J|A|C|S
ARTICLES
Published on Web 10/19/2004

Spectroscopic Identification of Different Types of Copper Centers Generated in Synthetic Four-Helix Bundle Proteins

Robert Schnepf,*[,†,‡,||] Wolfgang Haehnel,[‡] Karl Wieghardt,[†] and Peter Hildebrandt*[,†,§]

Contribution from the Max-Planck-Institut für Bioanorganische Chemie, Stiftstrasse 34-36, D-45470 Mülheim a. d. Ruhr, Germany, Albert-Ludwigs-Universität Freiburg, Institut für Biologie II/Biochemie, Schänzlestrasse 1, D-79104 Freiburg, Germany, and Technische Universität Berlin, Institut für Chemie, Max-Volmer-Laboratorium, Sekr. PC 14, Strasse des 17. Juni 135, D-10623 Berlin, Germany

Received March 18, 2004; E-mail: hildebrandt@chem.tu-berlin.de; rs@complex-biosystems.com

Abstract: Using a combined rational-combinatorial approach, stable copper binding sites were implemented in template-assembled synthetic four-helix bundle proteins constructed by three different helices with only 16 amino acid residues. These peptides include two histidines and one cysteine at positions appropriate for coordinating a copper ion. Sequence variations of the helices were made in the second coordination shell or even more remote from the copper binding site (i) to increase the overall stability of the metalloproteins and (ii) to fine-tune the structure and properties of the copper center. As a result, ca. 90% of the 180 proteins that were synthesized were capable to bind copper with a substantially higher specificity than those obtained in the first design cycle (Schnepf, R.; Hörth, P.; Bill, E.; Wieghardt, K.; Hildebrandt, P.; Haehnel, W. *J. Am. Chem. Soc.* **2001**, *123*, 2186−2195). Furthermore, the stabilities of the copper protein complexes were increased by up to 2 orders of magnitude and thus allowed a UV−vis absorption, resonance Raman, electron paramagnetic resonance, and (magnetic) circular dichroism spectroscopic identification and characterization of three different types of copper binding sites. It could be shown that particularly steric perturbations in the vicinity of the His_2Cys ligand set control the formation of either a tetragonal (type II) or a tetrahedral (type I) copper binding site. With the introduction of two methionine residues above the histidine ligands, a mixed-valent dinuclear copper binding site was generated with spectroscopic properties that are very similar to those of Cu_A sites in natural proteins. The results of the present study demonstrate for the first time that structurally different metal binding sites can be formed and stabilized in four-helix bundle proteins.

Introduction

The biological function of a large number of enzymes is controlled by the structure and reactivity of a metal site.[1] In many cases, the same metal ion can adopt different coordination spheres and structures, which in turn contribute to the specific biological function. Typical examples are copper proteins that form various metal binding sites with a trigonal (type I) or tetragonal (type II) coordination sphere or even may form a binuclear binding site (e.g., Cu_A).[2,3] Proteins involving structurally different Cu binding sites may exert quite different functions, for example, as electron shuttles in electron transfer chains or as catalysts in the oxidation of substrates. Elucidation of relationships between the metal site structure and the function of the protein, which are still far from being fully understood, represents a challenge in contemporary life science both with respect to a profound analysis of biological processes on a molecular level and in view of the design of new enzymes with tailor-made functional properties of potential biotechnological relevance.[4] In this respect, the de novo synthesis of metalloproteins represents a promising approach that may provide new insight into correlations between amino acid sequence, protein fold and the structure, stability, and reactivity of the metal site.[5−9]

We have used spot synthesis[10] combined with the TASP (template assembled synthetic proteins) concept[11] to perform a combinatorial synthesis on cellulose membranes.[12] This combined rational-combinatorial approach was used to create a

† Max-Planck-Institut für Bioanorganische Chemie.
‡ Albert-Ludwigs-Universität Freiburg.
§ Technische Universität Berlin.
|| Present address: Complex Biosystems GmbH, Im Neuenheimer Feld 584, 69120 Heidelberg, Germany.

(1) *Handbook of Metalloproteins*; Messerschmidt, A.; Huber, R.; Poulos, T.; Wieghardt, K., Eds.; John Wiley & Sons: New York, 2001.
(2) Sykes, A. G. *Adv. Inorg. Chem.* **1991**, *36*, 377−408.
(3) Adman, E. T. *Adv. Protein Chem.* **1991**, *42*, 144−197.
(4) Gilardi, G.; Fantuzzi, A. *Trends Biotechnol*. **2001**, *19*, 468−476.
(5) Hellinga, H. W. *J. Am. Chem. Soc.* **1998**, *120*, 10055−10066.
(6) Lombardi, A.; Summa, C. M.; Germina, S.; Randaccio, L.; Pavone, V.; DeGrado, W. F. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 6298−6305.
(7) Lu, Y.; Berry, S. M.; Pfister, T. D. *Chem. Rev.* **2001**, *101*, 3047−3080.
(8) Marsh, E. N. G.; DeGrado, W. F. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 5150−5154.
(9) Maglio, O.; Nastri, F.; Pavone, V.; Lombardi, A.; DeGrado, W. F. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 3772−3777.
(10) Frank, R. *Tetrahedron* **1992**, *48*, 9217−9232.
(11) Mutter, M.; Altmann, E.; Altmann, K. H.; Hersperger, R.; Koziej, P.; Nebel, K.; Tuchscherer, G.; Vuilleumier, S. *Helv. Chim. Acta* **1988**, *71*, 835−847.
(12) Rau, H. K.; DeJonge, N.; Haehnel, W. *Angew. Chem., Int. Ed.* **2000**, *39*, 250−253.

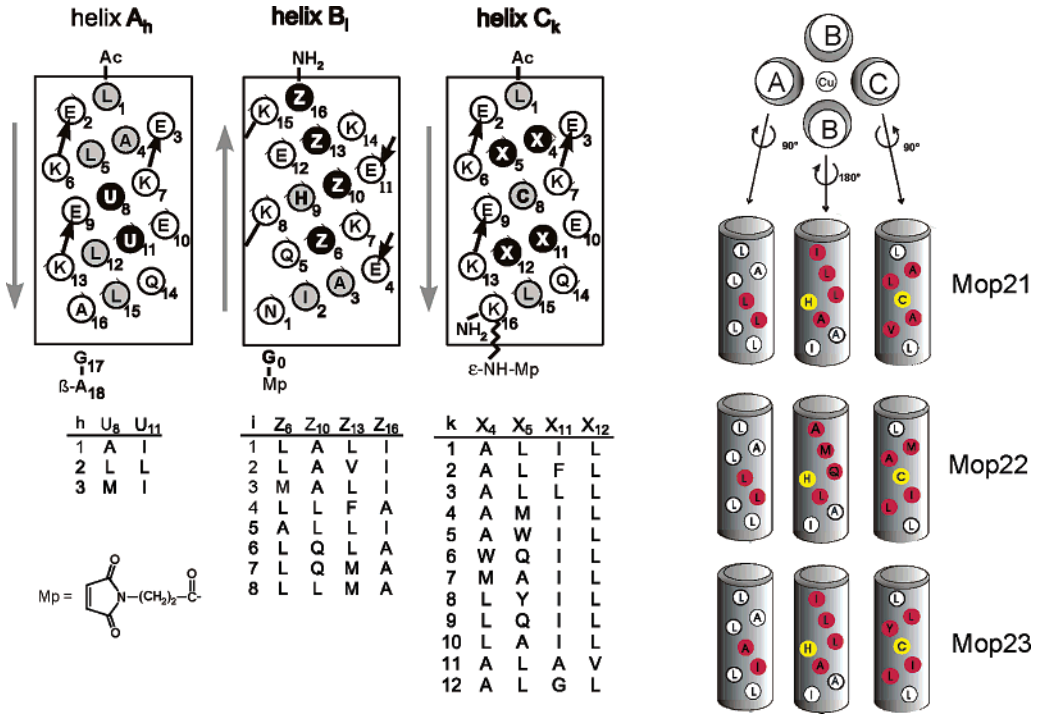

Figure 1. (Left) Helical net representation of the helices A_h, B_i, and C_k. The black filled circles (U, Z, and X) in the helical nets refer to the positions that were varied according to the matrixes displayed below. Invariant residues assumed to form the hydrophobic core and exposed polar residues are indicated by gray filled and open circles, respectively. (Right) Schematic representation of the helices constituting the four-helix bundle proteins Mop21 ($A_2(B_5)_2C_{11}$), Mop22 ($A_2(B_7)_2C_4$), and Mop23 ($A_1(B_5)_2C_8$). The hydrophobic sides of the three different helices are shown with the Cu ligands highlighted in yellow. The positions that have been varied in the synthesis are indicated in red.

cysteine-coordinated copper site within a four-helix bundle structure.[13,14] The synthesis of such metal proteins is a particular challenge, since an optimum arrangement of the putative metal-ligating residues has to be achieved to avoid the competing oxidation of the cysteinate thiol by Cu(II), which is thermodynamically favored as compared to cysteine-Cu(II) ligation.

The TASP concept has shown promising results in the past for the synthesis of heme, Zn, and Co proteins.[12,15−17] In our previous studies, we have successfully expanded the template structure to a design that allows forming cellulose-bound four-helix bundles containing three different helices.[13] This was essential to avoid the presence of two adjacent cysteines that are likely to form an intramolecular disulfide bridge. Also the positioning of the copper site within the shielded hydrophobic core was a prerequisite to prevent cysteine oxidation.

The goal of the first design cycle was to find a scaffold that provides the fundamental requirements for a copper site within a four-helix bundle structure.[13] A structural model was created based on rational considerations. The helix design was guided by the principles known to stabilize coiled coil four-helix bundle formation.[18] Within the framework of a putative four-helix bundle, His residues and a single Cys residue as potential copper ligands were positioned next to each other but at various positions in the sequence to create copper sites in the hydrophobic core. Residues adjacent to the potential ligands were varied in size to optimize the packing around the copper site. Following these design principles we have previously synthesized 96 proteins, among which a few proteins indeed fulfilled the steric requirements for copper coordination.[13] Particularly, one four-helix bundle copper protein, denoted as Cu-Mop5, showed moderate stability with a lifetime of a couple of minutes and the ability to form a homogeneous tetragonal copper complex, a so-called type II copper site.

The scaffold of Mop5 constituted the starting point for the second design cycle in the present work which aimed to enhance the stability of the copper site and to explore the structural potential of the fold for creating different types of copper sites. In this design cycle we have synthesized further 180 proteins, which were screened with respect to their copper binding capabilities. Representative examples were synthesized as soluble proteins and characterized by various spectroscopic techniques.

Experimental Section

General. Materials, solvents, instrumentation and general methods, as well as procedures for peptide synthesis, followed the literature from our laboratory.[12,13] Therefore, only a brief description of procedures will be given below, except for modifications that will be outlined in more detail.

Syntheses. Syntheses of the peptides A_h, B_i, and C_k (Figure 1) were performed with a multiple batch solid-phase synthesizer (Advanced ChemTech Model 396) using standard 9-fluorenylmethoxycarbonyl (Fmoc)/2-(1*H*-benzotriazole-1-yl)-1,1,3,3-tetramethyluronium tetrafluoroborate (TBTU) protection/activation chemistry as described.[12,13]

The template-helix building block T4-A_h for the cellulose-based protein assembly was synthesized as described with allyl (All) and 1-(4,4-dimethyl-2,6-dioxocyclohex-2-ylidene)ethyl (Dde) as protecting groups of Lys and the α-carboxyl group of Glu, respectively.[13] Glu-(OH)-OAll was attached to the resin by its γ-carboxyl group, and the

(13) Schnepf, R.; Hörth, P.; Bill, E.; Wieghardt, K.; Hildebrandt, P.; Haehnel, W. *J. Am. Chem. Soc.* **2001**, *123*, 2186−2195.
(14) Schnepf, R. Ph.D. Thesis, Universität Bochum, 2000.
(15) Mutter, M.; Tuchscherer, G. *Cell. Mol. Life Sci.* **1997**, *53*, 851−863.
(16) Rau, H. K.; DeJonge, N.; Haehnel, W. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 11526−11531.
(17) Rau, H. K.; Haehnel, W. *J. Am. Chem. Soc.* **1998**, *120*, 468−476.
(18) Baltzer, L.; Nilsson, H.; Nilsson, J. *Chem. Rev.* **2001**, *101*, 3153−3163.

amino acids were coupled according to the sequence [K(Dde)-P-G-C(SBu^t)-C(Trt)-C(Acm)-P-G-C(SBu^t)-E(OAll)]. The cleavage of the Fmoc group from the last amino acid K(Dde) of the linear peptide was followed by that of the allyl group of E(OAll) using Pd($(C_6H_5)_3P)_4$ in 37/2/1 (v/v/v) $CHCl_3$/acetic acid/*N*-methylmorpholine to allow for the head-to-tail cyclization of the decapeptide at the resin.[19] After removal of the protecting group of K(Dde), helix A_h was synthesized onto ϵ-N(K) and acetylated, and the template-helix construct was cleaved from the resin and purified by HPLC.[13] The synthesis of the cyclic template T2 (cyclo-[C(SBu^t)-A-C(Trt)-P-G-C(SBu^t)-A-C(Acm)-P-G]) for the proteins in solution has been described previously.[16]

For the assembly of the modular proteins (Mop) the previous procedure[13] was modified as follows. Whatman 1CHR cellulose sheets (7.4 cm × 11.6 cm) were derivatized with β-alanine as described by Frank.[10] The activated modified Rink linker was applied to 48 spots per sheet using a patterned metal plate as a position marker. Activation was achieved by preparing the *N*-hydroxy-9-azabenzotriazole (HOAt) ester of the linker at a concentration of 0.3 M in DMF via reaction with 1.1 equiv of HOAt and 1.1 equiv of *N*,*N*′-diisopropylcarbodiimid (DIC) for 5 min. Acetylation of free amino groups, washing, and Fmoc cleavage were carried out as described by Frank.[10] 3-Maleimidopropionic acid was attached to the Rink linker via the symmetrical anhydride in DMF.[13] To monitor the progress of the synthesis, additional control peptides were prepared on each sheet.

Aliquots (3 μL) of a freshly prepared stock solution of 12.5 mg/mL building block T4-A_h in 2/1 (v/v) 0.25 M sodium phosphate/acetonitrile were given onto each spot at a loading density of approximately 40 nmol/cm^2. After 15 min the excess 3-maleimidopropionyl groups were deactivated by incubation in sodium phosphate buffer (pH 7.5) containing 250 μL of β-mercaptoethanol for 30 min. Subsequently, the SBu^t groups of the building blocks T4-A_h were cleaved by a 1 h treatment in 45/45/10 *n*-propanol/0.1 ammoniumacetate (pH 7.8)/tris-*n*-butylphosphin. The sheets were washed 3 times with *n*-propanol and 3 times with 3/97 (v/v) acetic acid/methanol. After drying, the amount of bound building block T4-A_h was determined by thiol analysis of a few spots.[20] The helices B_i and C_k were attached by chemoselective coupling of their 3-maleimidopropionyl group to the cysteinyl sulfur of the as described.[13]

Aliquots (3 μL) of a solution of 50 mg/mL helix B_i in 2/1 (v/v) 0.25 M sodium phosphate/acetonitrile (pH 8) were transferred to the spots of the sheet and allowed to react for 10 min. This procedure was repeated 1 time. The membranes were washed 3 times with 3/97 (v/v) acetic acid in water, 3 times in 1.5/48.5/50 (v/v/v) acetic acid/water/methanol and 2 times with methanol and dried under reduced pressure.

The Acm protecting group of Cys at the template was cleaved by shaking the membranes in a solution of 75 mg mercuric acetate in 15 mL ammonium acetate (pH 4) for 30 min. Then the sheet was shaken for 3 h with 150 mg of dithiothreitol (DTT) in 0.15 M sodium phosphate (pH 7) and subsequently washed and dried as described.[13] The building blocks C_k were added to the individual spots of the sheet analogous to the coupling of the other helices.

Deprotection of the Acm-protected Cys of helices C_k was performed as described above.

All reactions involving reagents with free thiol groups (DTT and β-mercaptoethanol) were carried out under argon in an airtight container to avoid oxidation. To monitor the assembly of the de novo proteins by HPLC and mass spectrometry, spots with precursors and final products were cut out and cleaved from the membrane by incubation in a solution of 50/45/4/1 (v/v/v/v) TFA/CH_2Cl_2/water/triisopropylsilane, precipitated, and analyzed. To incorporate copper, the membranes were shaken in a solution of 1/1 (v/v) water/acetonitrile, washed 3 times with degassed 100 mM Tris/HCl (pH 7.5), and incubated with a degassed solution of 100 μM $CuCl_2$ in 100 mM NaCl and 100 mM Tris/HCl (pH 7.5).

The soluble variants of the modular four-helix bundle proteins Mop21, Mop22, and Mop23 were synthesized from template T_2 and the corresponding helices A_h, B_i, and C_k as described and purified by HPLC.[13] Metal incorporation was achieved by addition of defined amounts of a Cu(II) solution.

Size-exclusion chromatography was performed essentially as described previously.[13]

UV−vis Absorption Spectroscopy and Screening. UV−vis absorption spectra of the cellulose bound metalloproteins in 100μM $CuCl_2$, 100 mM NaCl and 100 mM Tris/HCl (pH 7.5) were recorded from 200 to 1000 nm at different times after metal ion addition. The experimental setup included a photodiode array spectrometer (J&M, Germany) and a deuterium and halogen light source that were coupled to the membrane by optical fibers. UV−vis absorption spectra of the copper proteins in solution were recorded from a microcuvette of 100 μL volume.

Extinction coefficients of the copper proteins Cu-Mop21, Cu-Mop22, and Cu-Mop23 and the dissociation constant of Cu-Mop21 were determined by metal titration experiments monitoring the absorbance A_{Cu} of the strongest absorption band. Aliquots of a concentrated $CuCl_2$ solution were added to a protein solution of 80 μM in degassed 100 mM Tris/HCl (pH 7.5). At the beginning of the titration, i.e., at an excess of protein, A_{Cu} increases linearly and the extinction coefficient can be derived from the first term of eq 1

$$A_{Cu} = \frac{\epsilon([P] + [M] + K_D)}{2} + \frac{\epsilon\sqrt{([P] + [M] + K_D)^2 - 4[P][M]}}{2} \quad (1)$$

where [P] and [M] denote the concentrations of the protein and the metal ion, respectively, and K_D is the dissociation constant. This latter value was obtained from the entire titration curve by a fit of eq 1.

Electron Paramagnetic Resonance and Resonance Raman Spectroscopy. X-band and S-band electron paramagnetic resonance (EPR) spectra of frozen solutions of the copper proteins were recorded on a Bruker ESP 300E spectrometer equipped with a helium flow cryostat (Oxford Instruments ESR 910) at 60 K in a quartz cell (d = 3 mm, X-band; d = 4 mm, S−Band).

Resonance Raman (RR) spectra were recorded with a double monochromator (Spex 1403, 2400/mm holographic gratings) equipped with a photomultiplier using the 413, 476, 513, and 569 nm lines of a krypton ion laser. The spectral bandwidth was 2.8 cm^{-1}, and the wavenumber increment, 0.5 cm^{-1}. The copper protein complexes were contained in an EPR tube (d = 3 mm) deposited in a cryostat. The spectra were measured at −140 °C with a laser power of less than 40 mW at the sample using a 135° scattering geometry.

Magnetic Circular Dichroism and Circular Dichroism Spectroscopy. MCD and CD spectra were recorded with a Jasco 700 spectrometer equipped with a cryostat (Oxford Instruments, Spectromag) in case of MCD measurements. The samples with a protein concentration of 10 μM in 10 mM Tris/HCl at pH 7.5 were contained in a quartz cuvette of a 1 mm path length. The mean residue molar ellipticity was calculated by dividing the molar ellipticity by the number of residues.

Denaturation was performed by guanidinium hydrochloride (GuHCl) as described previously,[13] and monitored by CD spectroscopy. The free energy of folding was estimated from a fit of ΔG_{obs} (= ΔG_{H_2O} − m[GuHCl]) to the experimental data.

Results

Design and Synthesis of the Proteins. As the starting point, we have chosen the sequence of Cu-Mop5, the most stable copper protein obtained in the first design cycle.[13] Mop5 contains a minimum coordination set of two histidines and one cysteine in the center of the four-helix bundle that allows the

(19) Kates, S. A.; Daniels, S. B.; Albericio, F. *Anal. Biochem.* **1993**, *212*, 303−310.
(20) Riddles, P. W.; Blakeley, R. L.; Zerner, B. *Methods Enzymol.* **1983**, *91*, 49−61.

perfect control and localization of the copper site. This primary coordination sphere proved essentially capable of forming copper sites and was therefore preserved in this work. The same is true for most residues distant from the metal site as well as those on the hydrophilic faces of the peptides. Only a few intrahelical salt bridges were introduced to enhance the stability of the fold and to reduce the net charge from +4 to 0 at pH 7.

In the second design cycle we have focused on the variation of the second coordination sphere, which has a crucial influence on the exact geometry and stability of the metal site as demonstrated by previous studies on copper and zinc proteins.[13] Therefore, we have varied the structure directly above and below the copper site by introducing residues of different polarity, bulkiness, and flexibility (i) to optimize packing interactions and thus the stability of the copper protein and (ii) to impose steric strains in the copper site that may lead to alternative coordination geometries. Furthermore, we have introduced Met in some sequences as an additional ligand to eliminate solvent molecules from the coordination sphere as well as to open a way toward different coordination geometries since Met-ligation is frequently found in natural copper proteins (e.g., type I, Cu_A site). Prior to these sequence changes, we have tried to assess the possible impact of individual substitutions on the basis of structural models that were constructed from the backbone coordinates of ROP, a natural supercoiled four-helix bundle.[21] Following the design criteria outlined above, we synthesized 12 helices $C_{2.k}$, 8 helices $B_{2.i}$, and 3 helices $A_{2.h}$ which were combined to 180 four-helix bundle proteins according to the pattern ABCB (Figure 1). For a cyclic (ABCB) arrangement of the helices, orthogonal protecting groups and linkers were used to assemble the four-helix bundle on a cellulose bound template as described in detail previously.[13] Briefly, a helix A-template conjugate T4-A_h is bound to the cellulose membrane and consecutively coupled to two helices B and one helix C in alternating orientation. Following this strategy, helix A_1 was combined with both sets of helices B and C to 96 cellulose-bound proteins according to the pattern ABCB. To check the correct composition of the bundles, some proteins were cleaved from the cellulose membrane and characterized by ESI-MS as described.[13] Based on the copper binding capabilities of the investigated proteins, a second set of promising four helix bundles were synthesized by combining helices A_2 and A_3 with B_i (i = 1−3, 5−7) and C_k (k = 1, 3, 4, 6, 8, 10,11). Further details of the synthesis strategy are given elsewhere.[13,14]

Screening. Copper binding of the cellulose-bound proteins was achieved by addition of 100 μM $CuCl_2$ in Tris buffer (pH 7.5). Figure 2 shows on a single cellulose sheet spots with three different colors of the copper proteins, purple, brown, and yellow. The UV−vis absorption spectra were recorded directly on the membrane.[13] This technique allows us to categorize cysteinate-containing copper sites by their relatively strong (S−Cu) CT absorptions between 400 and 800 nm into several classes called type I, type II, or Cu_A.[22−24] Therefore, the absorption

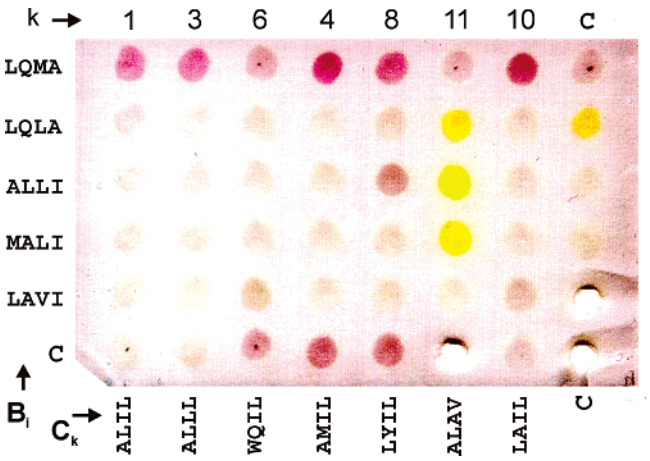

Figure 2. Cellulose membrane with immobilized proteins, taken 40 min after addition of a 100 μM $CuCl_2$ solution (100 mM NaCl, 50 mM Tris, pH 7.5). All proteins include helix A_3. The amino acids varied in helices B_i and C_k are indicated (cf. Figure 1). C indicates control proteins. In the lower right corner three spots have been punched out for control by mass spectrometry.

spectra of the cellulose bound proteins were measured at certain time intervals after $CuCl_2$ addition in the range from 200 to 1000 nm directly from the cellulose membrane. The absorption bands and the time-dependent decays reveal information on the type of copper site and the stability of the complex. The results obtained by combination of helix $A_{2.1}$ with helices B and C are summarized in Figure 3. The spectra obtained for the proteins containing helix $A_{2.2}$ and $A_{2.3}$ are almost identical to those of the proteins including helix $A_{2.1}$ (data not shown).

The arrangement of the matrixes displays the effect of the individual amino acid substitutions in the helices on the properties of the metal site. Altogether 90% of the synthesized proteins exhibit strong absorption bands characteristic of copper binding. In addition, the stability as monitored by the time-dependent decay of the absorption bands (Figure 3) has increased by 2 orders of magnitude compared to the first synthesis cycle, which reflects the progress achieved in the second cycle. The most unexpected result, however, is the diversity of absorption spectra observed for the different proteins, even though the position of the core coordination set His_2Cys remained unchanged.

A careful analysis of all spectra suggests three classes of copper sites that can be represented by the proteins Cu-Mop21 (yellow), Cu-Mop22 (purple), and Cu-Mop23 (brown or blue) (Table 1, Figures 1, 4). These proteins were synthesized in their soluble form to allow for a thorough spectroscopic characterization. The sequences of the soluble proteins were identical to those of the cellulose-bound forms except for two minor alterations which had been introduced for practical reasons.[13] First, the linker between helix A and the template T4 was substituted by a linker of similar length. Second, a different template T2 has been used that lacks the anchor for cellulose binding.

UV−vis Absorption Spectroscopy. The UV−vis spectra of the copper complexes of the soluble proteins Mop21, Mop22, and Mop23 (Figure 4) are essentially identical to those of the immobilized counterparts which further proves that spectroscopic screening on the solid phase is a reliable method to identify novel copper sites in de novo proteins.[13] Only for Cu-Mop23, minor differences between the solution and solid-phase

(21) Banner, D. W.; Kokkinidis, M.; Tsernoglou, D. *J. Mol. Biol.* **1987**, *196*, 657−675.
(22) Gamelin, D. R.; Randall, D. W.; Hay, M. T.; Houser, R. P.; Mulder, T. C.; Canters, G. W.; de Vries, S.; Tolman, W. B.; Lu, Y.; Solomon, E. I. *J. Am. Chem. Soc.* **1998**, *120*, 5246−5263.
(23) LaCroix, L. B.; Randall, D. W.; Nersissian, A. M.; Hoitink, C. W. G.; Canters, G. W.; Valentine, J. S.; Solomon, E. I. *J. Am. Chem. Soc.* **1998**, *120*, 9621−9631.
(24) Solomon, E. I.; Baldwin, M. J.; Lowery, M. D. *Chem. Rev.* **1992**, *92*, 521−542.

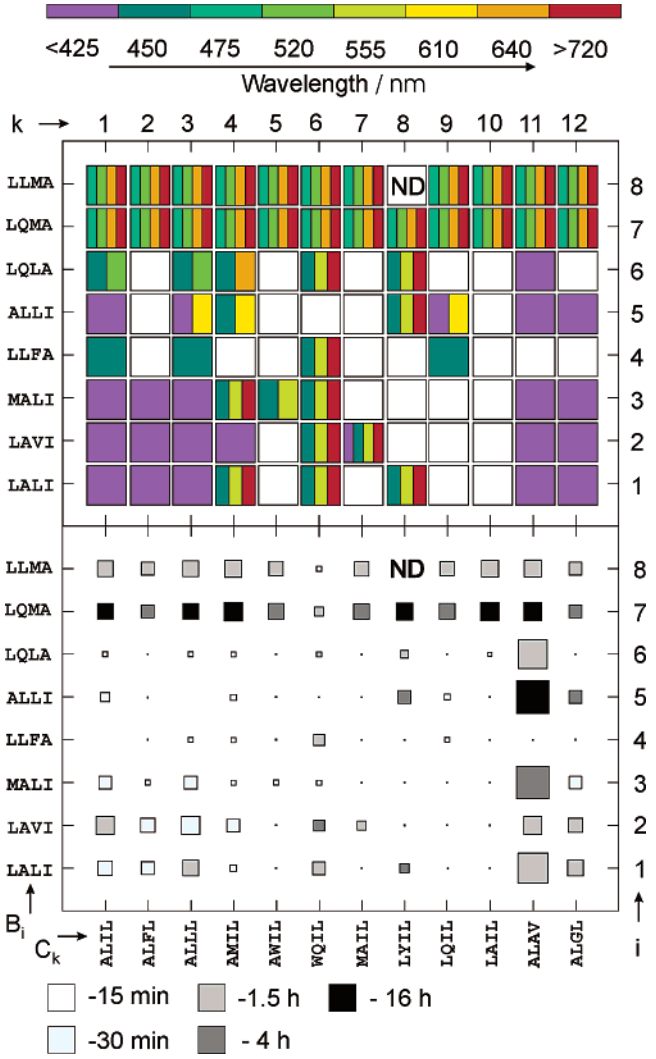

Figure 3. Matrix representation of the absorption properties of the copper protein complexes including helix A_1 (ND, not determined). In the upper matrix the colors reflect the position of the detectable absorption bands in the visible spectrum. Various colors within one square indicate several absorption bands. The size of the squares in the lower matrix is proportional to the intensity of the S→Cu CT band, and the gray scale provides a semiquantitative description of the temporal stability of the copper complexes. The approximate "lifetimes" refer to the time at which the intensity of the absorption bands has decreased to 75% of the original value. The residues that are varied are indicated at the ordinate (B_i) and abscissa (C_k) (cf. Figure 1).

Table 1. Molecular Masses, UV−vis Absorption Bands, and Copper Binding Constants of Mop21, Mop22, and Mop23

protein	Mop21	Mop22	Mop23
assembled modules	$T_2A_2(B_5)_2C_{11}$	$T_2A_2(B_7)_2C_4$	$T_2A_1(B_5)_2C_8$
calcd av mass	9229.5	9369.7	9335.7
exptl mass[a]	9229.0 ± 1.0	9374.6 ± 2.6	9337.9 ± 0.7
absorption bands,[b] λ/nm	400	374, *477*, 525, 637, 774	429, *571*
ϵ/M^{-1} cm^{-1} [b]	6100	4040	1000
K_{dis}/nM[c]	6	nd[d]	nd[d]

[a] Determined for the apoproteins by ESI mass spectrometry, deconvoluted masses. [b] The strongest absorption bands of the holoproteins to which the extinction coefficients refer are italicized. [c] Expressed in terms of the dissociation constant of the copper protein complex. [d] Not determined.

spectra were noted which indicates some flexibility of the metal site. This is also reflected by a somewhat lower temporal stability of Cu-Mop23 in solution, whereas Cu-Mop21 and Cu-Mop22 display the same stability in solution as in the solid phase.

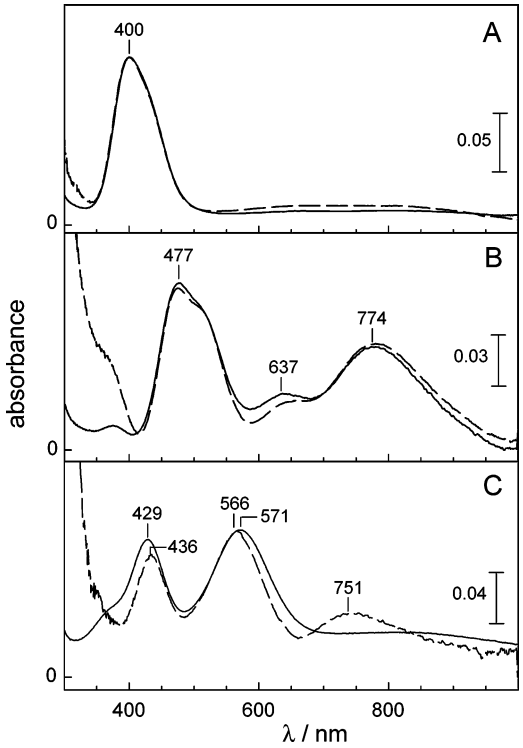

Figure 4. UV−vis absorption spectra of the Cu(II) complexes of (A) Mop21 (50 μM protein + 25 μM $Cu(II)Cl_2$), (B) Mop22 (50 μM protein + 50 μM $Cu(II)Cl_2$), and (C) Mop23 (250 μM protein and 120 μM $Cu(II)Cl_2$) in solution in 100 mM NaCl and 50 mM Tris (pH 7.5, 25 °C). The extinction coefficients for the strongest bands of the Cu(II) complexes have been estimated to be 6100, 4040, and 1000 M^{-1} cm^{-1} for Mop21, Mop22, and Mop23, respectively. The solid and dashed lines refer to the copper protein complexes in solution and in the solid phase, respectively.

However, the most striking result refers to the variations in the absorption spectra which are observed for very similar structural motifs. The differences in the absorption spectra can be related to different copper site geometries on the basis of empirical data accumulated for natural copper proteins.[2,3]

Like Cu-Mop5,[13] the starting protein of this study, the yellow Cu-Mop21 shows a single strong absorption band at ca. 400 nm indicating a dominant (S→Cu) σ→p* CT transition and thus a tetragonal Cu site. Presumably, a solvent molecule or possibly the carbonyl function of an adjacent amino acid completes the implemented His_2Cys coordination set.

The absorption spectrum of the bluish Cu−Mop23 reveals two equally strong bands at 429 and 571 nm which can be assigned to a (S→Cu) σ→p* CT and a (S→Cu) π→p* CT, respectively. This spectral motif is reminiscent of so-called type I copper proteins such as nitrite reductase and can be ascribed to a distorted tetrahedral coordination geometry.[25] A fourth ligand might weakly interact with the Cu(II) ion although it cannot yet be identified. The Tyr residue of helix C_8 (Figure 1) can be ruled out due to steric hindrance as can be shown by the structural model.

The most complex absorption spectrum was obtained for the purple copper complex of Mop22, which is unmatched by

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

natural as well synthetic monocopper sites. However, the spectrum reveals far-reaching similarities with that of the mixed valent dinuclear Cu_A site of cytochrome *c* oxidase. Each absorption band of Cu-Mop22 has its counterpart in the spectrum of the Cu_A site with respect to both the absorption maximum and the relative intensity. RR spectroscopic and theoretical studies consistently indicate that the absorption band at 770 nm arises from a $\Psi \rightarrow Y$ transition in the Cu−Cu moiety, whereas the intense transitions around 480−550 nm originate from (S→Cu) CT transitions.[26−28]

In contrast to the Cu_A sites in natural proteins, the putative $Met_2CysHis_2$ ligand set in Mop22 includes only one bridging cysteine. However, further evidences for assignment of the copper center in Cu-Mop22 to a Cu_A site will be given below. For the sake of simplicity, this copper protein is now referred to as Cu_2-Mop22.

Dissociation Constants and Extinction Coefficients. Cu(II)-titration experiments of the copper proteins Cu-Mop21, Cu_2-Mop22, and Cu-Mop23 were performed by stepwise addition of $CuCl_2$ aliquots to the apoprotein solution and monitored by UV−vis absorption spectroscopy. The extinction coefficients derived from these experiments are in the range of values for natural and engineered copper proteins (Table 1).[29] During titration absolute band intensities increase, whereas no changes of energies and relative band intensities were observed indicating the formation of homogeneous complexes.

The protein concentration determined on the basis of the estimated extinction coefficient was considerable lower than the protein concentration derived from the thiol-selective Ellmans assay. For Mop21 and Mop23 one could imagine that an initial redox reaction of Cu(II) competes with copper coordination. This assumption is supported by the observation that the inconsistencies described above are only noted in the case of Cu(II) but not upon binding with redox inactive metals such as Co(II). Since the mixed-valent Cu_A site includes a Cu(II) and a Cu(I) center, an additional redox reaction must be involved in the case of Mop22. In the absence of any other reducing agent, sacrificial cysteine oxidation might take place as it has been shown for an engineered Cu_A site in azurin by the addition of either Cu(II) or a mixture of Cu(I) and Cu(II).[30,31] Indeed, for Cu_2-Mop22 the maximum extinction is 80% higher in case of an addition of Cu(I)/Cu(II) compared to Cu(II), whereas for Cu-Mop21 and Cu-Mop23 such an effect was not observed. Due to the complex binding kinetics for Cu_2-Mop22 and the lower stability of Cu-Mop23 in solution, no reliable determination of the dissociation constant was possible for these complexes. However, the value determined for Cu-Mop21 (Table 1) compares well with those obtained for natural copper proteins.[29]

RR Spectroscopy. Excitation in resonance with the (S→Cu) LMCT transition provides a selective enhancement of the vibrational modes involving internal coordinates of the copper-cysteinate moiety (Figure 5).[32] The strongest resonance enhancement is expected for the mode with the largest Cu−S stretching contribution, which represents a marker for the Cu−S bond distance and thus the type of copper site.

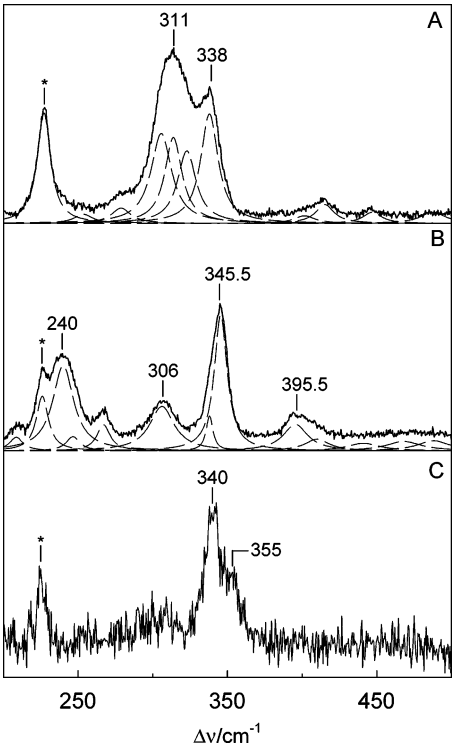

Figure 5. RR spectra of the copper complexes of (A) Mop21, (B) Mop22, and (C) Mop23 at a concentration of ca. 1.5 mM in 100 mM NaCl and 100 mM Tris (pH 7.5) measured at −140 °C. The excitation line was 413, 482, and 568 nm in A, B, and C, respectively. Raman bands of ice are marked by asterisks. The dashed lines refer to fitted Lorentzian line shapes.

The RR spectrum of the Cu−Mop21, obtained with excitation at 413 nm, reveals several bands ranging from 305 to 340 cm^{-1} that can be assigned to modes with significant contributions of the Cu−S stretching. Further RR bands at 278, 415, and 447 cm^{-1} are relatively weak and most likely originate from modes involving internal coordinates of the cysteine ligand. The center of the Cu−S stretching motif can be estimated to be at 323 cm^{-1} indicating a rather weak Cu−S bond as it is found in tetragonal type II copper sites. Based on empirical relationships derived from natural copper proteins, this frequency corresponds to a bond length of approximately 2.23 Å.[33]

For Cu-Mop23, the 568 nm excitation line was chosen that is in resonance with the low-frequency LMCT absorption band. The RR spectrum obtained in this way is dominated by two overlapping bands at 340 and 355 cm^{-1}. This marked frequency upshift compared to Cu-Mop21 can be attributed to a strengthening of the Cu−S bond, which is in line with a change from tetragonal (type II) to tetrahedrally distorted (type I) and the conclusion derived from the UV−vis absorption spectra. The peak center is at ca. 345 cm^{-1}, corresponding to a bond length of 2.205 Å.[33] Additional weaker bands are not discussed further due to the relative low signal-to-noise ratio of the RR spectrum.

A detailed analysis of the results obtained from the monocopper proteins in this and in our previous study reveals that

(26) Hay, M.; Richards, J. H.; Lu, Y. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 461−464.
(27) Andrew, C. R.; Han, J.; DeVries, S.; Van de Oost, J.; Averill, B. A.; Loehr, T. M.; Sanders-Loehr, J. *J. Am. Chem. Soc.* **1994**, *116*, 10805−10806.
(28) Lappalainen, P.; Aasa, R.; Malmstrom, B. G.; Saraste, M. *J. Biol. Chem.* **1993**, *268*, 26416−26421.
(29) Andrew, C. R.; Yeom, Y.; Valentine, J. S.; Karlsson, G.; Bonander, N.; van Pouderoyen, G.; Canters, G. W.; Loehr, T. M.; Sanders-Loehr, J. *J. Am. Chem. Soc.* **1994**, *116*, 11489−11498.
(30) Wang, X. T.; Ang, M. C.; Lu, Y. *J. Am. Chem. Soc.* **1999**, *121*, 2947−2948.
(31) Hay, M. T.; Ang, M. C.; Gamelin, D. R.; Solomon, E. I.; Antholine, W. E.; Ralle, M.; Blackburn, N. J.; Massey, P. D.; Wang, X. T.; Kwon, A. H.; Lu, Y. *Inorg. Chem.* **1998**, *37*, 191−198.
(32) Qiu, D.; Dasgupta, S.; Kozlowski, P. M.; Goddard, W. A.; Spiro, T. G. *J. Am. Chem. Soc.* **1998**, *120*, 12791−12797.
(33) Andrew, C. R.; Sanders-Loehr, J. *Acc. Chem. Res.* **1996**, *29*, 365−372.

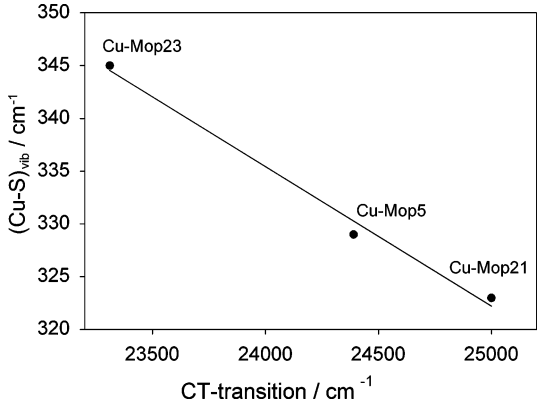

Figure 6. Correlation between the Cu−S stretching frequency and the S→Cu CT transition in de novo four-helix bundle copper proteins.

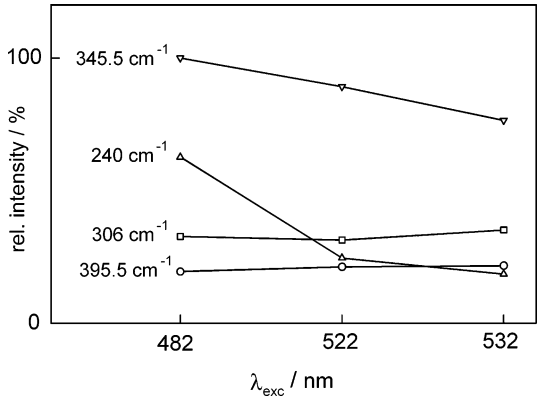

Figure 7. Excitation profiles of RR bands of Cu_2-Mop22. Band intensities were determined relative to the Raman band of ice at 226 cm^{-1} and normalized to the strongest RR band at 482 nm excitation.

the observed Cu−S stretching frequency increases with the position of the high energy S−Cu LMCT band (Figure 6). This correlation, although it is only based on three data points, is in line with the results obtained for natural copper proteins and indicates that the geometry of this copper site is reflected by the electronic and vibrational spectra.[33]

The RR spectrum of the Cu_2-Mop22 is distinctly different from the monocopper proteins Cu-Mop21 and Cu-Mop23, since it displays several bands of comparable RR intensity in a relatively wide frequency range, i.e., at 240, 306, 345.5, and 395.5 cm^{-1}. In fact, natural Cu_A sites as well as model compounds exhibit spectra with a similar distribution of several intense RR bands in this region.[33−35] Among them, there is a characteristic marker band at ca. 340 cm^{-1} that is well matched by the 345.5 cm^{-1} band of Cu_2-Mop22. Such a good agreement in the frequencies is not found for the other bands, which however is not surprising since there is only one bridging cysteine ligand in Cu_2-Mop22. There is no doubt that the differences in geometry affect the nature and position of vibrational transitions. To gain more insight in the nature of the absorptions, we have measured the RR spectra at different excitation lines (522, 482, and 532 nm). The corresponding excitation profiles of the four most intense peaks (Figure 7) resemble those obtained from natural Cu_A sites,[27,33] suggesting that the absorption bands at 477 and 525 nm originate from the same transitions as those of natural Cu_A sites, which are (S→Cu) LMCT transitions.

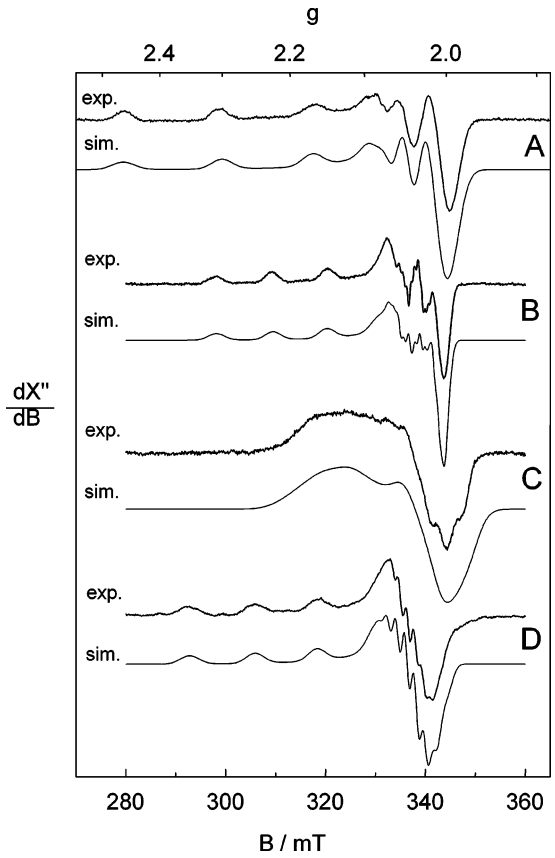

Figure 8. X-band EPR spectra of (A) $Cu(II)Cl_2$, (B) Cu-Mop21, (C) Cu_2-Mop22, and (D) Cu-Mop23 at a concentration of (A) 50 μM, (B) 80 μM, (C) 35 μM, and (D) 50 μM in 50 mM Tris (pH 7.5) and 100 mM NaCl at 60 K. The microwave frequencies and the modulation amplitudes were 9.634 19 GHz and 11.4 G in spectrum A, 9.634 73 GHz and 4.0 G in spectrum B, 9.633 62 GHz and 11.4 G in spectrum C, and 9.623347 GHz and 8 G in spectrum D. The microwave power was 5 mW in all experiments. The thin lines represent the simulated spectra. Simulation parameters are given in Table 2.

Table 2. EPR Spectroscopic Parameters[a] of Copper Protein Complexes and $CuCl_2$ in Solution

parameter	Cu-Mop21	Cu_2-Mop22	Cu-Mop23	$CuCl_2$(Tris)
g_x	2.032	2.010	2.040	2.037
g_y	2.036	2.035	2.052	2.050
g_z	2.190	2.155	2.210	2.237
$A_{x,Cu}$	24	18	15	35
$A_{y,Cu}$	21	32	13	32
$A_{z,Cu}$	107	30	123	184
$A_{x,N}$	7		17	
$A_{y,N}$	9		17	
$A_{z,N}$	7		10	

[a] Data obtained from simulations of X- and S-band EPR spectra. Hyperfine coupling constants are given in 10^{-4} cm^{-1}. For Cu-Mop21 and Cu-Mop23, coupling with two equivalent N atoms was included.

EPR Spectroscopy. The experimental and simulated X-band EPR spectra of the Cu(II) complexes of the copper proteins Cu-Mop21, Cu_2-Mop22, and Cu-Mop23 are distinctly different from that of $CuCl_2$ dissolved in Tris buffer (Figure 8, Table 2), further confirming specific binding sites in the individual proteins. In particular, we note substantially decreased A_z values that suggest a pronounced covalent interaction like a Cu−S bond. In addition,

(34) Wallace-Williams, S. E.; James, C. A.; DeVries, S.; Saraste, M.; Lappalainen, P.; Van de Oost, J.; Fabian, M.; Palmer, G.; Woodruff, W. H. *J. Am. Chem. Soc.* **1996**, *118*, 3986−3987.
(35) Andrew, C. R.; Fraczkiewicz, R.; Czernuszewicz, R. S.; Lappalainen, P.; Saraste, M.; Sanders-Loehr, J. *J. Am. Chem. Soc.* **1996**, *118*, 10436−10445.

both monocopper sites in Cu-Mop21 and Cu-Mop23 reveal marked super hyperfine splitting in the g_x and g_y regions. To reproduce the line shape of those spectra in detail, two equivalent N nuclei and one copper nucleus had to be included in the simulation. Thus, these results agree nicely with the proposed minimal ligand set of His_2Cys. The calculated A_z values were 123×10^{-4} cm^{-1} for Cu-Mop23 and 107×10^{-4} cm^{-1} for Cu-Mop21, which are in the boundary region of the A_z values reported to be characteristic of type I ($<100 \times 10^{-4}$ cm^{-1}) and type II sites ($>130 \times 10^{-4}$ cm^{-1}), respectively.[29,36,37] Although the EPR data alone do not allow an unambiguous identification of the type of copper sites, they are consistent with conclusions drawn from UV−vis absorption and RR spectra that indicate a tetragonal and distorted tetrahedral structure for Cu-Mop21 and Cu-Mop23, respectively.

Simulation of the EPR spectrum of the dinuclear copper complex Cu_2-Mop22 should allow a decision of whether the unpaired electron is localized on one copper or delocalized over two copper atoms. In the latter case, the *g* values are expected to split into a characteristic seven-band pattern according to $2(^3/_2)$, which should be readily detectable for the g_z region. However, the poor resolution of the X-band spectra of Cu_2-Mop22 does not permit a final conclusion. To obtain a more dispersed hyperfine structure, the EPR spectra were monitored at lower frequency in the S-band. In fact, the formerly shapeless g_z region is resolved into distinct bands (Figure 9). Simultaneous simulation of the two S-band and the X-band spectra was performed on the basis of two alternative assumptions, i.e., a localized and a delocalized unpaired electron. For the hypothesis of a localized spin we included two ^{14}N nuclei in the parameter set to reproduce the multiplicity of the EPR bands but a simulation of all spectra with the same set of parameters was not possible. In addition to this inconsistency, a rather unusually high coupling constant $A_{z,N}$ was required, which rules out the hypothesis of a localized spin. The second hypothesis appears to be more realistic as one set of parameters including two equivalent copper nuclei was sufficient to reproduce all three experimental spectra with respect to multiplicity, position, and intensity in a satisfactory manner. Furthermore, also the relatively small A_z value of 30×10^{-4} cm^{-1} (Table 2) indicates a delocalization of the electron.[38] Some minor deviations might be due to differences in the local environment of the two copper atoms. Nonetheless, these findings strongly support the idea of a delocalized unpaired electron in a dicopper site as it is also found in natural proteins.[38]

MCD Spectroscopy. MCD spectroscopy is a powerful method for the elucidation of absorption spectra of transition metal complexes. In general, the MCD signals can be of positive and negative signs and intensities for LMCT, and ligand field transitions frequently differ from those detected by absorption spectroscopy, thereby facilitating the assignments of the absorption bands to individual electronic transitions. This is also documented by the MCD spectra of Cu-Mop21 inasmuch as the spectrum exhibits numerous signals of opposite signs that show intensities remarkably different from those of the absorption spectrum (Figure 10). For instance, the strong MCD signal at 15 500 cm^{-1} is hardly detectable in the absorption spectrum and, thus, is assigned to a ligand field type transition. The long wavelength shoulder of the 400 nm absorption band consists of two individual transitions, as indicated by the two MCD signals of opposite sign. In the region from 13 000 to 27 000 cm^{-1}, a consistent fit to the experimental spectra was achieved on the basis on eight bands, each of them having the same frequency but different bandwidths and intensities in the MCD and the UV−vis absorption spectra. Considering theoretical and empirical data on type I and type II copper proteins, we provide a preliminary assignment of the absorption bands (Table 3)[23,39] that is consistent with a type II center, as it has been inferred from UV−vis absorption, RR, and EPR data. Specifically, the strong absorption bands at ca. 400 nm, designated as transitions 5 and 7, were both assigned to S→Cu CT transitions.

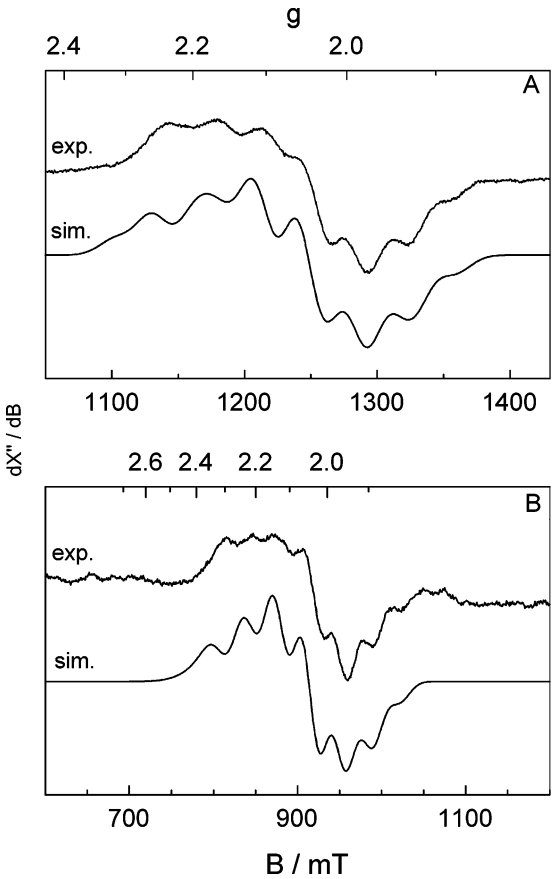

Figure 9. S-band EPR spectra of Cu_2Mop22 at a concentration of 400 μM in 50 mM Tris (pH 7.5) and 100 mM NaCl at 60 K. The microwave frequencies were 3.574 63 and 2.618 65 GHz in parts A and B, respectively. The microwave power and modulation amplitude were 400 μW and 9.6 G, respectively. The thin lines represent the simulated spectra. Simulation parameters are given in Table 2.

Cu_2−Mop22 reveals a complex MCD spectrum (Figure 11) that bears a remarkable resemblance to those spectra obtained from natural Cu_A sites (inset).[38] Even the signs of the key motif, the double-banded feature at ca. 500 nm, are nicely reproduced suggesting a similar electronic structure. The band fitting analysis of the MCD and UV−vis absorption spectra reveals nine bands. Since MCD and UV−vis absorption spectra have been recorded at different temperatures, in this case also

(36) Palmer, A. E.; Randell, D. W.; Xu, F.; Solomon, E. I. *J. Am. Chem. Soc.* **1999**, *121*, 7138−7149.
(37) 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*, 5907−5908.
(38) Farrar, J. A.; Neese, F.; Lappalainen, P.; Kroneck, P. M. H.; Saraste, M.; Zumft, W. G.; Thomson, A. J. *J. Am. Chem. Soc.* **1996**, *118*, 11501−11514.
(39) Lu, Y.; Roe, J. A.; Bender, C. J.; Peisach, J.; Banci, L.; Bertini, I.; Gralla, E. B.; Valentine, J. S. *Inorg. Chem.* **1996**, *35*, 1692−1700.

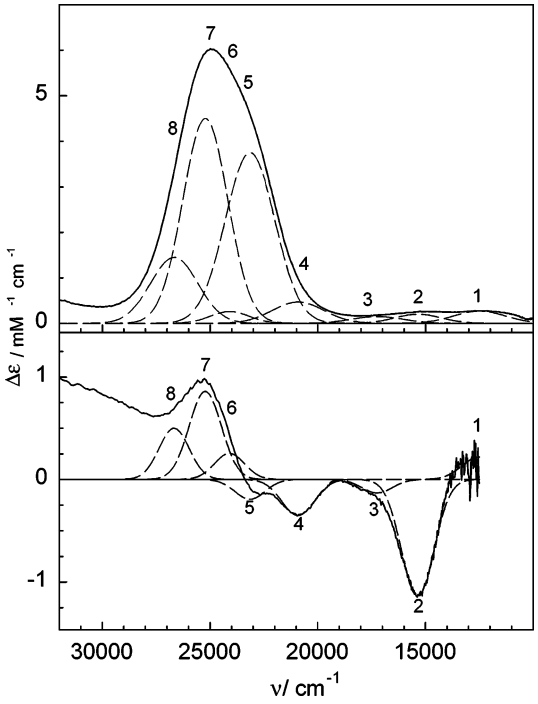

Figure 10. UV−vis absorption spectrum measured at 298 K (top) and MCD spectrum at 6.3 K and 1.8 T (bottom) of Cu-Mop21. MCD experiments were carried out with a protein concentration of 380 μM in 100 mM Tris (pH 7.5) and 50% v/v glycerol. Experimental conditions for the UV−vis absorption spectra were the same as those in Figure 4. The dashed lines represent Gaussian line shapes obtained in a fit to both the UV−vis absorption and MCD spectra as described in the text.

Table 3. Electronic Transitions of Cu-Mop21[a]

band no.	band position/cm^{-1}	assignment
1	12 500	Cu d→d
2	15 361	Cu d→d
3	17 241	Cys-S$_\pi$→Cu?
4	20 921	
5	23 148	Cys-S$_\sigma$→Cu
6	24 096	His-N→Cu?
8	26 667	
7	25 221	Cys-S$_\sigma$→Cu

[a] Assignments of the Gaussian bands that were fitted to the UV−vis absorption and MCD spectra were made on the basis of empirical data of type II copper proteins. Further details are given in the text.

variation of the band maxima was allowed to take into account temperature-induced frequency shifts. These frequency differences between MCD and UV−vis absorption spectra were found to be comparable to those observed in the UV−vis absorption spectra measured at different temperatures (results not shown). Each of the nine electronic transitions determined for Cu_2-Mop22 in this way has its counterpart of very similar energy in the spectra Cu_A sites in natural proteins (Table 4). This finding is even more remarkable in view of the different ligand set in Cu_2-Mop22.

CD Spectroscopy and Size Exclusion Chromatography. Size exclusion chromatography and CD spectroscopy in combination with chemically induced denaturation studies are common methods to assess protein secondary structure and stability. The apo- and copper-proteins that were investigated show characteristic α-helical CD spectra with minima at 208 and 222 nm in agreement with the proposed structural model (Table 5). Upon copper complex formation, no significant

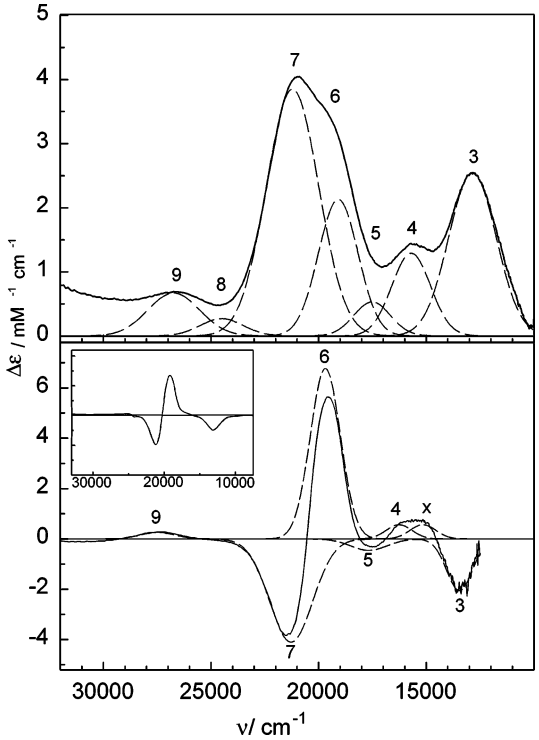

Figure 11. UV−vis absorption spectrum measured at 298 K (top) and MCD spectrum at 5.8 K and 1.8 T (bottom) of Cu_2-Mop22. MCD experiments were carried out with a protein concentration of 380 μM in 100 mM Tris (pH 7.5) and 50% v/v glycerol. Experimental conditions for the UV−vis absorption spectra were the same as those in Figure 4. The dashed lines represent Gaussian line shapes obtained in a fit to both the UV−vis absorption and MCD spectra as described in the text. The inset shows the MCD spectrum of the Cu_A site of NO reductase.[38]

Table 4. Electronic Transitions of Cu_2-Mop22[a]

band no.	Cu_2-Mop22 (UV–vis)	Cu_2-Mop22 (MCD)	Cu_A, PdII[b]	Cu_A, COX, *B. subtilis*[c]	Cu_A, azurin[d]
3	12 853	13 459	12 690	13 175	13 670
x		15 152			
4	15 699	16 260		15 000	15 400
5	17 513	17 668	17 000	16 500	17 700
6	19 084	19 685	19 080	18 800	18 860
7	21 209	21 277	20 870	21 060	20 670
8	24 510			23 000	
9	26 738	27 473	(26 055)	26 500	27 500

[a] Assignments of the Gaussian bands that were fitted to the UV−vis absorption and MCD spectra were made on the basis of empirical data of Cu_A proteins. Further details are given in the text. [b] Cu_A site of subunit II from cytochrome *c* oxidase from *P. denitrificans*.[38] [c] Cu_A site of subunit II from cytochrome *c* oxidase from *B. subtilis*.[35] [d] Cu_A site obtained by site-directed mutagenesis of azurin.[31]

Table 5. CD Spectroscopic Properties of Copper Protein Complexes and Apoproteins in Solution

protein	θ_{222}/ deg cm^2 dmol^{-1}	% helix, exptl[a]	% helix, theor
Mop21	−24 375	68−76	80
Cu-Mop21	−24 232	68−75	80
Mop22	−22 855	64−71	80
Cu_2-Mop22	−23 375	65−73	80

[a] Determination of the helix content is based on $\theta_{222} = -35\,700$ and $-32\,200$ deg cm^2 dmol^{-1}.[40,41]

spectral change was detectable ruling out significant alterations of the secondary structure.

GuHCl-induced denaturation, monitored by CD spectroscopy, allows determining the free energy of folding ΔG. These

Table 6. Apparent Molar Masses[a] and Thermodynamic Data of GuHCL Induced Unfolding of Apoproteins and Zn-Protein Complexes

protein	apparent molar mass/ kDa	ΔG/J mol^{-1}	m/J mol^{-1}
Mop21	14.6	6826	2853
Zn-Mop21	14.1	7939	3029
Mop22	16.0	8442	3027
Zn-Mop22	13.9	7905	3015
Mop23	14.1	nd[b]	nd[b]
Zn-Mop23	14.3	nd[b]	nd[b]

[a] Determined by size exclusion chromatography. [b] Not determined.

experiments were carried out with the Zn(II)-protein complexes to avoid interference by Cu(II)-induced cysteine oxidation. The values were found to lie in the range of natural proteins exhibiting a small hydrophobic core (Table 6). Possible aggregation was analyzed by size exclusion chromatography for the apoproteins and Zn(II)-protein complexes. Apo- and holoproteins eluted in single peaks with apparent masses around 14−16 kD (Table 6), which is in good agreement with a monomeric state. The deviation from the actual mass is attributed to the nonglobular shape of the protein and the loosely packed linker region.

Binding Capabilities of Divalent Metals. In synthetic inorganic chemistry, the stability of complexes of divalent metal ions are largely determined by electrostatic effects and can be described by the Irving−Williams series with an increasing stability in the order $Co^{2+} < Cu^{2+} < Zn^{2+}$.[42] This tendency was also observed for Mop5 in our previous study.[13] Mop21 and Mop22, however, display a different behavior. Incubating the apoproteins stepwise with equimolar equivalents of two different divalent metal ions demonstrates that Cu^{2+} was the only metal ion that is capable of replacing the other divalent metal ions. These findings indicate that for Mop21 and Mop22 the metal binding affinity cannot solely be attributed to electrostatic effect, but additional steric and electronic factors are likely to be involved. Such deviations from the Irving−Williams series are also known for natural proteins and reflect the metal specificity of binding sites. Thus the results also imply that the second design cycle did yield copper proteins with not only higher stability but also higher metal specificity.

Discussion

Implementation of a specific thiolate-coordinated copper binding site into a four-helix bundle protein imposes high demands on the stability of the protein fold and the appropriate preorientation of the ligands. This task is particular challenging, since no natural four-helix bundle protein exists that could serve as a structural model. Only the minimum coordination set His_2/Cys was adopted from natural β-sheet-dominated type I copper proteins. In this respect, the design of the present copper proteins was truly de novo. The results of the previous[13] and present work further demonstrate that specifically in absence of a structural model a purely rational design will remain unsuccessful in generating stable metalloproteins. Moreover, repetitive combinatorial design cycles provide the clue for the optimization of the desired functions. This conclusion is nicely illustrated by the increasing success of the two design cycles. The first cycle comprising 96 proteins yielded only ca. 30 proteins that were capable of binding copper, and among them only one was sufficiently stable at low temperature for a spectroscopic characterization.[13] In the second cycle, 90% of the 180 proteins formed copper complexes, and nearly all of them exhibit a stability higher by 2 orders of magnitude. The sequence modifications made in the second cycle were directed to improve the design in two ways.

At first, we attempted to enhance the stability of the overall protein fold inter alia via revising the hydrophilic faces of the peptides. These changes had no notable effect on the GuHCl-induced denaturation inasmuch as the unfolding energy did not increase compared to Mop5 of the first design cycle. However, these modifications may also contribute to the increased stability of the copper binding site itself which was the second objective of this design cycle.

While keeping the original positions of the ligating residues His_2Cys unchanged, amino acids in the vicinity of the first coordination sphere were exchanged resulting in variations of the steric demands (Leu vs Val), polarity (Leu vs Glu), or flexibility (Leu vs Ile).

The minimum His_2Cys ligation set for a Cu binding site allows generating structurally different Cu complexes upon altering the secondary coordination sphere. The detailed spectroscopic analysis of representative proteins demonstrates that it is possible to generate three types of copper sites that are similar to those in natural proteins.

Copper Sites of Type Cu_A. Cu_A-like copper sites were obtained by substituting Met at position Z13, four residues apart from a coordinating His (Figure 1). This intrahelical distance of two ligating residues is a familiar motif in helical metalloproteins. Evidence for Met as a copper ligand is provided by the fact that exclusively the proteins with Met at this position formed a stable copper site with the unique spectral properties. These properties are completely lost upon isosteric replacement of Met through Leu. The most stable proteins are obtained by the helical combinations of helix B_7 and helix C_4, whereas combinations including helix C_{11} and B_7 (B_8) form rather unstable complexes, even though helix C_{11} forms the most stable type II sites (Figure 3). Evidently, the steric demands are different for the various copper sites. Indeed, natural four-helix bundles containing a dicopper diiron site are characterized by a broadening of the bundle that requires medium-sized residues to fill the gaps above and below the site.[43,44] Residues that are too bulky, however, as in the case of X11-Phe in C_2 may also have a deleterious effect on complex stability.

Copper Sites of Type II. Even more sensitive to the steric environment are the type II copper sites. Here, the interplay of helices B and C is necessary to create conditions for an appropriate fold. A careful inspection of the matrixes in Figure 2 reveals two areas where this type of complex is formed: in the lower left part in columns 1 to 4 and in the lower right part in columns 11 to 12. The initial protein of this study, Mop5, falls into the first group, and its hydrophobic core composition is represented by the bundle $T4(B_1)_2C_1$, disregarding the

(40) Lau, S. Y. M.; Taneja, A. K.; Hodges, R. S. *J. Biol. Chem.* **1984**, *259*, 13253−13261.
(41) Chen, Y. H.; Yang, J. T.; Chau, K. H. *Biochemistry* **1974**, *13*, 3350−3359.
(42) Lippard, S. J.; Berg, J. M. *Bioinorganic Chemistry*; John Wiley & Sons: New York, 1998.
(43) Sheriff, S.; Hendrickson, W. A.; Smith, J. L. *J. Mol. Biol.* **1987**, *197*, 273−296.
(44) Volbeda, A.; Hol, W. G. J. *J. Mol. Biol.* **1989**, *209*, 249−279.

substitutions on the hydrophilic sides of the peptides made in the second design cycle. A significantly improved stability was achieved by the replacement of Leu by Val at position Z13 in $T4(B_2)_2C_1$. The steric effect of this substitution is particularly pronounced since helix B occurs 2-fold in the bundle. Residues at Z13 that are too small like Ala have a destabilizing effect as shown in our previous work. The isosteric replacement of Leu by Ile at position X11 has no major influence on the stability, whereas the introduction of sterically demanding side chains at position X4 leads to the formation of different copper site geometries or the total loss of the copper site.

The most stable copper complexes on the right side of the matrix are characterized by several structural parameters. A bulky residue at position Z10 and a small residue at position Z6 are of utmost importance as exemplified by Mop21 ($T4(B_5)_2C_{11}$). Furthermore, small residues at positions X11 and X12 right below the copper site are favorable. The markedly different contributions to the copper complex stability of the individual helices indicate a major difference in the steric requirements. Apparently the copper site is relocated within the bundle structure. The small residues Ala and Val at position X11 and X12 in helix C_{11} create a gap with respect to the two bulky Leu residues in helix C_3. The copper might shift down to this site which is facilitated by the fact that only three amino acid ligands restrict the mobility of the copper. This hypothesis could explain the unexpected high stability of $T_4(B_i)_2C_{11}$ ($i = 1, 3, 5, 6$) compared to $T_4(B_2)_2C_{11}$. The substitution of the medium-size Val at Z13 by a bulky Leu presumably stabilizes the packing right above the metal site, whereas in Mop5-like copper sites complex formation is impaired by too bulky residues at Z13.

Copper Sites of Type I. According to the matrixes in Figure 3, proteins in the lower part of columns 4 to 8 reveal strong CT transitions in the visible regions indicative for a transformation to (distorted) type I copper sites. It appears to be that one origin is a sterically demanding residue at position C4. However, due to the heterogeneity of the class of proteins exhibiting sites that fall into the type I category and the few examples, no clear-cut relationships can be extracted. Yet unprecedented is the fact that the formation of the copper site in Mop23 $T_4(B_5)_2C_8$, the most stable representative of this class, is particularly influenced by helix A, which in turn plays no significant role for the formation and stability of the two other classes of proteins.

Conclusions

1. Within a combined rational-combinatorial approach, successive design cycles allow creating copper binding sites in four-helix bundle proteins with increasing stability and binding specificity. On the basis of a minimum ligation set, consisting of two His and one Cys residue, structurally different copper binding sites are obtained upon alterations in the environment of the coordination sphere.

2. Employing a variety of complementary spectroscopic techniques, detailed information about the copper site structures is obtained that, upon comparison with empirical data, allows classifying the proteins in terms of the naturally occurring type I, type II, and type Cu_A copper proteins.

3. Substitutions in the secondary coordination sphere that are likely to alter steric interactions were found to be the origin for the conversion of the tetragonal type II copper protein (Cu-Mop5) to a tetrahedral type I center. Introducing Met at various positions that were considered to be appropriate for forming an additional ligand−metal coordination may cause an expansion of the coordination sphere and may lead to the transformation to a dinuclear mixed valent copper center similar to a Cu_A site.

4. The present iterative synthesis strategy is not necessarily restricted to the design of copper proteins but may represent a general approach for creating novel metal binding sites in four-helix bundle proteins. It may even be extended to proteins involving secondary structure elements other than α-helix.

Acknowledgment. Dr. E. Bill's support in measuring and interpreting the EPR spectra is gratefully acknowledged.

JA0484294