

Molecular Design of Specific Metal-Binding Peptide Sequences from Protein Fragments: Theory and Experiment

Milan Kožíšek,^[a, b] Aleš Svatoš,^[c] Miloš Buděšínský,^[a] Alexander Muck,^[c] Mikael C. Bauer,^[d] Pavel Kotrba,^[e] Tomáš Ruml,^[e] Zdeněk Havlas,^[a] Sara Linse,^[d] and Lubomír Rulíšek^{*[a]}

Abstract: A novel strategy is presented for designing peptides with specific metal-ion chelation sites, based on linking computationally predicted ion-specific combinations of amino acid side chains coordinated at the vertices of the desired coordination polyhedron into a single polypeptide chain. With this aim, a series of computer programs have been written that 1) creates a structural combinatorial library containing $Z_i-(X)_n-Z_j$ sequences ($n=0-14$; Z: amino acid that binds the metal through the side chain; X: any amino acid) from the existing protein structures in the non-redundant Protein Data Bank; 2) merges these fragments into a single $Z_1-(X)_{n_1}-Z_2-(X)_{n_2}-Z_3-$

$(X)_{n_3}-\dots-Z_j$ polypeptide chain; and 3) automatically performs two simple molecular mechanics calculations that make it possible to estimate the internal strain in the newly designed peptide. The application of this procedure for the most M^{2+} -specific combinations of amino acid side chains (M: metal; see L. Rulíšek, Z. Havlas *J. Phys. Chem. B* **2003**, *107*, 2376–2385) yielded several peptide sequences (with lengths of 6–20 amino acids) with the potential for specific binding with six metal ions

(Co^{2+} , Ni^{2+} , Cu^{2+} , Zn^{2+} , Cd^{2+} and Hg^{2+}). The gas-phase association constants of the studied metal ions with these de novo designed peptides were experimentally determined by MALDI mass spectrometry by using 3,4,5-trihydroxyacetophenone as a matrix, whereas the thermodynamic parameters of the metal-ion coordination in the condensed phase were measured by isothermal titration calorimetry (ITC), chelatometry and NMR spectroscopy methods. The data indicate that some of the computationally predicted peptides are potential M^{2+} -specific metal-ion chelators.

Keywords: ab initio calculations • mass spectrometry • metal-ion chelation • molecular design • peptides

Introduction

The design of novel binding sites in peptides with a high specificity for a particular metal ion is a highly attractive

goal for two reasons: it may provide a deeper understanding of the molecular basis for metal-ion specificity in proteins and peptides on the one hand and binders for biomedical and technical applications on the other. These applications


[a] M. Kožíšek, M. Buděšínský, Dr. Z. Havlas, Dr. L. Rulíšek
Gilead Sciences & IOCB Research Centre
Institute of Organic Chemistry and Biochemistry
Academy of Sciences of the Czech Republic
Flemingovo náměstí 2, 16610 Prague 6 (Czech Republic)
Fax: (+420) 220-183-578
E-mail: lubos@uochb.cas.cz

[b] M. Kožíšek
Department of Biochemistry
Faculty of Science, Charles University
Hlavova 8, 128 93 Prague 2 (Czech Republic)

[c] Dr. A. Svatoš, Dr. A. Muck
Mass Spectrometry Research Group
Max Planck Institute for Chemical Ecology
Hans-Knöll-Strasse 8, 07745 Jena (Germany)

[d] M. C. Bauer, Prof. S. Linse
Department of Biophysical Chemistry
Lund University, Chemical Centre
P.O. Box 124, 22100 Lund (Sweden)

[e] Dr. P. Kotrba, Prof. T. Ruml
Department of Biochemistry and Microbiology
Faculty of Food and Biochemical Technology
Institute of Chemical Technology
Technická 5, 16628 Prague 6 (Czech Republic)

 Supporting information for this article (detailed information about the NMR measurements) is available on the WWW under <http://dx.doi.org/10.1002/chem.200800178>.

include the removal of metals from polluted environments, either by bacterial strains^[1–5] or other biotechnological techniques,^[6–8] the design of novel biosensors,^[9–11] the redesign of proteins (to provide new building blocks or maquettes)^[12,13] or the design of new chelating compounds for medicinal chemistry.^[14]

Numerous successful attempts have recently been made to alter the metal-binding properties of selected proteins with the aims of modifying protein function,^[15–20] engineering novel metalloproteins or metal-binding sites in natural metalloproteins^[21–26] or tuning metalloprotein or peptide specificity towards a particular metal ion.^[27–31] While most studies employ the standard repertoire of amino acids, one approach has been to include unnatural amino acids or non-native metal-containing cofactors.^[32]

In many cases, experimental efforts in protein redesign have been supported by new computational algorithms, which have yielded the best candidate residues for mutation, such that a new metal-binding site may appear in the protein.^[33–35] In general, the level of selectivity enhancement achieved is dependent on the ability to predict and control the specificity of the site (that is, through an optimal combination of metal-binding residues and the preferred coordination number).^[36] The predictions are frequently based on the observed frequency of occupation of coordination sites in metalloproteins^[37] or use semiquantitative concepts, like the HSAB theory of Parr and Pearson.^[38] Only recently has the design been assisted by advances and accumulated experience in quantum (theoretical) bioinorganic chemistry. By using high-level quantum chemical methods coupled with efficient algorithms for solving the Poisson–Boltzmann equation to account for the effect of environment, many important properties of first- and second-shell ligand binding have been calculated,^[39,40] for example, the effect of the carboxylate-binding mode on metal binding/selectivity and function in proteins.^[41]

We have contributed to the development of novel and modified metal-ion binding sites through several computational studies, in which the affinities of all of the metal-binding amino acid (AA) side chains towards selected transition metal ions (Co^{2+} , Ni^{2+} , Cu^{2+} , Zn^{2+} , Cd^{2+} and Hg^{2+}) were evaluated,^[42,43] the cooperative effect accompanying the simultaneous binding of more than one AA side chain was quantitatively characterised^[44,45] and, finally, “ M^{2+} -specific” combinations of amino acid side chains were proposed for each of these metal ions.^[46] The predicted M^{2+} -specific combinations can be considered, on the basis of large basis set DFT/B3LYP calculations of the complexation energies of the metal ions in $[\text{M}(\text{Y}_1)\cdots(\text{Y}_n)]^{c+}$ complexes (with Y_i being an amino acid side chain), to be the best candidates for highly selective metal-binding sites in proteins or peptides. The specificity can be further increased by taking into account the preferred coordination geometry for each metal ion, based on an analysis of Protein Data Bank (PDB) and Cambridge Structural Database (CSD) structures.^[47] Consequently, the Co^{2+} - and Ni^{2+} -specific combinations are defined as $[\text{Co}(\text{Y}_1)\cdots(\text{Y}_6)]^{c+}$ (or $[\text{Ni}(\text{Y}_1)\cdots(\text{Y}_6)]^{c+}$) complexes

(with Y_i representing “optimum” amino acid side chains) in an octahedral coordination geometry. For the Cd^{2+} and Zn^{2+} ions, the $[\text{Zn}(\text{Y}_1)\cdots(\text{Y}_4)]^{c+}$ or $[\text{Cd}(\text{Y}_1)\cdots(\text{Y}_4)]^{c+}$ sites are tetrahedral, whereas the sites are square-planar for the Cu^{2+} ions and linear for the Hg^{2+} ions (that is, $[\text{Hg}(\text{Y}_1)(\text{Y}_2)]^{c+}$).

The aim of the current work is to develop and evaluate a computational strategy that can be used to link the isolated amino acid side chains into a single polypeptide chain by 1) creating a library of $\text{Z}_i-(\text{X})_n-\text{Z}_j$ sequences ($n=0\text{--}14$; Z : metal-binding amino acids; X : any amino acid) from the non-redundant PDB database, 2) merging these fragments into a single $\text{Z}_1-(\text{X})_{n_1}-\text{Z}_2-(\text{X})_{n_2}-\text{Z}_3-(\text{X})_{n_3}-\cdots-\text{Z}_j$ polypeptide chain and 3) estimating the strain energy by calculating the difference between the folded and unfolded peptide with molecular mechanics (MM) methods (Figure 1). It is an ab initio design of new peptide sequences and, if successful, it may become one of the very few examples of the predictive power of computational chemistry in the literature.^[48]

Results

Hg^{2+} -specific metal-binding peptides (linear coordination geometry): Design and theoretical calculations: Based on the earlier results (PDB and CSD analyses),^[47] the mercury(II) ion has been assumed to prefer a linear coordination geometry. The four most Hg^{2+} -specific sites are:^[46]

- 1) $[\text{Hg}(\text{Imi})_2]^{2+}$ (Imi: imidazole), with the optimum distance being 2.11 Å (Hg–N).
- 2) $[\text{Hg}(\text{Imi})(\text{H}_2\text{S})]^{2+}$, with the optimum distances being 2.13 Å (Hg–N) and 2.48 Å (Hg–S).
- 3) $[\text{Hg}(\text{H}_2\text{S})_2]^{2+}$, with the optimum distance being 2.50 Å (Hg–S).
- 4) $[\text{Hg}(\text{SH})_2]$, with the optimum distance being 2.38 Å (Hg–S).

Each combination (1–4) is compatible with only one site, that is, $[\text{Hg}(\text{His})_2]^{2+}$, $[\text{Hg}(\text{His})(\text{Met})]^{2+}$, $[\text{Hg}(\text{Met})_2]^{2+}$ and $[\text{Hg}(\text{Cys})_2]$, respectively; Cys is assumed to be in the deprotonated state (that is, CH_2S^-). To limit the number of fragments found, we have used stricter criteria for accepting the fragment than in cases with the other metal ions: d_{read} (Line 6 of the input file, see below) = 0.012 Å (only fragments with d_{actual} differing by less than 1.2% from the d_{ideal} above were considered) and $3 < n+2 < 9$ (that is, tetra- to octapeptides were considered). The numbers of sites found for each combination are listed in Table 1.

Table 2 shows the five sequences for each type of site (20 in total) that had the most favourable values of ΔE_F and mutual positions of the binding residues (visual analysis). These can be considered as the best candidates for Hg^{2+} -specific sites.

It should be mentioned that in the case of the linear coordination geometry, the peptides are fragments found in proteins. Consequently, the ΔE_F value is very small because many of the fragments represent stable parts of protein structures.

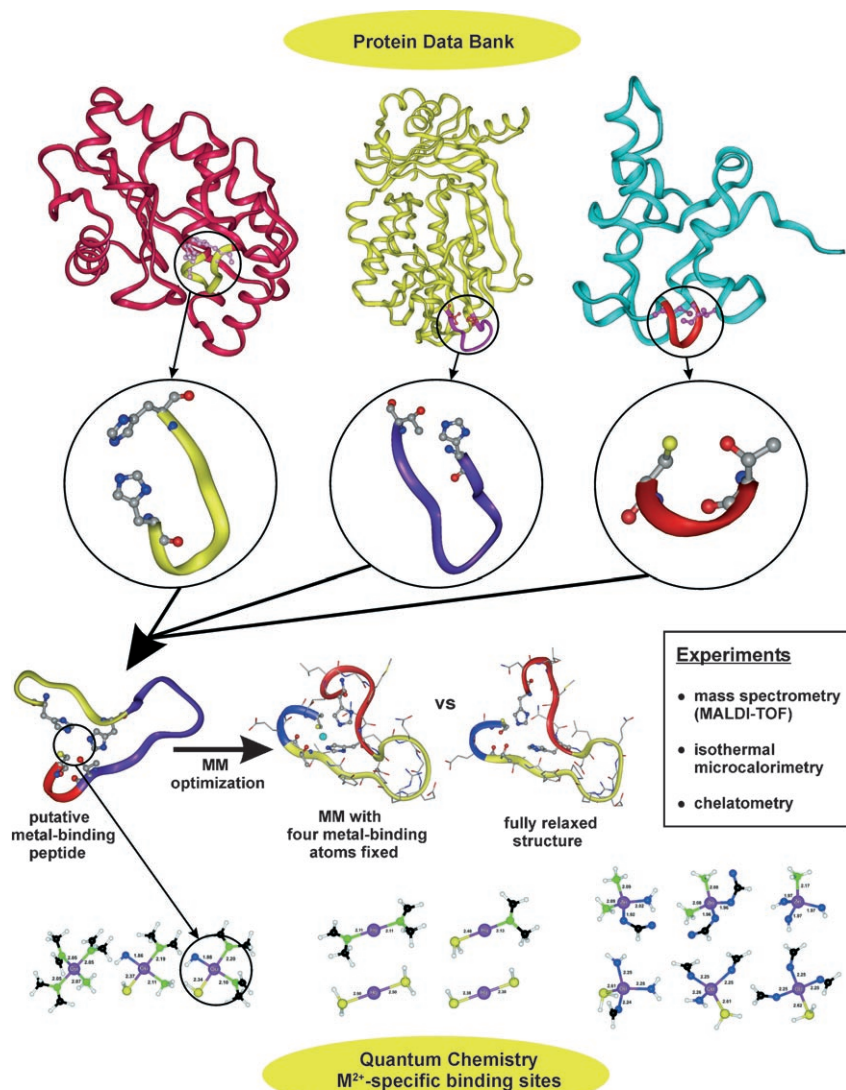


Figure 1. A depiction of the key steps in the de novo design strategy adopted in this work.

Table 1. The numbers of peptide fragments (4–8 amino acids), N_{FRG} , found for the four most Hg^{2+} -specific sites (with the number of non-equivalent entries, $N_{\text{FRG}}(\text{neq})$, in parentheses). The numbers of entries with strain energy, ΔE_{F} , smaller than 4 (or 6) kJ mol^{-1} are listed in the last column.

	$N_{\text{FRG}}(\text{neq})$	$\Delta E_{\text{F}} [\text{kJ mol}^{-1}]$	N_{FRG} with $\Delta E_{\text{F}} < 4$ (or 6) kJ mol^{-1}
His- X_n -His	62 (39)	< 6	18
Met- X_n -His	31 (19)	< 4	8
His- X_n -Met	44 (34)	< 4	14
Met- X_n -Met	41 (35)	< 4	14
Cys- X_n -Cys	30 (28)	< 4	16
Total	208 (155)		70

Zn²⁺-specific metal-binding peptides (tetrahedral coordination geometry): Design and theoretical calculations: The zinc(II) ion has been assumed to prefer a tetrahedral coordination geometry.^[47] The three most Zn²⁺-specific sites are:^[46]

- 1) $[\text{Zn}(\text{NH}_3)_2(\text{H}_2\text{O})(\text{HCOO})]^+$, with the optimum distances being 2.09 (Zn–N_{NH₃}), 2.02 (Zn–O_{H₂O}) and 1.92 Å (Zn–O_{HCOO}).
- 2) $[\text{Zn}(\text{NH}_3)_2(\text{HCOO})_2]$, with the optimum distances being 2.08 (Zn–N_{NH₃}) and 1.96 Å (Zn–O_{HCOO}).
- 3) $[\text{Zn}(\text{OH})_3(\text{NH}_3)]^-$, with the optimum distances being 1.97 (Zn–O) and 2.17 Å (Zn–N).

12 combinations of amino acid side chains compatible with these three sites are given in Table 3, to which three more combinations containing His (which had been modelled by Imi in an earlier study^[46]) instead of Lys (which had been modelled by NH₃) were added. All of these combinations were used in the input files of *build_peptide*. The set of input parameters (denoted as the standard set) depicted in Figure 6 (Lines 6–11; in the Computational Details section) was essentially used for all tetra-coordinate sites (that is, tetrahedral Zn²⁺ and Cd²⁺ ions and square-planar Cu²⁺ ions). In certain cases, the parameters have been slightly modified to increase or decrease the number of entries found for a particular combination of metal-binding AA residues.

Table 2. The best candidates for Hg^{2+} -specific metal-binding peptides selected on the basis of a visual inspection of 70 species with a favourable ΔE_{F} value.

Sequence type			
His- X_n -His	His- X_n -Met; Met- X_n -His	Met- X_n -Met	Cys- X_n -Cys
H-IP-H	H-DR-M	M-NK-M	C-NVI-C
H-EFG-H	H-KDL-M	M-RAQ-M	C-LAG-C
H-QRA-H	H-YVE-M	M-FMH-M	C-MNRK-C
H-YKQ-H	M-DQT-H	M-DNR-M	C-KSNC-C
H-NLGK-H	M-CDVEGN-H	M-RKEG-M	C-SGAQM-C

The numbers of non-equivalent converged structures are listed in Table 3. It should be noted here that peptides are considered to be equivalent when their sequences are 100% identical and the energies obtained from the pre-screening

Table 3. The numbers of peptide fragments (<22 amino acids) found for the four most Zn^{2+} -specific sites (with the numbers of converged and non-equivalent entries in parentheses).

Type	N_{FRG} (neq)	Type	N_{FRG} (neq)
{Lys,Lys,Glu,Ser}	6 (0)	{Lys,Lys,Asp,Asp}	28 (7)
{Lys,Lys,Asp,Ser}	12 (2)	{Lys,Lys,Glu,Glu}	24 (3)
{Lys,Lys,Glu,Thr}	6 (3)	{Lys,Tyr,Tyr,Tyr}	0 (0)
{Lys,Lys,Asp,Thr}	32 (9)	{Lys,Ser,Tyr,Tyr}	4 (0)
{Lys,Lys,Asp,Glu}	18 (6)	{His,Lys,Asp,Glu}	6 (3)
{His,Lys,Asp,Ser}	11 (11)	{Lys,Ser,Ser,Tyr}	6 (3)
{His,Lys,Glu,Ser}	0 (0)	{Lys,Ser,Thr,Tyr}	2 (2)
{Lys,Thr,Thr,Tyr}	4 (2)		
Total		159 (51)	

full optimisation of the peptide structures (by using the fast first-order optimisation procedure; see the Computational Details section) are almost identical (difference of <8 kJ mol⁻¹).

The converged peptides were sorted based on the ΔE_{F} values. The initial and final structures of the entries with ΔE_{F} values of less than 50 kJ mol⁻¹ were visually inspected. Preference was given to entries with less initial crowding and a more favourable mutual orientation of binding residues. It was also taken into account that each of the three site types should be represented in the final set of specific metal-binding peptides. In Table 4, the nine most favourable sequences are summarised, along with their ΔE_{F} values.

Table 4. The nine best candidates for Zn^{2+} -specific metal-binding peptides selected on the basis of the ΔE_{F} values, a visual inspection and the (non-)crowding in the initial structures (that is, the output from the *build_peptide* program).

Sequence (length in AAs)	Peptide ID	ΔE_{F} [kJ mol ⁻¹]
KTEYVDERSKSLTVDLTK (18)	KDTK_T016	7.5
KFFKDFRHKPATELTHED (18)	KDKD_T003	9.6
YLFLGMHPDLSSLSK (15)	YSSK_T001	10.9
KIYKEKKYGKEPQVAKT (17)	KEKT_T002	22.6
DAAGKVEGKDDNK (13)	KDKD_T025	25.9
KTASGGIESYNDSH (15)	KDSH_T003	32.2
KEIRESAFGKSVE (13)	KEKE_T005	35.9
KKFMDFRHKPATELTHED (18)	KDKD_T006	36.4
SFTSGIYTLWNDQIVK (16)	STYK_T001	48.9

Cd^{2+} -specific metal-binding peptides (tetrahedral coordination geometry): Design and theoretical calculations: The cadmium(II) ion has been assumed to prefer a tetrahedral coordination geometry.^[47] The three most Cd^{2+} -specific sites are:^[46]

- 1) $[\text{Cd}(\text{H}_2\text{O})_2(\text{H}_2\text{S})(\text{H}_2\text{CO})]^{2+}$, with the optimum distances equal to 2.25 (Cd–O_{H₂O}), 2.24 (Cd–O_{H₂CO}) and 2.61 Å (Cd–S).
- 2) $[\text{Cd}(\text{H}_2\text{CO})_2(\text{H}_2\text{O})(\text{H}_2\text{S})]^{2+}$, with the optimum distances equal to 2.25 (Cd–O_{H₂CO}), 2.26 (Cd–O_{H₂O}) and 2.61 Å (Cd–S).

- 3) $[\text{Cd}(\text{H}_2\text{CO})_3(\text{H}_2\text{S})]^{2+}$, with the optimum distances equal to 2.25 (Cu–O) and 2.62 Å (Cu–S).

16 combinations of amino acid side chains compatible with these three sites, used in the input files of *build_peptide*, are given in Table 5, along with the numbers of resulting converged and non-equivalent structures.

Table 5. The numbers of peptide fragments (<22 amino acids) found for the four most Cd^{2+} -specific sites (with the numbers of converged and non-equivalent entries in parentheses).

Type	N_{FRG} (neq)	Type	N_{FRG} (neq)
{Met,Asn,Ser,Ser}	12 (5)	{Met,Gln,Gln,Ser}	2 (1)
{Met,Gln,Ser,Ser}	6 (1)	{Met,Asn,Asn,Thr}	6 (2)
{Met,Asn,Ser,Thr}	11 (8)	{Met,Asn,Gln,Thr}	5 (3)
{Met,Gln,Ser,Thr}	6 (4)	{Met,Gln,Gln,Thr}	2 (1)
{Met,Asn,Thr,Thr}	10 (3)	{Met,Asn,Asn,Asn}	0 (0)
{Met,Gln,Thr,Thr}	6 (2)	{Met,Asn,Asn,Gln}	0 (0)
{Met,Asn,Asn,Ser}	4 (1)	{Met,Asn,Gln,Gln}	0 (0)
{Met,Asn,Gln,Ser}	8 (2)	{Met,Gln,Gln,Gln}	0 (0)
Total		78 (33)	

By using the same approach as that described above, the nine most favourable candidates for Cd^{2+} -specific sites were chosen; they are listed in Table 6.

Table 6. The nine best candidates for Cd^{2+} -specific metal-binding peptides selected on the basis of the ΔE_{F} values, a visual inspection and the (non-)crowding in the initial structures (that is, the output from the *build_peptide* program).

Sequence (length in AAs)	Peptide ID	ΔE_{F} [kJ mol ⁻¹]
TLRSGSTEDQMDIVGFSQEEQ (21)	TSMQ_T005	14.6
NGLTKNSAACRAAKLQDCTM (20)	NTSM_T006	22.6
QVTPTSTFGSM (11)	QSSM_T001	45.1
MSVGOKITAEDGTQ (14)	MQTQ_T001	49.3
MRSSWSGN (8)	MSSN_T002	56.8
QNACSEHGATLSLKM (15)	QTSM_T002	62.3
NSRSGSTEDQMDIVGFSQEEQ (20)	NSMQ_T007	66.5
MGNTRLGLN (9)	MNTN_T001	73.2
QSEDIKTYTTRVRCM (15)	QSQM_T001	74.4

Cu^{2+} -specific metal-binding peptides (square-planar coordination geometry): Design and theoretical calculations: The copper(II) ion has been assumed to prefer a square-planar coordination geometry.^[47] The three most Cu^{2+} -specific sites are:^[46]

- 1) $[\text{Cu}(\text{Imi})_3(\text{NH}_3)]^{2+}$, with the optimum distances equal to 2.05 (Cu–N_{imi}) and 2.07 Å (Cu–N_{NH₃}).
- 2) $[\text{Cu}(\text{Imi})(\text{NH}_3)(\text{OH})(\text{SH})]$, with OH⁻ and NH₃ in the *trans* positions, and with the optimum distances equal to 1.86 (Cu–O), 2.37 (Cu–S), 2.19 (Cu–N_{imi}) and 2.11 Å (Cu–N_{NH₃}).
- 3) $[\text{Cu}(\text{Imi})_2(\text{OH})(\text{SH})]$, with (Imi,Imi) in the *cis* positions, and with the optimum distances equal to 1.88 (Cu–O), 2.34 (Cu–S), 2.10 (Cu–N_{imi}, *trans* to OH⁻) and 2.20 Å (Cu–N_{imi}, *cis* to OH⁻).

Seven combinations of amino acid side chains compatible with these three sites, used in the input files of *build_peptide*, are listed in Table 7, along with the numbers of non-equivalent and converged peptides.

Table 7. The numbers of peptide fragments (<22 amino acids) found for the three most Cu²⁺-specific sites (with the numbers of converged and non-equivalent entries in parentheses).

Type	N _{FRG} (neq)	Type	N _{FRG} (neq)
{His,His,His,Lys}	2 (1)	{His,His,Cys,Ser}	9 (3)
{His,Lys,Cys,Ser}	2 (1)	{His,His,Cys,Thr}	2 (1)
{His,Lys,Cys,Thr}	3 (2)	{His,His,Cys,Tyr}	2 (1)
{His,Lys,Cys,Tyr}	1 (1)		
		Total	21 (10)

By using the same approach as that described above, the six most favourable candidates for Cu²⁺-specific sites were chosen; they are listed in Table 8. It should be noted here that only 11 structures (out of a total of 21) converged in the molecular mechanics optimisation. For five of them, $\Delta E_F > 150 \text{ kJ mol}^{-1}$ and, therefore, only the six peptides listed in Table 8 can be considered as suitable candidates for Cu²⁺-specific sites.

The remaining two M²⁺-specific sites, the octahedral Ni²⁺-specific and Co²⁺-specific sites, have not yet been subjected to any attempt to design M²⁺-specific peptides because we consider it unlikely that a single peptide could fold around the metal ion and provide six amino acid side chains (if we exclude the bidentate coordination mode of carboxylic side chains) at the vertices of the coordination polyhedron. A different strategy will be adopted in a forthcoming study in which it is taken into consideration that some of the coordination positions may be occupied by water molecules.

Two selected model structures, as predicted by the *build_peptide* program and simple MM minimisation, are depicted in Figure 2.

MALDI-TOF MS experiments:

Electrospray-ionisation mass spectrometry (ESIMS) has been used to elucidate the gas-phase binding of metals to proteins^[49] and for quantitative measurements of peptide cationisation with metal ions.^[50,51] Until now, the gas-phase attachment of heavy metals by using

Table 8. The six best candidates for Cu²⁺-specific metal-binding peptides selected on the basis of the ΔE_F values, a visual inspection and the (non-)crowding in the initial structures (that is, the output from the *build_peptide* program).

Sequence (length in AAs)	Peptide ID	ΔE_F [kJ mol ⁻¹]
HNLGMNHDLDQGERPYVTEGC (20)	HHTC_S002	47.7
CPSEDHVSQDK (11)	CHSK_S001	59.4
CFNCGKEGHVSTAARH (16)	CHSH_S001	64.4
CLLQHAVTKRKTIRYDKEAK (20)	CHTK_S003	80.7
CPFEDHSHFSKKNKAYVH (18)	CHYH_S001	87.4
CPFEDHSHFSKKNKAYNSSTK (21)	CHYK_S001	108.3

matrix-assisted desorption/ionisation (MALDI) has not been reported in the literature, despite the fact that the formation of the ions after laser irradiation of a solid sample under a high vacuum without the presence of solvents^[52] is a condition favourable for complex peptide-cation formation. The low-pressure (approximately 5×10^{-8} mbar) MALDI source seems to mimic well the conditions under which the binding of peptides with divalent ions was optimised. When MALDI matrices designed for peptides, such as α -cyano, 2,5-dihydroxybenzoic and sinapic acid, were used, preferential protonation of the peptide resulted in a low yield of the metal complexes (data not shown). The protonation can be diminished by using matrices without an acidic functionality.^[52] We have successfully utilised 3,4,5-trihydroxyacetophenone (THAP) for the metal-complexation study. The acid

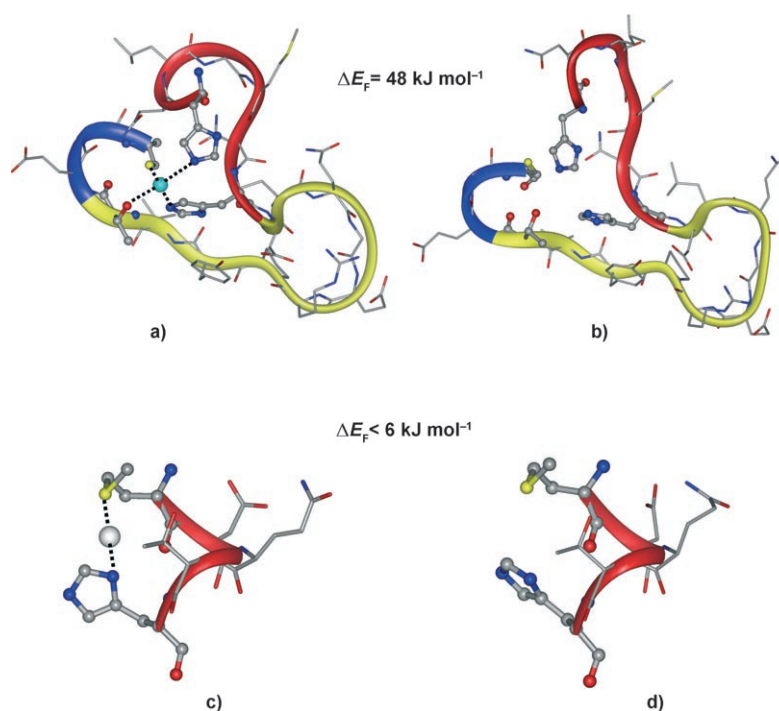


Figure 2. The model structures in a predicted folded state around the central metal ion: a) the HHTC 20-mer peptide in a folded state with a metal ion bound by two His, one Thr and one Cys residues in a square-planar coordination geometry; b) the HHTC 20-mer peptide in an unfolded state (the local minimum nearest to a folded state in the absence of a metal ion); c) the MDQTH peptide in a folded state binding to a metal ion through Met and His terminal residues in a linear coordination geometry; d) the MDQTH peptide in an unfolded state.

cannot be fully avoided as the peptides were only soluble by using trifluoroacetic acid (TFA; 1% w/w) to form stock solutions; however, further dilutions were made by using MilliQ water. The metal ions were added as sulphate salt solutions in tenfold excess over the peptide concentration and, after a short incubation, the mixtures were combined with the matrix solution and co-applied on the MALDI target by using the dry droplet technique.

Not a single case has been observed that would correspond to the exclusive formation of a peptide–metal complex in the gas phase. Instead, the signals of protonated $[M+H]^+$ and metal-cationised peptides $[M+\text{metal}-H]^+$ of different abundance were found (Figure 3; results are analogous to previous studies with electrospray ionisation)^[50,51] along with signals of divalent metal–peptide adducts and, in a few cases, also weak signals of bi- and trimetal complexes. For the MALDI study, representative peptides were selected to cover both the “best” and “poor” candidates from the theoretical study. Unlike the theoretical screening, all samples were treated with all divalent metal ions, applied as sulphates. In the cases of Hg^{2+} and Cd^{2+} , no signals of cationised peptides were noted, presumably due to the precipitation of the formed complexes or to polymerisation preventing evaporation during the MALDI process. Additionally, Ni^{2+} was used to investigate the influence of methionine and histidine on gas-phase binding. Table 9 summarises the presence/absence (+/–) of cationised peptides in a MALDI-TOF mass spectrum and the ratio of $[M+H]^+/[M+\text{metal}-H]^+$ if higher than 110:1. In conclusion, we consider MALDI experiments to be an elegant method for the initial scanning of peptide binding, although they do not yield any structural information about possible binding modes.

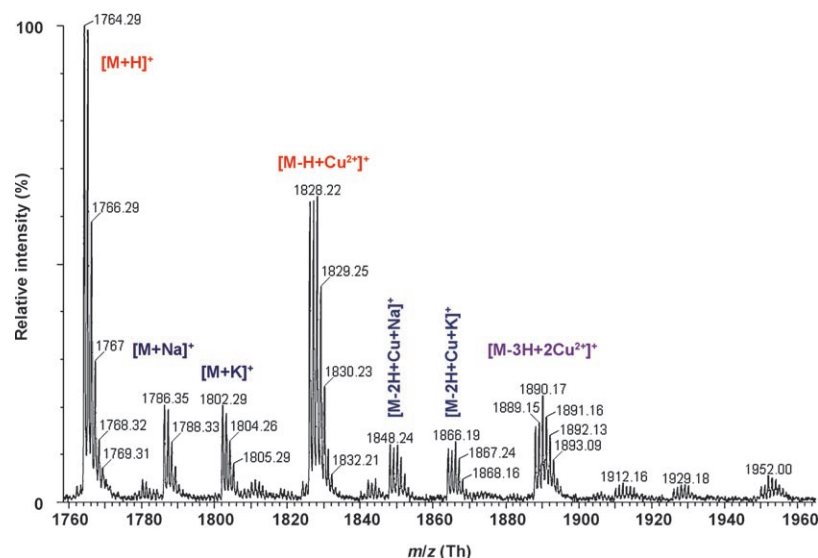


Figure 3. A section of the MALDI-TOF mass spectrum of a mixture of the YSSK peptide with CuSO_4 (1:10 molar ratio) mixed with a 3,4,5-trihydroxyacetophenone (THAP) matrix.

Table 9. A summary of the observed MALDI-TOF MS signals of protonated and divalent metal-cationised peptides and the ratio of their intensities.

	Zn^{2+}	Cu^{2+}	Cd^{2+}	Hg^{2+}	Ni^{2+}
CHSK	–/–	–/–	+/–	–/–	+/–
HEFGH	+/+ 1:2.5	+/+ 1:4.8	+/–	+/–	+/–
MDQTH	–/–	+/+ 2:1	+/–	–/–	+/+ 4.5:1
MFMHM	+/+ 8.8:1	–/–	+/–	–/–	+/+ 2.6:1
HHTC	+/+ 6.4:1	+/+ 4:1	+/–	–/–	+/+ 31.2:1
KDKD	+/+ 8.8:1	+/+ 9.7:1	+/–	–/–	+/–
MNTN	+/–	+/–	+/–	–/–	+/–
MSSN	+/–	+/+ 3.4:1	+/–	–/–	+/+ 3.5:1
NTSM	+/–	+/–	+/–	–/–	+/–
QSSM	–/–	+/–	–/–	–/–	+/–
TMSQ	+/+ 19.2:1	+/–	+/–	+/–	+/–
YSSK	+/+ 15.2:1	+/+ 1:1.3	+/–	+/–	+/–
KEKE	+/+ 89:1	+/–	+/–	+/–	+/+ 103:1
KSDH	+/–	+/+ 17:1	+/–	–/–	+/–

ITC measurements: The thermodynamic properties of metal-ion binding in solution were determined by using two methods: isothermal titration calorimetry (ITC) and titration in the presence of a fluorescent chelator (see below). Based on the MALDI-TOF MS results, six peptides were selected for the ITC experiments. For five of them (KEKE, CHSK, HEFGH, MSSN, YSSK), ITC did not reveal any interactions characterised by an association constant (K_a) higher than $\approx 10^4 \text{ M}^{-1}$. Higher concentrations of the peptides would be necessary to enable the testing of the interaction below this limit. Nevertheless, due to the low solubility of the peptides, it was not possible to determine their binding characteristics by ITC. The higher affinity binding was observed for the HHTC peptide, a result that will be discussed in greater detail below.

ITC was used to monitor the binding of divalent copper, zinc and cadmium ions to the peptide HNLGMNHDLDQGERPYV-TEGC (HHTC). The results are summarised in Table 10 and the titration performed in *N*-(2-acetamido)-2-aminoethanesulfonic acid (ACES) buffer is shown in Figure 4. When titrations were performed in buffers with different enthalpies of ionisation (ACES: $\Delta H_{\text{ion}} = 31.4 \text{ kJ mol}^{-1}$; 2-(*N*-morpholino)-ethanesulfonic acid (MES): $\Delta H_{\text{ion}} = 15.5 \text{ kJ mol}^{-1}$), the experiments resulted in the same binding enthalpies; this indicates that metal-ion binding is not associated with any proton transfers^[53] (the data for the MES buffer are not shown).

From the titration curves, the stoichiometries of both ions were estimated to be 1.0. The

Table 10. The thermodynamic parameters of the copper(II), zinc(II) and cadmium(II) ions binding to the HNLGMNHDLQGERPYVTEGC peptide. The ΔG^0 and ΔS^0 values were calculated for a standard concentration of 1 M.

Ion	ΔG^0 [kJ mol ⁻¹]	ΔH^0 [kJ mol ⁻¹]	$-T\Delta S^0$ [kJ mol ⁻¹]	K_a [M ⁻¹]	K_d [μ M]
Cu ²⁺	-36.4 ± 0.4	-28.5 ± 0.4	-8.0 ± 0.4	$(2.4 \pm 0.5) \times 10^6$	0.42 ± 0.09
Zn ²⁺	-28.6 ± 0.1	-10.9 ± 0.1	-17.7 ± 0.2	$(1.03 \pm 0.04) \times 10^5$	9.7 ± 0.4
Cd ²⁺	-25.6 ± 0.1	-32.0 ± 0.5	6.4 ± 0.5	$(3.00 \pm 0.08) \times 10^4$	33 ± 1

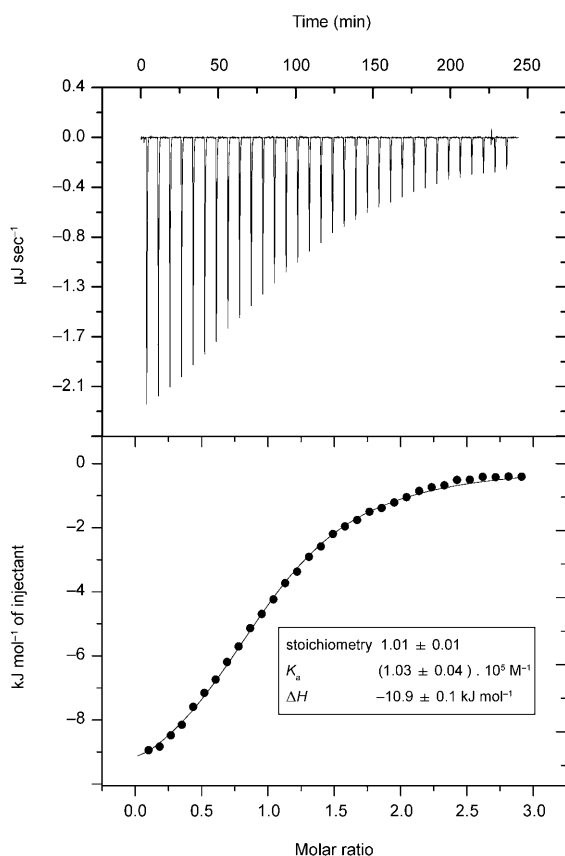


Figure 4. Isothermal titration of the peptide HNLGMNHDLQGERPYVTEGC with zinc chloride. The titration was performed at 25 °C in 20 mM ACES at pH 7.0. Upper graph: experimental data. Lower graph: Fitted (line) to the integrated heats (●) from each injection, corrected for the heat of dilution of the peptide.

Cu²⁺ ion binds to the peptide with a value of $K_a = (2.4 \pm 0.5) \times 10^6 \text{ M}^{-1}$, which corresponds to a total Gibbs free energy change (ΔG^0) of $-36.4 \text{ kJ mol}^{-1}$. The calorimetric measurement shows that, for Cu²⁺ ions, most of the free energy change under the selected standard conditions (1 M) is of enthalpic ($\Delta H^0 = -28.5 \text{ kJ mol}^{-1}$) rather than entropic ($-T \times \Delta S^0 = -8.0 \text{ kJ mol}^{-1}$) origin. The entropic term includes both favourable contributions from the desolvation of the ion and peptide groups on the one hand and unfavourable contributions from the organisation of the peptide around the metal ion on the other. The zinc ion binds to this peptide approximately 20-fold less tightly; the association constant for Zn²⁺ ions was determined to be $K_a = (1.03 \pm 0.04) \times 10^5 \text{ M}^{-1}$, which corresponds to a standard Gibbs free energy

change of $-28.6 \text{ kJ mol}^{-1}$. This is divided into a favourable enthalpic ($\Delta H^0 = -10.9 \text{ kJ mol}^{-1}$) contribution and a favourable entropic ($-T \times \Delta S^0 = -17.7 \text{ kJ mol}^{-1}$) contribution under the standard conditions (1 M). The association constant for the cadmium(II) ion ($3 \times 10^4 \text{ M}^{-1}$) was found to be the smallest of the three and corresponds to a Gibbs free energy of binding of $-25.6 \text{ kJ mol}^{-1}$. The cadmium(II) ion binds to the HHTC peptide with a favourable enthalpy change ($\Delta H^0 = -32.0 \text{ kJ mol}^{-1}$), which is, however, compensated for by an unfavourable entropic contribution ($-T \times \Delta S^0 = 6.4 \text{ kJ mol}^{-1}$).

Chelatometry: The titration of the Zn²⁺-specific chelator FluoZin-1 in the presence of peptides in different concentrations was used to determine the Zn²⁺ affinity for the peptides. Figure 5 shows a titration of FluoZin-1 and the HHTC peptide. The fitted dissociation constant was found to be 50 μM , which is in very good agreement with the value obtained by using ITC.

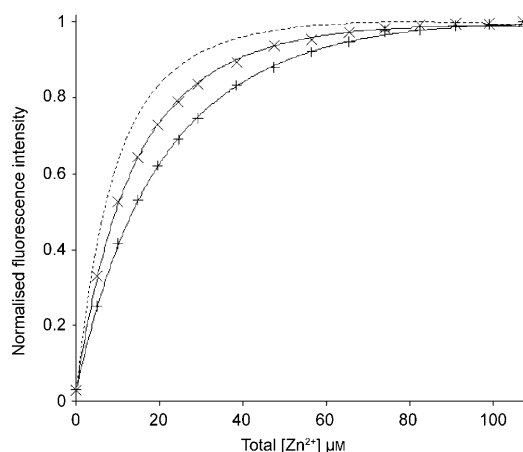


Figure 5. Fluorescence spectrum of 5 μM FluoZin-1 in the presence of 40 μM (x) or 80 μM (+) HHTC in 20 mM MES at pH 6.5 during titration with Zn²⁺. The solid lines show the best fits to the data for each concentration by using a one-site model. The dotted line is the fit of a titration of 5 μM FluoZin-1 alone.

NMR spectroscopy: The ¹H NMR spectrum of the HHTC peptide showed the presence of two species in a ratio of about 85:15, due to isomerism at the Arg-Pro peptide bond. Only the NMR spectra of the major species with a *trans*-Arg-Pro amide bond could be analysed in detail. For the structural assignment, a 2.77 mM solution of peptide HHTC (3.2 mg) in H₂O/D₂O (9:1; 0.5 mL) was used and a standard strategy based on a combination of COSY, TOCSY and NOESY spectra was applied. The 1D ¹H NMR spectrum is shown in Figure S1 and the proton chemical shifts are given in Table S1, both in the Supporting Information.

The NMR spectra measured at 280, 300 and 320 K showed a strong temperature dependence for all of the

amide NH protons. The observed temperature coefficients in the range -5 to -10 ppb (Table S1 in the Supporting Information) exclude the presence of strong intramolecular H-bonds and indicate high flexibility of the peptide molecule. This is also supported by the chemical shifts of the H_{α} protons and $J(\text{NH}, H_{\alpha})$ values that are close to random-coil values.

The dependence of the NMR spectra on the concentration of peptide HHTC was tested with 2.77, 1.3 and 0.65 mM solutions. The observed chemical shifts changed in the range from 0 to 0.06 ppm. To eliminate even small dilution shifts and to prevent possible precipitation, the effect of peptide complexation with metal cations was studied with a peptide solution of about 1.5 mM and with a two-step addition of the solid salt (CuCl_2 (molar ratio M/P = 6:1 and 20:1), ZnCl_2 (M/P = 18:1 and 68:1) and/or CdSO_4 (M/P = 4:1 and 12:1)).

[Peptide+CuCl₂]: Even after the first addition of CuCl_2 , a strong non-selective line broadening was observed. It was still pronounced after the second addition of the salt (Figure S2 in the Supporting Information), probably due to the paramagnetic properties of CuCl_2 . No information about potential binding sites could be obtained.

[Peptide+ZnCl₂]: Only weak line broadening and non-selective small induced up-field chemical shifts (Figure S3 in the Supporting Information) were observed after the stepwise addition of the Zn salt. Somewhat larger induced chemical shifts (Figure S4 in the Supporting Information) appeared for the signals of the ring hydrogen atoms of both His residues.

[Peptide+CdSO₄]: Similar to the results after addition of ZnCl_2 , only weak line broadening and non-selective very small induced chemical shifts (Figure S5 in the Supporting Information) were observed during stepwise addition of the Cd salt. In contrast to the previous case, the signals of the ring hydrogen atoms of both His residue are not influenced by addition of CdSO_4 (Figure S6 in the Supporting Information).

We may conclude that NMR experiments did not provide reliable supporting data for the existence of peptide complexation with metal cations at the expected binding sites (His1, His7, Thr17 and Cys20). Only the behaviour of the two histidine residues after addition of ZnCl_2 could indicate formation of a weak complex.

Discussion

In designing metal-specific sites, several factors play a role. The most important ones are likely to be the coordination geometry and the structure of the first coordination sphere (that is, the metal-binding residues). In our previous work, we have, by employing a purely theoretical approach, devised the best candidates for M^{2+} -specific combinations of amino acid side chains. With His and Cys being the most

common residues in metal-binding sites (and, along with carboxylate groups, the strongest chelators), a His- and/or Cys-rich peptide sequence that would certainly bind divalent metal ions with high affinity can perhaps be designed without the need to employ quantum chemical calculations. However, the control of metal-ion specificity is much lower in such a brute-force approach. In our theoretical attempt, we sacrificed the overall affinity to achieve metal-ion specificity and had no restraints on the minimum number of "good chelators" in the binding site. Still, the question arose as to whether we would be able to link the isolated "ideal combination" of metal-binding side chains into a single polypeptide chain. Existing fragments from the PDB were used as elementary building blocks spanning the neighbouring positions in the coordination polyhedra. These fragments certainly may not be the global minima of the corresponding short peptide sequences; nevertheless, the energy penalty of their folding into the conformation that they adopt in proteins should not be prohibitively large. Quite frequently, they represent stable parts of the protein structure, such as helices or β turns.

By merging the fragments capable of spanning two positions in the coordination polyhedron (repeatedly, if the coordination number is four), it may be possible to provide a single peptide that is potentially capable of folding around the metal ion to provide a set of optimum amino acid side chains at the vertices of a given coordination polyhedron. This is an original procedure for the design of novel metal-binding sites. In principle, it may not be limited to short peptides, although problems can be anticipated in the case of longer sequences (which are likely not to fold into the desired geometry). This approach is currently being used to design a short peptide sequence mimicking the active site of purple acid phosphatase.^[54]

The affinity of the resulting peptide sequences in solution is slightly lower than expected and definitely less than the Cys-rich IGA 16 peptide (four Cys residues) reported to bind zinc(II) with femtomolar affinity,^[30] our best binding constants are in the micromolar range, which is the penalty for using less typical metal-binding residues. The affinity may be improved by lowering the loss in conformational entropy upon chelation by making the peptides cyclic.

Unfortunately, the low solubility of the HHTC peptide and its metal-ion complexes has hampered efforts to crystallise this peptide with the Cu^{2+} ion.

In this report, we wish to share the new procedure and computer program and to demonstrate that they can be used to design peptides binding selected metal ions, both in the gas phase (by using a set of original MALDI/TOF experiments) and in solution. One can think of many improvements that may be achieved through designing a future generation of peptide sequences with higher metal-binding affinities. In the first stage, that is, the quantum chemical calculations of the complexation energies of all of the possible combinations of metal-binding side chains,^[42,44,46] several approximations had to be invoked. It was not possible to screen all of the ≈ 10000 combinations within a reasonable

timeframe. Both methodological progress (for example, RI-DFT methods with modern functionals) and an enormous increase in computational power have made this task accomplishable. In our study, we have mostly used the gas-phase values of the important descriptors, like the complexation energy or strain energy of the peptide chain (ΔE_F). These can be improved by taking into account solvation models, at least the implicit formulations, such as PCM or COSMO. Moreover, the best candidates can be subjected to combined quantum mechanics/molecular mechanics (QM/MM) calculations to obtain an accurate structure with the complexed metal ion (possibly in a surrounding sphere of water molecules, by use of the explicit solvation model).

These improvements can further increase the potential of the described procedure and lead to a new generation of short and specific metal-binding sequences.

Conclusion

The molecular design of M^{2+} -specific peptides has been achieved by using the *build_peptide* program, the results from earlier quantum chemical calculations and the simple molecular mechanics optimisation of both folded and unfolded states of the designed peptides to estimate the strain energy of the peptide chain. This process yielded more than 40 candidates for specific metal-binding peptides. Experimental evaluation combining mass spectrometry, ITC, NMR spectroscopy, and chelatometry techniques revealed that the devised strategy is not a merely computational exercise but can result in novel selective metal chelators and can be considered as a promising tool in metal-binding peptide design.

Experimental Section

Computational details

Fragment library: The fragment library is a collection of the 3D structures of all $Z_i-(X)_n-Z_j$ sequences (protein fragments) from the experimentally determined protein structures as available in the PDB and satisfying the following criteria: 1) both Z_i and Z_j are residues with metal-binding side chains (Asn, Asp, Cys, Gln, Glu, His, Lys, Met, Ser, Thr or Tyr); 2) $D_{ij} < 6.5$ Å, when D_{ij} is the distance between the metal-binding atoms of the first (Z_i) and last (Z_j) amino acid; 3) $0 \leq n \leq 14$. To avoid redundancy in the fragment library, the PDB was reduced to a non-redundant database, which contains only different chains (that is, we excluded structures that have 100% sequence identity with another structure of higher resolution). In total, 734 423 fragments were obtained, sorted into $11 \times 11 \times 15 = 1815$ “Zi-Zj_n.pdb” files (n is the number of amino acids in the $(X)_n$ chain).

The *build_peptide* program: The *build_peptide* program constructs the $Z_1-(X)_{n_1}-Z_2-(X)_{n_2}-Z_3-(X)_{n_3}-\dots-Z_j$ polypeptide chain by joining the $Z_i-(X)_{n_i}-Z_j$ fragments. The strain energy in the folded structure of the resulting peptide chain is then estimated in molecular mechanics calculations by using the Gaussian 98 program suite^[55] and the (standard AMBER) force field of Cornell et al.^[56] The partial charges for all residues were taken from the *all amino02.in* file in standard Amber libraries. The source code of the *build_peptide* program (written in Fortran 77) is available upon request; the main features and program flow are briefly described below. In Figure 6, part of the structure of the *build_peptide* input file is shown.

```
<Line1> TH
      {Comment: coordination geometry}
<Line2> GLU 1.98
<Line3> HIS 2.10
<Line4> CYS 2.33
<Line5> SER 2.05
<Line6> 0.033
      {Comment: all fragments satisfying  $(1-0.033)*d_{ij} <$ 
 $d_{actual} < (1+0.033)*d_{ij}$ , where  $d_{ij}$  is the ideal distance
between a pair of metal-binding atoms  $L^i$  and  $L^j$ , will be
read and processed}
<Line7> 0.050
      {parameter  $p_1$ , see next page}
<Line8> 0.080
      {parameter  $p_2$ }
<Line9> 22.0
      {parameter  $p_3$ }
<Line10> 0.12
      {parameter  $p_4$ }
<Line11> 12.0
      {parameter  $p_5$ }
```

Figure 6. An example of a *build_peptide* input file; the $Z_1-(X)_{n_1}-Z_2-(X)_{n_2}-Z_3-(X)_{n_3}-\dots-Z_4$ sites will be constructed, in which $\{Z_1, Z_2, Z_3, Z_4\}$ is any permutation of {Glu, His, Cys, Ser} with the $M-L_i$ optimum distances given. The parameters are explained below.

The program starts with reading of the input file and calculates all the distances between the k metal-binding atoms from the specified coordination geometry and metal-ligand distances.

It reads all the necessary fragments from the fragment library. The number of read fragments is determined by the allowed deviance (Line 6 of the example input file in Figure 6) between the actual and ideal (computed in the preceding step) distances of the corresponding metal-binding atoms. For linear coordination geometry, each read fragment is the final peptide sequence and the Gaussian 98 files are written (see below). For other coordination geometries, the program attempts to link the $(k-1)$ fragments (with k being the coordination number) into a single peptide chain for each permutation of coordinating residues.

For a given permutation, such as $\{Z_1, Z_2, Z_3, Z_4\}$, the program tries to superimpose fragment 1 (F1; $Z_1-(X)_{n_1}-Z_2$) and fragment 2 (F2; $Z_2-(X)_{n_2}-Z_3$) so as to achieve the maximum overlap of selected (important) atoms of the common residue Z_2 . More specifically, it first checks whether the conformations of both side chains and their orientation with respect to the peptide bond are within the limits specified by the parameters p_3 , p_4 and p_5 (see Figure 6). p_3 is the maximum tolerance for the difference between $N_{12}-C_{\alpha 12}-C_{\text{pept}12}-O_{\text{pept}12}$ and $N_{21}-C_{\alpha 21}-C_{\text{pept}21}-O_{\text{pept}21}$ dihedral angles (subscript 12 stands for the C-terminal amino acid of F1, whereas subscript 21 indicates the N-terminal amino acid of F2). p_4 is the maximum allowed difference between the $C_{\alpha 12}-L_{12}$ and $C_{\alpha 21}-L_{21}$ distances (with L being the metal-binding atom). p_5 is the maximum tolerance in the $L_{12}-C_{\alpha 12}-C_{\text{pept}12}$ versus $L_{21}-C_{\alpha 21}-C_{\text{pept}21}$ and the $L_{12}-C_{\alpha 12}-N_{\text{pept}12}$ versus $L_{21}-C_{\alpha 21}-N_{\text{pept}21}$ angles. Subsequently, the program translates both fragments, so that L_{12} and L_{21} are at the origin of Cartesian space. Afterwards, it makes the $L_{12}-C_{\alpha 12}$ and $L_{21}-C_{\alpha 21}$ bonds colinear and, finally, it makes the $L_{12}-C_{\alpha 12}-C_{\text{pept}12}$ and $L_{21}-C_{\alpha 21}-C_{\text{pept}21}$ planes coplanar.

The distance between the atoms L_{11} (the N-terminal metal-binding atom of F1) and L_{22} (the C-terminal metal-binding atom of F2) in the merged fragments is determined. It should be close to the ideal distance of the specific ion coordination geometry, as computed in the first step. If the distance is within the allowed deviation as determined by parameter p_1 (Line 7 in Figure 6), the program proceeds to the next step.

The new single fragment 1–2 (F12), that is, $Z_1-(X)_{n_1}-Z_2-(X)_{n_2}-Z_3$, is merged with fragment 3 (F3; $Z_3-(X)_{n_3}-Z_4$) through a procedure analogous to that described above. After these two fragments have been connected, two distances, $L_{11}-L_{32}$ and $L_{21}-L_{32}$, are examined. The tolerance is specified by the single parameter p_2 . For tetracoordinated species (tetrahedral, square-planar), this is the end of the main body of the program; for an octahedral geometry, the program continues to add fragments 4 and 5 in a similar way.

Finally, the resulting peptide sequence, for example, $Z_1-(X)_{n_1}-Z_2-(X)_{n_2}-Z_3-(X)_{n_3}-Z_4$, is written in the PDB format and four Gaussian 98/Amber input files are written, preceded by an addition of the hydrogen atoms to the final structures (this subroutine is a part of the program) and explicit specification of the connectivities.

A1A2A3A4_Tnnf.inp is the Gaussian 98/Amber input file for full optimisation by using the first-order optimiser in Gaussian 98 (used for pre-screening).

A1A2A3A4_Tnna.inp is the Gaussian 98/Amber input file for N steps (with N being the number of atoms) of partial optimisation in Cartesian coordinates with metal-binding atoms fixed in Cartesian space by using the steepest descent method.

A1A2A3A4_Tnnb.inp is the Gaussian 98/Amber input file for partial optimisation in redundant internal coordinates with the distances between all metal-binding atoms fixed by using the Berny optimisation algorithm and starting from the geometry "pre-optimised" in the preceding step.

A1A2A3A4_Tnnc.inp is the Gaussian 98/Amber input file for full optimisation by using the Berny optimisation algorithm and starting from the fully converged geometry of the previous partial optimisation. An example of each type of input file is included in the Supporting Information. Successful completion of the molecular mechanics optimisations yields an estimate of the value of the strain energy (ΔE_F), which is the most important descriptor of the accessibility of the folded state of the given peptide. The lower the value of ΔE_F , the better the peptide for specific metal binding of the particular transition-metal ion.

Peptide synthesis and purification: Peptides with selected sequences were synthesised by using a solid-phase synthesiser ABI433A (Applied Biosystems, Foster City, CA, USA) by stepwise coupling of the corresponding 9-fluorenylmethoxycarbonyl (Fmoc) amino acids to the growing chain on Rink amide resin (100–200 mesh, 0.69 mmol g⁻¹). The fully protected peptide was synthesised on the resin according to a standard procedure involving 1) cleavage of the N^α-Fmoc protecting group with 20% piperidine in *N*-methylpyrrolidine (NMP) and 2) a coupling mediated by mixtures of the coupling reagents *O*-(benzotriazol-1-yl)-*N,N,N',N'*-tetramethyluronium hexafluorophosphate (HBTU)/1-hydroxy-1*H*-benzotriazole (HOBT) in DMF. Upon completion of the syntheses, the deprotection and detachment of the linear peptides from the resin were carried out simultaneously by using a TFA/thioanisole/ethanethiol/anisole (90:5:3:2) cleaving mixture. The resin was washed with dichloromethane and the combined TFA filtrates were evaporated at room temperature. The peptides were precipitated with cold *tert*-butyl methyl ether, collected by suction, dissolved in acetonitrile/H₂O (60:40) and dried. The peptides were then purified by reversed-phase HPLC to >95% purity and the correct sequence were verified by using mass spectrometry.

MALDI-TOF MS experiments: The peptides obtained as described above were weighed (approximately 1 mg) and dissolved in 0.1% aq TFA. The stock solution was diluted by the same solvent to provide 1 pmol μL⁻¹ solutions. In the cases of not fully dissolved peptides in the stock solutions, a fine suspension obtained after sonication was used. The formed solution was clear and presumably all of the peptides were dissolved. The peptide solution (1.1 μL of 1 μM) was mixed with the metal sulphate solution (1.1 μL of 10 μM) in a 1 mL Eppendorf tube. After approximately 1–2 min of incubation, 3,4,5-trihydroxyacetophenone (THAP) matrix was added (1.1 μL) and, after thorough mixing, 1.1 μL of the solution were deposited on the metallic MALDI target. The identity of the peptides was examined in the same way by using filtered water (0.2 μM MilliQ filter) instead of sulphate solution. The dried spots were analysed by MALDI micro MX (Waters/Micromass, Manchester, UK) in automatic mode by using reflectron settings. The measured masses were calibrated by means of a bovine serum albumin trypsin digest; human Glu-1-fibrinopeptide B (*m/z*: 1570.6447) was used as an external lock mass.

Chelator experiments: Chelator experiments that monitor the competition between a peptide and a fluorescent chelator for Zn²⁺ ions were performed with solutions of 5 μM FluoZin-1 (Molecular Probes, Eugene, OR, USA) and a peptide at 40 or 80 μM in 20 mM MES (pH 6.5). Zn²⁺ ions were added in a stepwise manner from a zinc acetate stock solution in

the same buffer and the fluorescence intensity of FluoZin-1 was recorded by using a Perkin-Elmer LS 50 B fluorescence spectrometer connected to a Julabo F25 thermostatic water bath. All of the measurements were performed at 25°C, with excitation and emission wavelengths of 495 and 515 nm, respectively. Binding curves were fitted to the data (fluorescence intensity versus total Zn²⁺ concentration) by using CaLigator software.^[57] The p*K*_a value for the chelator–Zn²⁺ complex was determined to be 5.2; this value was used in fitting to titration data for chelator–peptide mixtures.

Isothermal microcalorimetry: The binding of metal ions (Cu²⁺, Zn²⁺ and Cd²⁺) to the peptide HNLGMNHDLQGERPYVTEGC (HHTC) was monitored by using a VP-ITC microcalorimeter (MicroCal, Northampton, USA) at 25°C. The solutions of the reactants were prepared in 20 mM ACES (or MES) at pH 7.0 and the exact concentrations of peptides were determined by an amino acid analysis. Typically, 9 μL aliquots of 650 μM CuCl₂ (or ZnCl₂) were injected stepwise into the sample cell containing 50 μM peptide (1.43 mL) until saturation was achieved. The solutions used for the titration of the cadmium ions were 1.25 mM CdSO₄ and 70 μM HHTC peptide. The experiment was accompanied by a control experiment in which the salts of the metal ions were injected into the buffer alone. The thermodynamic parameters and association constants were estimated by fitting a 1:1 binding model to the data by using the MicroCal Origin software.

NMR spectroscopy: ¹H NMR spectra were measured on a Bruker AVANCE II instrument at 600.13 MHz with a triple-resonance cryoprobe (5 mm CPTCI ¹H–¹³C/¹⁵N/D Z-GRD) in a H₂O/D₂O (9:1) solution at the natural pH value of 3.34. Dioxane was used as the internal reference and the chemical shifts were recalculated by using δ(dioxane) = 3.75 ppm. For structural assignment of the hydrogen atoms, a series of 2D-H,H-COSY, 2D-H,H-TOCSY and 2D-H,H-NOESY spectra was collected.

Acknowledgements

We gratefully acknowledge financial support by the Ministry of Education, Youth and Sports of the Czech Republic (research project nos.: Z4055905, LC512, 1M0508, MSM6046137305 and 1M6837805002), the Max Planck Society and the Swedish Research Council, VR.

- [1] P. Kotrba, P. Pospíšil, V. de Lorenzo, T. Ruml, *J. Recept. Signal Transduction Res.* **1999**, *19*, 703–715.
- [2] P. Kotrba, L. Dolečková, V. de Lorenzo, T. Ruml, *Appl. Environ. Microbiol.* **1999**, *65*, 1092–1098.
- [3] C. Sousa, P. Kotrba, T. Ruml, A. Cebolla, V. de Lorenzo, *J. Bacteriol.* **1998**, *180*, 2280–2284.
- [4] C. Sousa, A. Cebolla, V. de Lorenzo, *Nat. Biotechnol.* **1996**, *14*, 1017–1020.
- [5] T. Macek, P. Kotrba, A. Svatos, M. Nováková, K. Demnerova, M. Mackova, *Trends Biotechnol.* **2008**, *26*, 146–152.
- [6] S. Brown, *Nat. Biotechnol.* **1997**, *15*, 269–272.
- [7] B. I. Dahiyat, S. L. Mayo, *Science* **1997**, *278*, 82–87.
- [8] J. Kostal, A. Mulchandani, W. Chen, *Macromolecules* **2001**, *34*, 2257–2261.
- [9] S. Ramanathan, M. Ensor, S. Daunert, *Trends Biotechnol.* **1997**, *15*, 500–506.
- [10] J. S. Marvin, H. W. Hellinga, *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 4955–4960.
- [11] I. Bontidean, A. Kumar, E. Csoregi, I. Y. Galaev, B. Mattiasson, *Angew. Chem.* **2001**, *113*, 2748–2750; *Angew. Chem. Int. Ed.* **2001**, *40*, 2676–2678.
- [12] B. R. Gibney, F. Rabanal, J. J. Skalicky, A. J. Wand, P. L. Dutton, *J. Am. Chem. Soc.* **1999**, *121*, 4952–4960.
- [13] X. Chen, B. M. Discher, D. L. Pilloud, B. R. Gibney, C. C. Moser, P. L. Dutton, *J. Phys. Chem. B* **2002**, *106*, 617–624.
- [14] *Handbook of Metal-Ligand Interactions in Biological Fluids*, Vol. 2 (Ed.: G. Berthoin), Marcel Dekker, New York, **1995**.

- [15] T. Hayashi, Y. Hisaeda, *Acc. Chem. Res.* **2002**, 35, 35–43.
- [16] D. W. Low, M. G. Hill, *J. Am. Chem. Soc.* **2000**, 122, 11039–11040.
- [17] S. M. Berry, M. D. Gieselman, M. J. Nilges, W. A. van der Donk, Y. Lu, *J. Am. Chem. Soc.* **2002**, 124, 2084–2085.
- [18] T. D. Pfister, A. Y. Mirarefi, A. J. Gengenbach, X. Zhao, C. Danstrom, N. Conatser, Y. G. Gao, H. Robinson, C. F. Zukoski, A. H. J. Wang, Y. Lu, *J. Biol. Inorg. Chem.* **2007**, 12, 126–137.
- [19] J. Kaplan, W. F. DeGrado, *Proc. Natl. Acad. Sci. USA* **2004**, 101, 11566–11570.
- [20] M. Allert, M. A. Dwyer, H. W. Hellinga, *J. Mol. Biol.* **2007**, 366, 945–953.
- [21] Y. Lu, J. S. Valentine, *Curr. Opin. Struct. Biol.* **1997**, 7, 495–500.
- [22] D. E. Benson, M. S. Wisz, H. W. Hellinga, *Curr. Opin. Biotechnol.* **1998**, 9, 370–376.
- [23] W. F. DeGrado, C. M. Summa, V. Pavone, F. Nastro, A. Lombardi, *Annu. Rev. Biochem.* **1999**, 68, 779–819.
- [24] D. Jantz, B. T. Amann, G. J. Gatto, Jr, J. M. Berg, *Chem. Rev.* **2004**, 104, 789–799.
- [25] C. J. Reedy, B. R. Gibney, *Chem. Rev.* **2004**, 104, 617–649.
- [26] S. S. Huang, B. R. Gibney, S. Stayrook, P. L. Dutton, M. Lewis, *J. Mol. Biol.* **2003**, 326, 1219–1225.
- [27] B. T. Farrer, N. P. Harris, K. E. Balchus, V. L. Pecoraro, *Biochemistry* **2001**, 40, 14696–14705.
- [28] B. Farrer, C. McClure, J. E. Penner-Hahn, V. L. Pecoraro, *Inorg. Chem.* **2000**, 39, 5422–5423.
- [29] M. Matzapetakis, B. T. Farrer, T.-C. Weng, L. Hemmingsen, J. E. Penner-Hahn, V. L. Pecoraro, *J. Am. Chem. Soc.* **2002**, 124, 8042–8054.
- [30] A. K. Petros, A. R. Reddi, M. L. Kennedy, A. G. Hyslop, B. R. Gibney, *Inorg. Chem.* **2006**, 45, 9941–9958.
- [31] M. Matzapetakis, D. Ghosh, T. C. Weng, J. E. Penner-Hahn, V. L. Pecoraro, *J. Biol. Inorg. Chem.* **2006**, 11, 876–890.
- [32] Y. Lu, *Curr. Opin. Chem. Biol.* **2005**, 9, 118–126.
- [33] D. E. Benson, M. S. Wisz, H. W. Hellinga, *Proc. Natl. Acad. Sci. USA* **2000**, 97, 6292–6297.
- [34] H. W. Hellinga, *J. Am. Chem. Soc.* **1998**, 120, 10055–10066.
- [35] H. W. Hellinga, *Nat. Struct. Biol.* **1998**, 5, 525–527.
- [36] K. H. Lee, M. Matzapetakis, S. Mitra, E. N. G. Marsh, V. L. Pecoraro, *J. Am. Chem. Soc.* **2004**, 126, 9178–9179.
- [37] J. P. Glusker, *Adv. Protein Chem.* **1991**, 42, 1–76.
- [38] a) R. G. Pearson, *J. Am. Chem. Soc.* **1963**, 85, 3533–3539; b) R. G. Parr, R. G. Pearson, *J. Am. Chem. Soc.* **1983**, 105, 7512–7516.
- [39] T. Dudev, C. Lim, *J. Am. Chem. Soc.* **2002**, 124, 6759–6766.
- [40] T. Dudev, C. Lim, *J. Phys. Chem. B* **2001**, 105, 10709–10714.
- [41] T. Dudev, C. Lim, *Acc. Chem. Res.* **2007**, 40, 85–93.
- [42] L. Rulíšek, Z. Havlas, *J. Am. Chem. Soc.* **2000**, 122, 10428–10439.
- [43] L. Rulíšek, Z. Havlas, *J. Chem. Phys.* **2000**, 112, 149–157.
- [44] L. Rulíšek, Z. Havlas, *J. Phys. Chem. A* **2002**, 106, 3855–3866.
- [45] L. Rulíšek, Z. Havlas, *Int. J. Quantum Chem.* **2003**, 91, 504–510.
- [46] L. Rulíšek, Z. Havlas, *J. Phys. Chem. B* **2003**, 107, 2376–2385.
- [47] L. Rulíšek, J. Vondrášek, *J. Inorg. Biochem.* **1998**, 71, 115–127.
- [48] H. Yin, J. S. Slusky, B. W. Berger, R. S. Walter, G. Vilière, R. I. Litvinov, J. D. Lear, G. A. Caputo, J. S. Bennett, W. F. DeGrado, *Science* **2007**, 315, 1817–1822.
- [49] R. Feng, A. L. Castelhan, R. Billedeau, Y. Zhengyu, *J. Am. Soc. Mass Spectrom.* **1995**, 6, 1105–1111.
- [50] M. Dabrio, G. Van Vyncht, G. Bordin, A. R. Rodriguez, *Anal. Chim. Acta* **2001**, 435, 319–330.
- [51] P.-R. Sudhir, H.-F. Wu, Z.-C. Zhou, *Rapid Commun. Mass Spectrom.* **2005**, 19, 1517–1521.
- [52] B. Breuker, R. Knochenmuss, J. Zhang, A. Stortelder, R. Zenobi, *Int. J. Mass Spectrom.* **2003**, 226, 211–222.
- [53] H. Fukada, K. Takahashi, *Proteins* **1998**, 33, 159–166.
- [54] I. N. Jakab, Z. Jenei, B. Gyurcsik, T. Gajda, T. Körtvélyesi, A. Mikulová, L. Rulíšek, T. Raskó, A. Kiss in *Achievements in Coordination, Bioinorganic and Applied Inorganic Chemistry* (Eds.: M. Melník, J. Šima, M. Tatarko), Slovak Technical University Press, Bratislava, **2007**, pp. 80–90.
- [55] Gaussian 98, Revision A.9, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, V. G. Zakrzewski, J. A. Montgomery, Jr., R. E. Stratmann, J. C. Burant, S. Dapprich, J. M. Millam, A. D. Daniels, K. N. Kudin, M. C. Strain, O. Farkas, J. Tomasi, V. Barone, M. Cossi, R. Cammi, B. Mennucci, C. Pomelli, C. Adamo, S. Clifford, J. Ochterski, G. A. Petersson, P. Y. Ayala, Q. Cui, K. Morokuma, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. Cioslowski, J. V. Ortiz, A. G. Baboul, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. Gomperts, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, J. L. Andres, C. Gonzalez, M. Head-Gordon, E. S. Replogle, J. A. Pople, Gaussian, Inc., Pittsburgh PA, **1998**.
- [56] W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, Jr., D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, P. A. Kollman, *J. Am. Chem. Soc.* **1995**, 117, 5179–5197.
- [57] I. Andre, S. Linse, *Anal. Biochem.* **2002**, 305, 195–205.

Received: January 29, 2008
Published online: July 16, 2008