DALTON
FULL PAPER

Co-ordination ability towards Cu^{II} of the 29-amino acid residue trypsin inhibitor of squash and two of its analogues †

Piotr Mlynarz,[a] Daniela Valensin,[b] Henryk Kozlowski,*[a] Teresa Kowalik-Jankowska,[a] Jacek Otlewski,[c] Gianni Valensin[b] and Nicola Gaggelli[b]

[a] *Faculty of Chemistry, University of Wroclaw, F.Joliot-Curie 14, 50-383 Wroclaw, Poland. E-mail: henrykoz@wchuwr.chem.uni.wroc.pl*
[b] *Dipartimento di Chimica, Universita di Siena, via Aldo Moro, 53100 Siena, Italy*
[c] *Institute of Biochemistry and Molecular Biology, University of Wroclaw, Tamka 2, 50-137, Poland*

Received 1st November 2000, Accepted 20th December 2000
First published as an Advance Article on the web 13th February 2001

Successful application of the potentiometric method together with NMR, EPR, CD and absorption spectroscopy yielded accurate data concerning the stabilities of the complexes formed and their binding modes between Cu^{II} and squash trypsin inhibitor. The major residue involved in the metal ion co-ordination is the His-25 imidazole side chain, which acts as an anchoring donor and is bound to metal ion over the whole pH range (3–11.5) studied. The 3N complex with $\{N_{imid}, N^-_{His25}, N^-_{Glu24}\}$ binding mode dominates at physiological pH. The data obtained indicate that the protein after a particular mutation could be useful to model metal centres of large proteins.

Introduction

In short peptides a single histidine could play the role of an anchoring site using its imidazole side chain nitrogen as a donor atom. There is a second efficient anchoring site placed at the N-terminal amino group nitrogen.[1–3] In oligopeptides both sites may compete to bind such metal ions as Cu^{II} or Ni^{II}.[3–6] Earlier results have shown that the imidazole nitrogen is the primary anchoring site both for copper(II)[4] and nickel(II)[5] ions. When the His residue is close to the N-terminal amino acid the next nitrogen to bind may be that of the terminal amino group, forming a macrocyclic chelate ring.[3–7]

CMTI-I, Arg-Val-Cys-Pro-Arg-Ile-Leu-Met-Glu-Cys-Lys-Lys-Asp-Ser-Asp-Cys-Leu-Ala-Glu-Cys-Val-Cys-Leu-Glu-His-Gly-Tyr-Cys-Gly, a serine proteinase inhibitor from pumpkin seeds (*Cucurbita maxima* trypsin inhibitor-I), is a small protein containing 29 amino acids, including three disulfide linkages. These disulfide bridges induce high stability and rigidity. CMTI-I contains a single His residue at the carbon terminus at position 25, which is separated very well from the N-terminal segment (Fig. 1). There is only one side chain donor system in the vicinity of His-25, a carboxylate of Glu-24. Two adjacent cysteines, Cys-22 and Cys-28, are involved in a disulfide bridge formation (Cys10–Cys22 and Cys16–Cys28). These bridges cause the peptide backbone in this fragment of protein to be bent and more rigid than that in the linear oligopeptides. Thus, the formation of complexes with the His imidazole and the consecutive amide nitrogens may be distinctly affected by the steric hindrance of the peptide backbone. This could allow the N-terminal sequence, Arg-Val-Cys-Pro-Arg, to be more competitive with the His-25 site in metal ion binding. In oligopeptides the anchoring site at His means that successive complexes are formed between Cu^{II} and the peptide, *i.e.* 1N, 2N, and 3N species with one, two or three nitrogens all containing the imidazole nitrogen as well as adjacent, one or more, amide nitrogens, respectively. At high pH the 4N species with $\{NH_2, 3N^-_{amid}\}$ binding mode is usually formed at the nitrogen terminus.[3–7] However, in CMTI-I the Pro-4 residue acts as the break-point and the 4N complex cannot be formed, as the secondary Pro nitrogen involved in the peptide bond is unable to co-ordinate to metal ions.[2,3,9] This particular situation creates competition between the His imidazole and N-terminal amino group. A study on Cu^{II}–CMTI-I may then add new information on to this particular problem.

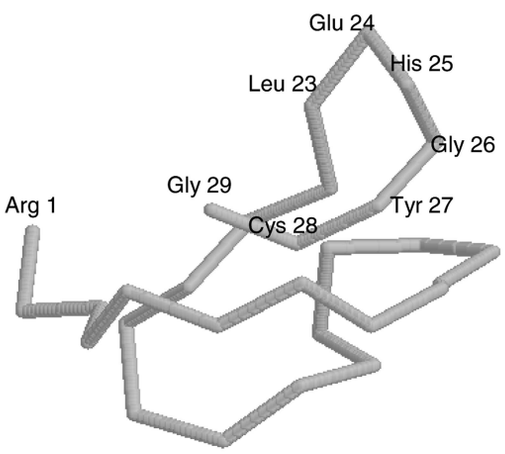

Fig. 1 Structure of the CMTI-I back bone according to ref. 8.

Although a single His is not a very effective binding site to model real metalloproteins, small proteins with well defined structure can be used for engineering of specific metalloprotein sites. Appropriate mutations around the His residue, *e.g.* adding further His or other residues, may allow one to model the particular metal binding environment found in natural metalloproteins.[10]

Thus, studies performed on the Cu^{II}–CMTI-I system will lead, on the one hand, to a better understanding of the competition between the nitrogen terminus and His sites in binding of Cu^{II} by natural peptides and on the other will allow one to evaluate the potential of this small protein as a model for further studies on metal binding sites in metalloproteins. To verify our data we have studied also two analogues CMTI*-I,

† Electronic supplementary information (ESI) available: low-field ^1H NMR spectra, temperature dependence of R_{1p} values and NOESY spectra of the CMTI-I + Cu^{II} system, EPR spectra of the CPTI*-II + Cu^{II} system. See http://www.rsc.org/suppdata/dt/b0/b008790o/

DOI: 10.1039/b008790o

Table 1 Protonation and stability constants for three small proteins CMTI-I, CMTI*-I and CPTI*-II and their copper(II) complexes at 298 K and $I = 0.1$ mol dm^{-3} (KNO_3)

	CMTI-I			CMTI*-I			CPTI*-II		
Species	log β	log K		log β	log K		log β	log K	
HL	10.98 ± 0.02	10.98	Lys	11.06 ± 0.01	11.06	Lys	11.06 ± 0.02	11.06	Lys
H_2L	20.97 ± 0.02	9.99	Tyr	21.21 ± 0.01	10.15	Tyr	21.71 ± 0.02	10.65	Lys
H_3L	28.63 ± 0.02	7.66	$NH_{2(term)}$	29.40 ± 0.01	8.19	$NH_{2(term2)}$	31.79 ± 0.01	10.08	Tyr
H_4L	34.55 ± 0.03	5.92	$N_{imidazole}$	36.45 ± 0.01	7.05	$NH_{2(term1)}$	40.13 ± 0.02	8.34	$NH_{2(term2)}$
H_5L	39.25 ± 0.03	4.70	Glu	41.80 ± 0.01	5.35	$N_{imidazole}$	47.56 ± 0.02	7.43	$NH_{2(term1)}$
H_6L	43.43 ± 0.03	4.18	Glu	46.07 ± 0.01	4.27	Glu	53.51 ± 0.02	5.95	$N_{imidazole}$
H_7L	47.11 ± 0.04	3.68	Glu	50.05 ± 0.02	3.98	Glu	58.38 ± 0.02	4.87	Glu
H_8L	50.05 ± 0.04	2.94	Asp	53.37 ± 0.02	3.32	Glu	62.68 ± 0.03	4.30	Glu
H_9L	52.85 ± 0.04	2.80	Asp	56.35 ± 0.03	2.98	Asp	66.90 ± 0.03	4.22	Glu
$H_{10}L$				59.01 ± 0.02	2.65	Asp	—	—	
$H_{11}L$							74.01 ± 0.03	3.55*2	Asp
CuH_5L									
CuH_4L							45.99 ± 0.02 (2N)		
CuH_3L	32.92 ± 0.03 (1N)			35.03 ± 0.03 (2N)			39.59 ± 0.02 (3N)	6.40	
CuH_2L	27.46 ± 0.02 (2N)	5.46		29.10 ± 0.03 (3N)	5.93		32.04 ± 0.03 (3N)	7.55	
CuHL	20.80 ± 0.02 (3N)	6.66		22.15 ± 0.04 (3N)	6.95		23.54 ± 0.03 (3N)	8.50	
CuL	12.23 ± 0.03 (3N)	8.57		13.87 ± 0.05 (3N)	8.28		13.83 ± 0.04 (4N)	9.71	
$CuH_{-1}L$	2.53 ± 0.03 (4N)	9.70		4.25 ± 0.06 (4N)	9.62		3.33 ± 0.07 (4N)	10.50	
$CuH_{-2}L$	−8.19 ± 0.04 (4N)	10.72		−6.26 ± 0.08 (4N)	10.51		−7.41 ± 0.09 (4N)	10.74	
$CuH_{-3}L$	−19.50 ± 0.04 (4N)	11.31		−17.19 ± 0.09 (4N)	10.93		−18.43 ± 0.06 (4N)	11.02	

i.e. CMTI-I with a hydrolysed peptide bond between Arg-5 and Ile-6, and CPTI*-II, which compared to CMTI*-I contains two mutations Arg-5 ⟶ Lys and Val-21 ⟶ Ile. This work was also performed to evaluate the usefulness of the potentiometric method for studying relatively large metal–peptide systems in order to obtain good quality stability constants as a precise measure of the metal–ligand interactions.

Experimental

Potentiometric studies

Stability constants for proton and copper(II) complexes were calculated from titrations carried out at 25 °C using total volumes of 1.2–1.35 cm^3. Alkali was added from a 0.250 cm^3 micrometer syringe calibrated by both weight titration and the titration of standard materials. The ligand concentration was (0.5–1.0) × 10^{-3} mol dm^{-3} and the metal-to-ligand ratios were 1 : 1.2 and 1 : 3. The pH-metric titrations were performed at 25 °C in 0.1 mol dm^{-3} KNO_3 on a MOLSPIN pH-meter system using a Russel CMAW 711 semi-micro combined electrode calibrated in hydrogen ion concentrations using HNO_3.[11] Duplicate titrations were performed for each molar ratio in the pH range 3–10.7, and the SUPERQUAD computer program was used for stability constant calculations.[12] Standard deviations quoted were computed by SUPERQUAD, and refer to random errors only. They are, however, a good indication of the importance of a particular species in the equilibrium.

Spectroscopic studies

EPR spectra were recorded on a Bruker ESP 300E spectrometer at X-band frequency (9.3 GHz) at 120 K, in a quartz tube about 2 mm diameter, using 33% ethylene glycol, 0.2 × 10^{-3} mol dm^{-3} Cu^{II} and a metal to ligand ratio of 1 : 1.2 in the pH range 3–11.5. Absorption spectra were recorded on a Beckman DU 650 spectrophotometer, circular dichroism (CD) spectra on a JASCO J 715 spectropolarimeter in the 800–240 nm range. The metal concentration in CD and UV-VIS spectroscopic measurements was adjusted in the range (1–0.24) × 10^{-3} mol dm^{-3} and metal to ligand ratios were 1 : 1 and 1 : 1.2. The spectroscopic parameters were obtained at the maximum concentration of the particular species from the potentiometric calculations.

NMR experiments were carried out on a Bruker Avance DRX 600 spectrometer at controlled temperature (±0.2 K). Solutions were made up in deionized water containing 8% deuterium oxide and carefully deoxygenated by freezing–thawing cycles. Chemical shifts were referenced to internal 3-trimethylsilyl[$^2H_4$]propanesulfonate (TSP-d_4). The pH was adjusted to the desired value with either DCl or NaOD. A stock solution of copper nitrate in H_2O–D_2O was used to obtain the desired copper concentration in the CMTI-I solution.

Spin–lattice relaxation rates were measured with standard inversion recovery pulse sequences and calculated by a exponential regression analysis of the recovery curves of longitudinal magnetization components. TOCSY experiments were run with standard pulse sequences by using a total spin-locking time of 75 ms given by an MLEV-17 mixing sequence. NOESY spectra were obtained with standard pulse sequences using 100 ms as optimized mixing time. Suppression of the water signal was obtained by using the WATERGATE pulse sequence.

Inhibitor preparation

Squash inhibitors CMTI-I and CPTI*-II were isolated from figleaf gourd and summer squash seeds, respectively, and purified as described earlier.[13] CMTI*-I was prepared by 48 h incubation of the intact form CMTI-I with trypsin at pH 2.8, followed by ion exchange chromatography and desalting.[14]

Results

Protonation constants

CMTI-I has potentially thirteen protonation sites including those at the N-terminal amino group, six carboxylates (carbon terminal, three glutamates and two aspartates), histidine imidazole, tyrosine phenolate, two lysine side chain amino groups and two guanidines of arginines. Calculations based on the potentiometric titrations gave nine protonation constants, which can be assigned as shown in Table 1. The large number of protonation constants means that the extreme values, very high and very low, are more difficult to evaluate precisely from the data obtained in the pH range studied; *e.g.* the pK values for the second Lys and two Arg should distinctly exceed 11 and their evaluation cannot be convincingly precise. The values of the

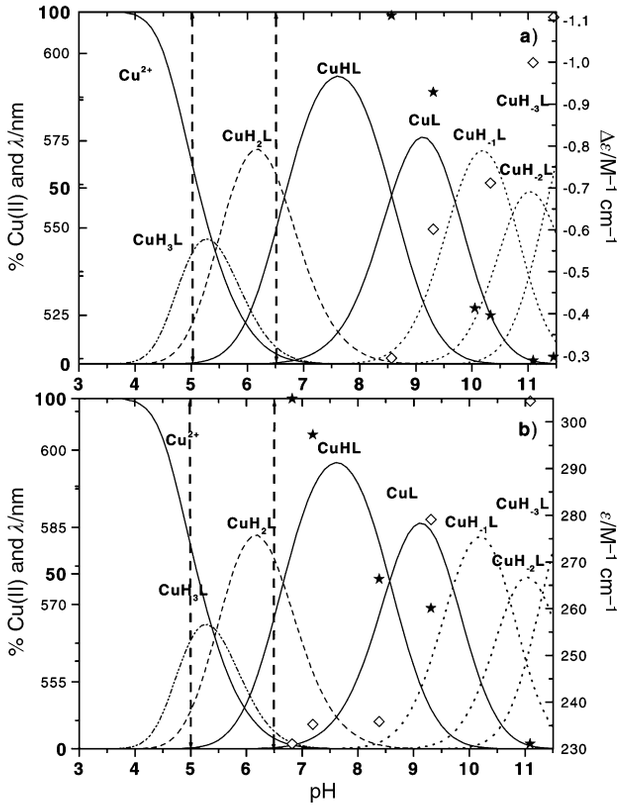

Fig. 2 Species distribution curves for Cu^{II}–CMTI-I, metal to ligand molar ratio 1 : 1.2, $[Cu^{II}] = 5 \times 10^{-4}$ mol dm^{-3}. ◇, $\Delta\varepsilon$ (a) or ε (b) values of the d–d transition taken from CD or absorption spectra; *, the respective λ_{nm}. The dashed dotted line corresponds to 1N, dashed line to 2N, solid line to 3N and dotted line to 4N species. The areas of precipitation of the complexes are indicated.

respective pK and their assignment are rather characteristic and do not differ from those presented in the literature.[15] The analogue CMTI*-I with a hydrolysed peptide bond between Arg-5 and Ile-6 exhibits clearly an additional protonation constant for the NH_2 group of the Ile residue (Table 1). Insertion of the Lys residue instead of Arg-5 for the hydrolysed CPTI*-II between the same residues as for CMTI*-I results in two additional protonation constants including those at pK = 10.65 and 8.34 corresponding to protonations of the amino groups of the Lys-5 and Ile-6, respectively (Table 1).

Cu^{II}–CMTI-I complexes

Potentiometric and spectroscopic data. Solutions containing Cu^{II} and inhibitor show foaming at pH 5–6.5 indicating the possibility of precipitation in the titrated samples, due most likely to formation of minor species with metal ion bridging two protein molecules. Copper(II) ion bound to imidazole of exposed His-25 residue has the potential ability to bridge two molecules of CMTI-I leading to precipitation of $Cu(CMTI\text{-}I)_2$ species. Outside this pH region titration could be done precisely and the calculations based on the potentiometric data indicate formation of the species collected in Table 1. Addition of SDS to the Cu^{II}–CMTI-I solution improves slightly the solubility problem at pH 5–6.5 and the spectroscopic data obtained for SDS (sodium dodecyl sulfate) samples are similar to those obtained for pure aqueous solution. In the case of the two analogues, CMTI*-I and CPTI*-II, solutions were clear over the whole pH range studied.

Co-ordination of Cu^{II} to CMTI-I. Computer analysis of the potentiometric data revealed the presence of the following species in the Cu^{2+}–CMTI system: CuH_3L, CuH_2L, CuHL, CuL, $CuH_{-1}L$, $CuH_{-2}L$ and $CuH_{-3}L$ (Table 1, Fig. 2). Accord-

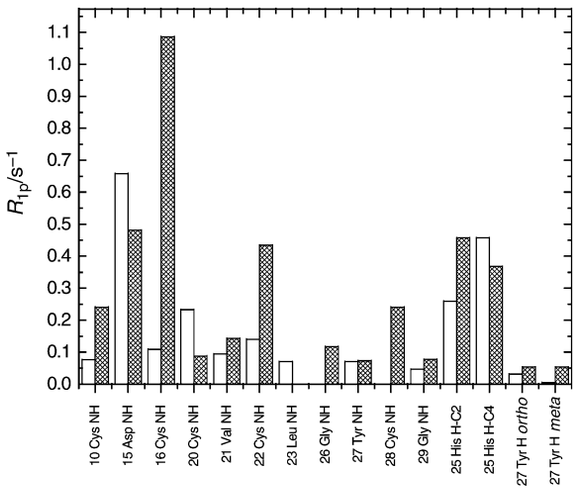

Fig. 3 Histogram of R_{1p} values measured for selected protons of 2.3×10^{-3} mol dm^{-3} CMTI-I in 92% H_2O–8% D_2O at pH 4.8 (blank column) and 10.3 (pattern column) in the presence of Cu^{II} (4.26×10^{-6} mol dm^{-3}) at 303 K.

ing to the EPR parameters $A_{\parallel} = 159$ G, $g_{\parallel} = 2.296$, (Table 2) the complex formed below pH 6 (CuH_2L species) is the 2N complex with likely involvement of imidazole and amide nitrogens in co-ordination.[2] The NMR relaxation measurements (see below) clearly support this indication. At pH 6.5–8 the energy of the d–d transition moves to 610 nm indicating binding to a third nitrogen resulting in the $\{N_{imid}, 2N^-\}$ co-ordination mode.[2] Increase of pH above 11 moves a d–d band to 543 nm. This again indicates successive nitrogen binding and formation of the 4N species.[2] The third amide nitrogen, in the 4N species, should derive, however, from Leu-23 being next to Cys-22 involved in the disulfide bridge formation. The formation of the Cys10–Cys22 disulfide unit induces a very rigid bent structure and, as a result, a distinct hindrance to the residues placed around Cys-22. Thus, the involvement of the Leu-23 amide nitrogen in the copper(II) ion co-ordination could be less favourable. A very similar set of complexes was obtained for the two analogues CMTI*-I and CPTI*-II (Table 1). The calculations of the stability constants for the latter systems were easier as no precipitation was observed during pH variation.

Owing to the high complexity of the systems studied, very careful correlation of the species distribution with spectroscopic data was performed (Fig. 2a,b). The pH dependence of the d–d transition energy, $\Delta\varepsilon$ and λ values fits very well the species distribution pattern indicating that the chemical model obtained from the potentiometric data calculations is very likely to be correct.

NMR spectra of CMTI-I in water–deuterium oxide were recorded at pH 4.8 and 10.3. Assignments were performed by use of TOCSY spectra (ESI supplementary data Table S1). The measured chemical shifts were in substantial agreement with previous data at pH 4.7.[8,16] It is noticeable that several signals were missing in the amide region of NMR spectra obtained at pH 10.3. This is most likely determined by enhanced exchange rates among conformational isomers. Upon addition of copper (ESI supplementary data Fig. S1) NMR spectra underwent extensive line broadening and spin–lattice relaxation rates were selectively affected as shown in Fig. 3. The reported R_{1p} is defined in eqn. (1) where $1/T_1^f$ and $1/T_1^b$ are the spin–lattice

$$R_{1p} = \frac{1}{T_1^{obs}} - \frac{1}{T_1^f} \approx \frac{p_b}{T_1^b + \tau_b} \quad (1)$$

relaxation rates of protons in the free and metal-bound states, $1/T_1^{obs}$ is the rate measured in the presence of chemical exchange between the two environments, p_b the fraction of metal-bound ligand and τ_b^{-1} the off-rate constant. Delineation

Table 2 Spectroscopic data for complexes of Cu^{2+} ions with the three small proteins CMTI-I, CMTI*-II and CPTI*-II (values for 3N complexes are in bold)

	CMTI-I					CMTI*-I					CPTI*-II				
	CD	EPR		UV-Vis		CD	EPR		UV-Vis		CD	EPR		UV-Vis	
Species	λ/nm ($\Delta\varepsilon$/M^{-1} cm^{-1})	$A_{\parallel}$/G	$g_{\parallel}$	λ/nm	ε/M^{-1} cm^{-1}	λ/nm ($\Delta\varepsilon$/M^{-1} cm^{-1})	$A_{\parallel}$/G	$g_{\parallel}$	λ/nm	ε/M^{-1} cm^{-1}	λ/nm ($\Delta\varepsilon$/M^{-1} cm^{-1})	$A_{\parallel}$/G	$g_{\parallel}$	λ/nm	ε/M^{-1} cm^{-1}
CuH_4L												151	2.325	657[a] 324[b]	96 475
CuH_3L							158	2.317	643[a] 321[b]	84 520	**613 (−0.325)**[a] **323 (+1.953)**[c] **278 (−3.834)**[d]	**172**	**2.221**	**604**[a] **323**[b]	**103** **732**
CuH_2L		159	2.296			**622 (−0.441)**[a] **322 (+1.939)**[c] **278 (−4.741)**[d]	**167**	**2.224**	**619**[a] **313**[b]	**116** **755**	**603 (−0.405)**[a] **324 (+1.967)**[c] **277 (−4.342)**[d]	**177**	**2.215**	**598**[a] **322**[b]	**112** **781**
CuHL	**570 (+0.090)**[a] **683 (−0.397)**[a] **331 (+0.743)**[c] **273 (−7.803)**[d]	**170**	**2.225**	**610**[a]	**231**	**609 (−0.687)**[a] **323 (+3.271)**[c] **278 (−6.147)**[d]	**172**	**2.220**	**611**[a] **313**[b]	**117** **824**	**607 (−0.407)**[a] **326 (+1.972)**[c] **277 (−4.829)**[d]	**183**	**2.208**	**583**[a] **322**[b]	**112** **852**
CuL	**611 (−0.294)**[a] **331 (+1.082)**[c] **277 (−7.659)**[d]	**176**	**2.219**	**575**[a]	**235**	**594 (−0.702)**[a] **328 (+3.252)**[c] **279 (−6.916)**[d]	**179**	**2.217**	**593**[a] **318**[b]	**117** **913**	508 (−0.918)[a] 327 (+1.521)[c] 278 (−4.329)[d]	193	2.197	523[a] 319[b]	175 1386
$CuH_{-1}L$	525 (−0.714)[a] 334 (+0.737)[c] 277 (−8.84)[d] 247 (+2.867)[e]					521 (−0.101)[a] 326 (+3.184)[c] 280 (−6.691)[d] 249 (+0.289)[e]	193	2.201	523[a] 324[b]	156 1190					
$CuH_{-2}L$	513 (−1.119)[a] 328 (+0.892)[c] 279 (−6.641)[d] 246 (+10.217)[e]	196	2.196	543[a]	304										
$CuH_{-3}L$						519 (−1,462)[a] 325 (+2.489)[c] 280 (−6.221)[d] 248 (+7.643)[e]	199	2.180	514[a] 322[b]	182 1390	512 (−1.453)[a] 327 (+0.968)[c] 279 (−3.493)[d]	204	2.180	510[a] 324[b]	224 1523

[a] d–d transition. [b] $N \rightarrow Cu^{2+}$ charge transfer transition. [c] $N_{imidazole} \rightarrow Cu^{2+} + N^{-} \rightarrow Cu^{2+}$ charge transfer transition. [d] Intraligand transition + $N_{imidazole} \rightarrow Cu^{2+}$. [e] Intraligand transition.

of copper binding sites is only possible when the R_{1p} is mainly contributed by $1/T_1^b$, *i.e.* in the conditions of fast exchange. The fast exchange was found by measuring the temperature dependence of paramagnetic relaxation rates (ESI supplementary data Fig. S2) at both pH values. The relaxation rate in the metal-bound state is accounted for by the simplified Solomon equation [17–19] (2) where γ_H and γ_e are the proton and

$$\frac{1}{T_1^b} = \frac{1}{10}\frac{\gamma_H^2\gamma_e^2\hbar^2}{r^6}\left\{\frac{3\tau_R}{1+\omega_H^2\tau_R^2}\right\} \qquad (2)$$

electron magnetogyric ratios, r is the metal–proton distance, ω_H the proton Larmor frequency and τ_R the rotational correlation time of the metal complex. Structural delineation of the CMTI-I–copper adduct is however not an easy task because, upon binding the metal, (i) the amide nitrogen undergoes loss of the bound hydrogen which no longer contributes to the nuclear relaxation rate and (ii) the crowded spectrum prevents determination of H_α spin–lattice relaxation rates. The observed effects suggest binding of the imidazole ring at both pH values, the measured R_{1p} being consistent with copper distances from ^{25}His 2C-H in the range 0.17–0.20 nm and from ^{25}His 4C-H in the range 0.18–0.20 nm. The distance limits were calculated by considering $p_b = [Cu^{2+}]/[CMTI\text{-}I]$ and a motional correlation time in the range 0.1–1.0 ns at 303 K. Quite interestingly large paramagnetic effects were measured for NMR signals far from the metal binding site, namely on ^{15}Asp NH, ^{16}Cys NH and ^{8}Met NH, the last two being sizeable at pH 10.3 only. The R_{1p} of ^{8}Met NH was not included in Fig. 3 because it was much larger than the other values and also because the extensive line broadening yielded a large experimental error in the measurement. The occurrence of dipole–dipole interactions between Cu^{II} and such relatively distant protons is easily interpreted in the light of the observed folding of the peptide metal binding site towards the N-terminal residues. NOESY spectra showed strong connectivities between protons belonging to different moieties (ESI supplementary data Fig. S3), in the absence as well as in the presence of copper(II) ions, including among others: ^{22}Cys NH–^{13}Asp H_α, ^{23}Leu NH–^{28}Cys H_α, ^{10}Cys NH–^{27}Tyr H_α and aromatic protons of ^{27}Tyr with ^{9}Glu H_α, ^{7}Leu Hγ and ^{7}Leu H_δ. The interaction between protein fragments is also seen in the crystal structure of CMTI-I with bovine β-trypsin.[20]

NOESY maps were also taken for further characterization of the metal complexes in solution by focussing on the NH–H_α connectivities as shown in Fig. 4a. It is well known that the intensity of a cross peak depends upon the efficiency of the exchange process, namely the dipolar cross relaxation rate and the transverse relaxation rate of the involved NMR resonances.[21] As a consequence, the addition of paramagnetic metal ions is expected selectively to wash out the cross-peaks connecting protons experiencing dipolar or scalar local fields induced by the electron magnetic moment.[22] The occurence of chemical exchange of peptide molecules between the free and copper-bound environments affects the observed NOESY maps provided the rate of exchange is faster than the proton relaxation rates. The selective attenuation of cross-peaks (Fig. 4b) may then demonstrate the vicinity of the involved H_α protons to the paramagnetic center. From this point of view it is worth noticing that the most affected cross-peaks, besides those involving ^{25}His H_α, ^{24}Glu H_α, ^{23}Leu H_α, were those belonging to the indicated proton pairs located towards the N-terminal region.

According to the potentiometric data (Fig. 2a,b) there are two species present at pH around 5 (CuH_3L and CuH_2L). It should be mentioned here that potentiometric data were obtained for stoichiometric molar ratios, while NMR relaxation measurements were performed for a high excess of protein. Although NMR data may not be directly transferred into the stoichiometric conditions they strongly indicate the

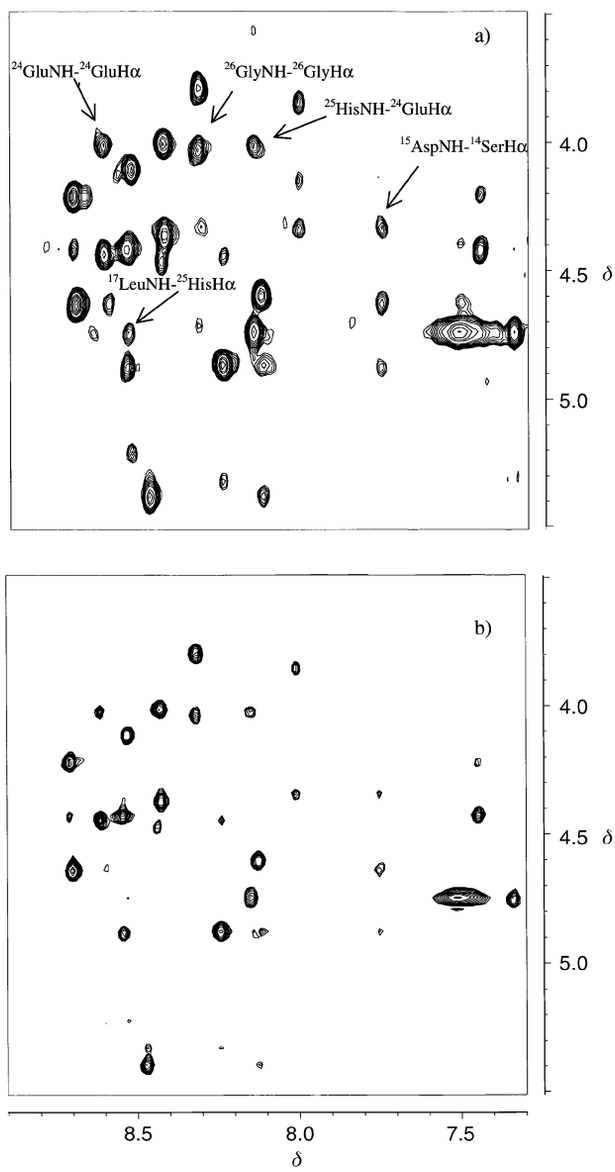

Fig. 4 Effect of Cu^{II} on NH–H_α connectivities in the NOESY spectrum of 23 mmol dm^{-3} CMTI-I in 92% H_2O–8% D_2O at pH 4.8, 303 K; ligand "free" (a) and (b) in the presence of Cu^{II} (5.75×10^{-6} mol dm^3).

His-25 side-chain is an anchoring donor set for Cu^{II}. Thus, the NMR data support the suggestion given above that CuH_3L (1N) and CuH_2L (2N) complexes are formed at the His-25 site involving $N_{imidazole}$ (1N) or $\{N_{imid}, N^-\}$ (2N) donor sets. This co-ordination mode has also some minor impact on the amide protons around the binding site, *e.g.* Gly-26, Glu-24 and Leu-23.

For the CuHL, 3N, species the presence (in CD spectra) of the charge transfer transition of $N_{Im} \rightarrow Cu^{2+}$ at 331 nm,[22] the EPR parameters $A_\parallel = 170$ G, $g_\parallel = 2.225$ and the d–d transition energy at 610 nm suggest $\{N_{imid}, 2N^-\}$ co-ordination[2] with involvement of His-25 and Glu-24 deprotonated amide nitrogens (Table 2).

The further increase of pH results in deprotonation of the CuHL complex to CuL species. This does not lead to any distinct variation in the spectral parameters. The pK value of 8.57 of this reaction may indicate deprotonation of the N-terminal NH_3^+ group (Table 1). Above pH 8.5 the d–d transition is shifted to much higher energy with a band centred at 543 nm (Table 2). Both the CD and EPR parameters indicate formation of the 4N complex with four nitrogens involved in the metal ion co-ordination. A further two deprotonations, $CuH_{-1}L \longrightarrow CuH_{-2}L \longrightarrow CuH_{-3}L$, correspond to deproton-

ations of Tyr-27 and one of the Lys side chains. The $CuH_{-1}L$ and $CuH_{-2}L$ species have very similar CD parameters supporting no variation in the co-ordination mode during these two deprotonations. The formation of the 4N complex in the Cu^{II}–CMTI-I system (data obtained for pH 10.3) results in an increase in relaxation rates of the amide protons of Tyr-27, Cys-28 and Gly-29. A very high increase of the relaxation rates is observed also for Met-8, Asp-15 and Cys-16 amide protons. The latter residue forms a disulfide bridge with Cys-28. The other two considerably affected amide protons are those of Cys-22. These variations of the relaxation rates may indicate distinct variations in the protein conformation upon metal ion binding at the His-25 imidazole ring nitrogen when 4N species is formed. Also the high p*K* value of the reaction CuL(3N) ⟶ $CuH_{-1}L$(4N) of 9.70 (Table 1) when compared to those for other similar peptides (≈8.47) seems to support a major structural variation when the 3N species is transformed into the 4N complex.[23] The assumption of the formation of the 4N complex with involvement of the amide nitrogens on the C-terminal side of His-25 (Gly-26, Tyr-27 and Cys-28) could explain *e.g.* the extremely large change of the Met amide proton relaxation rate in the presence of copper(II) ions. According to the earlier crystal structure of CMTI-I with bovine β-trypsin,[20] the Met-8 carbonyl oxygen is hydrogen bonded to the Cys-28 amide nitrogen. Cys-28 is also connected with Cys-16 *via* the disulfide bridge. This could explain influence of Cu^{II} on the Cys-16 and Asp-15 amide protons. However, the $\{N_{im}, 3N^-\}$ binding to a His-Xaa-Yaa-Zaa sequence is rather unlikely[24] and involvement of the fourth amide nitrogen from Leu-23 has to be assumed. The involvement of Leu-23 in the metal ion binding in the 4N species is reflected in the ^{23}Leu H_α proton, whose relaxation rate increases distinctly when Cu^{II} is added only at pH 10.3. However, the overlapping of the signals of this proton with those of ^{21}Val H_α and ^{17}Leu H_α does not allow one to quantify this change in relaxation rate.

Copper(II) binding to CMTI*-I. Cleavage of the peptide bond between Arg-5 and Ile-6 in CMTI-I results in formation of a protein still having 29 amino acids due to the presence of the Cys3–Cys20 disulfide bridge. There are two N-terminal amino groups at Arg-1 and Ile-6 with p*K* = 7.05 and 8.19, respectively (Table 1).

From the calculations of the potentiometric titration data the following species were found: CuH_3L, CuH_2L, CuHL, CuL, $CuH_{-1}L$, $CuH_{-2}L$ and $CuH_{-3}L$. The CuH_3L species may be regarded as comparable to the CuH_2L complex of CMTI-I with the additional protonated amino group of Ile-6. The complex formation pattern is similar to that of CMTI-I although the first observed complex, CuH_3L, is the 2N species rather than the 1N complex (Table 1, Fig. 5a,b). The absorption band of the d–d transition at 643 nm and the EPR parameters $A_\parallel = 158$ G, $g_\parallel = 2.317$ strongly support such a binding mode (Table 2).[2] The deprotonation reaction $CuH_3L \longrightarrow CuH_2L$ is accompanied by a shift of the d-d band to ≈600 nm and decrease of $g_\parallel$ to 2.224 in the EPR spectra suggesting binding of the third nitrogen. Thus, the CuH_2L species is a 3N complex with a $\{N_{imid}, 2N^-\}$ co-ordination mode.[2] The p*K* value of the second amide deprotonation of 5.93 is slightly higher than the respective value for CMTI-I (Table 1). The two other deprotonations leading to CuHL and CuL species with p*K* of 6.95 and 8.28 can be assigned to proton dissociations at the amino groups of Arg-1 and Ile-6, respectively. The CD spectra of all three 3N species, CuH_2L, CuHL and CuL, are very close to each other supporting the latter suggestion.

The three other complexes, $CuH_{-1}L$, $CuH_{-2}L$ and $CuH_{-3}L$, with d–d transitions around 514–523 nm are clearly 4N complexes. The deprotonation reactions $CuH_{-1}L \longrightarrow CuH_{-2}L + H^+$ and $CuH_{-2}L \longrightarrow CuH_{-3}L + H^+$ with p*K* 10.51 and 10.93 can be assigned to proton dissociation from Tyr-27 and one of the Lys side chain amino groups.

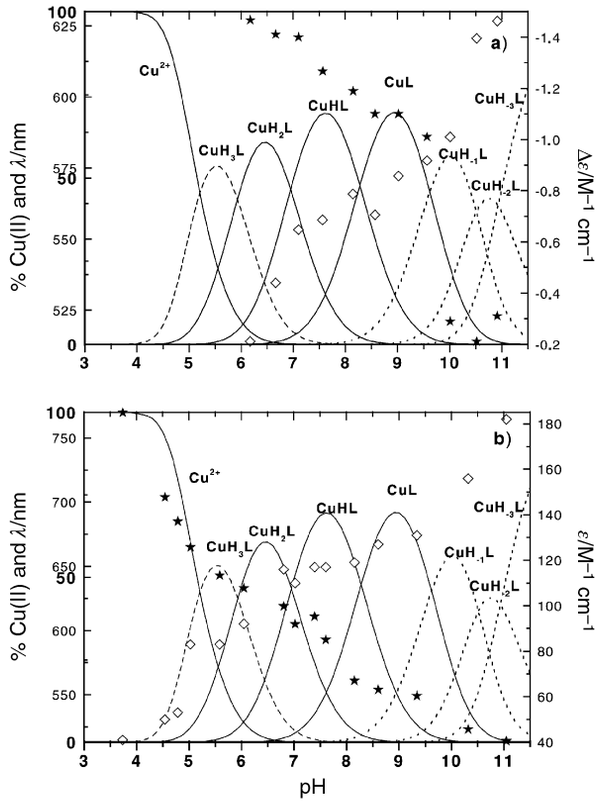

Fig. 5 Species distribution curves for Cu^{II}–CMTI*-I, metal to ligand molar ratio 1 : 1.3, $[Cu^{II}] = 5.5 \times 10^{-4}$ mol dm^{-3}. ◇, Δε (a) or ε (b) values of the d–d transition taken from CD or absorption spectra; *, respective λ_{nm}. Dashed line corresponds to 2N, solid line to 3N and dotted line to 4N species.

Copper(II) binding to CPTI*-II. Both potentiometric and spectroscopic data (Tables 1, 2) indicate that co-ordination of Cu^{II} to this protein is very similar to that discussed above for CMTI-I and CMTI*-I. It starts at pH around 4.5 with formation of the CuH_4L species, which according to the EPR and UV-Vis data is a 2N complex (Table 2). Successive deprotonation results in formation of the CuH_3L complex with 3N co-ordination. The CuH_2L and CuHL complexes are formed after deprotonation of N-terminal NH_3^+ and lateral Ile-6 amino groups (Table 1, Fig. 6a,b). The protonation constants of the CuH_2L and CuHL complexes are 7.55 and 8.50, respectively, and they are comparable to those of the "free" ligand (Table 1). The spectroscopic data with d–d bands at 604–583 nm and EPR parameters $A_\parallel = 172$–183 G and $g_\parallel = 2.221$–2.208 are characteristic for $\{N_{imid}, 2N^-\}$ co-ordination.[2]

Above pH 9 the formation of the CuL complex is accompanied by a strong shift of the d–d band towards higher energy (510–523 nm). The EPR parameters $A_\parallel = 204$ G, $g_\parallel = 2.180$ and the presence of nine superhyperfine lines correspond very well to the co-ordination of four nitrogens to Cu^{II} (Table 2, ESI supplementary data Fig. S4). These variations of the spectroscopic data clearly indicate the formation of the 4N complex as was found for the proteins discussed above.

Discussion

The potentiometric as well as the spectroscopic data obtained for Cu^{II} with three small proteins show very similar patterns of metal ion binding. The measurements of the relaxation rates performed for the Cu^{II}–CMTI-I system show the His-25 imidazole nitrogen is the only anchoring site in the whole pH range studied. All three proteins form very easily the 3N species with the $\{N_{imid}, 2N^-\}$ binding mode, involving imidazole and two successively deprotonated amide groups of His-25 and Glu-24. This is a typical co-ordination mode found for

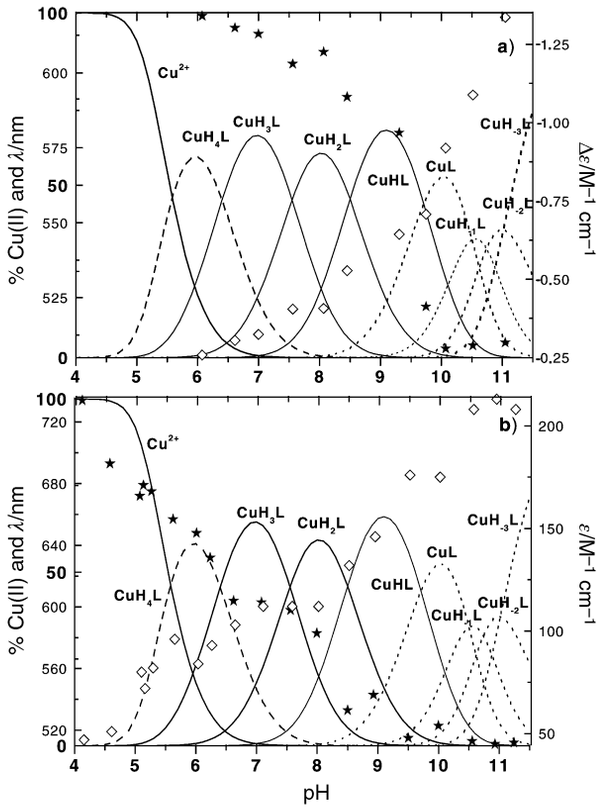

Fig. 6 Species distribution curves for Cu^{II}–CPTI*-II, metal to ligand molar ratio 1 : 1.3, $[Cu^{II}] = 4.6 \times 10^{-4}$ mol dm^{-3}. Details as in Fig. 5.

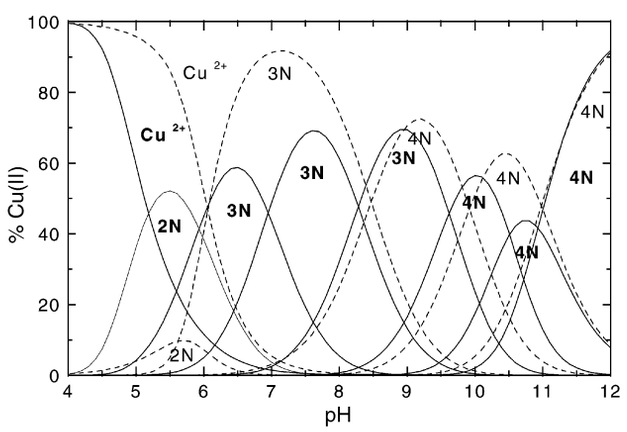

Fig. 7 Comparison of the distribution diagrams for Cu^{II}–CMTI*-I (solid lines) and Cu^{II}–Ac-Ala-Lys-Arg-His-Arg-Lys-Am (dashed line, Am corresponds to the blocked amide group) (ref. 24). $[Cu^{II}] = 1 \times 10^{-3}$ mol dm^{-3}, metal to ligand ratio 1 : 1.

oligopeptides having a His residue inserted into the peptide sequence with a blocked N-terminal amino group.[24] The presence of an amino group at the nitrogen terminus in the oligopeptide allows the formation of large chelates involving both imidazole and N-terminal amino nitrogens.[3–7] Such macrocyclic chelates are thought to induce the formation of the 4N complex at the nitrogen terminus without involvement of imidazole nitrogen, when His is inserted far away from the terminus.

The binding of Cu^{II} by the three inhibitors studied here is more effective than by the oligopeptide protected at the nitrogen terminus (Fig. 7, CMTI*-I was chosen as no precipitation was observed during the experiment). However, the formation of the 3N species with the protein occurs at the same pH as that for a short peptide. It is clear that $\{N_{imid}, 2N^-\}$ binding is very efficient and exists in the whole biologically relevant pH range. It is interesting, however, that the formation of the 4N complex is much more difficult for the proteins studied here than, *e.g.*, in the case of a hexapeptide. This may strongly indicate that the formation of the 4N species of CMTI-I and its analogues requires a major reorganization of the donor set or structure around the metal ion binding site. As is shown above, the His-25 imidazole is involved in the metal ion binding also in the 4N species. This excludes metal ion binding to amide nitrogens directed towards the C-terminal part of the ligand. As the disulfide bridge between Cys-22 and Cys-10 introduces a very rigid structure of the peptide backbone around Cys-22, the binding to the amide nitrogen of Leu-23 is less favoured and thus the formation of the 4N complex is more difficult than in the case of linear peptides by one log unit.

The distinct changes in the relaxation rates of the Cys-28, Cys-16 (linked to Cys-28 by disulfide) and Met-8 (hydrogen bonded to Cys-28) residues strongly suggest that copper(II) co-ordination to His-25 and three adjacent amide nitrogens of His-25, Glu-24 and Leu-23 does not change dramatically the interactions between the particular amino acid residues observed in the metal-free protein. The changes of the relaxation rates at Cys-28 at pH above 10 support the involvement of the Leu-23 amide, which according to NOESY spectra is linked to ^{28}Cys H_α (see above).

An interesting finding is that showing no competitiveness of the nitrogen terminus to co-ordinate Cu^{II}, even when the formation of the 4N species at His-25 is not favourable. The major reason for Cu^{II} remaining at His-25 in the 4N species could be the distinctly higher ability to form 1N–3N complexes by the inhibitor and its analogues when compared to that of linear oligopeptides. It should be mentioned that the large distance between the N-terminal amino group and His-25 (about 18 Å) (Fig. 1, ref. 8) does not allow a macrochelate structure. Thus, competition between the N-terminal amino group and distant His residue to bind metal ions is of meaning only for short, structurally not organized oligopeptides when macrocyclic co-ordination may occur.

The effective binding of metal ion to His-25 indicates that modification of the amino acid residues around this site, *e.g.* their substitution by additional His, may allow the creation of effective and specific sites to model large protein–metal binding sites, both catalytic and structural ones. It is also interesting that the cases discussed above are the largest peptides to which precise potentiometric measurements have been applied. This allowed comparison of the binding of copper(II) ions by small proteins having well organized structure with that by simple linear oligopeptides.

Acknowledgements

This work was supported by the Polish State Committee for Scientific Research (KBN 3T09A10514) and COST D8 0018/97.

References

1 H. Sigel and R. B. Martin, *Chem. Rev.*, 1982, **82**, 385.
2 L. D. Pettit, J. E. Gregor and H. Kozlowski, in *Perspectives on Bioinorganic Chemistry*, eds. R. W. Hay, J. R. Dilworth and K. B. Nolan, JAI Press, London, 1991, vol. 1, p. 1041.
3 H. Kozlowski, W. Bal, M. Dyba and T. Kowalik-Jankowska, *Coord. Chem. Rev.*, 1998, **184**, 319.
4 L. D. Pettit, S. Pyburn, W. Bal, H. Kozlowski and M. Bataille, *J. Chem. Soc., Dalton Trans.*, 1990, 3565.
5 W. Bal, H. Kozlowski, R. Robbins and L. D. Pettit, *Inorg. Chim. Acta*, 1995, **231**, 7.
6 B. Decock Le Reverend, F. Liman, C. Livera, L. D. Pettit, S. Pyburn and H. Kozlowski, *J. Chem. Soc., Dalton Trans.*, 1988, 887.
7 K. Varnagy, J. Szabo, I. Sovago, G. Malandrinos, N. Hadjiliadis, D. Sanna and G. Micera, *J. Chem. Soc., Dalton Trans.*, 2000, 467.
8 J. Otlewski, T. A. Holak, D. Gondol and T. Wilusz, *J. Mol. Biol.*, 1989, **210**, 635.
9 L. Chruscinski, M. Dyba, M. Jezowska-Bojczuk, H. Kozlowski,

G. Kupryszewski, Z. Mackiewicz and A. Majewska, *J. Inorg. Biochem.*, 1996, **63**, 49 and refs. therein.
10 C. Vita, Ch. Roumestand, F. Toma and A. Menez, *Proc. Natl. Acad. Sci. U.S.A.*, 1995, **92**, 6404.
11 H. M. Irving, M. H. Miles and L. D. Pettit, *Anal. Chim. Acta*, 1967, **38**, 475.
12 P. Gans, A. Sabatini and A. Vacca, *J. Chem. Soc.*, *Dalton Trans.*, 1985, 1995.
13 J. Otlewski, A. Polanowski, J. Leluk and T. Wilusz, *Acta Biochim. Pol.*, 1984, **31**, 267.
14 J. Otlewski and T. Zbyryt, *Biochemistry*, 1984, **33**, 200.
15 IUPAC Stability Constants Database, Academic Software, Release 4, 1999. E-mail: scdbase@acadsoft.co.uk
16 R. Krishnamoorthi, C.-L. Sun Lin and D. VanderVelde, *Biochemistry*, 1992, **31**, 4965.
17 I. Solomon, *Phys. Rev.*, 1955, **99**, 559.
18 G. Valensin, R. Basosi, W. E. Antholine and E. Gaggelli, *J. Inorg. Biochem.*, 1985, **23**, 125.
19 I. Bertini and C. Luchinat, *Coord. Chem. Rev.*, 1996, **150**, 1.
20 W. Bode, H. J. Greyling, R. Huber, J. Otlewski and T. Wilusz, *FEBS Lett.*, 1989, **242**, 285.
21 J. Jeener, B. H. Meier, P. Bachmann and R. R. Ernst, *J. Chem. Phys.*, 1979, **71**, 4546.
22 E. Gaggelli, A. Maccotta and G. Valensin, *Inorg. Chem.*, 1993, **32**, 2788.
23 T. Kowalik-Jankowska, M. Jesionowski and L. Lankiewicz, *J. Inorg. Biochem.*,1999, **76**, 63.
24 M. A. Zoroddu, T. Kowalik-Jankowska, H. Kozlowski, H. Molinari, K. Salnikow, L. Broday and M. Costa, *Biochim. Biophys. Acta*, 2000, **1475**, 163.