

Article

Unravelling the Structure of the Tetrahedral Metal-Binding Site in METP3 through an Experimental and Computational Approach

Salvatore La Gatta ^{1,†}, Linda Leone ^{1,†}, Ornella Maglio ^{1,2}, Maria De Fenza ¹, Flavia Natri ¹, Vincenzo Pavone ¹, Marco Chino ^{1,*} and Angela Lombardi ^{1,*}

¹ Department of Chemical Sciences, University of Napoli Federico II, Via Cintia, 80126 Napoli, Italy; salvatore.lagatta@unina.it (S.L.G.); linda.leone@unina.it (L.L.); ornella.maglio@unina.it (O.M.); maria.defenza@unina.it (M.D.F.); flavia.natri@unina.it (F.N.); vincenzo.pavone@unina.it (V.P.)

² Istituto di Biostrutture e Bioimmagini (IBB), National Research Council (CNR), Via Mezzocannone 16, 80134 Napoli, Italy

* Correspondence: marco.chino@unina.it (M.C.); alombard@unina.it (A.L.); Tel.: +39-081-674418 (M.C. & A.L.)

† These authors contributed equally to this work.

Citation: La Gatta, S.; Leone, L.; Maglio, O.; De Fenza, M.; Natri, F.; Pavone, V.; Chino, M.; Lombardi, A. Unravelling the Structure of the Tetrahedral Metal-Binding Site in METP3 through an Experimental and Computational Approach. *Molecules* **2021**, *26*, 5221. <https://doi.org/10.3390/molecules26175221>

Academic Editors: Antonio Rosato, Francesco Musiani and Claudia Andreini

Received: 19 July 2021

Accepted: 26 August 2021

Published: 28 August 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

Abstract: Understanding the structural determinants for metal ion coordination in metalloproteins is a fundamental issue for designing metal binding sites with predetermined geometry and activity. In order to achieve this, we report in this paper the design, synthesis and metal binding properties of METP3, a homodimer made up of a small peptide, which self assembles in the presence of tetrahedrally coordinating metal ions. METP3 was obtained through a redesign approach, starting from the previously developed METP molecule. The undecapeptide sequence of METP, which dimerizes to house a Cys₄ tetrahedral binding site, was redesigned in order to accommodate a Cys₂His₂ site. The binding properties of METP3 were determined toward different metal ions. Successful assembly of METP3 with Co(II), Zn(II) and Cd(II), in the expected 2:1 stoichiometry and tetrahedral geometry was proven by UV-visible spectroscopy. CD measurements on both the free and metal-bound forms revealed that the metal coordination drives the peptide chain to fold into a turned conformation. Finally, NMR data of the Zn(II)-METP3 complex, together with a retrostructural analysis of the Cys-X-X-His motif in metalloproteins, allowed us to define the model structure. All the results establish the suitability of the short METP sequence for accommodating tetrahedral metal binding sites, regardless of the first coordination ligands.

Keywords: miniaturized proteins; tetrahedral metal binding site; metalloprotein models; spectroscopic characterization; NMR structure; bioinformatics

1. Introduction

Metalloproteins employ a limited number of metal ions to accomplish countless different functions, including electron transfer, signal transduction and enzymatic catalysis [1–3]. To achieve this incredible variety, a wide array of protein sites has been selected over billions of years of evolution, each purposed towards a specific task [4]. In doing this, Nature has taken advantage of the unique physicochemical properties of metal ions, such as their small volume, Lewis acidity, coordination geometry preferences, specific ligand affinity, variable oxidation and spin states, which can be finely modulated by the protein matrix [5]. In fact, proteins and metal cofactors are in a synergic relationship, which enables metalloproteins to be both folded and functional [6]. The selection of the first coordination sphere ligands usually drives the selectivity toward a specific metal ion in a certain oxidation and spin state, thus controlling its reactivity [5]. Concurrently, the binding of a metal ion imposes structural constraints on the protein-derived ligands, which in turn

adopt proper conformations to satisfy the metal coordination requirements. In some cases, protein folding occurs upon metal ion coordination, while in others, tight metal binding is assured by the protein, which imparts its own structural preferences to the metal ion [3]. Zinc-fingers are excellent examples of the first phenomenon, since the assembly of a tetrahedral Zn(II) binding site is essential to stabilize the global fold of the protein [7]. Conversely, in the blue copper proteins, a rigid protein environment dictates the coordination geometry of the metal ion [8]. The trigonal His₂Cys coordination site observed in azurin is not optimal for either the Cu(II) or the Cu(I) oxidation states, and only minimal structural changes occur upon redox cycling. This “entatic” or “strained” state has the effect of enhancing electron transfer mediated by copper sites [9].

Owing to the limited collection of ligands available in protein scaffolds, the sole composition of the first coordination sphere could not account for the wide variability of functions observed among metalloproteins. Second-shell interactions strongly contribute to defining the specificity and reactivity of the metal-binding site, by orienting the first shell-ligands in the proper geometry and/or altering their electronic properties [10]. Several fundamental aspects of metal cofactor assembly and functions have been clarified by studying natural metalloproteins and their site-directed mutants [11,12]. In addition, numerous model systems have been developed by engineering metal-binding sites into artificial protein scaffolds [13–20]. They represent useful platforms for structure–activity relationship studies, allowing for a fast and easy screening of several mutants [21,22]. Furthermore, they are suitable candidates for biomedical, pharmaceutical or environmental applications [23–27]. Progress in this field has recently been accelerated by tremendous advances in computational methods and structural biology [28–32], and several mono-, di- or poly-nuclear metal binding sites have been successfully implemented into designed protein scaffolds [33–39].

One possible approach to metalloprotein design involves a miniaturization process. It starts with the analysis of a metalloprotein three-dimensional structure and aims to detect the lowest number of interactions needed to rebuild the metal binding site out of the whole protein [40]. The presence of geometric relationships between the elements of secondary structure that define the active site aids the design process, since the target structure can be generated by applying symmetry operations to a minimal fragment. We have successfully applied this approach to the development of miniaturized metalloproteins [41–47]. In particular, the retrostructural analysis carried out on rubredoxin (Rds) has allowed us to develop METP [48], made up of a β -hairpin undecapeptide with two properly spaced Cys residues, that self-assembles in the presence of a coordinating metal ion, to form a C₂ symmetric tetrathiolate binding site. The characterization by several spectroscopic techniques showed that METP retains the main features of the parent protein. It binds iron, closely resembling the rubredoxin active site spectroscopic features. Encouraged by these results, we focused our attention on the development of METP analogues. By systematically varying the number and chemical nature of the ligands, our work was aimed at verifying the feasibility of METP to serve as a general scaffold for engineering tetrahedral metal-binding sites. Even though tetrahedral binding sites have been implanted into de novo designed proteins [48–52], METP holds the advantage of being the smallest scaffold reported to date.

Here, we report the design, synthesis, spectroscopic and structural characterization of METP3, a new analogue conceived to reproduce a Cys₂His₂ metal binding site. Only small changes in the undecapeptide sequence of METP were needed for preserving the tetrahedral geometry of the metal site upon modification of the first-coordination sphere ligands. The correctness of the design was proven by solution spectroscopic characterization. A survey of zinc-containing metalloprotein structures in the PDB was also performed, enabling the dimeric structure of Zn(II)-METP3 to be deciphered, thus offering new insights for subsequent design.

2. Results

2.1. Design

The design of METP3 peptide was based on the previously characterized METP prototype [48]. To convert the Cys₄ (Figure 1a) into a Cys₂His₂ (Figure 1c) metal binding site, the Cys residue at position four in each METP chain was mutated to His (Figure 1b). Additional substitutions were necessary to complement the changes in the first coordination sphere. Indeed, Cys⁴His substitution resulted in clashes of His⁴ with Lys³ and Ala⁶ side chains (Figure 1b). To accommodate the larger His ligand into the site, Lys³ and Ala⁶ were replaced by Gly. These substitutions altered the pattern of the inter-chain hydrogen bond interactions with respect to the prototype model, where Lys^{3(3′)} of one peptide chain was hydrogen bonded to Glu^{10′(10)} of the other. Interestingly, Lys³Gly substitution provided the appropriate distance for a second-shell-hydrogen bond between His^{4(4′)} N and Glu^{10′(10)}, which could stabilize the metal binding geometry (Figure 1c). The hydrogen bonds between Glu^{10′(10)} and Asn^{7(7′)} residues are conserved as in the parent model. Moreover, Gly^{5(5′)} of METP was replaced with a Ser, whose side chain forms a hydrogen bond with the carbonyl backbone of itself and may also improve water solubility of the complex. This extended hydrogen bond network, together with hydrophobic Ile¹¹-Ile^{11′} interaction, is expected to drive the homodimerization of METP3. Furthermore, such interchain interactions are expected to stabilize only a single topology of the dimer. Symmetrical dimerization (via a C₂ axis) in the presence of a tetrahedrally coordinating metal ion may in principle give rise to Δ and Λ enantiomers. They are actually diastereomers because of the chirality of the amino acids [48]. As in the case of METP, the designed interactions may stabilize a Λ configuration, as found in natural Rds [48]. The amino acid sequences of the newly designed peptide METP3 and of the prototype METP are reported in Figure 1d.

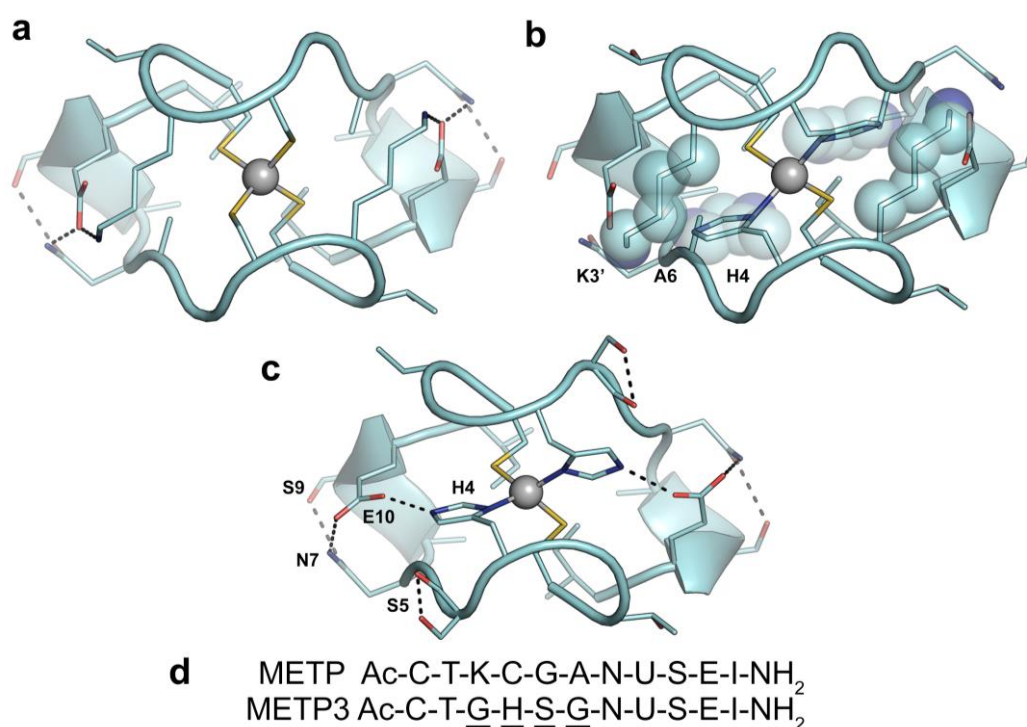


Figure 1. Redesign of the Cys₄ METP binding site into Cys₂His₂. (a) Energy minimized structure of Zn(II)-METP complex. The coordinating and hydrogen bonded residues are shown as sticks. (b) Cys₂His₂ Zn(II)-binding site housed in the unoptimized Zn(II)-METP model structure. The steric clashes of His⁴ with Lys³ and Ala⁶ side chains are evidenced as spheres. (c) Energy-minimized model of Zn(II)-METP3 complex. In all the panels, the Zn(II) ion is reported as a gray sphere. Peptide backbone is represented as a ribbon. (d) Comparison of the monomeric peptide sequence of METP and METP3. U represents 2-aminoisobutyric acid (Aib).

2.2. Synthesis

The monomeric MEPT3 peptide was obtained in good yield (90%) by solid-phase synthesis, using standard Fmoc/tBu protocols. The crude peptide was first reduced and purified by preparative RP-HPLC (purity > 95%). The homogeneity and identity of the final product was confirmed by analytical RP-HPLC and MALDI-TOF mass spectrometry (Figures S1 and S2).

2.3. Metal Binding Properties

METP3 metal binding properties were studied in the presence of different divalent metal ions, in particular cobalt, zinc and cadmium ions.

Cobalt was selected for initial binding studies, because Co(II) complexes show distinctive UV-visible spectra, which are sensitive to the metal coordination geometry and to the nature of the ligands [53]. To determine the optimal pH conditions for coordination, a UV-vis pH titration of METP3 in the presence of Co(II) ions was performed in the pH range 3.8–9.2 (Figure 2a). An increase in the absorption band intensities was observed upon raising the pH from 6 to 8.2. This indicates ligands deprotonation and concomitant coordination, with a midpoint transition at pH 7.1. Based on these results, all the spectroscopic studies were performed at pH 8.2.

Figure 2b shows the UV-visible absorption spectrum obtained upon addition of a 10% molar excess of CoCl₂ to the fully reduced peptide in aqueous solution.

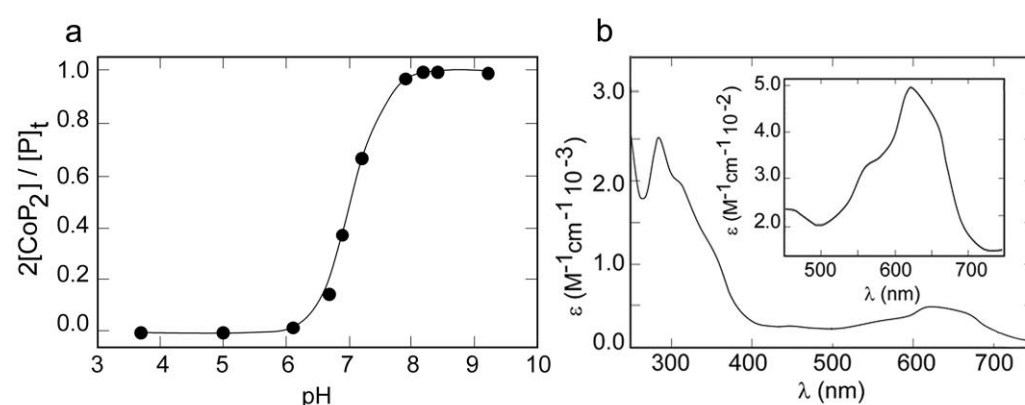


Figure 2. (a) pH titration of Co(II)-METP3 complex in the presence of 10% excess of Co(II). (b) UV-visible absorption spectrum of Co(II)-METP3 complex (10 mM TRIS buffer pH 8.2).

Table 1 compares the UV-visible spectral data of the Co(II) complexes of METP3 and METP [48]. The spectral data of Co(II)-GGG, a minimal peptide model of a zinc finger protein [54], and of Co(II)-CP-1, a 26 amino acid peptide based on the consensus sequence of 131 zinc finger domains [55], containing a Cys₂His₂ and a Cys₃His coordination site, respectively, are also reported for comparison.

Table 1. UV-visible spectral data of Co(II) complexes of METP3.

Compound	Wavelength [nm] ($\epsilon \times 10^{-3}$ [$M^{-1}cm^{-1}$])				Coordination Site
METP3	308 (2.10)	564 (0.313)	624 (0.470)	649 (0.440)	Cys ₂ His ₂
GGG	285 (3.4)	550 (0.220)	625 (0.500)	660 (0.430)	Cys ₂ His ₂
CP-1	310 (2.50)	570 (0.300)		640 (0.900)	Cys ₃ His
METP	353 (3.57)	624 (0.420)	685 (0.617)	735 (0.582)	Cys ₄

The spectrum of Co(II)-METP3 shows an intense absorption band in the near-ultra-violet at 308 nm, and an absorption envelope in the visible region (500–700 nm). The former band can be assigned to S→Co(II) ligand-to-metal charge-transfer (LMCT) transitions, indicative of cysteine coordination [49]. The observed molar absorptivity ($\epsilon_{308 \text{ nm}}$) value of $2100 \text{ M}^{-1}\text{cm}^{-1}$ indicates the presence of two thiolate ligands in the first coordination sphere, in agreement with the average value of $\sim 900\text{--}1200 \text{ M}^{-1}\text{cm}^{-1}$ per thiolate-cobalt bond, as reported in the literature [56].

The position and intensity of the *d-d* electronic transitions in the visible region give further insights into the Co(II) coordination environment. The absorption envelope displays a maximum at 624 nm ($\epsilon = 470 \text{ M}^{-1}\text{cm}^{-1}$) and shoulders at 649 nm ($\epsilon = 440 \text{ M}^{-1}\text{cm}^{-1}$) and at 564 nm ($\epsilon = 313 \text{ M}^{-1}\text{cm}^{-1}$). The intensities of these bands ($\epsilon \geq 300 \text{ M}^{-1}\text{cm}^{-1}$) are distinctive of a tetrahedral or distorted tetrahedral geometry [57] and rule out any higher coordination number. Indeed, lower molar absorptivity values are expected for both penta-coordinate and octahedral Co(II) sites [53,55,58]. The energy of *d-d* transitions is also sensitive to the number of thiolate ligands. Based on the position of imidazole and thiolate in the spectrochemical series, the tetrahedral ligand field splitting (Δ_t) increases almost linearly as thiolates are replaced by imidazoles [59], with a consequent shift of the *d-d* transition at higher energies. In agreement with these observations, a hypsochromic shift of all the absorption bands is observed when comparing the spectral data of Co(II)-METP3 with those of Co(II)-METP (Table 1). Moreover, the spectral features of Co(II)-METP3 are also consistent with those reported in the literature for Co(II)-Cys₂His₂ binding sites in proteins and model systems [54,59].

The stoichiometry of the Co(II)-METP3 complex formation and the dissociation constant (K_D) were determined by direct spectrophotometric titration. The saturation fraction $2[\text{CoP}_2]/[\text{P}]_t$ was plotted against the concentration ratio $[\text{Co}]_t/[\text{P}]_t$ (Figure 3a) and the experimental data were fitted to a binding isotherm (see experimental section, Equation (1)). The best fit gave a peptide:Co(II) ratio of 2:1 and a K_D of $(85 \pm 4) \mu\text{M}$.

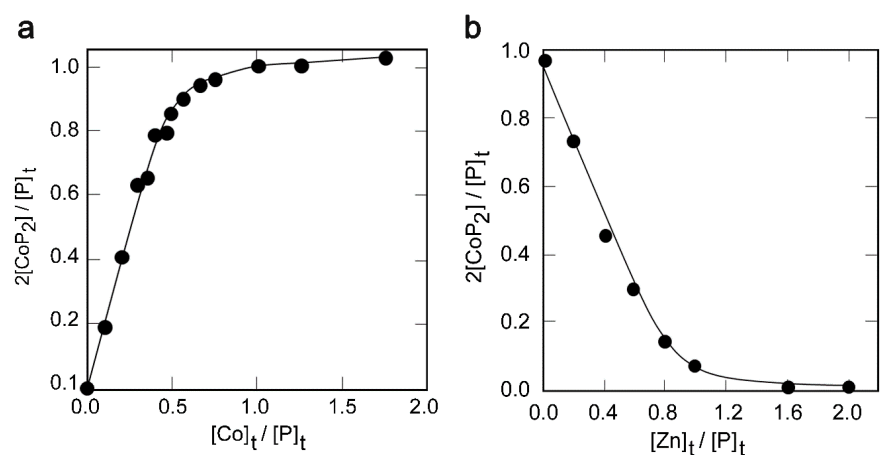


Figure 3. (a) Spectrophotometric titration of METP3 peptide with Co(II). The saturation fraction $2[\text{CoP}_2]/[\text{P}]_t$ is plotted against the concentration ratio $[\text{Co}]_t/[\text{P}]_t$. Solid line shows the best fit of data points to Equation (1). (b) Spectrophotometric competition experiment of Co(II)-METP3 complex with Zn(II). The saturation fraction $2[\text{CoP}_2]/[\text{P}]_t$ is plotted against the concentration ratio $[\text{Zn}]_t/[\text{P}]_t$. Solid line shows the best fit of data points to equation (2).

Zn(II) binding to METP3 was investigated by competition experiments, because Zn(II) complexes are spectroscopically silent. Zn(II) addition to a solution of the peptide containing 3.5 eq. of Co(II) led to a sharp decrease in the UV-visible bands of the cobalt complex (Figure S3), indicating that Zn(II) readily displaces Co(II). In this condition, a K_D value of $(5.6 \pm 0.2) \mu\text{M}$ was obtained for Zn(II) binding to METP3 (Figure 3b), proving that METP3 binds Zn(II) more tightly than Co(II).

Similar competition experiments were also performed with Cd(II), in the same experimental conditions used for Zn(II). Displacement of Co(II) with Cd(II) resulted in fading of the visible bands of the Co(II)-METP3 complex (Figure S4) and a K_D of (5.0 ± 1.0) μM was determined for the cadmium complex.

Analogously to METP [48], the observed lower affinity of METP3 peptide for Co(II), with respect to Zn(II) and Cd(II) (Table 2), is fully consistent with a tetrahedral coordination site [59]. In addition, the comparable affinity of METP3 toward Cd(II) and Zn(II) is expected because of the same metal d^{10} configuration.

Table 2. K_D values for METP and METP3 peptides toward different metal ions.

Metal ion	K_D (μM) METP ^a	K_D (μM) METP3
Co(II)	(53.5 ± 2.8)	(85 ± 4)
Zn(II)	(2.7 ± 0.1)	(5.6 ± 0.2)
Cd(II)	/	(5.0 ± 1.0)

^a Data for METP are taken from [48].

Coordination of metal ions to the METP3 binding site involves deprotonation of the two cysteine to thiolate. Therefore, undesired oxidation products could be formed upon metal complex dissociation, via disulfide bond formation. RP-HPLC analyses were performed to investigate the stability of METP3 metal complexes towards Cys oxidation. A reference sample of the intermolecular oxidized form, corresponding to the disulfide bridged dimer, was prepared by oxidation of the fully reduced pure peptide with DMSO (see experimental and Figure S5). The reduced monomeric peptide METP3 and the dimeric intermolecular oxidized form display different retention times (Figure S1 and S5), thus being easily distinguishable by HPLC analysis.

Interestingly, RP-HPLC analysis of a Zn(II)-METP3 solution, performed 4 days after the preparation, revealed the peptide chain to be in the almost fully reduced state (main peak at $R_t = 13.9$ min, Figure S6a), with only a small percentage of the intermolecular oxidized product ($R_t = 15.6$ min, Figure S6a). Conversely, the peptide was almost fully oxidized (87%) only 2 hours after the preparation of the Co(II)-METP3 complex (Figure S6b). These results highlight that the higher affinity of METP3 toward Zn(II) with respect to Co(II) correlates with a lower propensity of the Cys residues to form disulfide bridge.

2.4. CD Spectroscopy

The effects of the solvent and metal binding on METP3 folding were examined by CD spectroscopy in the far UV region (Figure 4). The CD spectrum of the fully reduced peptide (0.2 mM) in aqueous solution indicates an unfolded peptide [60], exhibiting a strong negative band at 198 nm and a small shoulder at 220 nm (Figure 4a). The coordination of Zn(II) causes a decrease in the intensity of the minimum at 198 nm, together with a slight increase in the minimum at 220 nm. CD spectra of apo- and Zn(II)-METP3 were also acquired in 2,2,2-trifluoroethanol (TFE), due to the stabilizing effect of this solvent on both α and β folds [61]. In the spectrum of apo-METP3 in TFE (Figure 4b), the lower wavelength minimum appears less intense and slightly red-shifted (200 nm) with respect to the spectrum recorded in aqueous solution (198 nm). A slight decrease in the intensity of the minimum at 200 nm, with a concomitant increase in the minimum at 220 nm is observed upon Zn(II) binding. The spectral changes observed upon metal binding indicate that the Zn(II) coordination induces structural stabilization, both in aqueous and TFE solutions [60]. However, the peptide chain is better folded in TFE, where it assumes a turned conformation even in the apo state.

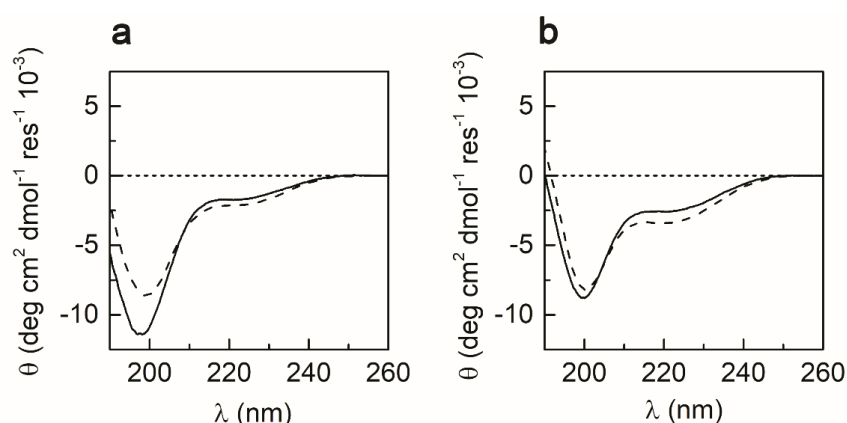


Figure 4. Far-UV CD spectra at 0.2 mM peptide concentration of fully reduced METP3 (solid lines) and Zn(II)-METP3 complex (dashed lines) in: (a) H₂O at pH 7.5 or (b) TFE.

2.5. NMR analysis

The solution structure of Zn(II)-METP3 was investigated by NMR spectroscopy. A complete characterization in aqueous solution was hampered by the rapid exchange of the amide protons with the solvent under the alkaline conditions required for metal binding. Different solvent systems were explored for the NMR analysis, including methanol, as used for the prototype METP [48]. Among them, TFE was selected as the best solvent, as ascertained by CD spectroscopy. The use of TFE, compared to methanol, allowed for better spectral resolution and subsequent peak assignment.

First, 1D ¹H NMR spectra were recorded to qualitatively ascertain the extent of structural organization of the molecule in solution, both in the presence and in the absence of Zn(II). Figure 5 shows the amide regions of ¹H NMR spectra of the fully reduced apo-METP3 and of the Zn(II)-METP3 complex in deuterated TFE (CF₃CD₂OH).

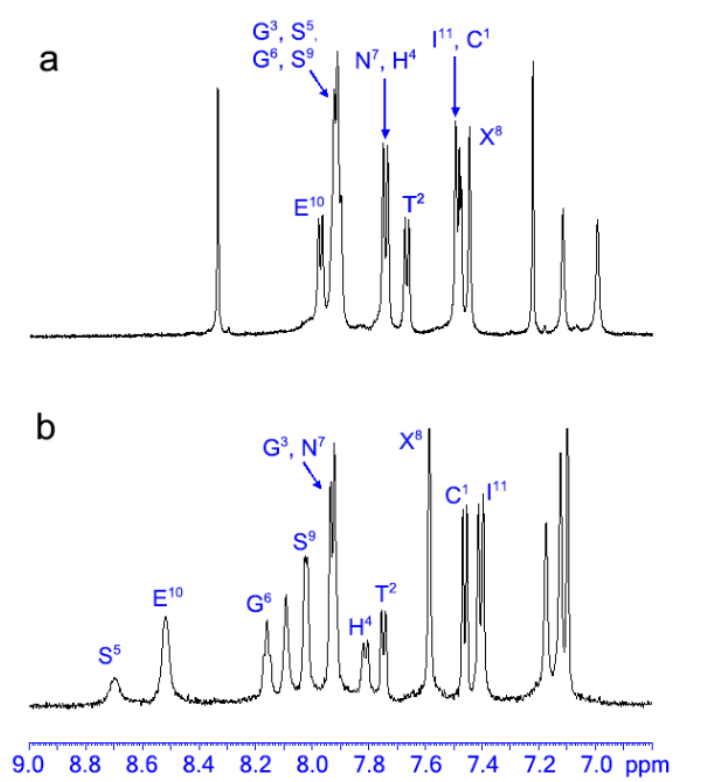


Figure 5. Amide proton region of the 1D ¹H NMR spectra of: (a) fully reduced METP3 (b) Zn(II)-METP3 complex. Both samples were prepared in CF₃CD₂OH.

In the absence of Zn(II), the ^1H NMR spectrum of the peptide shows poor peak dispersion and severe resonance overlap in the amide region, suggesting the absence of a stable conformation (Figure 5a). Upon addition of 1.1 equivalents of Zn(II) to the peptide, the amide proton resonances become well-dispersed, indicative of a well-defined structure (Figure 5b). Furthermore, a single set of resonances was observed for each residue, clearly indicating that the complex is a symmetric dimer.

Additionally, 2D NMR experiments (TOCSY and NOESY) were performed to obtain structural information. Examination of Nuclear Overhauser Effect (NOE) effects, temperature coefficient and $^3J_{\text{NH}-\alpha\text{CH}}$ coupling constant values allowed us to delineate METP3 secondary structure. A total of 10 restraints per residue were found, and only few NOEs were observed in the segment spanning residues 5–7. Figure 6 summarizes the sequential and medium-range NOE connectivities.

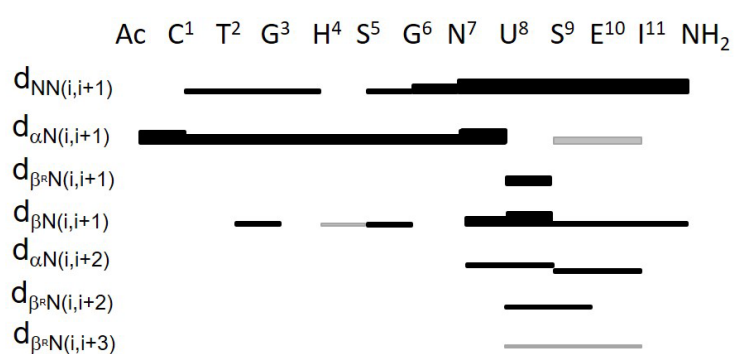


Figure 6. Summary of short- and medium-range NOE connectivities observed for Zn(II)-METP3 complex. The thickness of solid bars indicates the relative intensity of the NOEs, i.e. weak, medium and strong. Grey bars indicate spectral overlap.

The presence and the relative intensities of the $\text{NH}_i\text{-NH}_{i+1}$ NOE connectivities in the Asn⁷-Glu¹⁰ segment suggest an incipient 3_{10} -helix, characterized by a type III β -turn enclosing Aib⁸-Ser⁹ residues. The turned conformation is further supported by the presence of $\beta\text{CH}_i\text{-NH}_{i+1}$, $\beta\text{CH}_i\text{-NH}_{i+2}$ medium range contacts together with weaker NOE effects $\alpha\text{CH}_i\text{-NH}_{i+2}$. The presence of NOE contacts between Asn⁷- αCH and Ser⁹-NH, Ser⁹- αCH and Ile¹¹-NH, Aib⁸- $\beta\text{CH}_3^{\text{proR}}$ and Glu¹⁰-NH (see Figure 6), and the concomitant absence of $\alpha\text{CH}_i\text{-NH}_{i+4}$ ($d_{\alpha\text{N}(i,i+4)} \sim 5.5$ Å for 3_{10} helix) also support this conformation [62,63]. The $^3J_{\text{NH}-\alpha\text{CH}}$ coupling constant value of Ser⁹ (4.2 Hz) (Table S1) is very close to the theoretical value expected for a corner residue (position $i+2$) of a type III β -turn [63].

In the N-terminal region, the presence of $\text{NH}_i\text{-NH}_{i+1}$ connectivities suggests an α -turn conformation in the Cys¹-Ser⁵ segment. An extended conformation can be inferred in the Gly⁶-Asn⁷ segment, because the $^3J_{\text{NH}-\alpha\text{CH}}$ coupling constant values were found as 8.2 and 7.7 Hz for Gly⁶ and Asn⁷, respectively, and the ϕ solution of the Karplus equation [64] was -180° for both residues.

The conformational analysis carried out on this molecule is in striking agreement with the secondary structure elements of the designed METP3 model. The latter was thus used as a template to construct a plausible initial model of the monomeric peptide chain, that was subsequently subjected to restrained molecular dynamic (RMD) simulation, using the experimental NOE contacts [65]. Figure S7 shows the backbone superposition of an ensemble of 30 representative NMR conformers along the RMD trajectory. The root mean square deviation (RMSD), derived from the backbone atom superposition between each member of the ensemble and the average structure, is (1.5 ± 0.2) Å.

In all structures, the N-terminal region (Cys¹-His⁴) adopts an α -turned conformation, and the C-terminal segment can be described as a single 3_{10} helical turn. The RMSDs for the backbone atom superposition are (0.5 ± 0.1) Å and (0.4 ± 0.1) Å, for the N- and C-terminal segment, respectively (Figure S8). These two segments display a considerable

conformational similarity with the designed monomer. The RMSDs, obtained from the N- and C-segment backbone atom superposition of the average RMD simulation derived model and the designed model, are (0.8 ± 0.2) Å and (0.3 ± 0.2) Å, respectively.

The linker segment (Ser⁵-Gly⁶) displays a higher conformational freedom, preventing the N- and C-terminal regions from being uniquely oriented. Furthermore, only one connectivity can be assigned as inter-chain. In fact, a careful analysis of the NOESY spectrum showed the presence of a NOE involving the ϵ^1 and amide protons of His^{4(4')} (8.10/7.92 ppm) (Figure S9). Such contact can only be assigned as an inter-chain correlation, because these two histidine protons are far away from each other in the monomeric structure ($d > 6$ Å), but could be close in space, when the two monomers would be arranged in the homodimer. Additional inter-chain connectivities, needed to unequivocally define the correct positioning of the two monomers and to delineate the overall topology of the homodimer were not observed in the explored experimental conditions. Thus, a retrostructural analysis was performed for deciphering the structure of the Zn(II)-METP3 complex.

2.6. Retrostructural Analysis of Pseudo-Symmetric Dimers Containing CXXHX Motif

To help to solve the dimeric structure, compatible with the experimental NMR data, a search for pseudo-symmetric dimers, containing at least one CXXHX coordination environment, was carried out in the Protein Data Bank (PDB).

To create an unbiased data set, a detailed analysis of the segment sequence and conformation was performed. Similar structures were removed from the initial database (see experimental section), obtaining a set of 129 structures (out of ≈ 2100 protein structures deposited in the PDB that satisfy the initial filters). Furthermore, only those containing the Zn binding-site in a C-X₁-X₂-H-X₃ sequence with an α -turn conformation were manually selected (Table S2).

The Zn-sites were clustered according to their structural similarity, as measured by RMSD. This process resulted in four clusters and three outlier structures that could not be attributed to any clusters within the RMSD cutoff (Figure 7). An overall average RMSD of 0.5 Å and a maximum RMSD of 1.0 Å were observed, whereas the highest averaged distance, observed between any given centroid and the members of its cluster, was 0.3 Å (clusters 2 and 4).

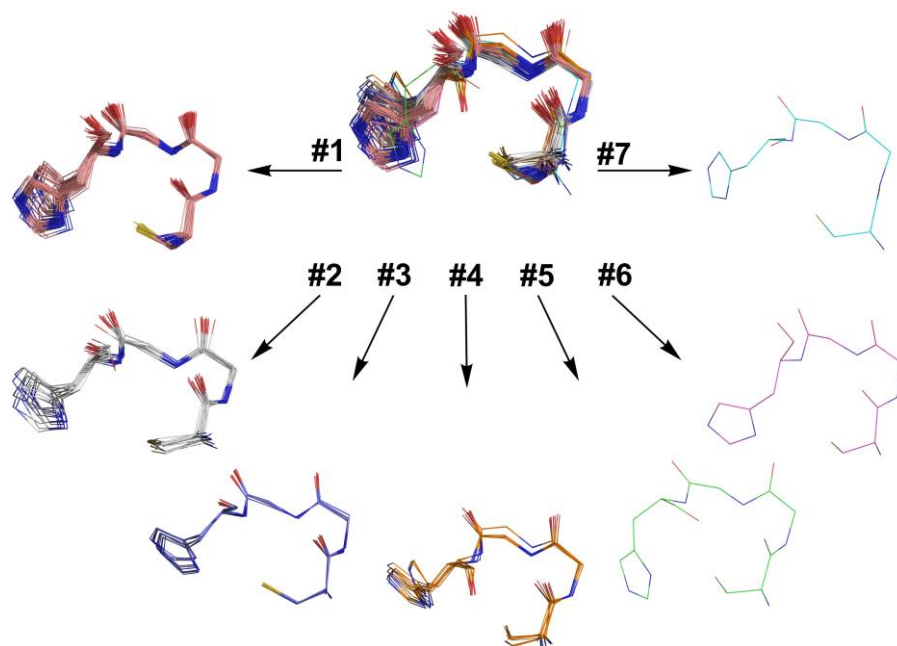


Figure 7. Superimposed C-X₁-X₂-H-X₃ motifs as obtained from the clusterization procedure.

The sequence motif did not appear to be related to a specific function, as more than 16 different functions have been assigned to the evaluated proteins (Figure S10). A first analysis of the primary structure of the non-redundant data set (Table S3) did not show prominent residue preferences for positions X_1 , X_2 and X_3 (Table 3 and Figure S11).

Table 3. Occurrences (first line) and propensities ^a (second line) for the proteogenic amino acids from the non-redundant -C- X_1 - X_2 -H- X_3 - structural database.

Residue	X_1	X_2	X_3
A	3 0.74	3 0.74	1 0.25
C	0 0	0 0	0 0
D	1 0.41	1 0.41	0 0
E	6 2.20	6 2.20	8 2.93
F	0 0	1 0.58	1 0.58
G	0 0	0 0	7 2.17
H	1 1.03	4 4.13	3 3.10
I	1 0.41	3 1.23	0 0
K	5 2.32	2 0.93	4 1.86
L	2 0.46	0 0	2 0.46
M	0 0	0 0	0 0
N	2 1.21	3 1.81	2 1.21
P	8 3.69	3 1.38	3 1.38
Q	4 2.41	0 0	0 0
R	6 2.34	3 1.17	1 0.39
S	0 0	3 1.02	2 0.68
T	3 1.23	3 1.23	3 1.23
V	1 0.33	4 1.31	3 0.98
W	1 1.75	1 1.75	0 0
Y	0 0	4 3.16	4 3.16

^a Propensities are calculated as reported in [66]

However, inspection of Table S3 shows that: (i) Pro residue is highly recurrent in position X_1 in clusters 2 and 3, whereas it is never found in this position in clusters 1 and 4 with only one exception; (ii) Gly residue is never found in position X_1 and X_2 , while it is distinctive of position X_3 in cluster 2; (iii) Thr is very frequent in position X_3 in cluster 1, as it is involved into a C-capping hydrogen bond.

A detailed analysis of the C- X_1 - X_2 -H- X_3 backbone torsion angles (ϕ and ψ), and of Cys and His χ_1 dihedral angles was performed for highlighting differences among the clusters (Figure 8 and Table 4).

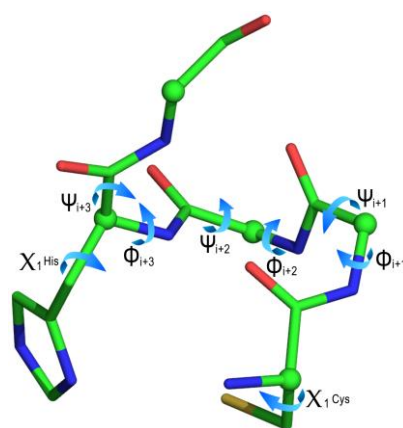


Figure 8. Prototypical C- X_1 - X_2 -H- X_3 motif, describing the ϕ , ψ and χ_1 dihedral angles.

Table 4. Average dihedral angles of clusters expressed in degrees. The number of structures per cluster is shown in the brackets of the first column. Standard deviations are reported in parentheses.

Cluster	ϕ_{i+1}	ψ_{i+1}	ϕ_{i+2}	ψ_{i+2}	ϕ_{i+3}	ψ_{i+3}	χ_1 Cys	χ_1 His
1 (86)	-59(5)	-37(8)	-65(7)	-47(12)	-79(18)	-21(19)	72(4)	-75(8)
2 (16)	-63(6)	-26(11)	-91(13)	-34(12)	-116(12)	10(32)	177(6)	-61(6)
3 (13)	-63(3)	-31(10)	-79(3)	-41(3)	-116(7)	-6(2)	67(3)	-53(4)
4 (11)	-62(8)	-40(11)	-83(9)	-33(13)	-119(12)	78(9)	-178(3)	-51(6)
5 (1)	-86	-3	-110	-24	-135	158	-175	93
6 (1)	-58	-48	-59	-54	-63	-45	-177	-81
7(1)	-121	-37	-57	-30	-115	18	178	-50

In particular, the analysis carried out on main chain torsion angles confirmed that the selected motif adopted an α -turn conformation (type I α RS) in all the clusters [67,68]. Few structures did not show the expected $i, i+4$ hydrogen bond (seven of them possessing a Pro residue in the X_3 position), as reported in Table S4. Moreover, the four clusters and the three outliers differ for the side chain conformation of the coordinating residues. The analysis of the Cys and His χ_1 angles allowed the main differences to be delineated among the clusters and to further group them into three subsets. In particular, the Cys and His χ_1 angles adopts the following conformations, respectively: (i) g^+ , g^- in the first subset, including clusters 1 and 3 (Cys $\chi_1 = 72^\circ \pm 5^\circ$ and His $\chi_1 = -72^\circ \pm 10^\circ$); (ii) *trans*, g^- in the second subset, comprising clusters 2 and 4 and outliers 6 and 7 (Cys $\chi_1 = 179^\circ \pm 6^\circ$ and His $\chi_1 = -58^\circ \pm 8^\circ$); (iii) *trans*, g^+ in the last, represented by outlier 5 (Cys $\chi_1 = -175^\circ$ and His $\chi_1 = 93^\circ$). In this perspective, clusters 3 and 4 can be considered as special cases of clusters 1 and 2, respectively. In these clusters, Cys residue adopts pre-Pro conformation and this finding is supported by the high occurrence of Pro residue at X_1 position.

Concerning zinc coordination, the histidine residue utilizes its $N^{\delta 1}$ atom to bind the zinc ion in all clusters. The binding site is characterized by $S\gamma$ -Zn and N^{δ} -Zn distances of

(2.34 ± 0.07) Å and (2.1 ± 0.1) Å, respectively, in agreement with the bond lengths reported in previous surveys of zinc coordinating proteins [69,70].

In order to gain information about the homodimer topology, the database was further analyzed, searching for the propensity in symmetric dimer formation among the clusters. Unfortunately, no Cys₂His₂-containing dimer was found among the analyzed structures, and only four entries, belonging to clusters 2, 4 and outlier 5 (Table 5), displayed almost symmetrical Cys₃His dimers, in which symmetry-related coordinating residues adopt the same χ_1 torsion angles.

Table 5. Pseudo-symmetrical Cys₃His dimers observed in the database.

Entry	#Clust.	Sequence	χ_1 Cys ^a	χ_1 His ^a	Sequence	χ_1 Cys ^a	χ_1 Cys ^a
2v0c_A	4	¹⁷⁶ C-W-R-H-E ¹⁸⁰	174	-59	¹⁵⁹ C-P-K-C-Q ¹⁶³	177	-65
4ijd_A	2	²¹⁶ C-A-A-H-G ²²⁰	174	-62	²⁰⁵ C-E-M-C-Q ²⁰⁹	179	-61
4ijd_B	2	²¹⁶ C-A-A-H-G ²²⁰	-176	-56	²⁰⁵ C-E-M-C-Q ²⁰⁹	177	-57
2au3_A	5	³⁵ C-P-F-H-P ³⁹	-175	93	⁵⁶ C-F-G-C-G ⁶⁰	-172	67

^a χ_1 dihedral angles are expressed in degrees.

Two considerations originate from this analysis: (i) in all dimers, also comprising nonsymmetrical ones, the configuration around the metal ion is Λ ; (ii) pseudo-symmetrical dimers were observed only when the χ_1 torsion angles of Cys and His belongs to the second (*trans*; *g*-) or the third (*trans*; *g*+) subset. The interpretation of the structural determinants, responsible of the observed behavior in natural dimers, is premature at this stage; it will be subject of further study once more data from a complete database, comprising all the possible dimer combinations and metal ions, will be available.

2.7. Dimer Selection and Model Validation

The information deduced from the structural database analysis was used to select and construct a suitable dimeric structural model for the Zn(II)-METP3 complex, which may account for the NMR data. To this end, the crystallographic coordinates of the C-X₁-X₂-H-X₃ motifs, extracted from 4ijd (belonging to cluster 2) and 2au3 (outlier 5) proteins, were used (Figure 9a). Applying the appropriate C₂ symmetry operation, two possible dimeric structures, housing the Cys₂His₂ binding motif, were built (Figure 9b).

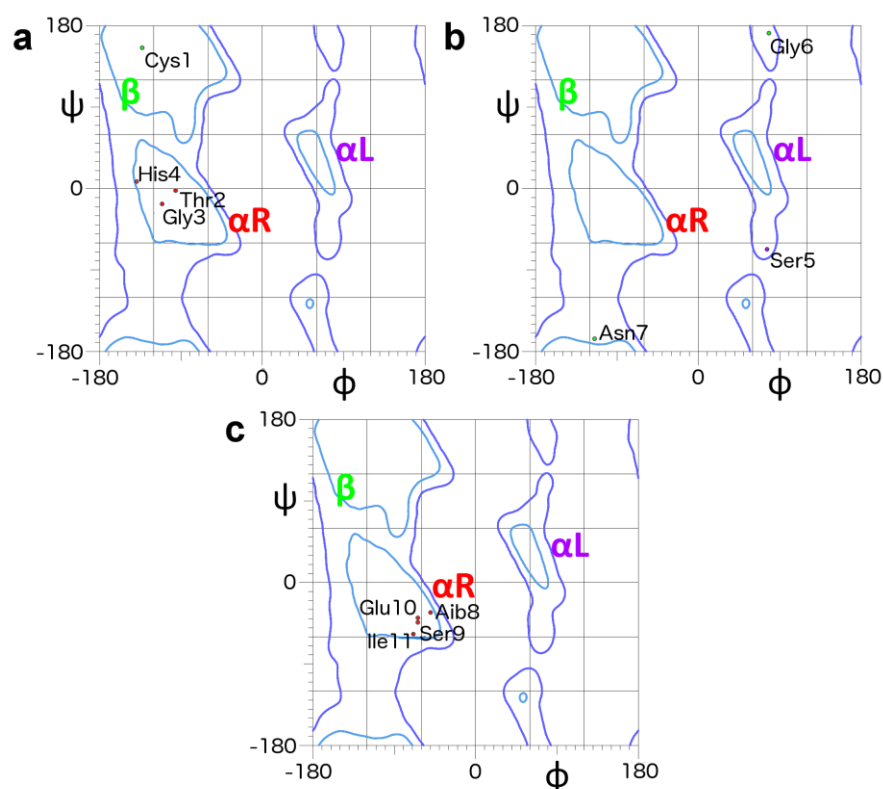


Figure 10. Ramachandran plots of the refined METP3 model. (a) Region from Cys¹ to His⁴ encompassing a type I α RS turn. (b) Ser⁵ to Asn⁷ in the extended region. (c) Terminal helical region from Aib⁸ to Ile¹¹. Color-coded labels define the most representative areas of the Ramachandran plot: right-handed α -helix/red, β -strand/green, left-handed α -helix/violet. ϕ , ψ values of chain A are reported. Molprobity [71] has been used to generate the plot (<http://molprobity.biochem.duke.edu/>).

The second region comprising the segment Ser⁵-Gly⁶, is characterized by dihedral angles typical of an extended conformation (Figure 10b). The values of dihedral angles for the residues belonging to this peptide segment are: ϕ_5 , $\psi_5 = 75^\circ$, -75° ; ϕ_6 , $\psi_6 = 81^\circ$, 179° for chain A and ϕ_5 , $\psi_5 = 76^\circ$, -68° ; ϕ_6 , $\psi_6 = 78^\circ$, 172° for chain B. However, the resulting arrangement of this regions could not be unambiguous, due to the low density of restraints.

The third region, corresponding to the C-terminal segment (Aib⁸-Ile¹¹), is characterized by high structural organization and can be described as a C-capping motif (β -helix) [72], which caps an incipient helix (Aib⁸-Ile¹¹) (Figure 9c blue box and Figure 10c). This structure is stabilized by an intramolecular hydrogen bond between the CO of Asn⁷ and the NH of Glu¹⁰, as well as the *i*, *i*+4 hydrogen bond between the CO of Asn⁷ and the NH of Ile¹¹. Further stabilization is given by a strong hydrogen bond interaction involving the side chain carbonyl of Asn⁷ and the amide of Ser⁹ (Figure 9c). The Glu¹⁰-His⁴ hydrogen bond, which was intended as an intermolecular interaction in the designed model, is retained as intramolecular in the refined model derived from the cluster analysis (Figure 9c). The full hydrogen bond network is reported in Table S6.

The overall fold of Zn(II)-METP3 is slightly looser with respect to the designed model, and the monomer poorly interacts with the symmetric counterpart. This finding is in agreement with the lack of strong intermolecular cross-peaks in the NOESY spectrum. Nevertheless, the single inter-chain NOE connectivity involving the ϵ^1 and amide protons of His^{4(4')} would be satisfied in the current refined model, where a distance of 4.0 Å is found between these two protons (Figure 9c). As a final remark, the overall looseness of the final model comes with no surprise. The TFE has a well-documented effect on protein folding when mixed with water, generally driving the increase in secondary structure content. However, its low dielectric constant (8.55) undoubtedly affects the entropically

driven hydrophobic collapse of the dimer when used as the sole solvent. In this respect, the role exerted by the metal coordination is crucial in the peptide dimerization, as intended in the design. A more compact structure is therefore expected in water, and future studies will address this aspect.

3. Discussion

The aim of this work was to assess whether a short peptide sequence could be used as scaffold for developing metal binding sites capable of binding different metal ions with identical coordination geometries. This system could represent a simple and versatile platform for studying metal-coordination properties, evaluating the physical properties of the bound ions and the selectivity of the metals toward different coordination environments.

METP3 was obtained by grafting a Cys₂His₂ metal binding site into the previously developed METP model, a small peptide containing a tetrahedral Cys₄ coordination site. Similarly to its precursor, METP3 is a homodimer consisting of two undecapeptide chains, self-assembling in the presence of divalent metal ions. Three substitutions (Lys³Ala/Gly⁵Ser/Ala⁶Gly) proved necessary to complement the changes in the first coordination sphere, resulting in a small, dimeric system, able to host a tetrahedral Cys₂His₂ metal binding site. Indeed, the spectroscopic characterization demonstrated that METP3 binds transition metal ions with the predicted 2:1 peptide/metal stoichiometry and the desired tetrahedral geometry. In particular, the UV-visible spectral features of the Co(II)-METP3 complex are distinctive of a tetrahedral Cys₂His₂ metal binding environment [56,73] (Figure 2 and Table 1). The K_D values (Table 2) for METP3 toward different metal ions further confirm a tetrahedral arrangement of the ligands around the metal site [57,58,74,75]. Indeed, Co(II) ions experience a loss of ligand field stabilization energy (LFSE) in the transition from the octahedral hexa-aquo complex Co(H₂O)₆²⁺ to the tetrahedral Co(II)-peptide complex, whereas *d*¹⁰ ions (as Zn(II) and Cd(II)) have no LFSE for any coordination geometry [59]. Zn(II) is more efficient than Co(II) in preventing cysteine oxidation, further evidencing a tight and specific Zn(II)-peptide binding.

Conformational analysis by CD and NMR spectroscopies indicates that peptide fold is driven by metal ion coordination. In this respect, METP3 behaves as a minimal model of Zn-finger proteins. The observed CD spectral pattern of Zn(II)-METP3 complex can be attributed to a turned conformation such as α -helical turns, ₃₁₀ helical turns, or type I (III) β -turns, similarly to previous findings for the parent Zn(II)-METP complex [48].

The presence of turned structures was further confirmed by NMR studies of the Zn(II) complex. Even though high-resolution structural information could not be obtained solely from NMR data, a considerable help in outlining the dimer topology came from the analysis of zinc-containing protein structures in the PDB. The examined database was limited to high-resolution crystal structures of zinc-proteins with a Cys_xHis_y ($x + y = 4$; $x, y \neq 0$) coordination environment in pseudo-symmetric dimers, where x and y refer to the number of Cys and His, respectively, located in the first coordination sphere. This analysis was not only helpful in solving the Zn(II)-METP3 structure, but also provided useful information on the molecular architecture of the CXXHX motif, both in terms of peptide conformation and zinc coordination geometry.

The CXXHX motif has been previously identified in several studies [76–78], as implicated in the formation of DNA binding domains (zinc fingers), or protein dimerization domains, although it has not been systematically studied in terms of its conformational preferences. Herein, by a search in the PDB database, we demonstrate that when Cys and His residues are both coordinated to zinc ion, the CXXHX motif predominantly adopts an α -turn conformation (type I α_{RS}). The selected Zn-sites were grouped into four clusters and three outlier structures (Figure 7), that mainly differ from each other for the side chain conformation of the coordinating residues. The position-dependent occurrence and propensity, at the X₁, X₂ and X₃ positions, for each amino acid in the C-X₁-X₂-H-X₃ α -turn motif, were also examined (Table 3). Gly is not allowed at the X₁ and X₂ positions, because they correspond to the *i*+1 position of type III and I β -turns composing the α -turn [79–81].

In contrast, Pro displays the highest propensity at the χ_1 position. Furthermore, the χ_3 position is optimal for C-cap interaction with the backbone carbonyl group of the Cys residue. These findings may provide useful information to guide future designs. Analysis of the structural database also provided insights into zinc coordination. All structures within our dataset exhibited a tetrahedral coordination around the metal ion, with the His binding through the N^δ atom. The preference of N^δ over N^ϵ histidine coordination is highly frequent among Zn-fingers, but the functional and structural significance of these two diverse binding modes remains not fully clarified [82,83].

Overall, the database analysis revealed that: (i) all the dimeric structures have a Λ configuration around the metal ion; (ii) no symmetric Cys₂His₂, but only Cys₃His sites are found in pseudo-symmetrical dimers (C_2 -symmetry-related α -turns both showing the same χ_1 conformation of coordinating residues); (iii) in all the symmetrical dimers, the Cys χ_1 is *trans* conformation, while either a g^+ or g^- conformation is adopted by the His χ_1 .

Interestingly, this analysis highlights that different χ_1 angles adopted by the His residue ($\chi_1 = 60.0^\circ$ and -60.0°) may be responsible for different interactions involving His $\epsilon 1$ proton, and only one side chain rotamer justifies the observed inter-monomer NOE contact found in the NMR data. Based on these results, we generated a Zn(II)-MEPT3 symmetrical homodimer, consistent with the NMR data, using the 2au3_A protein motif as template. The minimized model is characterized by three regions with different complexity and structural organization (Figure 10), and it deviates from the backbone conformation of the original designed model only in the His⁴-Ser⁶ segment.

In conclusion, the use of a combined approach of experimental data and statistical methods allowed to define the structural model of METP3, an artificial metalloprotein housing a Cys₂His₂ zinc binding site, unprecedented in nature.

4. Materials and Methods

All 9-fluorenylmethoxycarbonyl (Fmoc) protected amino acids, Rink Amide MBHA resin and coupling reagents: N-hydroxybenzotriazole (HOBt) and benzotriazole-1-oxyl-tris-pyrrolidino-phosphonium hexafluorophosphate (PyBOP) were purchased from NovaBiochem (EMD Biosciences, La Jolla, CA). All solvents used in the peptide synthesis and purification were anhydrous and High-Performance Liquid Chromatography (HPLC) grade, respectively, and were supplied by Romil. Piperidine, 3-(N-morpholino) propanesulfonic acid (MOPS), dimethylsulfoxide (DMSO) and scavengers (ethanedithiol, triisopropylsilane) were obtained from Sigma-Aldrich. N,N-Diisopropylethylamine (DIEA), trifluoroacetic acid (TFA) were supplied from Applied Biosystems. N,N-dimethylformamide (DMF), dichloromethane (DCM), pyridine, ethanol, methanol and 1-methyl-2-pyrrolidone (NMP) were supplied by Romil.

4.1. Instrumentations

Peptide synthesis was performed on an ABI 433 automatic peptide synthesizer using Fmoc solid phase peptide synthesis (SPPS). HPLC analysis were performed with a Shimadzu LC-10ADvp equipped with an SPDMS10Avp diode-array detector.

Mass spectra were acquired on a MALDI-TOF Voyager PerSeptive BioSystem.

UV-vis spectra were recorded on a Cary Varian 50 Spectrophotometer, equipped with a thermostated cell compartment (Varian, Palo Alto, CA, USA); CD analysis were performed using a J-815 spectropolarimeter equipped with a thermostated cell holder (JASCO, Easton, MD, USA).

NMR experiments were acquired at 298 K on a Varian Inova spectrometer, operating at 500 MHz.

4.2. Peptide Synthesis and Purification

The undecapeptide Ac-Cys-Thr-Gly-His-Ser-Gly-Asn-Aib-Ser-Glu-Ile-NH₂ was synthesized by standard Fmoc/tert-butyl (tBu) strategy, on a Rink Amide MBHA resin (0.65

mmol/g resin substitution). The synthetic procedure can be summarized as follows: The α -Fmoc group, was removed at the beginning of every cycle with a 20% piperidine solution in NMP. After deprotection, the resin was washed with NMP to remove the piperidine. The peptide resin was then ready for coupling; single coupling was conducted for each amino acid. Acylation reactions were carried out using a 0.5 M HOBt solution in DMF. In the coupling step, the activated Fmoc-amino acid reacts with the amino-terminal group of the growing peptide chain to form a peptide bond. Capping reaction was performed after each coupling step, using Ac₂O/HOBt/DIEA solution in NMP. Deprotection, coupling and capping steps were repeated with each subsequent amino acid, until the chain assembly was completed. When the coupling was complete, the resin was washed with NMP. Peptides N-terminal amino groups were acetylated with Ac₂O/HOBt/DIEA solution in NMP. Cleavage from the resin and sidechain deprotection was achieved with ethanedithiol/H₂O/trisopropylsilane/TFA 0.25:0.25:0.1:9.5 (*v/v/v/v*) for 2 h. Extraction of the scavengers and precipitation of the crude product was achieved by addition of cold diethyl ether. The peptide was reduced by 3 h incubation in aqueous solution at pH 8, with 10-fold excess DTT. Purification by preparative RP-HPLC gave the desired pure product, as assayed by analytical HPLC (Vydac C8 column, using a gradient of acetonitrile in 0.1% aqueous TFA, 50% to 90% over 20 min, at a flow rate of 1.0 mL/min) and matrix-assisted laser desorption ionization time-of-flight spectrometry [molecular weight: calculated 1130.4 atomic mass units (amu); obs 1131 ± 1 amu] (Figures S1 and S2).

The inter-molecularly oxidized peptide was prepared by dissolving the reduced peptide (1 mM peptide concentration) in a MOPS/DMSO (90:10 *v/v*) solution at pH 6.5. The formation of the oxidized peptide was confirmed by RP-HPLC analysis (Figure S5).

4.3. Metal-Binding Experiments

All manipulations were performed under strictly anaerobic conditions. Both peptide and metal stock solutions were freshly prepared.

The Co(II) and Zn(II) complexes were prepared by addition of CoCl₂ or ZnCl₂, in 10% molar excess, to an aqueous solution of the reduced METP3 monomer at pH ≈ 4 (1 mM peptide concentration); the pH was then raised to 8.2 by addition of TRIS buffer (10 mM final buffer concentration).

The pH dependence of the Co(II)-METP3 complex formation was determined by titrating an aqueous solution of the reduced peptide at 1.0 mM concentration, and of CoCl₂ in 10% molar excess, with phosphate buffer, for the pH range 3.7–6.9 and with TRIS buffer for the pH range 7.2–9.2.

The binding affinity of Co(II) to METP3 peptide was measured by spectrophotometric titration of CoCl₂ into the reduced peptide at 8.0 × 10^{−4} M concentration (10 mM Tris buffer, pH 8.2), under anaerobic conditions. The observed absorbance intensities were fit (KALEIDAGRAPH, Synergy Software (Reading, PA); see Supplementary Material) to a binding isotherm [59,84] to estimate the dissociation constant *K_D* Co and the stoichiometry of the interaction *n*, according to the reaction:

$$1/n\text{CoP}_n = \text{P} + 1/n\text{Co} \quad K_{\text{D}}^{\text{Co}} = [\text{Co}]^{1/n}[\text{P}]_f / [\text{CoP}_n]^{1/n} \quad (1)$$

where [Co]_f, [P]_f, [CoP_n] are the concentrations of free cobalt, free peptide (in monomeric units) and the complex, respectively.

The affinity of Zn(II) was estimated by competition experiments on the Co(II) complex. Aliquots of ZnSO₄ aqueous solution (1.9 × 10^{−2} M) were titrated into a solution containing the reduced peptide at 1.0 × 10^{−3} M concentration, and 3.5-fold molar excess of Co(II) (10 mM Tris buffer, pH 8.2), under anaerobic conditions. The bleaching of the absorption spectrum was monitored.

The relative affinity of the interaction was obtained by fitting the data to the equilibrium [59]:



where $[ZnP_2]$, $[CoP_2]$, $[Co]_f$, $[Zn]_f$ are the concentrations of Zn(II) complex, Co(II) complex, free cobalt and free zinc ions, respectively. K_D^{Zn} was determined from K_D^{Co} and K_{ex} : $K_D^{Zn} = K_D^{Co}/(K_{ex})^{1/2}$.

A similar experiment was performed in order to determine the affinity of Cd(II). Aliquots of $CdCl_2$ aqueous solution (1.9×10^{-2} M) were titrated into a solution containing the reduced peptide at 1.0×10^{-3} M concentration, and 3.5-fold molar excess of Co(II) (10 mM Tris buffer, pH 8.2), under anaerobic conditions. The bleaching of the absorption spectrum was monitored. Similarly, to the experiment with Zn(II), the relative affinity of the interaction was obtained by fitting the data to the equilibrium:



where $[CdP_2]$, $[CoP_2]$, $[Co]_f$, $[Cd]_f$ are the concentrations of Cd(II) complex, Co(II) complex, free cobalt and free cadmium ions, respectively.

4.4. UV-Visible (UV-Vis) Spectroscopy

UV-vis spectra were recorded on a Cary Varian 50 Spectrophotometer, equipped with a thermostated cell compartment (Varian, Palo Alto, CA, USA) by using quartz cuvettes with a path length of 1.0 cm. Wavelength scans were performed at 25 °C from 200 to 800 nm, with a $600 \text{ nm} \cdot \text{min}^{-1}$ scan speed. Anaerobic spectra were recorded in rubber-sealed cuvettes under nitrogen.

4.5. Circular Dichroism (CD) Spectroscopy

Data were collected at 25 °C, from 260 to 190 nm at 0.2 nm intervals with a $20 \text{ nm} \cdot \text{min}^{-1}$ scan speed, at 2 nm band width and at 16 s response, and a 2 s response from 250 to 800 nm. Cells of 0.01 cm and 1 cm path length were used for the far UV and near UV-vis regions, respectively. Peptide concentrations were in the range 0.20–1.2 mM. CD intensities in the far UV and near UV-visible regions are expressed as mean residue ellipticities ($\text{deg cm}^2 \text{ dmol}^{-1} \text{ res}^{-1}$) and total molar ellipticities, respectively.

4.6. NMR Spectroscopy

NMR sample was prepared directly into an NMR tube by addition of NaOD (D_2O solution, 1.5 equivalents) to a CF_3CD_2OH (TFE) solution of reduced METP3 monomer (2.0 mM) and $ZnSO_4$ (aqueous solution, 1.1 equivalents). 3-(trimethylsilyl) propionic-2,2,3,3- D_4 acid sodium salt (TSP) was added as internal reference. The NMR tube was flushed with nitrogen and sealed.

Two-dimensional (2D) experiments, such as total correlation spectroscopy, nuclear Overhauser effect spectroscopy (NOESY) and double quantum filtered correlation spectroscopy (DQFCOSY) were recorded by the phase-sensitive States-Haberhorn method [85]. Total correlation spectroscopy experiments were acquired with a 70 ms mixing time, whereas NOESY experiments were acquired with 100, 200, and 300 ms mixing times.

According to Wüthrich [63], identification of amino acid spin systems was performed by comparison of TOCSY and DQFCOSY, while sequential assignment was obtained by the analysis of NOESY spectra. NOE intensities were evaluated by integration of cross-peaks in the 300 ms NOESY spectrum.

All two-dimensional data consisted of 4K data points in the direct dimension. There were 512 experiments recorded in the indirect dimension, using 64 scans per experiments. Raw data were multiplied in both dimensions using a Gaussian function. A polynomial base-line correction was applied in both dimensions.

Semi-quantitative information on inter-proton distances for structure calculation were obtained by analyzing the 300 ms NOESY spectrum that did not exhibit spin diffusion effects. $^3J_{NH-\alpha CH}$ coupling constant values were obtained from 1D and 2D spectra, and the related ϕ dihedral angles were computed according to Karplus [64].

The proton chemical shifts, coupling constants and temperature coefficients are reported in Table S1. The NOESY experiments yielded about 80 NOE contacts in positive regime.

4.7. Analysis of the Protein Structure Database

Analysis of conformational preferences of the Cys- X_1 - X_2 -His- X_3 sequence (X_1 , X_2 , X_3 = any amino acid) was performed by generating a protein structure database from the Protein Data Bank (September 2020, <http://www.pdb.org>), containing zinc-binding proteins sharing the following sequence pattern to prevent His-tag occurrence: XXXXXCXXHXXXXXX.

The database started with approximately 350 protein structures solved by X-ray crystallography, at a resolution of 2.4 Å or better, and with a sequence identity cut-off up to 50%. In the case of proteins having multiple chains with the same sequence, only one chain was considered.

As the first step of our database structural analysis, we found all the C- X_1 - X_2 -H- X_3 motifs, in which both Cys and His residues are bound to zinc, and we restricted our analysis to those adopting an α -turn conformation. Thus, we selected, from the PDB, only the motifs showing Ca_i - Ca_{i+4} distance ≤ 7 Å. We grouped together the Zn-sites, thus generating a structural database, which comprised 129 structures, 44 of which display a unique C- X_1 - X_2 -H- X_3 sequence (Table S2 and Table S3).

All structures were manually inspected to assess zinc coordinating ligands, metal-to-ligand bond lengths, and coordination geometries (see Table S4). The GROMOS clustering algorithm [86], implemented in *g_cluster*, was used to determine cluster memberships with a 0.04 nm RMSD as calculated on an 18 atoms subset as showed in Figure S12. The structure with the greatest number of neighbors was used as the center of the cluster.

4.8. METP3 Model Refinement

The program package InsightII/Discover with the Covalent Valence Force Field (CVFF) [87,88], was used for energy minimization and RMD simulation. The designed monomer model was used as an initial model; it was examined for inconsistencies with the experimental NMR data (interproton contacts and $^3J_{NH-\alpha CH}$ coupling constants), and then subjected to RMD calculations, in vacuo at 300 K. Three classes of distance restraints were adopted according to the relative intensity of the integrated NOE signals: 1.8 ± 2.5 (strong), 1.8 ± 3.3 (medium), 2.7 ± 5.0 (weak). RMD calculations were run replacing the Zn^{2+} ion with Na^+ , due to the lack of an implemented force field for Zn^{2+} in the Discover package. The coordination geometry was simulated as follows: (a) the position of Na^+ ion, Cys-S and His- $N^{\delta 1}$ atoms were fixed; (b) distance restraints were introduced between the $N^{\delta 1}$, S atoms and Na^+ ion by using a force constant stronger than that used for the interproton distances accounting for the NOE effects. The distances were allowed to deviate in a range of ± 0.2 Å. Reference distances were selected from the X-ray structures of Zn^{2+} coordinating proteins.

A distance dependent dielectric constant was used through all computations. Leapfrog integration algorithm was adopted, with a time step of 1 fs. The protocol consisted of an equilibration phase of 50 ps and of a simulation phase of 360 ps. A structure was saved every 100 fs during the simulation for analysis. The final average structure was checked for consistency with all observable NOE.

Final dimeric model was constructed as described in Section 2.7, under Accelrys Discovery Studio 3.0 environment. Briefly, 2au3_A CXXHX motif was used as template for Zn(II) complexation. A C_2 operation was performed to generate the symmetric tetrahedral site, and peptide sequence was mutated to the METP3 N-terminal sequence (from residue 1 to 5). Then, the C-terminal traits were grafted onto the freshly obtained dimer, by superimposing the average coordinates of the C-terminal turn (Gly⁶-Ile¹¹) from RMD simulation. Only minimal manual manipulation was needed to properly link the N- and C-terminal structural units. Then, a NOE restrained minimization was performed under

CHARMM27 forcefield parameterization, with a Newton–Raphson algorithm under an implicit Generalized Born solvent model [89]. Finally, an unrestrained minimization was performed and found only negligible deviation from the NOE-distance restrained model.

Supplementary Materials: The following are available online, Figure S1: RP-HPLC chromatogram at $\lambda = 210$ nm of pure METP3, Figure S2: MALDI-TOF spectrum of pure METP3, Figure S3: Visible spectra of Co(II)-METP3 complex in the absence and in the presence of 10% excess of Zn(II), pH=8.2, Figure S4: Spectrophotometric competition experiment of Co(II)-METP3 complex with Cd(II). The saturation fraction $2[\text{CdP}_2]/[\text{P}]t$ is plotted against the concentration ratio $[\text{Cd}]/[\text{P}]t$. Solid line shows the best fit of data points to equation (3), Figure S5: RP-HPLC chromatogram at $\lambda = 210$ nm of inter-molecularly oxidized METP3, Figure S6: RP-HPLC chromatograms at $\lambda = 210$ nm of METP3 metal complexes. (a) Zn(II)-METP3 complex 4 days after the preparation; (b) Co(II)-METP3 complex 2 hours after the preparation, Figure S7: Monomeric peptide chain of Zn(II)-METP3 complex: backbone atom superposition of an ensemble of 30 representative NMR conformers along the RMD trajectory, Figure S8: Monomeric peptide chain of Zn(II)-METP3 complex: backbone atom superposition of the N-terminal (a) and C-terminal (b) segments in the ensemble of 30 representative NMR conformers along the RMD trajectory, Figure S9: NOESY contour plot of the amide proton region of Zn(II)METP3 complex (mixing time 300 ms), Figure S10: Pie chart showing the functions of proteins making up the C-X₁-X₂-H-X₃ structural database, Figure S11: Web logo analysis of the C-X₁-X₂-H-X₃ motifs for: (a) the full dataset; (b) cluster 1; (c) cluster 2; (d) cluster 3; (e) cluster 4, Figure S12: Atom set used to superimpose the structures, Table S1: Proton Chemical Shifts, ³J Coupling Constants and Temperature Coefficients of Zn(II)-METP3 in TFE, at 298 K, Table S2: Protein structures containing the C-X₁-X₂-H-X₃ motif (X₁, X₂, X₃= any amino acid). The corresponding sequences are indicated as single Letter Codes. Protein function is also indicated, Table S3: Non-redundant C-X₁-X₂-H-X₃ motif and their classification (X₁, X₂, X₃ = any amino acid), Table S4: Database entries: conformational and metal coordination parameters for each C-X₁-X₂-H-X₃ motif (X₁, X₂, X₃ = any amino acid), Table S5: Torsion angles of Zn(II)-METP3 complex as obtained from REM (first row Chain A, second row Chain B) calculations *in vacuo* at 300 K, Table S6: Intramolecular hydrogen bonds of Zn(II)-METP3 complex as obtained from REM calculations *in vacuo* at 300K.

Author Contributions: Conceptualization, V.P., M.C. and A.L.; methodology, O.M., F.N. and M.C.; validation, O.M., F.N., M.C. and A.L.; formal analysis, O.M., F.N. and M.C.; investigation, S.L.G., L.L., M.D.F., O.M. and M.C.; writing—original draft preparation, S.L.G. and O.M.; writing—review and editing, L.L., F.N., M.C. and A.L.; Visualization, S.L.G. and L.L.; supervision, A.L.; project administration, O.M., F.N. and M.C.; funding acquisition, A.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by Programma Operativo Nazionale 2007-2013, PON02_00029.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available in Tables S1–S7 of the Supplementary Information. HPLC, UV-Vis, CD, and NMR data are available upon request. The examined protein structures are available in the Protein Data Bank (<https://www.rcsb.org>).

Institutional Review Board Statement: Not applicable.

Acknowledgments: The authors wish to thank Claudia Vicari and Fabio Pirro for their help in collecting preliminary data, and Monica Grasso for administrative support.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Crichton, R.R. *Biological Inorganic Chemistry: A New Introduction to Molecular Structure and Function*, 3rd ed.; Academic Press: London, UK, 2019.
2. Bertini, I.; Gray, H.B.; Stiefel, E.I.; Valentine, J.S. *Biological Inorganic Chemistry: Structure and Reactivity*; University Science Books: Sausalito, CA, USA, 2007.
3. Gomes, C.M.; Wittung-Stafshede, P. Protein Folding and Metal Ions: Mechanisms, Biology and Disease; CRC Press, 2016.
4. Andreini, C.; Bertini, I.; Cavallaro, G.; Holliday, G.L.; Thornton, J.M. Metal Ions in Biological Catalysis: From Enzyme Databases to General Principles. *J. Biol. Inorg. Chem.* **2008**, *13*, 1205–1218.

5. Maglio, O.; Natri, F.; Lombardi, A. Structural and Functional Aspects of Metal Binding Sites in Natural and Designed Metalloproteins. In: *Ionic Interactions in Natural and Synthetic Macromolecules*; Ciferri, A.; Perico, A. Eds.; John Wiley & Sons, Inc.: Hoboken, NJ, USA, **2012**; pp. 361–450.
6. Natri, F.; D'Alonzo, D.; Leone, L.; Zambrano, G.; Pavone, V.; Lombardi, A. Engineering Metalloprotein Functions in Designed and Native Scaffolds. *Trends Biochem. Sci.* **2019**, *44*, 1022–1040.
7. Kluska, K.; Adamczyk, J.; Krężel, A. Metal Binding Properties, Stability and Reactivity of Zinc Fingers. *Coord. Chem. Rev.* **2018**, *367*, 18–64.
8. Pérez-Henarejos, S.A.; Alcaraz, L.A.; Donaire, A. Blue Copper Proteins: A Rigid Machine for Efficient Electron Transfer, a Flexible Device for Metal Uptake. *Arch. Biochem. Biophys.* **2015**, *584*, 134–148.
9. Cascella, M.; Magistrato, A.; Tavernelli, I.; Carloni, P.; Rothlisberger, U. Role of Protein Frame and Solvent for the Redox Properties of Azurin from *Pseudomonas Aeruginosa*. *Proc. Natl. Acad. Sci.* **2006**, *103*, 19641–19646.
10. Dudev, T.; Lim, C. Metal Binding Affinity and Selectivity in Metalloproteins: Insights from Computational Studies. *Annu. Rev. Biophys.* **2008**, *37*, 97–116.
11. Mirts, E.N.; Bhagi-Damodaran, A.; Lu, Y. Understanding and Modulating Metalloenzymes with Unnatural Amino Acids, Non-Native Metal Ions, and Non-Native Metallocofactors. *Acc. Chem. Res.* **2019**, *52*, 935–944.
12. Ortmyer, M.; Fisher, K.; Basran, J.; Wolde-Michael, E.M.; Heyes, D.J.; Levy, C.; Lovelock, S.L.; Anderson, J.L.R.; Raven, E.L.; Hay, S.; et al. Rewiring the “Push-Pull” Catalytic Machinery of a Heme Enzyme Using an Expanded Genetic Code. *ACS Catal.* **2020**, *10*, 2735–2746.
13. Zastrow, M.L.; Pecoraro, V.L. Designing Functional Metalloproteins: From Structural to Catalytic Metal Sites. *Coord. Chem. Rev.* **2013**, *257*, 2565–2588.
14. Chino, M.; Maglio, O.; Natri, F.; Pavone, V.; DeGrado, W.F.; Lombardi, A. Artificial Diiron Enzymes with a De Novo Designed Four-Helix Bundle Structure. *Eur. J. Inorg. Chem.* **2015**, *2015*, 3371–3390.
15. Lombardi, A.; Pirro, F.; Maglio, O.; Chino, M.; DeGrado, W.F. De Novo Design of Four-Helix Bundle Metalloproteins: One Scaffold, Diverse Reactivities. *Acc. Chem. Res.* **2019**, *52*, 1148–1159.
16. Grayson, K.J.; Anderson, J.L.R. Designed for Life: Biocompatible de Novo Designed Proteins and Components. *J. R. Soc. Interface*, **2018**, *15*, 20180472.
17. Kakkis, A.; Gagnon, D.; Esselborn, J.; Britt, R.D.; Tezcan, F.A. Metal-Templated Design of Chemically Switchable Protein Assemblies with High-Affinity Coordination Sites. *Angew. Chem. Int. Ed.* **2020**, *59*, 21940–21944.
18. Basler, S.; Studer, S.; Zou, Y.; Mori, T.; Ota, Y.; Camus, A.; Bunzel, H.A.; Helgeson, R.C.; Houk, K.N.; Jiménez-Osés, G.; et al. Efficient Lewis Acid Catalysis of an Abiological Reaction in a de Novo Protein Scaffold. *Nat. Chem.* **2021**, *13*, 231–235.
19. Slope, L.N.; Hill, M.G.; Smith, C.F.; Teare, P.; Cogan, F.J. de; Britton, M.M.; Peacock, A.F.A. Tuning Coordination Chemistry through the Second Sphere in Designed Metallocoiled Coils. *Chem. Commun.* **2020**, *56*, 3729–3732.
20. Carvalho, H.F.; Branco, R.J.F.; Leite, F.A.S.; Matzapetakis, M.; Roque, A.C.A.; Iranzo, O. Hydrolytic Zinc Metallopeptides Using a Computational Multi-State Design Approach. *Catal. Sci. Technol.* **2019**, *9*, 6723–6736.
21. D'Souza, A.; Bhattacharjya, S. De Novo-Designed β -Sheet Heme Proteins. *Biochemistry*, **2021**, *60*, 431–439.
22. Chino, M.; Leone, L.; Zambrano, G.; Pirro, F.; D'Alonzo, D.; Firpo, V.; Aref, D.; Lista, L.; Maglio, O.; Natri, F.; et al. Oxidation Catalysis by Iron and Manganese Porphyrins within Enzyme-like Cages. *Biopolymers*, **2018**, *109*, e23107.
23. Siepi, M.; Oliva, R.; Petraccone, L.; Vecchio, P.D.; Ricca, E.; Istatico, R.; Lanzilli, M.; Maglio, O.; Lombardi, A.; Leone, L.; et al. Fluorescent Peptide DH3w: A Sensor for Environmental Monitoring of Mercury (II). *PLOS ONE*, **2018**, *13*, e0204164.
24. Bos, J.; Roelfes, G. Artificial Metalloenzymes for Enantioselective Catalysis. *Curr. Opin. Chem. Biol.* **2014**, *19*, 135–143.
25. Soler, M.; Feliu, L.; Planas, M.; Ribas, X.; Costas, M. Peptide-Mediated Vectorization of Metal Complexes: Conjugation Strategies and Biomedical Applications. *Dalton Trans.* **2016**, *45*, 12970–12982.
26. Leone, L.; D'Alonzo, D.; Maglio, O.; Pavone, V.; Natri, F.; Lombardi, A. Highly Selective Indole Oxidation Catalyzed by a Mn-Containing Artificial Mini-Enzyme. *ACS Catal.* **2021**, *11*, 9407–9417.
27. Quijano-Rubio, A.; Yeh, H.-W.; Park, J.; Lee, H.; Langan, R.A.; Boyken, S.E.; Lajoie, M.J.; Cao, L.; Chow, C.M.; Miranda, M.C.; et al. De Novo Design of Modular and Tunable Protein Biosensors. *Nature*, **2021**, *591*, 482–487.
28. Ferrando, J.; Solomon, L.A. Recent Progress Using De Novo Design to Study Protein Structure, Design and Binding Interactions. *Life*, **2021**, *11*, 225.
29. Bowman, S.E.J.; Bridwell-Rabb, J.; Drennan, C.L. Metalloprotein Crystallography: More than a Structure. *Acc. Chem. Res.* **2016**, *49*, 695–702.
30. Polizzi, N.F.; DeGrado, W.F. A Defined Structural Unit Enables de Novo Design of Small-Molecule-Binding Proteins. *Science*, **2020**, *369*, 1227–1233.
31. Sala, D.; Musiani, F.; Rosato, A. Application of Molecular Dynamics to the Investigation of Metalloproteins Involved in Metal Homeostasis. *Eur. J. Inorg. Chem.* **2018**, *2018*, 4661–4677.
32. Mylemans, B.; Killian, T.; Vandebroek, L.; Van Meervelt, L.; Tame, J.R.H.; Parac-Vogt, T.N.; Voet, A.R.D. Crystal Structures of Scone: Pseudosymmetric Folding of a Symmetric Designer Protein. *Acta Crystallogr. Sect. Struct. Biol.* **2021**, *77*, 933–942.
33. Zhang, S.-Q.; Chino, M.; Liu, L.; Tang, Y.; Hu, X.; DeGrado, W.F.; Lombardi, A. De Novo Design of Tetranuclear Transition Metal Clusters Stabilized by Hydrogen-Bonded Networks in Helical Bundles. *J. Am. Chem. Soc.* **2018**, *140*, 1294–1304.
34. Olson, T.L.; Espiritu, E.; Edwardraja, S.; Simmons, C.R.; Williams, J.C.; Ghirlanda, G.; Allen, J.P. Design of Dinuclear Manganese Cofactors for Bacterial Reaction Centers. *Biochim. Biophys. Acta BBA - Bioenerg.* **2016**, *1857*, 539–547.

35. Mirts, E.N.; Petrik, I.D.; Hosseinzadeh, P.; Nilges, M.J.; Lu, Y. A Designed Heme-[4Fe-4S] Metalloenzyme Catalyzes Sulfite Reduction like the Native Enzyme. *Science*, **2018**, *361*, 1098–1101.
36. Chino, M.; Leone, L.; Maglio, O.; Lombardi, A. Designing Covalently Linked Heterodimeric Four-Helix Bundles. *Methods Enzymol.* **2016**, *580*, 471–499.
37. Chino, M.; Leone, L.; Maglio, O.; D'Alonzo, D.; Pirro, F.; Pavone, V.; Natri, F.; Lombardi, A. A De Novo Heterodimeric Due Ferri Protein Minimizes the Release of Reactive Intermediates in Dioxygen-Dependent Oxidation. *Angew. Chem. Int. Ed.* **2017**, *56*, 15580–15583.
38. Mocny, C.S.; Pecoraro, V.L. De Novo Protein Design as a Methodology for Synthetic Bioinorganic Chemistry. *Acc. Chem. Res.* **2015**, *48*, 2388–2396.
39. Mutter, A.C.; Tyryshkin, A.M.; Campbell, I.J.; Poudel, S.; Bennett, G.N.; Silberg, J.J.; Nanda, V.; Falkowski, P.G. De Novo Design of Symmetric Ferredoxins That Shuttle Electrons in Vivo. *Proc. Natl. Acad. Sci.* **2019**, *116*, 14557–14562.
40. Leone, L.; Chino, M.; Natri, F.; Maglio, O.; Pavone, V.; Lombardi, A. Mimochrome, a Metalloporphyrin-Based Catalytic Swiss Knife†. *Biotechnol. Appl. Biochem.* **2020**, *67*, 495–515.
41. Faiella, M.; Maglio, O.; Natri, F.; Lombardi, A.; Lista, L.; Hagen, W.R.; Pavone, V. De Novo Design, Synthesis and Characterisation of MP3, A New Catalytic Four-Helix Bundle Hemeprotein. *Chem. – Eur. J.* **2012**, *18*, 15960–15971.
42. Caserta, G.; Chino, M.; Firpo, V.; Zambrano, G.; Leone, L.; D'Alonzo, D.; Natri, F.; Maglio, O.; Pavone, V.; Lombardi, A. Enhancement of Peroxidase Activity in Artificial Mimochrome VI Catalysts through Rational Design. *ChemBioChem*, **2018**, *19*, 1823–1826.
43. Leone, L.; D'Alonzo, D.; Balland, V.; Zambrano, G.; Chino, M.; Natri, F.; Maglio, O.; Pavone, V.; Lombardi, A. Mn-Mimochrome VI^a: An Artificial Metalloenzyme With Peroxygenase Activity. *Front. Chem.* **2018**, *6*.
44. Zambrano, G.; Ruggiero, E.; Malafronte, A.; Chino, M.; Maglio, O.; Pavone, V.; Natri, F.; Lombardi, A. Artificial Heme Enzymes for the Construction of Gold-Based Biomaterials. *Int. J. Mol. Sci.* **2018**, *19*, 2896.
45. Zambrano, G.; Natri, F.; Pavone, V.; Lombardi, A.; Chino, M. Use of an Artificial Miniaturized Enzyme in Hydrogen Peroxide Detection by Chemiluminescence. *Sensors*, **2020**, *20*, 3793.
46. Maglio, O.; Chino, M.; Vicari, C.; Pavone, V.; Louro, R.O.; Lombardi, A. Histidine Orientation in Artificial Peroxidase Regioisomers as Determined by Paramagnetic NMR Shifts. *Chem. Commun.* **2021**, *57*, 990–993.
47. Le, J.M.; Alachouzos, G.; Chino, M.; Frontier, A.J.; Lombardi, A.; Bren, K.L. Tuning Mechanism through Buffer Dependence of Hydrogen Evolution Catalyzed by a Cobalt Mini-Enzyme. *Biochemistry*, **2020**, *59*, 1289–1297.
48. Lombardi, A.; Marasco, D.; Maglio, O.; Di Costanzo, L.; Natri, F.; Pavone, V. Miniaturized Metalloproteins: Application to Iron–Sulfur Proteins. *Proc. Natl. Acad. Sci.* **2000**, *97*, 11922–11927.
49. Benson, D.E.; Wisz, M.S.; Liu, W.; Hellinga, H.W. Construction of a Novel Redox Protein by Rational Design: Conversion of a Disulfide Bridge into a Mononuclear Iron–Sulfur Center. *Biochemistry*, **1998**, *37*, 7070–7076.
50. Farinas, E.; Regan, L. The de Novo Design of a Rubredoxin-like Fe Site. *Protein Sci.* **1998**, *7*, 1939–1946.
51. Nanda, V.; Rosenblatt, M.M.; Osyczka, A.; Kono, H.; Getahun, Z.; Dutton, P.L.; Saven, J.G.; DeGrado, W.F. De Novo Design of a Redox-Active Minimal Rubredoxin Mimic. *J. Am. Chem. Soc.* **2005**, *127*, 5804–5805.
52. Tebo, A.G.; Pinter, T.B.J.; García-Serres, R.; Speelman, A.L.; Tard, C.; Sénéque, O.; Blondin, G.; Latour, J.-M.; Penner-Hahn, J.; Lehnert, N.; et al. Development of a Rubredoxin-Type Center Embedded in a de Novo-Designed Three-Helix Bundle. *Biochemistry*, **2018**, *57*, 2308–2316.
53. Bertini, I.; Luchinat, C. High Spin Cobalt(II) as a Probe for the Investigation of Metalloproteins. *Adv. Inorg. Biochem.* **1984**, *6*, 71–111.
54. Reddi, A.R.; Guzman, T.R.; Breece, R.M.; Tierney, D.L.; Gibney, B.R. Deducing the Energetic Cost of Protein Folding in Zinc Finger Proteins Using Designed Metallopeptides. *J. Am. Chem. Soc.* **2007**, *129*, 12815–12827.
55. N. Besold, A.; R. Widger, L.; Namuswe, F.; L. Michalek, J.; J. Michel, S.L.; P. Goldberg, D. Revisiting and Re-Engineering the Classical Zinc Finger Peptide: Consensus Peptide-1 (CP-1). *Mol. Biosyst.* **2016**, *12*, 1183–1193.
56. VanZile, M.L.; Cospier, N.J.; Scott, R.A.; Giedroc, D.P. The Zinc Metalloregulatory Protein *Synechococcus* PCC7942 SmtB Binds a Single Zinc Ion per Monomer with High Affinity in a Tetrahedral Coordination Geometry. *Biochemistry*, **2000**, *39*, 11818–11829.
57. Garmer, D.R.; Krauss, M. Ab Initio Quantum Chemical Study of the Cobalt D-d Spectroscopy of Several Substituted Zinc Enzymes. *J. Am. Chem. Soc.* **1993**, *115*, 10247–10257.
58. Regan, L.; Clarke, N.D. A Tetrahedral Zinc(II)-Binding Site Introduced into a Designed Protein. *Biochemistry*, **1990**, *29*, 10878–10883.
59. Krizek, B.A.; Merkle, D.L.; Berg, J.M. Ligand Variation and Metal Ion Binding Specificity in Zinc Finger Peptides. *Inorg. Chem.* **1993**, *32*, 937–940.
60. Sreerama, N.; Woody, R.W. Estimation of Protein Secondary Structure from Circular Dichroism Spectra: Comparison of CON-TIN, SELCON, and CDSSTR Methods with an Expanded Reference Set. *Anal. Biochem.* **2000**, *287*, 252–260.
61. Roccatano, D.; Colombo, G.; Fioroni, M.; Mark, A.E. Mechanism by Which 2,2,2-Trifluoroethanol/Water Mixtures Stabilize Secondary-Structure Formation in Peptides: A Molecular Dynamics Study. *Proc Natl Acad Sci USA*, **2002**, *99*, 12179–12184.
62. Wagner, G.; Neuhaus, D.; Wörgötter, E.; Vašák, M.; Kägi, J.H.R.; Wüthrich, K. Nuclear Magnetic Resonance Identification of “Half-Turn” and 310-Helix Secondary Structure in Rabbit Liver Metallothionein-2. *J. Mol. Biol.* **1986**, *187*, 131–135.
63. Wüthrich, K. *NMR of Proteins and Nucleic Acids*; Wiley: New York, NY, **1986**.
64. Karplus, M. Contact Electron-Spin Coupling of Nuclear Magnetic Moments. *J. Chem. Phys.* **1959**, *30*, 11–15.

65. Lombardi, A.; Natri, F.; Marasco, D.; Maglio, O.; Sanctis, G.D.; Sinibaldi, F.; Santucci, R.; Coletta, M.; Pavone, V. Design of a New Mimochrome with Unique Topology. *Chem. – Eur. J.* **2003**, *9*, 5643–5654.
66. Engel, D.E.; DeGrado, W.F. Amino Acid Propensities Are Position-Dependent Throughout the Length of α -Helices. *J. Mol. Biol.* **2004**, *337*, 1195–1205.
67. Pavone, V.; Gaeta, G.; Lombardi, A.; Natri, F.; Maglio, O.; Isernia, C.; Saviano, M. Discovering Protein Secondary Structures: Classification and Description of Isolated α -Turns. *Biopolymers*, **1996**, *38*, 705–721.
68. Ramakrishnan, C.; Nataraj, D.V. Energy Minimization Studies on α -Turns. *J. Pept. Sci.* **1998**, *4*, 239–252.
69. Andreini, C.; Bertini, I.; Cavallaro, G. Minimal Functional Sites Allow a Classification of Zinc Sites in Proteins. *PLoS ONE*, **2011**, *6*.
70. Ireland, S.M.; Martin, A.C.R. ZincBind—the Database of Zinc Binding Sites. *Database*, **2019**, *2019*, baz006.
71. Williams, C.J.; Headd, J.J.; Moriarty, N.W.; Prisant, M.G.; Videau, L.L.; Deis, L.N.; Verma, V.; Keedy, D.A.; Hintze, B.J.; Chen, V.B.; et al. MolProbity: More and Better Reference Data for Improved All-atom Structure Validation. *Protein Sci. Publ. Protein Soc.* **2018**, *27*, 293–315.
72. Hill, R.B.; Raleigh, D.P.; Lombardi, A.; DeGrado, W.F. De Novo Design of Helical Bundles as Models for Understanding Protein Folding and Function. *Acc. Chem. Res.* **2000**, *33*, 745–754.
73. Huang, C.-C.; Casey, P.J.; Fierke, C.A. Evidence for a Catalytic Role of Zinc in Protein Farnesyltransferase: Spectroscopy of Co^{2+} -Farnesyltransferase Indicates Metal Coordination of the Substrate Thiolate*. *J. Biol. Chem.* **1997**, *272*, 20–23.
74. Taylor, J.S.; Lau, C.Y.; Applebury, M.L.; Coleman, J.E. Escherichia Coli Co (II) Alkaline Phosphatase Absorption, Circular Dichroism, and Magnetic Circular Dichroism of the d-d Electronic Transitions. *J. Biol. Chem.* **1973**, *248*, 6216–6220.
75. Vallee, W.M.B.L. Cobalt as Probe and Label of Proteins. In: *Methods in Enzymology; Metallobiochemistry Part C: Spectroscopic and Physical Methods for Probing Metal Ion Environments in Metalloenzymes and Metalloproteins*; Academic Press, **1993**; Vol. 226, pp. 52–71.
76. Xu, S.; Gupta, Y.K. Natural Zinc Ribbon HNH Endonucleases and Engineered Zinc Finger Nicking Endonuclease. *Nucleic Acids Res.* **2013**, *41*, 378–390.
77. Haag, F.; Ahmed, L.; Reiss, K.; Block, E.; Batista, V.S.; Krautwurst, D. Copper-Mediated Thiol Potentiation and Mutagenesis-Guided Modeling Suggest a Highly Conserved Copper-Binding Motif in Human OR2M3. *Cell. Mol. Life Sci.* **2020**, *77*, 2157–2179.
78. Guérineau, M.; Bessa, L.; Moriau, S.; Lescop, E.; Bontems, F.; Mathy, N.; Guittet, E.; Bischerour, J.; Bétermier, M.; Morellet, N. The Unusual Structure of the PiggyMac Cysteine-Rich Domain Reveals Zinc Finger Diversity in PiggyBac-Related Transposases. *Mob. DNA*, **2021**, *12*, 12.
79. Chou, K.C. Prediction and Classification of Alpha-Turn Types. *Biopolymers*, **1997**, *42*, 837–853.
80. Ohage, E.C.; Graml, W.; Walter, M.M.; Steinbacher, S.; Steipe, B. Beta-Turn Propensities as Paradigms for the Analysis of Structural Motifs to Engineer Protein Stability. *Protein Sci.* **1997**, *6*, 233–241.
81. Hsu, H.-J.; Chang, H.-J.; Peng, H.-P.; Huang, S.-S.; Lin, M.-Y.; Yang, A.-S. Assessing Computational Amino Acid Beta-Turn Propensities with a Phage-Displayed Combinatorial Library and Directed Evolution. *Struct. Lond. Engl.* **1993**, *2006*, *14*, 1499–1510.
82. Laitaoja, M.; Valjakka, J.; Jänis, J. Zinc Coordination Spheres in Protein Structures. *Inorg. Chem.* **2013**, *52*, 10983–10991.
83. Maret, W.; Li, Y. Coordination Dynamics of Zinc in Proteins. *Chem. Rev.* **2009**, *109*, 4682–4707.
84. Berg, J.M.; Merkle, D.L. On the Metal Ion Specificity of Zinc Finger Proteins. *J. Am. Chem. Soc.* **1989**, *111*, 3759–3761.
85. Marion, D. An Introduction to Biological NMR Spectroscopy *. *Mol. Cell. Proteomics*, **2013**, *12*, 3006–3025.
86. Scott, W.R.P.; Hunenberger, P.H.; Tironi, I.G.; Mark, A.E.; Billeter, S.R.; Fennel, J.; Torda, A.E.; Huber, T.; Kruger, P.; Gunsteren, W.F. van. The GROMOS biomolecular simulation program package. *J. Phys. Chem. A*, **1999**, *103*, 3596–3607.
87. Barlow, S.; Rohl, A.L.; Shi, S.; Freeman, C.M.; O'Hare, D. Molecular Mechanics Study of Oligomeric Models for Poly(Ferrocenylsilanes) Using the Extensible Systematic Forcefield (ESFF). *J. Am. Chem. Soc.* **1996**, *118*, 7578–7592.
88. Lifson, S.; Hagler, A.T.; Dauber, P. Consistent Force Field Studies of Intermolecular Forces in Hydrogen-Bonded Crystals. 1. Carboxylic Acids, Amides, and the C:O.Cntdot..Cntdot..Cntdot.H- Hydrogen Bonds. *J. Am. Chem. Soc.* **1979**, *101*, 5111–5121.
89. Brooks, B.R.; Brooks, C.L.; Mackerell, A.D.; Nilsson, L.; Petrella, R.J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; et al. CHARMM: The Biomolecular Simulation Program. *J. Comput. Chem.* **2009**, *30*, 1545–1614.