵ plastocyanin from green alga (*E. prolifera*);²⁶ ascorbate oxidase from zucchini;²⁷ pseudoazurin from *A. faecalis*;²⁸ Met121Gln mutant of *A. denitrificans* azurin;²⁹ His35Gln mutant of *P. aeruginosa* azurin;³⁰ His35Leu mutant of *P. aeruginosa* azurin;³⁰ Phe114Ala mutant of *P. aeruginosa* azurin;³¹ Asn47Asp mutant of *P. aeruginosa* azurin;³² plastocyanin from *M. extorquens*;³³ plastocyanin from poplar;¹² plastocyanin from green alga (*C. reinhardtii*).³⁴
^b Symbols: S_c, cysteine sulfur atom; S_m, methionine sulfur atom; N1, the nitrogen atom from the histidine ligand farther from cysteine in sequence; and N2, the nitrogen atom from the histidine closer to the cysteine ligand. ^c Bond distances in angstroms; angles in degrees.

Table 2. Force Constants Associated with the Cu–Ligand Bonds^a

$K_r(\text{Cu–S}_{\text{Cys}})$	260 (1.86) ^b
$K_r(\text{Cu–S}_{\text{Met}})$	50 (0.36)
$K_r(\text{Cu–N}_{\text{His}})$	220 (1.57)
$K_\theta(\text{Cu–L–X})$	50 (0.36)
$K_\theta(\text{L–Cu–L})$	10 (0.07)
$K_{\theta\theta}[(\text{Cu–S–C})-(\text{S–C–C})]^\text{c}$	60 (0.42)

^a Units are kcal/(mol·Å²) or mdyne/Å for K_r ; and kcal/(mol·rad²) or (mdyne·Å)/rad² for K_θ , $K_{\theta\theta}$. ^b Value for plastocyanin and azurin; for the H46D and M121G azurin mutants, $K_r = 220$ kcal/(mol·Å²) or 1.57 mdyne/Å. ^c Only for plastocyanin.

comparison with experiment was that the calculated ³⁴S isotope shift be at least 0.3 cm⁻¹. This screen gave six modes observed for Pcy and five for Azu.

For most of the bands, the frequency match is excellent. Two of the calculated Azu modes are off by 5 and 7 cm⁻¹, and three of the Pcy modes deviate by 4–7 cm⁻¹, but all the other modes are calculated to within 3 cm⁻¹ of the observed values. Moreover, the calculated and observed isotope shift patterns are in decent overall agreement, except for ¹⁵N substitution for all N atoms, for which the calculated shifts are too high. Thus, the calculations give a fair account of the normal mode compositions of Pcy and Azu. Likewise, for the Azu mutants, the calculated frequencies and ³⁴S shifts are in satisfactory agreement with experiment (other isotopic data are unavailable) (Table 5).

A different view of the quality of the calculations is given in Figure 3, where the calculated frequencies are superimposed on the reported spectra of the four proteins under study, in natural abundance and ³⁴S isotopomers. The lengths of the bars represent the calculated relative intensities. There is good correspondence between observed and calculated relative intensities in some parts of the spectra, but not in others. Particularly gratifying is the reproduction of the ³⁴S effect on the spectral shape, with intensity shifting from higher to lower frequency components in the 400 cm⁻¹ region. The intensification of lower frequency components in the Azu mutants is also seen in the calculated spectra.

However, the calculated intensities are too low in the region below 400 cm⁻¹. This discrepancy probably reflects inadequacy of the assumption, that the only coordinate undergoing displacement in the resonant excited state is the Cu–S_{Cys} stretch. It is likely that angle bending displacements are not negligible for charge-transfer excitation and that these coordinates contribute selectively to the lower frequency modes, thereby accounting for some of the missing intensity.

3.3. Systematics of the Calculated Spectra. The final results reported in the preceding section were obtained after systematically varying the Cu–S stretching force constant, which has the most profound effect on the spectra of any of the FF elements. When this force constant was increased over a physically reasonable range^{4,10,11} (1.64–1.93 mdyne/Å), there was a systematic shift of ³⁴S sensitivity and of intensity from the lower to the higher frequency modes. Thus, the main effect of the force constant variation was to change the normal mode compositions. As the Cu–S_{Cys} force constant was increased, the Cu–S stretching contribution diminished for the lower frequency modes and increased for the higher frequency modes. These calculations support the interpretations of Azu mutant data by Dave et al.⁵ and by Andrew et al.,⁷ that Cu–S bond weakening leads to shifts in intensity to lower frequency modes, rather than to a frequency change per se. A Cu–S_{Cys} force constant of 1.86 mdyne/Å gives the best agreement with the observed intensity pattern and ³⁴S isotope shifts.

The same FF was applied to the two Azu mutants, modifying only the Cu–S_{Cys} force constant. To obtain agreement of comparable quality for the intensity and ³⁴S shift patterns of the Azu mutants, we found it necessary to lower $K_{\text{Cu–S}}$ to 1.57 mdyne/Å. This force constant decrease is reflected in the increased Cu–S_{Cys} bond length calculated in the energy-minimized structures, from 2.149 Å in Azu to 2.178 Å in both Azu mutants (Table 3). Application of Badger's rule³⁵ to these distances would suggest that the force constant be decreased from 1.85 to 1.77 mdyne/Å. This deviation from Badger's rule (which was derived for diatomics) is probably because the vibrational mode is not a pure stretch.

In the case of Pcy, the energy-minimized Cu–S_{Cys} bond length is the same as that for Azu (Table 3), and we found that the same force constant gave the best agreement with experiment. However, the quality of this agreement was not as good as that for Azu. We found it possible to improve the intensity and frequency match by introducing an interaction force constant between two angle bending coordinates, CuSC and SCC

(32) Sjölin, L.; Tsai, L. C.; Langer, V.; Pascher, T.; Karlsson, G.; Nordling, M.; Nar, H. *Acta Crystallogr.* **1993**, *D49*, 449–457.

(33) Inoue, T.; Kai, Y.; Harada, S.; Kasai, N.; Ohshiro, Y.; Suzuki, S.; Kohzuma, T.; Tobari, J. *Acta Crystallogr.* **1994**, *D50*, 317–328.

(34) Redinbo, M. R.; Cascio, D.; Choukair, M. K.; Merchant, T. O.; Yeates, T. O. *Biochemistry* **1993**, *32*, 10560–10567.

(35) (a) Badger, R. M. *J. Chem. Phys.* **1934**, *2*, 128. (b) Herschbach, D. R.; Laurie, V. W. *J. Chem. Phys.* **1961**, *35*, 458.

(36) The Caltech BioGraf (version 3.31) is based on the commercial version of POLYGRAF (version 3.21) distributed by Molecular Simulations Inc. (San Diego, CA) but enhanced with second derivative, vibrational frequency, and FF optimization (ref 17) models developed at Caltech.

Table 3. Energy-Minimized Structure of the Type 1 Copper Site

	plastocyanin, Pcy	<i>P. aeruginosa</i> azurin, Azu	H46D Azu	M121G Azu
rms deviation (Å) ^a	0.51	0.38	0.46	0.45
Cu–S _{cys} (Å)	2.148 (2.07) ^b	2.149 (2.27)	2.179	2.179
Cu–S _{Met} (Å)	3.002 (2.82)	3.176 (3.18)	3.091	2.197 ^c
Cu–N (Å)	1.988 (1.91)	2.037 (2.00)	1.906 ^d	2.017
Cu–N1 (Å)	2.104 (2.05)	2.085 (2.11)	2.091	2.092
∠Cu–S–C (deg)	113.5 (110.4)	107.8 (109.3)	102.9	108.5
∠S–C–C (deg)	111.4	111.0	110.3	108.8
∠C–C–C (deg)	111.0	109.5	109.8	108.5
τCu–S–C–C (deg)	–163.6 (–168.1)	–173.3 (–173.2)	–174.0	178.4
Cu–(N _{His} N _{His} S _{Cys}) ^e (Å)	0.06	0.06		0.46

^a Of all non-H atoms from the crystallographic coordinates. ^b Numbers shown in the parentheses indicate the data from the crystal structure. ^c Cu–OH₂ bond distance. ^d Cu–O (carboxylate) bond distance. ^e Cu displacement from the N_{His},N_{His},S_{Cys} coordination plane.

Table 4. Comparison of Observed and Calculated RR Frequencies and Isotopic Shifts^a for Azurin and Plastocyanin (cm⁻¹)

azurin		plastocyanin	
obsd ^a	calcd	obsd ^b	calcd
ν ($\Delta^{65}\text{Cu}$, $\Delta^{34}\text{S}$, ΔD_β , $\Delta^{15}\text{N}$)	ν ($\Delta^{65}\text{Cu}$, $\Delta^{34}\text{S}$, ΔD_β , $\Delta^{15}\text{N}$)	ν ($\Delta^{65}\text{Cu}$, $\Delta^{34}\text{S}$, ΔD_β , $\Delta^{15}\text{N}$, $\Delta^{15}\text{N-Cys}$)	ν ($\Delta^{65}\text{Cu}$, $\Delta^{34}\text{S}$, ΔD_β , $\Delta^{15}\text{N}$, $\Delta^{15}\text{N-Cys}$)
372 (0.6, 0.5, 3.8, 1.2)	364.8 (0.0, 0.8, 0.0, 8.9)	376 (0.1, 1.0, 1.8, 1.8, 1.0)	376.5 (0.1, 0.7, 1.4, 2.2, 0.9)
394 (0.2, 0.6, 1.6, 1.0)	398.9 (0.2, 0.7, 1.3, 4.2)	387 (0.2, 0.3, 2.3, 2.2, 1.4)	391.0 (0.1, 0.1, 1.6, 2.9, 0.7)
401 (0.4, 1.8, ~3.6, 2.1)	403.2 (0.2, 0.8, 0.3, 4.7)	403 (0.2, 0.8, 3.2, 1.5, 0.2)	397.5 (0.1, 0.3, 3.9, 1.7, 0.5)
409 (1.0, 3.8, 5.8, 2.0)	411.2 (0.6, 1.8, 4.1, 5.3)	420 (0.2, 2.2, 1.3, 2.4, 1.4)	413.6 (1.5, 2.8, 0.5, 1.3, 0.5)
428 (0.4, 1.4, 1.5, 4.3)	427.7 (0.8, 2.3, 1.1, 3.3)	429 (0.7, 2.3, 0.4, 1.6, 0.2)	432.0 (0.7, 2.3, 0.1, 0.9, 0.3)
		438 (1.0, 0.9, 1.1, 2.0, –0.2)	438.6 (0.1, 0.9, 0.6, 2.6, 0.1)
total (2.6, 7.8, 15.8, 10.6)	total (2.0, 5.6, 9.2, 10.7)	total (2.4, 7.5, 10.1, 11.5, 4.0)	total (2.6, 7.1, 8.1, 11.6, 3.0)

^a Natural abundance frequency minus frequency after substituting ⁶⁵Cu, ³⁴S, or ¹⁵N throughout the protein, or cysteine labeled with ¹⁵N ($\Delta^{15}\text{N-Cys}$), or with D at the C_β atoms (ΔD_β). Azu data for $\Delta^{65}\text{Cu}$, $\Delta^{34}\text{S}$, and ΔD_β from ref 8; data for $\Delta^{15}\text{N}$ from ref 7. ^b Data for $\Delta^{65}\text{Cu}$, $\Delta^{34}\text{S}$, and $\Delta^{15}\text{N}$ from ref 6; data for $\Delta^{15}\text{N-Cys}$ and ΔD_β from ref 9.

Table 5. Comparison of Observed and Calculated Vibration Modes (ν ($\Delta^{34}\text{S}$)) for Azurin Mutants^a

H46D Azu		M121G Azu	
obsd ^a	calcd	obsd ^b	calcd
375 (0.2)	371.0 (1.3)	371 (0.9)	375.7 (1.7)
393 (1.0)	396.7 (1.1)	393 (0.9)	392.6 (0.5)
400 (2.4)	398.5 (1.2)	400 (2.8)	402.7 (1.8)
410 (1.8)	406.3 (1.9)	408 (1.0)	409.9 (1.1)
427 (1.5)	426.1 (0.3)	414 (0.6)	410.7 (0.8)
		427 (1.2)	4223.5 (0.8)
total (6.9)	total (6.6)	total (7.4)	total (5.3)

^a Data from ref 5.

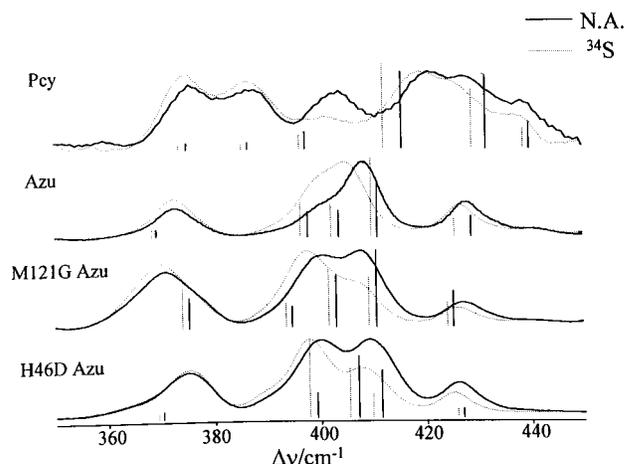
bending, which have the S–C bond in common. A value of 0.43 (mdyn·Å)/rad² gave the best agreement with experiment, although the agreement is still not as good as that for Azu (Figure 4).

3.4. Normal Mode Composition. We turn now to the issue of coordinate mixing among the vibrational modes represented in the RR spectra. One way to measure the extent of mixing is to examine the isotope shift pattern that emerges from experiment and calculation (Table 4):

(A) ⁶⁵Cu and ³⁴S. These shifts are spread throughout the calculated modes but are concentrated in two of them: 414- and 432-cm⁻¹ modes for Pcy, and 411- and 428-cm⁻¹ modes for Azu. This pair of bands carries most of the ³⁴S shift, qualifying them as predominantly Cu–S stretching modes.

(B) C_βD₂. These shifts show that the 414- and 432-cm⁻¹ modes for Pcy, and 411- and 428-cm⁻¹ modes for Azu, also contain contributions from the SCC bending coordinate; however, the shifts are much higher for the Cu–S modes of Azu than for those of Pcy.

(C) ¹⁵N. These shifts (all N atoms substituted) show behavior similar to C_βD₂. They are spread rather evenly among the Pcy modes, both computationally and experimentally. However, the

**Figure 3.** Comparison of calculated frequencies and relative RR intensities (lengths of the bars) with experimental RR spectra for natural abundance (NA) and ³⁴S-labeled Pcy (retraced from ref 6) and Azu and the Azu mutants M121G and H46D (retraced from ref 5).

calculated shifts are too high, indicating too much delocalization in the current FF.

The Cys-¹⁵N shifts determined for Pcy allow us to distinguish the contributions of angle bending coordinates at the beginning of the cysteine side chains. Again, these shifts are spread fairly evenly among the modes. Experimentally, they account for about one-third of the total ¹⁵N shifts, although the fractional contribution of the Cys-¹⁵N shifts varies widely from mode to mode. The relatively large fraction of the ¹⁵N shifts arising from N atoms other than the Cys-N atom implies appreciable contributions from non-Cys residues to the modes, confirming in some measure the conclusions of Urushimaya and Tobar.¹¹ Thus, we have a picture of substantial delocalization of the modes among bending coordinates of the Cys ligand and of

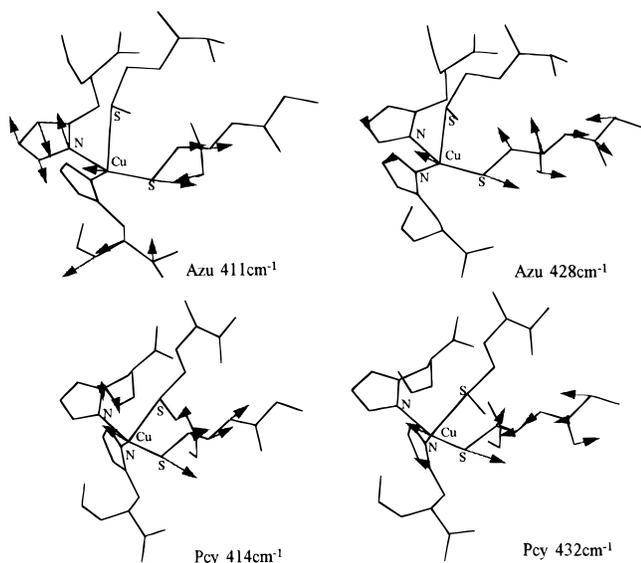


Figure 4. Eigenvectors for the Azu and Pcy modes having the greatest Cu–S displacements. Hydrogen atoms have been suppressed in the plots. Arrows indicate the direction and relative extent of atomic motions.

other residues as well, but with concentration of the Cu–S stretching coordinates in a pair of the higher frequency modes.

This view is confirmed by examining the eigenvectors from the calculation. The two modes with the largest Cu–S displacements are compared for the two proteins in Figure 4. It is evident that all four modes involve substantial deformations of the cysteine side chain and of the main chain. In the case of Pcy, the modes are limited to these atoms; other ligating side chains are not significantly involved. However, the Azu eigenvectors also involve some contributions from the histidine side chains. In particular, the 411-cm^{-1} Azu mode, which produces the strongest band in the RR spectrum, shows a substantial out-of-plane motion of the His46 imidazole ring. It is this motion that accounts for most of the ^{15}N shift.

However, this coupling to His found in the calculation is *not* supported by the available experimental data. Andrew et al.⁷ recently examined the H46G and H117G mutants of Azu, for

which the missing histidine ligands can be replaced with exogenous imidazole; the RR spectra of these adducts are nearly the same as that of native Azu. When ^{15}N -labeled imidazole was substituted in the adducts, there were no discernible isotope shifts in any of the $\sim 400\text{-cm}^{-1}$ bands. Thus, the FF should be modified to reduce the mixing of imidazole coordinates into the 411-cm^{-1} mode. We should emphasize that the current calculation used the AMBER FF for the His residue. No modifications were made.

4. Conclusions

Chromophore-in-protein modeling reproduces some of the complex features of type 1 Cu protein RR spectra, including the distribution of intensity and of ^{65}Cu and ^{34}S isotope shifts over several bands in the 400-cm^{-1} region. The spectral complexity reflects mixing of Cu–S stretching and cysteine angle bending coordinates in the normal modes, as has been speculated on the basis of experimental model compound studies. However, involvement of residues other than cysteine is also demonstrated by the larger isotope when ^{15}N is substituted throughout the protein than when it is substituted only in the cysteine peptide bond. The sensitivity of the spectra to delocalized aspects of the structure is seen in the quite different RR spectra observed and calculated for Pcy and Azu, despite the identical Cu–S force constant. Some features of the calculation, e.g., the imidazole involvement in Azu $\sim 400\text{-cm}^{-1}$ modes, and the total ^{15}N shift do not agree with experiment, indicating the need for FF refinement.

Acknowledgment. We thank Drs. Roman S. Czernuszewicz and Joann Sanders-Loehr for helpful discussions and for sharing data prior to publication, and Mr. Shoulian Dong for sharing his data. Supported by NIH Grant GM 13498 (to T.G.S.). The Caltech portion of the research was funded by DOE-BCTR (David Boron). The facilities of the MSC are also supported by grants from NSF (CHE 95-22179 and ASC 92-17368), Chevron Petroleum Technology, Asahi Chemical, Aramco, Owens Corning, Chevron Chemical Co., Asahi Glass, Chevron Research and Technology Co., BP Chemical, Hercules, Avery Dennison, and Beckman Institute.

JA964472I