UNIVERSITÀ DI PARMA

UNIVERSITA' DEGLI STUDI DI PARMA

DOTTORATO DI RICERCA IN
"Scienze Chimiche"

CICLO XXXIV

Copper/histidine sites in designed peptides and artificial proteins

Coordinatore:
Chiar.ma Prof. Alessia Bacchi

Tutore:
Chiar.mo Prof. Matteo Tegoni

Dottoranda: Silvia Gentili

Anni Accademici 2018/2019 – 2020/2021

To mum and dad

Table of contents

Introduction ***1***

COPPER IN BIOLOGICAL SYSTEMS: AN OVERVIEW 3

COORDINATION CHEMISTRY OF COPPER(II) 4

PROTEIN DESIGN 7

REFERENCES 8

Chapter 1 Models of the copper binding sites in tau protein ***11***

INTRODUCTION 13

BINDING OF COPPER(II) TO TAU FRAGMENTS 17

Peptide Protonation Equilibria 17

Copper(II) Complexation Studies 19

Cu^{2+}/R1 complex-formation equilibria 19

Cu^{2+}/R3 complex-formation equilibria 22

Cu^{2+}/R3A complex-formation equilibria. 25

Copper(II) binding affinities 27

BINDING OF COPPER(I) TO TAU FRAGMENTS 29

CONCLUSIONS 33

EXPERIMENTAL SECTION 34

Tau fragments 34

Copper(II) binding to Tau fragments 34

Copper(I) binding to Tau fragments 35

REFERENCES 37

Chapter 2 Bioinspired copper peptides: role of His sites in oxidation reactions ***43***

INTRODUCTION 45

BINDING OF COPPER(II) TO HIS-CONTAINING PEPTIDES 49

Peptide Protonation Equilibria 49

Copper(II) Complexation Studies 50

Cu^{2+}/PCtrl complex-formation equilibria 52
Cu^{2+}/PHis(-2) complex-formation equilibria 55
Cu^{2+}/PHis(-4) complex-formation equilibria 59
Cu^{2+}/PHis(-6) complex-formation equilibria. 62
Copper(II) binding affinities 65
OXIDATION OF 4-METHYL CATECHOL BY COPPER(II)/PEPTIDE COMPLEXES 67
CONCLUSIONS 73
EXPERIMENTAL SECTION 75
Peptide synthesis and purification 75
Copper(II) binding to His-containing peptides 76
Oxidation of 4-methyl catechol 77
REFERENCES 78
Chapter 3 Exploiting the SpyCatcher/SpyTag technology to design a new artificial metalloprotein 81
INTRODUCTION 83
Protein redesign: the Trojan horse approach 84
The SpyCatcher/SpyTag system 86
The ATCUN metal binding site 90
DESIGN OF THE NEW SPYCATCHER/SPYTAG-ATCUN SYSTEM 92
Design of SpyCatcher 93
Design of SpyTag 94
IN-SILICO STUDY OF THE MODEL OF THE DESIGNED SPYCATCHER/SPY-TAG-ATCUN SYSTEM (SC-MUT/ST-MUT1) 96
SPYTAG-ATCUN SEQUENCING AND RECOMBINATION STUDIES WITH SPYCATCHER 99
COPPER(II) BINDING TO THE SPYCATCHER/SPYTAG-ATCUN SYSTEM 103
Cu^{2+}/SpyCatcher (SC-mut) 103
Cu^{2+}/SpyTag-ATCUN (ST-mut1) 107
Cu^{2+}/SpyTag (ST-mut2) 111
Cu^{2+}/SpyCatcher/SpyTag-ATCUN (SC-mut/ST-mut1) 113

Hydrolysis of p-nitrophenyl phosphate by Cu(II)/SpyCatcher/SpyTag-ATCUN (SC-mut/ST-mut1) 115

Conclusions 119

Experimental section 121

Structural and computational studies 121

Protein expression and purification 121

Protein sequencing and recombination studies 121

Spectroscopic studies 122

Hydrolysis of p-nitrophenyl phosphate 123

References 124

Concluding remarks ***129***

Introduction

COPPER IN BIOLOGICAL SYSTEMS: AN OVERVIEW

Copper is a trace metal essential for all organisms living in oxygen-rich environments. The biological relevance of copper is associated to its role as structural and catalytic cofactor in a substantial number of metalloproteins and metalloenzymes. These copper-containing systems are involved in a remarkable variety of fundamental biological functions, among which respiration, ATP production, signalling, iron metabolism, antioxidant defence, and other processes required for organism growth and development.[1]

The redox chemistry at the basis of the biological tasks of most copper-dependent proteins entails cycling of copper ions between their oxidised (+2) and reduced (+1) states. Although this redox activity is essential for normal organism functions,[2] it may also produce harmful effects when copper dyshomeostasis occurs. Indeed, the participation of Cu^{2+} and Cu^{+} ions (either free or bound to biomolecules) in the production of reactive oxygen species (ROS) has been widely demonstrated and associated to a number of serious diseases.[3–9] It is therefore clear the necessity of organisms to tightly regulate copper homeostasis and metabolism.

Given the essential yet toxic nature of copper, living beings have developed several mechanisms to ensure controlled homeostasis of this metal and its transport throughout the organism. The efficiency of these regulation processes can be evinced by the fact that although the average amount of copper in human body is *ca.* 100 mg, the concentration of free copper in cells is exceptionally low ($< 10^{-18}$ M).[10,11] In fact, intracellular copper is bound by various chaperone proteins that prevent undesired reactions by regulating metal homeostasis.[12]

A general example of copper metabolism is reported in ***Figure I.1***. Extracellular copper is commonly found as Cu(II), whereas the reducing intracellular environment stabilises Cu(I). Before entering a cell, a metalloreductase reduces copper(II) to copper(I), which then crosses the cell membrane through a transmembrane protein (Copper Transporter 1 or 2, CTR1 or CTR2). Once in the cytosol, copper is bound by various chaperones that deliver it to specific metalloproteins, where it performs its biological function. For instance, copper is transported to Cu,Zn-superoxide dismutase (SOD1) by CCS (copper chaperone for SOD1), and cytochrome C oxidase chaperone (Cox17) transfers the metal to cytochrome C oxidase (CCO) in mitochondria. Also, antioxidant protein 1 (ATOX1) delivers copper to ATPase proteins that are found in the Golgi apparatus, which in turn can donate the metal to other copper-enzymes or eliminate it from the cell.[13] The latter Cu-ATPases are especially crucial in maintaining intracellular copper homeostasis, since they are responsible for the removal of metal excess from cells. Undesired mutation of these proteins actually leads to life-threatening disorders, *e.g.* Menkes' and Wilson's diseases.[14]

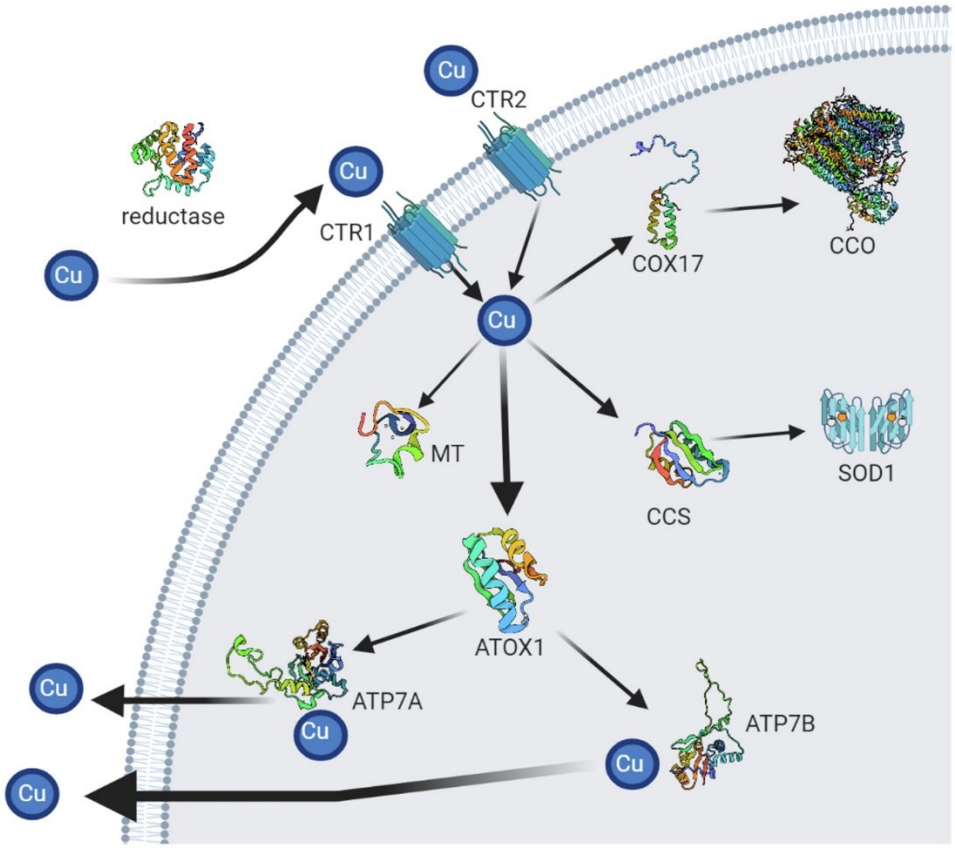


Figure I.1 *Generalised overview of copper metabolism in mammalian cells. Reprinted from Ref. 15. Copyright © 2021, with permission from Elsevier*

COORDINATION CHEMISTRY OF COPPER(II)

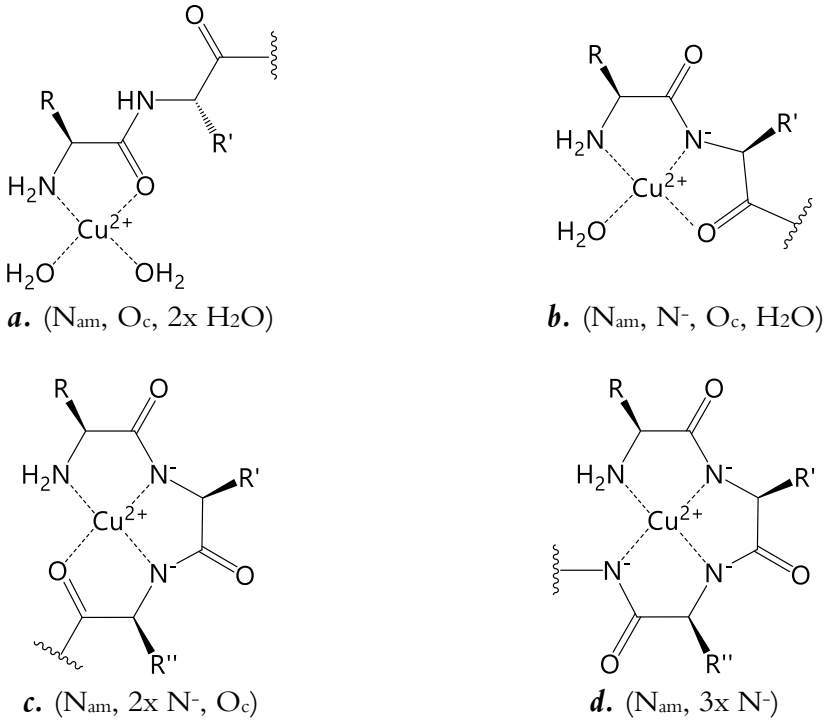
The interaction of copper with biomolecules is driven by its peculiar features in terms of coordination chemistry, which considerably depends on its oxidation state (+2 *vs.* +1). With respect to other bioavailable divalent transition metals, complexes formed by copper(II) are usually characterised by the strongest binding affinities, as postulated by Irving and Williams in 1953.[16] This tendency, which is reported in ***Scheme I.1***, is mostly dependent on metal ionic radius and on crystal field stabilisation energies (CFSE), whereas it is essentially independent of the ligand.

$$Mn(II) < Fe(II) < Co(II) < Ni(II) < \mathbf{Cu(II)} > Zn(II)$$

Scheme I.1 *The Irving-Williams order of stability of bivalent transition metal complexes.*

According to Pearson's Hard and Soft Acids and Bases Theory, Cu(II) is classified as borderline acid and its interaction with borderline/soft donor atoms, such as N and S, is therefore favoured.[17] Given its borderline properties, copper(II) can be coordinated also by harder O-donors, although with lower affinity. In addition, Ligand Field Theory is crucial to explain the coordination geometries of copper complexes. Cu(II) ions are often coordinated in square planar, square pyramidal, or distorted (by Jahn-Teller effect[18]) octahedral geometries as a result of Cu(II) preference for 4, 5, or 6 coordination numbers, respectively, which in turn arises from its incomplete d^9 configuration.[19–21] In the case of multidentate ligands, the formation of 5- or 6-membered chelation rings confers a remarkable stability to copper(II)-complexes.[22,23]

Altogether, the previously addressed features of copper(II) binding make biomolecules that are rich in N-, S-, and O-donors (*i.e.* amino acids, peptides, and proteins) ideal ligands for copper(II). Common binding motifs in peptides and proteins involve free N-terminal amino groups and carboxylate groups coordinated in the equatorial plane of copper(II), usually with the formation of 5- or 6-membered chelation rings. However, the additional coordination of deprotonated N^- atoms of the peptide backbone is most frequently observed (***Scheme I.2***), along with the participation of coordinative side chains of some amino acids such as histidine and cysteine.[24,25]

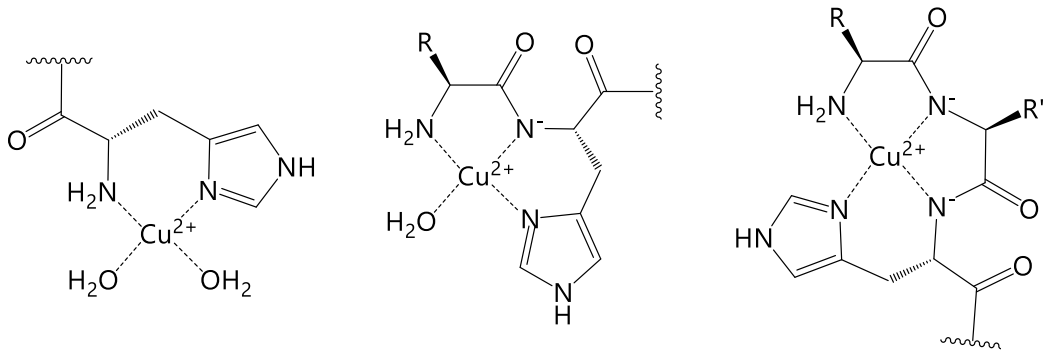


a. (N_{am}, O_c, 2x H_2O) ***b.*** (N_{am}, N^-, O_c, H_2O)

c. (N_{am}, 2x N^-, O_c) ***d.*** (N_{am}, 3x N^-)

Scheme I.2 *General representation of 1N3O (a), 2N2O (b), 3N1O (c) and 4N (d) coordination modes in peptides.*

Histidine side chain is characterised by the presence of an imidazole ring, which is more intermediate in its Hard-Soft behaviour than the amino or deprotonated amido-groups. His can coordinate copper(II) as a monodentate ligand, but very often its coordination occurs concomitantly to that of other functional groups. For instance, if His is in position 1, 2 or 3 in the peptide sequence, its coordination likely promotes the binding of both the terminal amino group and of the deprotonated backbone N^- atoms that separate the imidazole from NH_2 (***Scheme I.3***). The coordination to copper(II) of the deprotonated amino groups close to His, together with that of imidazole, occurs also if His is far from the N-terminal amino group or if the amino group is blocked/acetylated. This reveals that histidine is a peculiar ligand that allows copper(II) binding to basically all peptides or proteins bearing an accessible His residue. It is however necessary to underline that the imidazole, the amino, and the amido groups are weak acids and bases, and their coordination to copper(II) (*i.e.* coordination mode) is therefore pH dependent.[24,25]



Scheme I.3 *General representation of copper(II) N-terminal coordination by His residues in positions 1 (left), 2 (centre), and 3 (right).*

In the next chapters we will show examples of peptides that provide copper(II) with different coordination environments, together with various spectroscopic and catalytic features of the related copper(II) complexes, as a function of pH. In particular, in *Chapter 1* we will focus on the difference in the copper binding features of two model peptides of tau, a protein associated with neurodegeneration. These model peptides are characterised by the presence of either a single isolated His or two adjacent His-His residues, that differ for their coordination and thermodynamic properties. The binding features of both histidine motifs will be elucidated and will represent a starting point for the study described in *Chapter 2*. More precisely, in the second chapter we will examine the impact on copper binding and on pseudocatecholase activity of copper centres at His-His sites in the presence of a third His residue.

PROTEIN DESIGN

Metalloproteins promote some of the most complex biomolecular processes in Nature. This is one of the reasons why metalloprotein design has focused the attention of a great number of scientists around the world: creating new functional proteins means to impact at the molecular level the research in many fields that go from biochemistry and biotechnology to medicine, and beyond.

Designing an artificial metalloprotein means to design a biomolecular object at three levels: a protein scaffold with desired structure, a metal centre with a precise coordination geometry and donor atoms set, and a second-coordination sphere that may finely tune the functional properties of the macromolecule. The main idea is that if these three features are controlled and tuned, one may control with precision the function of the protein such as the binding affinity for a substrate or the catalytic properties toward specific reactions.[26] Examples of how a metal centre can be designed into purely artificial constructs in order to impart different catalytic properties and different stability is the "due-ferri" (di-iron) system, developed in the past two decades by Lombardi and DeGrado, or the TRI system, developed by Pecoraro and collaborators.[27–30] On the other hand, an example of how to control of the functional properties of a metal centre in a natural redesigned protein is the work of Yi Lu and co-workers on tuning of the redox potential of *blue copper* proteins by modification of the metal second coordination shell.[31–33]

It is evident that having tools for designing and isolating new artificial metalloproteins is of interest also in areas at the interface with biology. Nowadays, there is a huge interest for instance in the development of sensing devices for pollutants or biomarkers, or in the development of new catalysts that promote reactions not relevant in Nature but important to us for bio- and nanotechnological purposes. In all these fields, protein and metalloprotein design may provide with a great contribution to obtain cutting-edge devices that are biotechnology-based but not strictly tied to natural proteins.

Details of the concept and the strategies of protein design are given in *Chapter 3*. Here, one aspect is worth to be anticipated. In terms of strategies of preparation and isolation, peptides and proteins stand ideally at two different extremes. Peptides (until a few tenth of residues long) can be easily synthesised using solid-phase synthetic methods, and then purified using RP-HPLC in mild acidic conditions. Conversely, proteins need to be invariably expressed in microorganisms such as bacteria, and then purified using chromatography. The purification itself may be sometimes laborious: for instance, proteins can be expressed bearing tags needed for affinity column chromatography that are removed after a first purification workup, therefore requiring further purification steps. In *Chapter 3* the design of a new metalloprotein based on the SpyCatcher/SpyTag technology is presented. Through that construct, we aimed to move the complexity of the design and isolation of this new metalloprotein from the protein construct to shorter peptide sequences.

REFERENCES

1. Turski, M. L. & Thiele, D. J. New Roles for Copper Metabolism in Cell Proliferation, Signaling, and Disease. *J. Biol. Chem.* ***284***, 717–721 (2009).
2. Uauy, R., Olivares, M. & Gonzalez, M. Essentiality of copper in humans. *Am. J. Clin. Nutr.* ***67***, 952S-959S (1998).
3. Halliwell, B. & Gutteridge, J. M. C. Oxygen toxicity, oxygen radicals, transition metals and disease. *Biochem. J.* ***219***, 1 (1984).
4. Starkebaum, G. & Harlan, J. M. Endothelial cell injury due to copper-catalyzed hydrogen peroxide generation from homocysteine. *J. Clin. Invest.* **77**, 1370–1376 (1986).
5. Jomova, K. & Valko, M. Advances in metal-induced oxidative stress and human disease. *Toxicology* ***283***, 65–87 (2011).
6. Kanti Das, T., Wati, M. R. & Fatima-Shad, K. Oxidative Stress Gated by Fenton and Haber Weiss Reactions and Its Association With Alzheimer's Disease. *Arch. Neurosci.* ***2***, (2014).
7. Gaggelli, E., Kozlowski, H., Valensin, D. & Valensin, G. Copper Homeostasis and Neurodegenerative Disorders (Alzheimer's, Prion, and Parkinson's Diseases and Amyotrophic Lateral Sclerosis). *Chem. Rev.* ***106***, 1995–2044 (2006).
8. Kozlowski, H., Luczkowski, M., Remelli, M. & Valensin, D. Copper, zinc and iron in neurodegenerative diseases (Alzheimer's, Parkinson's and prion diseases). *Coord. Chem. Rev.* ***256***, 2129–2141 (2012).
9. Benedictis, C. A. De, Vilella, A. & Grabrucker, A. M. The Role of Trace Metals in Alzheimer's Disease. *Exon Publ.* 85–106 (2019) doi:10.15586/ALZHEIMERSDISEASE.2019.CH6.
10. Rae, T. D., Schmidt, P. J., Pufahl, R. A., Culotta, V. C. & O'Halloran, T. V. Undetectable intracellular free copper: The requirement of a copper chaperone for superoxide dismutase. *Science (80-.).* ***284***, 805–808 (1999).
11. Lippard, S. J. Free copper ions in the cell? *Science (80-.).* ***284***, 748–749 (1999).
12. Klomp, L. W. J. *et al.* Identification and Functional Expression of HAH1, a Novel Human Gene Involved in Copper Homeostasis. *J. Biol. Chem.* ***272***, 9221–9226 (1997).
13. da Silva, D. A. *et al.* Copper in tumors and the use of copper-based compounds in cancer treatment. *J. Inorg. Biochem.* ***226***, 111634 (2022).
14. Bull, P. C. & Cox, D. W. Wilson disease and Menkes disease: new handles on heavy-metal transport. *Trends Genet.* ***10***, 246–252 (1994).
15. Linder, M. C. & Hazegh-Azam, M. Copper biochemistry and molecular biology. *Am. J. Clin. Nutr.* ***63***, 797S-811S (1996).
16. Irving, H. & Williams, R. J. P. The stability of transition-metal complexes. *J. Chem. Soc.* 3192–3210 (1953) doi:10.1039/JR9530003192.
17. Pearson, R. G. Hard and Soft Acids and Bases. *J. Am. Chem. Soc.* ***85***, 3533–3539 (2002).

18. Englman, R. *The Jahn-Teller effect in molecules and crystals*. (John Wiley & Sons Ltd, 1972).
19. Vallee, B. L. & Williams, R. J. Metalloenzymes: the entatic nature of their active sites. *Proc. Natl. Acad. Sci. U. S. A.* ***59***, 498 (1968).
20. George, P., McClure, D. S., Griffith, J. S. & Orgel, L. E. Ligand Field Theory and the Stability of Transition Metal Compounds. *J. Chem. Phys.* ***24***, 1269 (2004).
21. Griffith, J. S. & Orgel, L. E. Ligand-field theory. *Q. Rev. Chem. Soc.* ***11***, 381–393 (1957).
22. Nagaj, J. *et al.* Selective control of Cu(II) complex stability in histidine peptides by β-alanine. *J. Inorg. Biochem.* ***119***, 85–89 (2013).
23. Wiloch, M. Z. *et al.* Copper(II) Complexes with ATCUN Peptide Analogues: Studies on Redox Activity in Different Solutions. *J. Electrochem. Soc.* ***164***, G77–G81 (2017).
24. Kozłowski, H., Bal, W., Dyba, M. & Kowalik-Jankowska, T. Specific structure–stability relations in metallopeptides. *Coord. Chem. Rev.* ***184***, 319–346 (1999).
25. Sigel, H. & Martin, R. B. Coordinating properties of the amide bond. Stability and structure of metal ion complexes of peptides and related ligands. *Chem. Rev.* ***82***, 385–426 (1982).
26. Yu, F. *et al.* Protein Design: Toward Functional Metalloenzymes. *Chem. Rev.* ***114***, 3495–3578 (2014).
27. Summa, C. M., Lombardi, A., Lewis, M. & DeGrado, W. F. Tertiary templates for the design of diiron proteins. *Curr. Opin. Struct. Biol.* ***9***, 500–508 (1999).
28. Calhoun, J. R. *et al.* Artificial diiron proteins: From structure to function. *Pept. Sci.* ***80***, 264–278 (2005).
29. Calhoun, J. R. *et al.* Oxygen reactivity of the biferrous site in the de novo designed four helix bundle peptide DFsc: Nature of the 'intermediate' and reaction mechanism. *J. Am. Chem. Soc.* ***130***, 9188–9189 (2008).
30. Chino, M. *et al.* Artificial Diiron Enzymes with a de Novo Designed Four-Helix Bundle Structure. *Eur. J. Inorg. Chem.* ***2015***, 3371–3390 (2015).
31. Garner, D. K. *et al.* Reduction potential tuning of the blue copper center in Pseudomonas aeruginosa azurin by the axial methionine as probed by unnatural amino acids. *J. Am. Chem. Soc.* ***128***, 15608–15617 (2006).
32. Marshall, N. M. *et al.* Rationally tuning the reduction potential of a single cupredoxin beyond the natural range. *Nature* ***462***, 113–116 (2009).
33. New, S. Y., Marshall, N. M., Hor, T. S. A., Xue, F. & Lu, Y. Redox tuning of two biological copper centers through non-covalent interactions: same trend but different magnitude. *Chem. Commun. (Camb).* ***48***, 4217–4219 (2012).

Chapter 1

Models of the copper binding sites in tau protein

INTRODUCTION

Tauopathies are a large group of neurodegenerative diseases that, although presenting very different phenotypic manifestations, share as a common feature the abnormal intraneuronal accumulation of the tubulin-associated unit (*tau* protein).[1] Among these disorders, Alzheimer's disease (AD) is the most common, worldwide accounting for 60%-80% of dementia cases, followed by Parkinson's disease (PD).[2,3] Progressive cognitive decline associated with memory loss, changes in personality, and impaired thinking ability is typical of AD. Conversely, people affected by PD experience rest tremors, movement slowness/rigidity and motor dysfunctions leading to physical disability, and possibly dementia. Both AD and PD conditions eventually lead to death. Despite some treatments (*e.g.* cholinesterase inhibitors and NMDA receptor antagonists[4]) succeeded in attenuating symptoms, an effective cure for these diseases has yet to be found.

Tau protein was first isolated in 1975 and it has been described as an essential factor for the assembly of microtubules (MTs), the major components of cytoskeleton.[5] Principally located within the axonal compartments of neurons,[6] the main role of tau protein lies in MTs stabilisation and in the regulation of MTs dynamics, although this function was recently questioned.[7] The involvement of tau in a variety of 'atypical' processes, such as mitochondrial function, was also reviewed.[8–10] Most (if not all) of these biochemical roles with impact in tauopathies are correlated to post-translational modifications of the protein, *e.g.* phosphorylation, nitration and acylation, which abnormally occur in pathological conditions.[11–15]

Tau is an intrinsically disordered protein that in adult human brain is encoded in six alternative splicing isoforms of 352-441 amino acids by a single gene located on chromosome 17.[16,17] Each isoform is characterised by an acidic and variable N-terminus that a proline-rich domain links to three or four pseudorepeats (R1-R4, depending on the isoform), and a basic C-terminus. The sequence of the R1-R4 region in the longest tau isoform ('4R', 441 aa.) is reported in ***Figure 1.1***. The R1-R4 pseudorepeats constitute the microtubule binding domain and consist of 31-34 amino acids, of which 18 are highly conserved residues among the different tau isoforms. In particular, the isoform containing all four pseudorepeats ('4R' tau, ***Figure 1.1***) was found to bind microtubules with high affinity (K_d *ca.* 100 nM) due to the presence of the KVQIINKK sequence placed between the R1 and R2 repeats, with the N-terminal region of tau possibly acting as a spacer while projecting from the microtubules.[18]

Despite the non-structured nature of tau protein, the propensity of two short sequences of the R2 and R3 repeats (^{275}VQIINK^{280} and ^{306}VQIVYK^{311}, respectively) to form β-structures was found to promote protein aggregation.[19] As previously mentioned, abnormal post-translational modification of tau occurs under pathological conditions and leads to

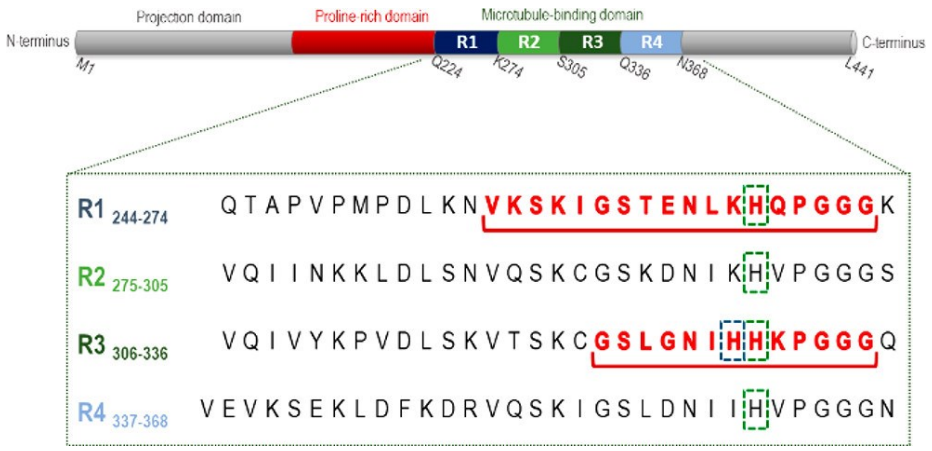


Figure 1.1 *Amino acid sequences of the four pseudorepeats in the '4R' isoform of tau protein. Each pseudorepeat consists of 31-32 amino acid residues and contains a highly conserved octadecapeptide (>60% homology). Reprinted with permission from Ref. 20. Copyright © 2019 American Chemical Society*

self-aggregation and oligomerisation of the protein. These processes eventually result in the formation of intraneuronal neurofibrillary tangles (NFTs), usually consisting of hyperphosphorylated tau (p-tau) and identified as hallmarks of several tauopathies. Recently, tau fibrils in diseased brains were structurally characterised by cryo-electron microscopy, which evidenced the involvement of the R3 and R4 regions in the formation of cross-β/β-helix structures.[21] Although both the formation mechanism and the true nature (toxic *vs.* protective) of NFTs remain unclear, it is known that tau aggregation and hyperphosphorylation result in impaired microtubule binding, which is one of the main causes for microtubule disassembly and neuronal death (***Figure 1.2***).[10,22]

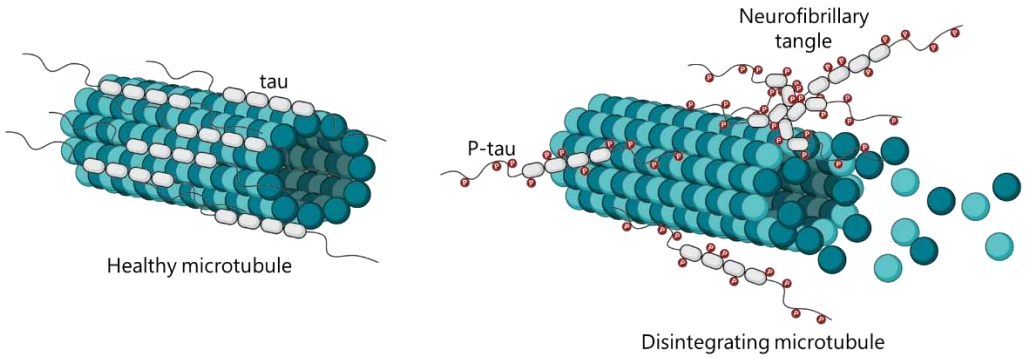


Figure 1.2 *Conceptual representation of the microtubule stabilising role of tau protein. In normal conditions (left), tau binds to the microtubule and contributes to maintaining its structure. In pathological conditions (right), tau protein undergoes hyperphosphorylation (p-tau, phosphate groups represented with red dots) and aggregates into NFTs, losing its function and resulting into microtubule disassembly.*

The formation of insoluble protein deposits such as NFTs is only one of several factors that contribute to the onset of neurodegenerative diseases. Dysregulation of essential metal ions and oxidative stress also play a central role in the progression of neurodegeneration. For instance, dyshomeostasis of redox active metals (*e.g.* Cu, Fe) may actually result in the production of reactive oxygen/nitrogen species (ROS, RNS) through Fenton and Haber-Weiss reactions.[23] Oxidative/nitrative stress is consequently promoted, with the formation of free radicals that damage proteins, lipids, DNA, *etc.*[24,25] Moreover, impaired regulation of metal ions often results into high metal concentrations being present at neuronal level, where they may interact with neuroproteins and promote their aggregation and other pathological processes.[26,27] In this field of investigation, the interaction of several metal ions with β-amyloid, α-synuclein and prion proteins (all involved in neurodegeneration) has been extensively studied and characterised.[28–39] Conversely, much less is known about tau protein interacting with metal ions, despite the growing number of studies.

The microtubule binding domain of tau protein contains several histidine residues located in the pseudorepeats sequences (His268 in R1, His299 in R2, His362 in R4, His329 and His330 in R3; ***Figure 1.1***), which thus represent potential metal binding sites. To achieve deeper insights on copper binding to tau, various fragments of the protein were studied as a model for full-length tau. In particular, these model peptides comprised portions of the R1-R4 pseudorepeats, which were reported to bind Cu(II) with 1:1 metal/protein stoichiometry and micromolar affinity, and from which emerged the key role of histidine residues as coordination sites.[40–45] However, given the presence of free amino and carboxyl groups at their N- and C-terminal regions, respectively, these peptides in fact do not model the coordination properties of full-length tau. For this reason, we decided to investigate N-acetylated and C-amidated fragments of R1 and R3, aiming at better defining the stoichiometry, structure, and redox properties of the related copper complexes. These specific fragments, with sequences reported in ***Scheme 1.1***, were chosen for the possibility to investigate the effect that the presence of one *vs.* two adjacent histidine residues may have on copper binding and redox activity of copper complexes. More in general, the idea was to explore the possibility to isolate the R1 and R3 binding sites and to study the copper binding properties and reactivity (R1 as a putative model also for R2 and R4, one His residue each).

R1 Ac-^{256}VKSKIGSTENLK**H**QPGGG^{273}-NH_2

R3 Ac-^{323}GSLGNI**HH**KPGGG^{335}-NH_2

R3A Ac-^{318}VTSK**A**GSLGNI**HH**KPGGG^{335}-NH_2

R3C Ac-^{318}VTSK**C**GSLGNI**HH**KPGGG^{335}-NH_2

Scheme 1.1 *Amino acid sequences of the fragments of the R1 and R3 pseudorepeats of tau protein studied in this work.*

The oxidase catalytical properties of the 'R1' and 'R3' model peptides in the presence of copper(II) were studied by the group of Prof. Luigi Casella and Prof. Simone Dell'Acqua at the University of Pavia, who examined catechol oxidation reactions of dopamine (DA) and 4-methyl catechol (MC) as substrates.[20] Given its neurotransmitter function, DA is a compound of interest which metabolic alteration is deeply correlated to neurodegeneration. MC, in turn, represents a DA model of easier characterisation. The kinetic profiles for the oxidation of these substrates are reported in ***Figure 1.3*** and clearly show the greater catalytic activity of Cu(II)/R3 complexes with respect to Cu(II)/R1.

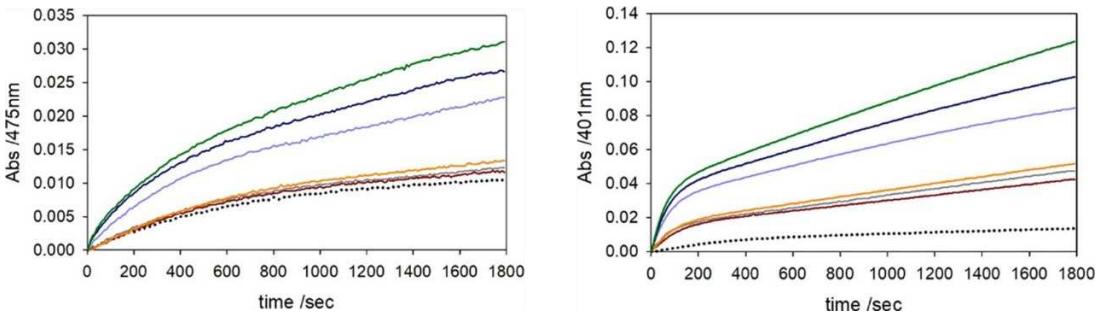


Figure 1.3 *Kinetic profiles (absorbance vs. time) of DA (left) and MC (right) oxidation in the presence of only copper(II) (25 μM, brown trace) and with 2 eq. (grey) and 4 eq. (orange) of the R1 peptide, or 1 eq. (light blue), 2 eq. (blue) and 4 eq. (green) of R3. Autoxidation of substrate (3 mM) is also shown (black dotted trace). For each data set: 3 mM substrate, 50 mM aqueous HEPES buffer, pH 7.4, T = 293.2 K. Adapted with permission from Ref. 20. Copyright © 2019 American Chemical Society*

The following reaction mechanism was previously proposed for the oxidation of catechols by copper adducts with amyloid beta, α-synuclein and prion peptide complexes.[46–48] The same mechanism was proposed also to take place for copper-tau peptide complexes, and it involves the following reaction steps:

1. Cu^{2+} + peptide ⇄ [Cu^{2+}-peptide]
2. [Cu^{2+}-peptide] + catechol → [Cu^{+}-peptide] + semiquinone$^{\bullet+}$
3. [Cu^{+}-peptide] + catechol ⇄ [Cu^{+}-peptide-catechol]
4. [Cu^{+}-peptide-catechol] + O_2 → [Cu-peptide-catechol-O_2]
5. [Cu-peptide-catechol-O_2] →→ [Cu^{2+}-peptide] + quinone

The first step of the reaction explains the first phase of the biphasic behavior of the kinetic profiles reported in ***Figure 1.3***, which for 4-methyl catechol is very clearly visible.

However, the rate-determining step is reaction 4 of the reduced form of Cu-peptide complex with dioxygen, that results into the formation of the ternary complex [Cu-peptide-catechol-O_2].

The hypothesis is that a more favourable copper(I/II) redox cycling occurs when the metal ion is coordinated to the R3 peptide. Although this hypothesis relates with an optimal coordination environment for copper in either or both its oxidation states that occurs at R3 but not at R1, no structural or thermodynamic evidence was yet provided. In this perspective, we have characterised the peptides in ***Scheme 1.1*** in terms of binding abilities to copper in both oxidation states. A particular emphasis was put on the elucidation of the speciation and coordination modes of the Cu(II)/R_{tau} systems, along with a detailed discussion of copper affinities and redox potentials. The latter will provide new insights on the previously shown oxidative behaviour of copper(II)-complexes of R1 and R3.

BINDING OF COPPER(II) TO TAU FRAGMENTS

Peptide Protonation Equilibria

The acid-base behaviour of tau fragments was explored through potentiometry, with the exception of R3C. This latter peptide is actually characterised by the presence of a cysteine residue that undergoes oxidation even at mild pH in aerobic conditions and in the presence of copper(II). Therefore, this peptide was excluded from those characterised by potentiometry.

The R1 peptide contains one acidic amino acid (Glu) which is protonated in its neutral form (HL), and other four basic residues (His, 3 Lys). Hence, when fully protonated, R1 behaves as a pentaprotic acid (H_5L^{4+}). The observed overall and stepwise protonation constants (***Table 1.1***) are in agreement with the dissociation of the carboxylic group of Glu, followed by the imidazolium group of His, and the three protonated amino groups of Lys residues, the latter ones occurring at pH > 9. However, the proton dissociation steps are macroconstants, and the dissociation processes on individual sites are partially overlapped and only marginally sequential. At physiological pH, the predominant form of R1 is H_3L^{2+}.

The R3 peptide contains three basic amino acids (2 His, 1 Lys) which are deprotonated in its neutral form (L). When fully protonated R3 therefore behaves as a triprotic acid (H_3L^{3+}). The calculated protonation constants (***Table 1.2***) are fully consistent with the dissociation of the two histidine imidazolium groups (in the pH range 4-8.5) followed by the protonated lysine side chain (at pH > 9). At pH 7.4, the predominant form of R3 is HL^+.

The R3A peptide is overall very similar to R3. It contains four basic amino acids (2 His, 2 Lys) which are deprotonated in its neutral form (L). Hence, R3A behaves as a tetraprotic acid when fully protonated (H_4L^{4+}). The calculated protonation constants (***Table 1.3***) are in agreement with the dissociation of the imidazolium group of the two His, followed by the two protonated amino groups of Lys residues, the latter ones occurring at pH > 9. At pH 7.4, the predominant form of R3A is H_2L^{2+}.

Table 1.1 *Overall protonation constants ($\log \beta_i^H$) and acid dissociation constants (pK_a) of R1 in aqueous solution (T = 298.2 K, I = 0.1 M in KCl). L^- represents the completely deprotonated form of the peptide. Standard deviations on the last significant figure are given in parenthesis.*

Reaction	$\log \beta_i^H$	Reaction	pK_a	*Diss. site*
$L^- + H^+ \rightleftarrows HL$	10.64(1)	$HL \rightleftarrows L^- + H^+$	10.64(1)	NH_3^+ Lys
$L^- + 2H^+ \rightleftarrows H_2L^+$	20.96(1)	$H_2L^+ \rightleftarrows HL + H^+$	10.32(1)	NH_3^+ Lys
$L^- + 3H^+ \rightleftarrows H_3L^{2+}$	30.65(1)	$H_3L^{2+} \rightleftarrows H_2L^+ + H^+$	9.69(1)	NH_3^+ Lys
$L^- + 4H^+ \rightleftarrows H_4L^{3+}$	37.05(2)	$H_4L^{3+} \rightleftarrows H_3L^{2+} + H^+$	6.40(2)	$N_{im}H^+$ His
$L^- + 5H^+ \rightleftarrows H_5L^{4+}$	41.21(3)	$H_5L^{4+} \rightleftarrows H_4L^{3+} + H^+$	4.16(3)	COOH Glu

Table 1.2 *Overall protonation constants ($\log \beta_i^H$) and acid dissociation constants (pK_a) of R3 in aqueous solution (T = 298.2 K, I = 0.1 M in KCl). L represents the completely deprotonated form of the peptide. Standard deviations on the last significant figure are given in parenthesis.*

Reaction	$\log \beta_i^H$	Reaction	pK_a	*Diss. site*
$L + H^+ \rightleftarrows HL^+$	10.19(1)	$HL^+ \rightleftarrows L + H^+$	10.19(1)	NH_3^+ Lys
$L + 2H^+ \rightleftarrows H_2L^{2+}$	16.94(3)	$H_2L^{2+} \rightleftarrows HL^+ + H^+$	6.75(3)	$N_{im}H^+$ His
$L + 3H^+ \rightleftarrows H_3L^{3+}$	22.72(4)	$H_3L^{3+} \rightleftarrows H_2L^{2+} + H^+$	5.78(4)	$N_{im}H^+$ His

Table 1.3 *Overall protonation constants ($\log \beta_i^H$) and acid dissociation constants (pK_a) of R3A in aqueous solution (T = 298.2 K, I = 0.1 M in KCl). L represents the completely deprotonated form of the peptide. Standard deviations on the last significant figure are given in parenthesis.*

Reaction	$\log \beta_i^H$	Reaction	pK_a	*Diss. site*
$L + H^+ \rightleftarrows HL^+$	10.50(1)	$HL^+ \rightleftarrows L + H^+$	10.50(1)	NH_3^+ Lys
$L + 2H^+ \rightleftarrows H_2L^{2+}$	20.33(1)	$H_2L^{2+} \rightleftarrows HL^+ + H^+$	9.83(1)	NH_3^+ Lys
$L + 3H^+ \rightleftarrows H_3L^{3+}$	26.88(1)	$H_3L^{3+} \rightleftarrows H_2L^{2+} + H^+$	6.56(1)	$N_{im}H^+$ His
$L + 4H^+ \rightleftarrows H_4L^{4+}$	32.55(1)	$H_4L^{4+} \rightleftarrows H_3L^{3+} + H^+$	5.67(1)	$N_{im}H^+$ His

Copper(II) Complexation Studies

The complexation of copper(II) with the tau fragments R1 and R3A was studied by potentiometric titrations in aqueous media. Analogous studies were performed on the R3 peptide and on the Cu^{2+}/R3 system by the group of Prof. Maurizio Remelli at the University of Ferrara. These latter results are presented here for the sake of completeness. Potentiometric data were complemented by visible absorption and circular dichroism characterisation of Cu^{2+}/peptide systems at different pH values, to provide information on the coordination environment of copper in the complex species. All CD spectra presented in this chapter, also those related to the R3 peptide, were collected in our laboratory. As previously anticipated, the same experiments could not be carried out for the R3C fragment since rapid oxidation of the peptide on the Cys residues was observed in the presence of copper(II) starting at low pH, as demonstrated experimentally by the marked drift observed in the e.m.f. values even at mild acidic pH conditions.

Cu^{2+}/R1 complex-formation equilibria

In the presence of copper(II), the tau fragment R1 forms seven complex species, which are reported in ***Table 1.4***. These are all characterized by a 1 to 1 copper/peptide stoichiometry regardless of the excess of ligand used and differ for their protonation states (***Table 1.4***).

Table 1.4 *Logarithms of overall formation constants (log β) and acid dissociation constants (pK_a) of the copper(II) complexes of R1 in aqueous solution (T = 298.2 K, I = 0.1 M in KCl). L^- represents the completely deprotonated form of the peptide. Standard deviations on the last significant figure are given in parenthesis.*

Reaction	*log β*	Reaction	*pK_a*
$Cu^{2+} + 3H^+ + L^- \rightleftarrows [CuH_3L]^{4+}$	34.84(8)	$[CuH_3L]^{4+} \rightleftarrows [CuH_2L]^{3+} + H^+$	6.74(9)
$Cu^{2+} + 2H^+ + L^- \rightleftarrows [CuH_2L]^{3+}$	28.1(3)	$[CuH_2L]^{3+} \rightleftarrows [CuHL]^{2+} + H^+$	5.40(6)
$Cu^{2+} + H^+ + L^- \rightleftarrows [CuHL]^{2+}$	22.70(5)	$[CuHL]^{2+} \rightleftarrows [CuL]^{+} + H^+$	8.62(9)
$Cu^{2+} + L^- \rightleftarrows [CuL]^{+}$	14.08(8)	$[CuL]^{+} \rightleftarrows Cu^{2+} + L^-$	14.08(8
$Cu^{2+} + L^- \rightleftarrows [CuLH_{-1}] + H^+$	4.46(7)	$[CuL]^{+} \rightleftarrows [CuLH_{-1}] + H^+$	9.6(1)
$Cu^{2+} + L^- \rightleftarrows [CuLH_{-2}]^- + 2H^+$	-5.85(9)	$[CuLH_{-1}] \rightleftarrows [CuLH_{-2}]^- + H^+$	10.3(1)
$Cu^{2+} + L^- \rightleftarrows [CuLH_{-3}]^{2-} + 3H^+$	-16.35(7)	$[CuLH_{-2}]^- \rightleftarrows [CuLH_{-3}]^{2-} + H^+$	10.5(1)

A representative speciation distribution diagram is reported in ***Figure 1.4***. The first complex species, $[CuH_3L]^{4+}$, is formed at pH 4 and reaches ca. 60% of total copper at pH 5.8. The lack of spectroscopic information (***Figure 1.5***) in this pH range allows only to assume that in this species the peptide acts as a monodentate ligand through the imidazole

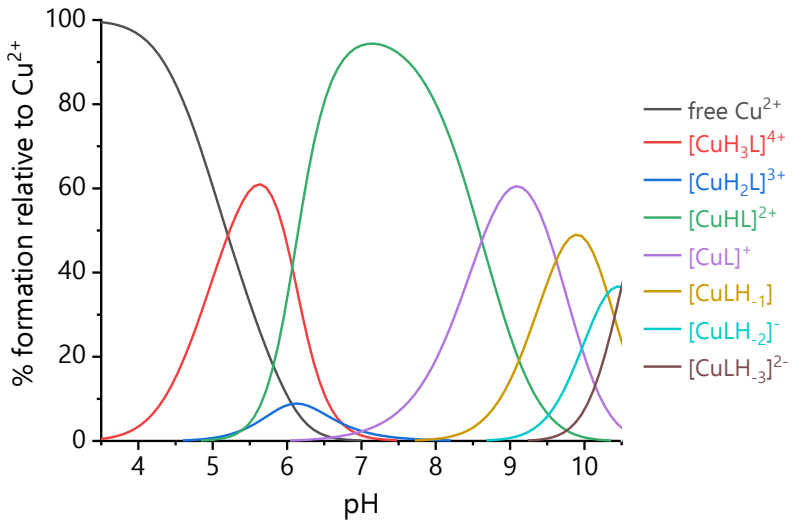


Figure 1.4 *Representative distribution diagram of the copper(II)/R1 (LH) system in aqueous solution (Cu^{2+}/R1 = 1:3, C_{Cu} = 0.47 mM, T = 298.2 K, I = 0.1 M in KCl).*

of His, with possible involvement of the carboxylate of Glu. At pH 5 two other species start forming: $[CuH_2L]^{3+}$ and $[CuHL]^{2+}$. However, while the former does not exceed 8%, $[CuHL]^{2+}$ reaches 95% of total copper at pH 7.2 and is predominant in the pH range 6-8.5. The spectra collected at pH 6.7 and 7.3 (***Figure 1.5***) show an absorption λ_{max} of *ca.* 584 nm. Exploiting Billo's method of the average environment,[49–51] this wavelength was ascribed to the presence in $[CuHL]^{2+}$ of one imidazole nitrogen, two deprotonated peptide nitrogen atoms and a water molecule coordinated on the equatorial plane of copper(II) (***Scheme 1.2***, left; calc. λ_{max} = 583 nm).[49–51] The involvement of the imidazole Nδ and of the deprotonated peptide nitrogen of His268 would result in a 6-membered chelation ring. On the contrary, a less stable 7-membered chelation ring would be obtained in the hypothesis of coordination of the peptide nitrogen of the successive Gln. The preceding Lys (Lys267) is thus more likely to provide the second peptide nitrogen coordinating Cu^{2+}. Two water molecules presumably occupy the axial positions of copper(II). On the basis of this proposed coordination mode, it is possible to put forward the hypothesis that the coordination of copper(II) in $[CuH_2L]^{3+}$ is (N_{im}, N^-, 2x H_2O).

When the pH is increased from 7.3 to 9.1, a shift from 584 to 523 nm of the ligand field transition maximum is observed (***Figure 1.5***). At this pH, as expected from the distribution diagram (***Figure 1.4***), $[CuL]^+$ is the most abundant species accounting for *ca.* 60% of total copper. However, *ca.* 25% of total copper is still in the form $[CuHL]^{2+}$. The shoulder at higher wavelengths is ascribable to the presence of the latter species, while the absorption maximum at 523 nm is very well explained by the presence in $[CuL]^+$ of one imidazole and three peptide nitrogen atoms on the equatorial plane of Cu^{2+} (***Scheme 1.2***, right; calc. λ_{max} = 523 nm)[49–51]. No shift of the absorption maximum and no significant

change in the shape of CD spectra were observed moving from pH 9.1 to 10.5 (***Figure 1.5***). This suggests that the coordination sphere of Cu^{2+} is the same in each of the species that are predominant above pH 8.6. Hence, it can be stated that the three deprotonation processes leading from $[CuL]^{+}$ to $[CuLH_{-3}]^{2+}$ mainly involve the three Lys residues, which remain uncoordinated.

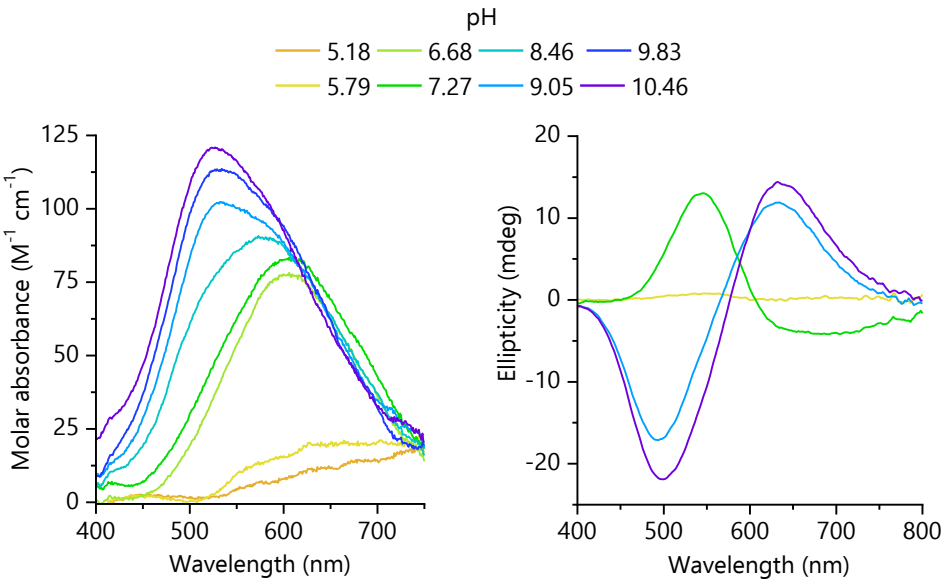
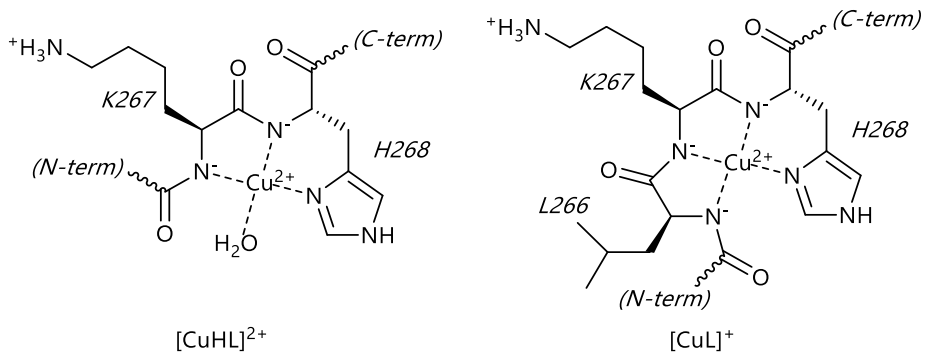


Figure 1.5 *Left: molar visible absorption spectra (absorbance/C_{Cu}) of solutions of copper(II) and R1 (Cu^{2+}/R1 = 1:1.5, C_{Cu} = 0.39 mM, T = 298.2 K, I = 0.1 M in KCl). Right: visible CD spectra of solutions of copper(II) and R1 (Cu^{2+}/R1 = 1:1.5, C_{Cu} = 0.52 mM, T = 298.2 K, I = 0.1 M in KCl).*



Scheme 1.2 *Schematic representation of copper(II) coordination in $[CuHL]^{2+}$ and $[CuL]^{+}$ species of the R1 peptide (R1 = LH). Adapted with permission from Ref. 20. Copyright © 2019 American Chemical Society*

Cu^{2+}/R3 complex-formation equilibria

In the presence of copper(II), the R3 peptide forms six complex species, all with a 1 to 1 copper/peptide stoichiometry and differing for their protonation states (***Table 1.5***). The analysis of the potentiometric curves suggests no formation of both 2:1 and 1:2 copper(II)/peptide species under the experimental conditions used.

Table 1.5 *Logarithms of overall formation constants (log β) and acid dissociation constants (pK_a) of the copper(II) complexes of R3 in aqueous solution (T = 298.2 K, I = 0.1 M in KCl). L represents the completely deprotonated form of the peptide. Standard deviations on the last significant figure are given in parenthesis.*

Reaction	*log β*	Reaction	*pK_a*
$Cu^{2+} + 2H^+ + L \rightleftarrows [CuH_2L]^{4+}$	20.48(7)	$[CuH_2L]^{4+} \rightleftarrows [CuHL]^{3+} + H^+$	4.67(7)
$Cu^{2+} + H^+ + L \rightleftarrows [CuHL]^{3+}$	15.81(2)	$[CuHL]^{3+} \rightleftarrows [CuL]^{2+} + H^+$	6.42(4)
$Cu^{2+} + L \rightleftarrows [CuL]^{2+}$	9.39(3)	$[CuL]^{2+} \rightleftarrows Cu^{2+} + L$	9.39(3)
$Cu^{2+} + L \rightleftarrows [CuLH_{-1}]^+ + H^+$	2.48(2)	$[CuL]^{2+} \rightleftarrows [CuLH_{-1}]^+ + H^+$	6.91(4)
$Cu^{2+} + L \rightleftarrows [CuLH_{-2}] + 2H^+$	-6.22(3)	$[CuLH_{-1}]^+ \rightleftarrows [CuLH_{-2}] + H^+$	8.70(4)
$Cu^{2+} + L \rightleftarrows [CuLH_{-3}]^- + 3H^+$	-16.31(4)	$[CuLH_{-2}] \rightleftarrows [CuLH_{-3}]^- + H^+$	10.09(5)

Figure 1.6 reports a representative distribution diagram for the Cu(II)/R3 system. The complexation of copper(II) starts at pH 3.5 with the formation of $[CuH_2L]^{4+}$, immediately followed by $[CuHL]^{3+}$ that reaches *ca.* 75% of total copper at pH 5.8. The broad band at *ca.* 690 nm observed in the visible absorption spectrum collected at pH 5.7 (***Figure 1.7***) is well accounted for a (2x N_{im}, 2x H_2O) donor set as calculated on the basis of the Billo's average environment theory (calc. λ_{max} = 692 nm).[49–51]

A blue shift of the d-d band to *ca.* 610 nm is observed when increasing the pH to 6.7, where the $[CuL]^{2+}$ species is predominant (48% of total copper). For this species it is reasonable to propose an equatorial coordination mode that involves one deprotonated peptide nitrogen, two imidazole donors and a water molecule, for which the expected λ_{max} is 608 nm (***Scheme 1.3***, top).

From pH 7 to 8.5 $[Cu(LH_{-1})]^+$ dominates the speciation, accounting for 80% of total copper at pH 7.8 (***Figure 1.6***). At this pH, an absorption λ_{max} of 585 nm is observed (***Figure 1.7***), which suggests a (N_{im}, 2x N^-, H_2O) coordination mode for copper(II) in this species (calc. λ_{max} = 584 nm; ***Scheme 1.3***, bottom left). One of the possible coordination environments proposed[20] involves the presence in the equatorial position of the imidazole and the peptide nitrogen atoms of His329, together with the peptide nitrogen of the preceding Ile328 residue. The second imidazole donor of His330 may be either coordinated to the axial position of Cu^{2+} or uncoordinated. The proposed changes in the coordination of Cu^{2+} from $[CuL]^{2+}$ to $[CuLH_{-1}]^+$ are supported by the major change in the CD spectrum

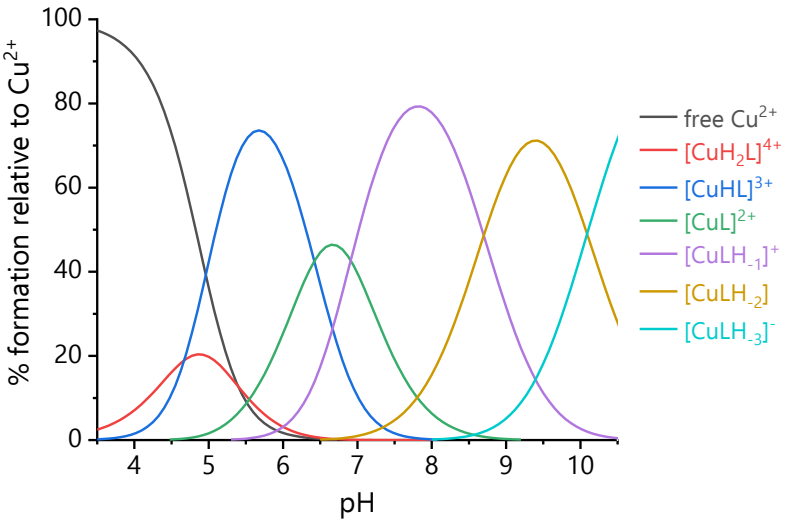


Figure 1.6 *Representative distribution diagram of the copper(II)/R3 (L) system in aqueous solution (Cu^{2+}/R3 = 1:3, C_{Cu} = 0.47 mM, T = 298.2 K, I = 0.1 M in KCl).*

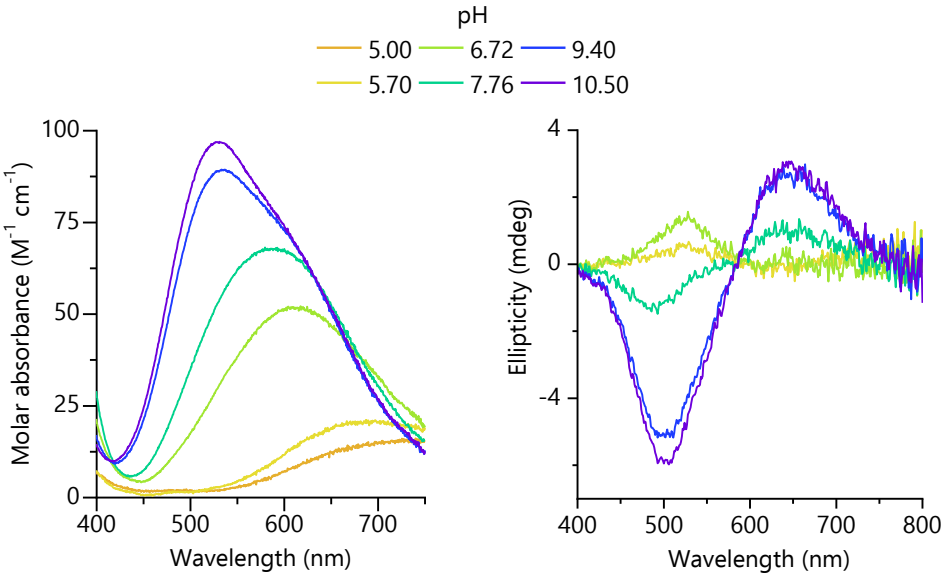
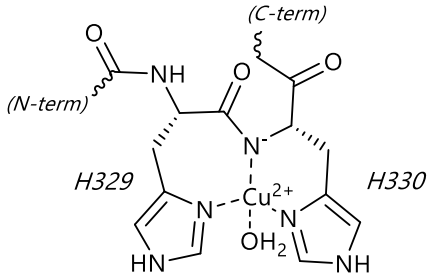


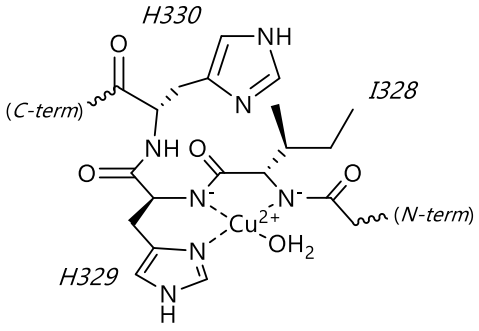
Figure 1.7 *Left: molar visible absorption spectra (absorbance/C_{Cu}) of solutions of copper(II) and R3 (Cu^{2+}/R3 = 1:1.25, C_{Cu} = 0.60 mM, T = 298.2 K, I = 0.1 M in KCl). Right: visible CD spectra of solutions of copper(II) and R3 (Cu^{2+}/R3 = 1:3, C_{Cu} = 0.51 mM, T = 298.2 K, I = 0.1 M in KCl).*

occurring in the pH range 6.7 to 7.8 (***Figure 1.7***). The $[CuLH_{-1}]^+$ species further dissociates to $[CuLH_{-2}]$ (*ca.* 70% of total copper at pH 9.5) with the formation of a (N_{im}, 3x N^-) equatorial coordination plane, where the coordinated groups are one imidazole donor and three deprotonated peptide nitrogen atoms (***Scheme 1.3***, bottom right). The observed absorption maximum at 533 nm is slightly higher than the one expected for the proposed coordination environment (calc. λ_{max} = 523 nm), suggesting that the second imidazole is coordinated to the axial position.

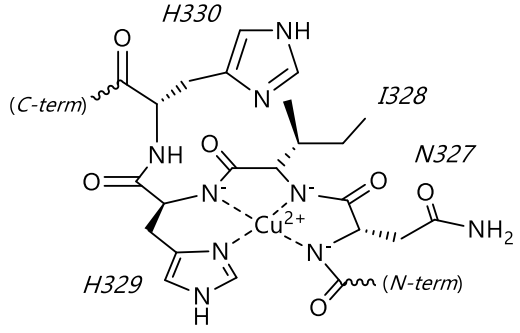
Moving to pH 10.5 no shift of the absorption maximum and no significant change in the pattern of CD spectra were observed, implying that the last deprotonation step that leads from $[CuLH_{-2}]$ to $[CuLH_{-3}]^-$ is associated to the deprotonation of the Lys residue, which remains uncoordinated.



$[CuL]^{2+}$ for R3, $[CuHL]^{3+}$ for R3A



$[CuLH_{-1}]^+$ for R3, $[CuL]^{2+}$ for R3A



$[CuLH_{-2}]$ for R3, $[CuLH_{-1}]^+$ for R3A

Scheme 1.3 *Schematic representation of copper(II) coordination in the* $[CuL]^{2+}$, $[CuLH_{-1}]^+$ *and* $[CuLH_{-2}]$ *species of the peptide R3 (R3 = L), and in the* $[CuHL]^{3+}$, $[CuL]^{2+}$ *and* $[CuLH_{-1}]^+$ *species of the peptide R3A (R3A = L). Adapted with permission from Ref. 20. Copyright © 2019 American Chemical Society*

Cu^{2+}/R3A complex-formation equilibria.

In the presence of copper(II), the R3A peptide forms six complex species, all with a 1 to 1 copper/peptide stoichiometry and differing for their protonation states (***Table 1.6***). The analysis of the potentiometric curves suggests no formation of both 2:1 and 1:2 copper(II)/peptide species under the experimental conditions used.

Table 1.6 *Logarithms of overall formation constants (log β) and acid dissociation constants (pK_a) of the copper(II) complexes of R3A in aqueous solution (T = 298.2 K, I = 0.1 M in KCl). L represents the completely deprotonated form of the peptide. Standard deviations on the last significant figure are given in parenthesis.*

Reaction	*log β*	Reaction	*pK_a*
$Cu^{2+} + 2H^+ + L \rightleftarrows [CuH_2L]^{4+}$	25.79(1)	$[CuH_2L]^{4+} \rightleftarrows [CuHL]^{3+} + H^+$	6.37(2)
$Cu^{2+} + H^+ + L \rightleftarrows [CuHL]^{3+}$	19.42(2)	$[CuHL]^{3+} \rightleftarrows [CuL]^{2+} + H^+$	6.74(2)
$Cu^{2+} + L \rightleftarrows [CuL]^{2+}$	12.68(2)	$[CuL]^{2+} \rightleftarrows Cu^{2+} + L$	12.68(2)
$Cu^{2+} + L \rightleftarrows [CuLH_{-1}]^+ + H^+$	4.09(3)	$[CuL]^{2+} \rightleftarrows [CuLH_{-1}]^+ + H^+$	8.59(2)
$Cu^{2+} + L \rightleftarrows [CuLH_{-2}] + 2H^+$	-5.84(3)	$[CuLH_{-1}]^+ \rightleftarrows [CuLH_{-2}] + H^+$	9.93(3)
$Cu^{2+} + L \rightleftarrows [CuLH_{-3}]^- + 3H^+$	-16.40(4)	$[CuLH_{-2}] \rightleftarrows [CuLH_{-3}]^- + H^+$	10.56(5)

As it can be seen from the distribution diagram reported in ***Figure 1.8***, the predominant species below pH 6.2 is $[CuH_2L]^{4+}$, which starts forming at pH 4 and reaches *ca.* 80% of total copper at pH 5.6. The visible absorption spectrum recorded at pH 5.7 (***Figure 1.9***) shows a broad band at *ca.* 660 nm, that is well accounted for a (N_{im}, N^-, 2x O_w) donor set as calculated on the basis of the Billo's average environment theory (calc. λ_{max} 661 nm).[49–51] This coordination environment may occur at both His329 or His330 binding sites, by the involvement of both imidazole and peptidic nitrogen atoms.

By increasing the pH, the species $[CuHL]^{3+}$ becomes predominant, and at pH 6.6 it accounts for *ca.* 40% of total copper. The observed λ_{max} at *ca.* 610 nm at pH 6.7 (***Figure 1.9***) suggests a (2x N_{im}, N^-, O_w) coordination mode for copper(II) in this species (calc. λ_{max} 608 nm; ***Scheme 1.3***, top). The absorption band at pH 6.7 is broad towards lower wavelengths due to the spectral contribution of the $[CuL]^{2+}$ species, which is present at *ca.* 30% of total copper and which absorbs at *ca.* 580 nm. This species dominates the speciation at pH 7.4 (*ca.* 75%) and reaches a maximum of *ca.* 80% of total copper at pH 7.6. The presence in the equatorial plane of copper(II) of one imidazole nitrogen, two deprotonated peptide nitrogen atoms and a water molecule (calc. λ_{max} = 584 nm) is entirely consistent with the experimental λ_{max} of *ca.* 584 nm observed at pH 7.6 (***Figure 1.9***). As previously hypothesised for R3, the imidazole coordinated to the equatorial site on copper(II) is likely that of His329, while the peptide nitrogen atoms are those of His329 and Ile328.

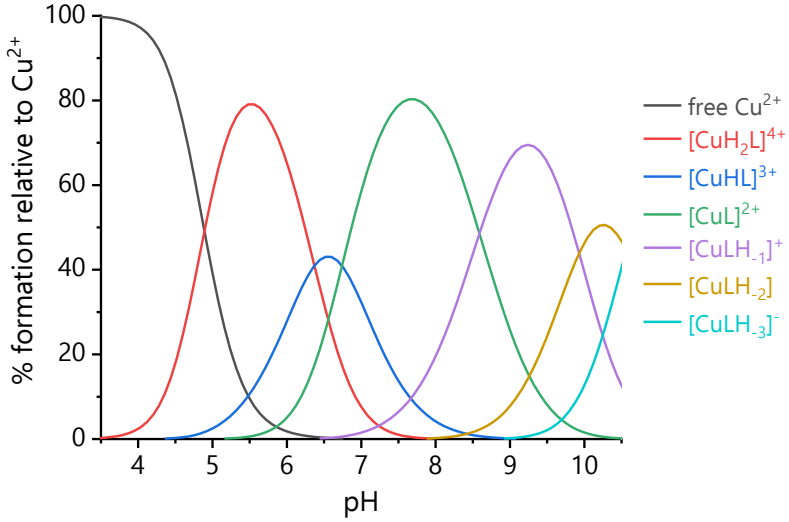


Figure 1.8 *Representative distribution diagram of the copper(II)/R3A (L) system in aqueous solution (Cu^{2+}/R3A = 1:2.1, C_{Cu} = 0.70 mM, T = 298.2 K, I = 0.1 M in KCl).*

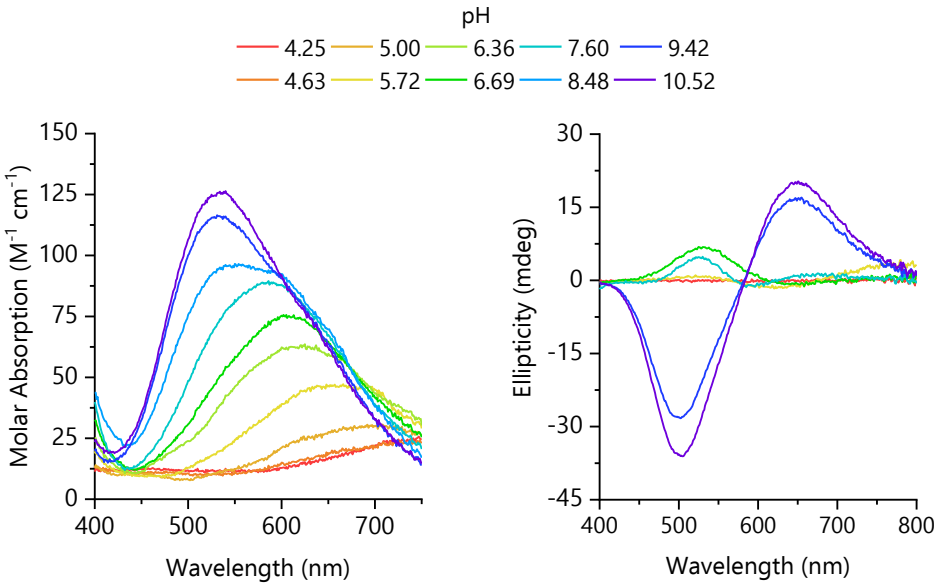


Figure 1.9 *Left: molar visible absorption spectra (absorbance/C_{Cu}) of solutions of copper(II) and R3A (Cu^{2+}/R3A = 1:2.1, C_{Cu} = 0.70 mM, T = 298.2 K, I = 0.1 M in KCl). Right: visible CD spectra of solutions of copper(II) and R3A (Cu^{2+}/ R3A = 1:2.1, C_{Cu} = 0.52 mM, T = 298.2 K, I = 0.1 M in KCl).*

This coordination mode also allows the imidazole of His330 to possibly coordinate copper(II) in an axial position (***Scheme 1.3***, bottom left). Moving from pH 7.6 to 9.4 the absorption maximum shifts to 525 nm with the appearance, at pH 8.5, of two overlapped absorption bands. These spectral data are consistent with a (N_{im}, 3x N^-) coordination mode of the $[CuLH_{-1}]^+$ species, for which a λ_{max} of 523 nm is actually expected (***Scheme 1.3***, bottom right). Importantly, no shift of the absorption maximum and no significant change in the pattern of CD spectra were observed increasing the pH from 9.4 to 10.5 (***Figure 1.9***). This suggests that the coordination sphere of Cu^{2+} is the same in each of the species that are predominant above pH 8.6. Therefore, the two deprotonation processes leading from $[CuLH_{-1}]^+$ to $[CuLH_{-3}]^-$ most likely involve the two Lys residues, which remain uncoordinated.

Finally, the speciation diagram in ***Figure 1.8*** parallels that previously reported for the shorter R3 fragment (***Figure 1.6***), which also lacks the cysteine residue. This strongly supports the hypothesis that the two peptides have the same behaviour in terms of copper(II) binding.

Copper(II) binding affinities

The speciation models obtained through potentiometry allowed to calculate the conditional formation or *binding* constants (K_f^{II}) for all Cu(II)/peptide systems. The K_f^{II} value (II refers to the +2 oxidation state of copper) is pH-depending, and it describes the binding capacity of a ligand for a metal under those pH conditions. It is defined as:

$$K_f^{II} = \frac{[\text{Cu(II) bound to L}]}{[\text{Cu(II) not bound to L}][\text{L not bound to Cu(II)}]}$$

where:

$$[\text{Cu(II) bound to L}] = [\text{CuL}] + [\text{CuHL}] + \ldots$$
$$[\text{Cu(II) not bound to L}] = [\text{Cu(II)}] + [\text{Cu(OH)}^+] + \ldots$$
$$[\text{L not bound to Cu(II)}] = [\text{L}] + [\text{HL}] + \ldots$$

In terms of affinity or dissociation constant:

$$K_d^{II} = \frac{1}{K_f^{II}}$$

The conditional binding constants (K_f^{II}) at pH 7.4 for the copper(II)/R_{tau} systems are reported in ***Table 1.7***.

Table 1.7 *Logarithms of conditional binding constants (log K_f^{II}) and conditional affinities (K_d^{II}) for Cu(II)/R1, Cu(II)/R3 and Cu(II)/R3A systems at pH 7.4, calculated with the Hyss2009 software*[52] *using the speciation models reported in* ***Table 1.4***, ***Table 1.5*** *and* ***Table 1.6***, *respectively.*

Species	log K_f^{II}	K_d^{II}
[Cu^{II}R1]	6.8(3)	150(10) nM
[Cu^{II}R3]	7.1(3)	71(5) nM
[Cu^{II}R3A]	6.9(3)	119(9) nM

The presence of two His donors in R3 results in a higher affinity for copper(II) with respect to R1. However, R3 is not selective in binding copper, as shown by the calculated competition diagram reported in ***Figure 1.10***. In a hypothetical system where copper(II), R1 and R3 are present in equimolar amounts, the relative amount of copper(II) bound to R3 and R1 at pH 7.4 is 58% *vs.* 42%, respectively. When the R1:R3 ratio increases to 3:1 as in tau (approximating the R2 and R4 sites equal to R1 in terms of copper(II) binding), the percentage of metal bound to R3 in a 1:3:1 Cu^{2+}:R1:R3 system decreases to 36%.

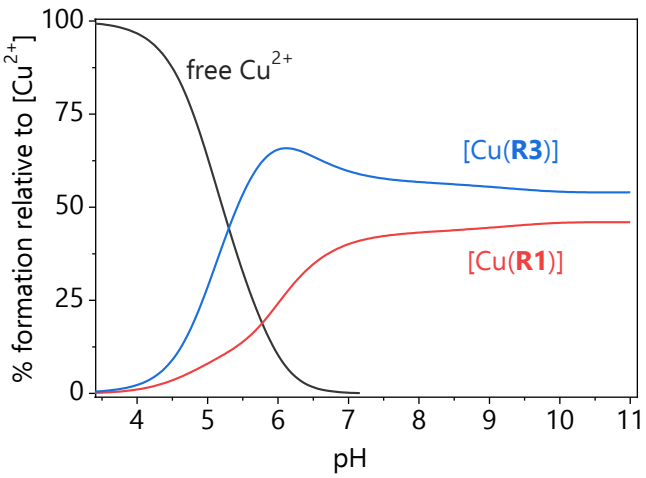


Figure 1.10 *Competition diagram for a Cu^{2+}/R1/R3 (1:1:1) system (C_{Cu} = 100 μM), calculated with the Hyss2009 software*[52] *using the speciation models reported in* ***Table 1.4*** *and* ***Table 1.5***. *[Cu(R)] = 100 x (Conc. of Cu^{2+} bound to R)/total C_{Cu}; R = R1 or R3.*

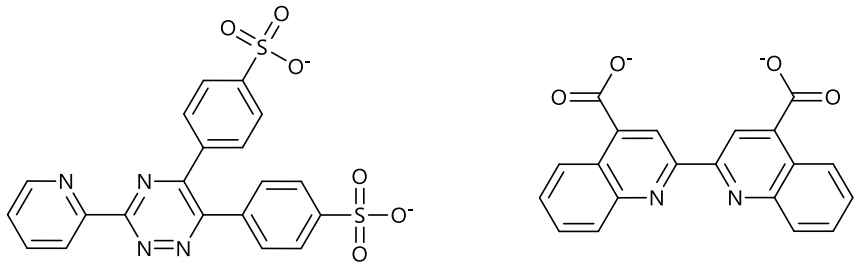
Interestingly, the affinity of the longer R3A peptide for Cu^{2+} is closer to that of R1 rather than R3 (***Table 1.7***). This suggests that the affinity for copper(II) of the His-His dyad of the R3 model peptide possibly overestimates that of R3 of the full-length tau, and that this affinity is likely to be sequence-dependent. Overall, these results show that the presence of a His-His tandem site (as in R3) is not sufficient to define a unique copper(II) binding site if another or some other sites that bear one His residue are present (such as

those in R1, R2 and R4 of tau). Therefore, the metal in tau is reasonably distributed over multiple sites, and the preference for a specific site suggested in the literature[53] may be a consequence of structural features that are not modelled using these R_{tau} peptides. Staying on the perspective of model peptides, however, it is evident that it is not the different stability of the copper(II)/R1 or R3 peptide adducts that modulates their redox activities. This is demonstrated by the thermodynamic data (log K_f^{II} and K_d^{II} values) which show that one equivalent of copper(II) is almost entirely bound to the peptides with similar stability, irrespective of the presence of one or two His residues. Hence, the attention was then focused on the binding to these fragments of copper in its reduced +1 oxidation state.

BINDING OF COPPER(I) TO TAU FRAGMENTS

The stability of copper(I) adducts with the tau fragments R1, R3, R3A and R3C was determined by spectrophotometric titrations of solutions containing $[Cu^{I}(CH_3CN)_4]^+$ using either ferrozine (Fz) or bicinchoninate (Bca, ***Scheme 1.4***). Both are metallochromic indicators for copper(I): in the presence of this metal ion, they form coloured complexes with 1:1 and 1:2 copper/indicator stoichiometries. Importantly, the complexes of Bca were found particularly stable, as reported in the literature.[54,55]

Experimentally, to determine the affinity of a peptide for copper(I), coloured solutions of Cu(I)/Fz or Cu(I)/Bca adducts can be back-titrated with the R_{tau} peptides.[54,55] Knowing the stability of the copper(I)/metallochromic indicator complexes, the stability of Cu^+/peptide adducts could be calculated. In the case of R_{tau} peptides, the use of ferrozine resulted appropriate for R1, R3 and R3A (log $\beta_{[CuFz2]}$ = 15.1). On the contrary, bicinchoninate was used for R3C, as the adduct of copper(I) with this peptide resulted of remarkable stability (log $\beta_{[CuBca2]}$ = 16.7).



Scheme 1.4 *Representation of the metallochromic indicators for copper(I) used in this study.*

The spectrophotometric titrations of Cu(I) solutions containing and Fz or Bca at pH 7.4 with the four different peptides in aqueous HEPES buffer are reported in ***Figure 1.11***. All datasets show a decrease in the intensity of the visible absorption band of the Cu(I)/indicator adducts. In the case of R3 and both its longer analogues, this is imputable to the competition between the indicators and the peptides in copper(I) binding. For R1, on the contrary, the decrease in absorbance is small and fully ascribable to dilution effects.

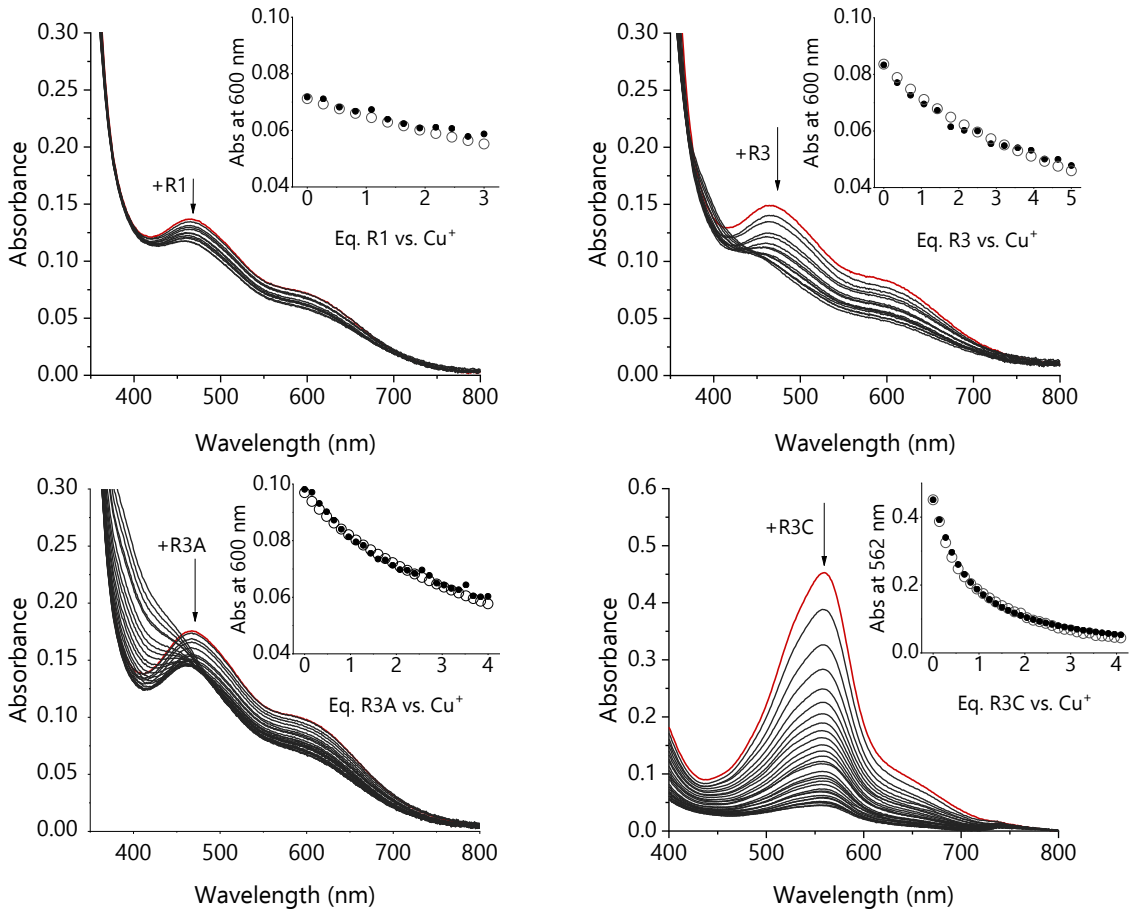


Figure 1.11 *Spectral data sets for the titration of solutions of $[Cu^{I}(CH_3CN)_4]BF_4$ and a metallochromic indicator (L) with: 3.0 eq. of R1 (top left: L = ferrozine, Cu^{I}/L = 1:2.15, C_{Cu} = 41 μM), 5.0 eq. of R3 (top right: L = ferrozine, Cu^{I}/L = 1:2.15, C_{Cu} = 41 μM), 4.0 eq. of R3A (bottom left: L = ferrozine, Cu^{I}/L = 1:2.15, C_{Cu} = 55 μM), and 4.2 eq. of R3C (bottom right: L = sodium bicinchoninate, Cu^{I}/L = 1:2.5, C_{Cu} = 55 μM). For all datasets: $C_{ascorbate}$ = 10 mM, 100 mM aqueous HEPES buffer solution, pH 7.4. Insets: observed (filled circles) and calculated (open circles) absorbance values at 600 nm (R1, R3, R3A) and 562 nm (R3C).*

Data treatment allowed to calculate the log K_f^I values of Cu(I)/peptide adducts, where the apex I indicates copper(I) (***Table 1.8***). More specifically, the dataset reported in ***Figure 1.11*** were treated taking into account the formation of 1:1 Cu(I)/peptide adducts, *i.e.* $Cu^+ + R_{tau} = [Cu^IR_{tau}]$ equilibria. To treat these data, the speciation of the Cu(I)/Fz and Cu(I)/BCA systems reported in the literature were used.[54,55]

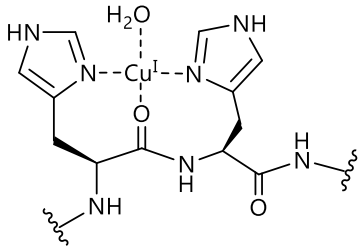
Table 1.8 *Logarithms of K_f^I values for the $Cu^+ + R_x = [Cu^IR_x]$ equilibria, and dissociation constants (K_d^I) of $[Cu^IR_x]$ complexes (R_x = R1, R3, R3A and R3C).*

Species	log K_f^I	K_d^I
$[Cu^IR1]$	nd	nd
$[Cu^IR3]$	10.2(1)[a]	$6.3\ 10^{-11}$ M[a]
$[Cu^IR3A]$	10.5(1)	$3.2\ 10^{-11}$ M
$[Cu^IR3C]$	13.4(1)	$4.0\ 10^{-14}$ M

a Re-evaluated values with respect to Ref. 20 following NMR standardization of R3 stock solution.

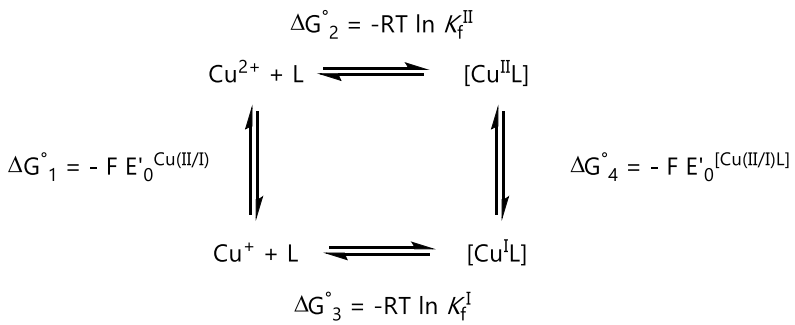
Since the experiments were carried out in aqueous HEPES buffer, these K_f^I are apparent constants. However, there is consensus in the literature for not considering HEPES as a strong competing ligand for copper(I).[56] Hence, it is reasonable to expect the values of these apparent formation constants to be not significantly different from conditional ones.

Values in ***Table 1.8*** suggest that R1, which bears a one-His site, does not compete with ferrozine in binding copper(I). This should result in a log K_f^I value lower than 5, estimated using the Hyss software. On the contrary, the affinity values obtained for R3 and its longer analogue R3A are in the high picomolar range and in agreement with those reported in the literature for systems containing a His-His tandem site, such as histatin peptides and models of Aβ and Ctr1.[56–59] These stability constants suggest that both the R3 and R3A peptides are likely to be coordinated to copper(I) by the two N_{im} of His329 and His330, with a possible weak interaction of the carbonyl oxygen of His329 (***Table 1.8***), as reported in the literature.[59] A representation of the coordination site is reported in ***Scheme 1.5***. Conversely, the high femtomolar affinity of R3C suggests that in the Cu(I)/R3C adducts the softer Cys322 residue is coordinated to copper(I) at pH 7.4, although no insights on the coordination environment of the metal ion can be obtained from current data.



Scheme 1.5 *Representation of the coordination of copper(I) by two N_{im} donors.*

Using the K_f^I and K_f^{II} constants it was possible to estimate the redox potential for the reduction of the adducts of Cu^{2+} with the R1, R3 and R3A peptides, to obtain the Cu^+/adducts. In this calculation, the thermodynamic box reported in ***Scheme 1.6*** was used. Different values with respect to those reported in Ref. [60] were obtained after a re-evaluation of the electrochemical reduction potential for aqueous Cu^{2+} using the values reported in the literature. $E'^{0}{}^{Cu(II/I)}$ is the reduction potential of free Cu^{2+} in aqueous solution, for which a value of +0.26 mV (*vs.* SHE) is reported in the literature,[61] consistently with the data in Ref. [62]. By using the Hess' law, the reduction potential for the formal adducts of Cu^{2+} with R1, R3 and R3A ($E'^{0}{}^{[Cu(II/I)L]}$) can be calculated. For the sake of clarity, [Cu^{II}L] in ***Scheme 1.6*** represents the ensemble of the complex species of copper(II) (in particular [CuHL] and [CuL], charges omitted). These species are present in concentrations defined by the speciation models reported in ***Table 1.4***, ***Table 1.5*** and ***Table 1.6***.



Scheme 1.6 *Representation of the thermodynamic box used for the calculation of the reduction potential of the Cu(II)/R_{tau} adducts to Cu(I)/R_{tau}.*

Table 1.9 *Reduction potentials (mV vs. NHE) for the reduction of the Cu^{2+} adduct with the R1, R3 and R3A peptides. Standard deviations are reported in parentheses.*

Species	$E'^{0}{}^{[Cu(II/I)L]}$
[Cu^{II}R1]	150(10) mV[a]
[Cu^{II}R3]	440(10) mV
[Cu^{II}R3A]	470(10) mV
[Cu^{II}tau]	340 mV[b]
MC	460 mV[c]

a log K_f^I put equal to 5 as the upper limit for R1 not competing with Fz. Therefore $E'^{0}{}^{[Cu(II/I)L]}$ for R1 represents an upper-limit value; b see Ref. 63; c see Ref. 64.

The values reported in ***Table 1.9*** reveal that the estimated reduction potential for the copper(II) ion coordinated to the R3 and R3A peptides exceeds the one determined for tau of *ca.* 100 mV. However, contrarily to our previously published data, the values in

Table 1.9 are more consistent with the observed capacity of the copper(II) adducts with R3 peptides only to promote the oxidation of MC, but not of that of copper(II) with R1. Although the reduction potential of copper(II) coordinated to the His-His site resulted about 100 mV higher than that of copper(II) coordinated to tau protein, it evidences that among the R1 (R2 or R4) and the R3 site, likely the redox Cu(II/I) cycle occurs at the latter repeat in the protein.

CONCLUSIONS

Despite the relevance of tau protein and its homeostasis in the onset and progression of neurodegenerative diseases, the interaction of tau with copper has been only partially characterised. In this work, four tau fragments that are comprised in the R1 and R3 pseudorepeats were thoroughly studied in terms of copper(II)/copper(I) complexes stability and structural models. Potentiometric analyses of the R1, R3 and R3A peptides allowed to obtain the speciation models for copper(II) binding. R3C was excluded from potentiometric characterisation, given the presence in its sequence of a cysteine residue that undergoes oxidation even at mild pH in aerobic conditions and in the presence of copper(II).

The most important result is the elucidation of the speciation at physiological pH. All the peptides prevalently bind copper(II) in a (N_{im}, 2x N^-, H_2O) coordination mode, although the peptides containing a His-His dyad (R3 and R3A) form an additional complex species in which both the imidazole groups of the dyad bind the metal in a (2x N_{im}, N^-, H_2O) coordination mode. Also, the study of copper(I) binding to R1, R3, R3A, and R3C suggested a much higher affinity for Cu(I) of the R3-like peptides with respect to R1.

Notably, it was successfully demonstrated that the presence of two adjacent His residues in R3 results in a strong binding of copper in both oxidation states, while the single His residue in R1 can only bind copper(II). In the case of R3 and R3A the K_d values are of the order of high nanomolar, which are however close to that of R1. The difference in the catalytic activity pivots around the affinity for copper(I). In fact, while R3 has an affinity for copper(I) in the order of picomolar, this affinity is negligible for R1. It is also interesting that a femtomolar copper(I) affinity of R3C was observed, ascribed to the involvement of the Cys residue in the coordination of the metal in its reduced oxidation state. Finally, the estimated redox potential for the $Cu^{II/I}$/adducts with the R1, R3 and R3A peptides, allowed to rationalise the catalytic behaviour of the Cu(II)/R1 and Cu(II)/R3 adducts in the oxidation of catechols.

Overall, this work suggests that in metalated tau the redox Cu(II/I) cycle likely occurs at the R3 repeat. Although the studied R_{tau} peptides cannot fully model the structural features of full-length tau they were useful for shedding light on still undisclosed features of copper/tau adducts.

EXPERIMENTAL SECTION

Tau fragments

R1	Ac-^{256}VKSKIGSTENLKHQPGGG^{273}-NH_2	mw 1878
R3	Ac-^{323}GSLGNIHHKPGGG^{335}-NH_2	mw 1271
R3A	Ac-^{318}VTSKAGSLGNIHHKPGGG^{335}-NH_2	mw 1758
R3C	Ac-^{318}VTSKCGSLGNIHHKPGGG^{335}-NH_2	mw 1790

All peptides were provided by Prof. L. Casella's research group (University of Pavia). Tau fragments were synthesised by solid-phase peptide synthesis using standard Fmoc strategy and purified through HPLC.[60]

Copper(II) binding to Tau fragments

Formation equilibria studies. Concentrated (*ca.* 1-10 mM) solutions of R1 and R3 were prepared by weight in high-purity D_2O, and the peptide concentration was determined by NMR using an external standard. The titre was consistent with what obtained by potentiometric titrations. The stock solution of R3A was prepared at 4.5 mM concentration in Milli-Q water with the addition of 1 eq. of HCl. The titre of the solution was checked potentiometrically. $CuCl_2$ stock solutions were prepared by weight from analytical grade metal salt and standardized by complexometric titrations with EDTA following standard protocols.[65] Standard KOH solutions were prepared by diluting carbon dioxide-free concentrated solutions in freshly boiled, Milli-Q water under N_2 atmosphere. The title of KOH solutions (*ca.* 0.05 M) was determined by potentiometric titrations against potassium hydrogen phthalate. Solutions of HCl in water were prepared by dilution of concentrated solutions of the acids and their titre determined by titrations with standard KOH using the Gran's method.[66]

Potentiometric measurements. Potentiometric titrations were carried out in aqueous solution at $T = 298.2$ K and $I = 0.1$ M (KCl) under a N_2 stream, using 1.5 to 2 mL samples. The potentiometric apparatus consisted of an automatic system combining a Metrohm Dosimat 665 automatic burette (for titrant addition) and an Orion 720A pH-meter provided with a Metrohm semimicro combined glass electrode (for pH measurements). A PC controlled the addition of titrant aliquots and monitored the equilibrium attainment through pH measurements *vs.* time. The semimicro combined glass electrode was calibrated in terms of $[H^+]$ by titrating HCl solutions with standard KOH solutions. The electrode parameters were determined using the Gran's method, and the pK_w value resulted to be 13.76(1).[66] Protonation studies of the peptides were carried out by alkalimetric titrations of the peptides in the pH range 3-11 ($C_{peptide}$ = 0.48-0.71 mM). Cu^{2+} complex formation constants were determined by alkalimetric titrations in the pH range 3-11 (Cu:R1 = 1:1.5-5.2, Cu:R3A = 1: 1.2-2.1, C_{Cu} = 0.11-0.47 mM). A minimum of two replicate titrations

were performed for each system. Overall equilibrium constants $\beta_{M_mH_hL_l}$ (being $\beta_{M_mH_hL_l} = [M_mH_hL_l]/[M]^m[H]^h[L]^l$) were calculated by fitting the potentiometric data from protonation or complexation titrations with the HyperQuad 2013 software.[67] The calculated constants were used to draw the species distribution curves with the Hyss2009 software[52].

UV-Visible spectroscopy. UV-Vis spectra were collected at $T = 298.2$ K and $I = 0.1$ M (KCl) on a Thermo Scientific Evolution 260 Bio spectrophotometer provided with a Peltier thermostat, using 1 cm path length black quartz cuvettes. Spectra of the copper(II)/peptide systems at different pH values were acquired in the spectral range 350-900 nm. Samples were prepared in doubly distilled water by mixing proper amounts of peptide and $CuCl_2$ stock solutions to obtain 0.33 mM to 1.22 mM final copper(II) concentrations. Copper(II):peptide ratios were 1:1.5 for R1 and in the 1:1.2-2.1 range for R3A. The pH was adjusted to the desired values by addition of concentrated KOH or HCl.

Circular dichroism (CD) spectroscopy. CD spectra were collected at $T = 298.2$ K and $I = 0.1$ M (KCl) with a Jasco J715 spectropolarimeter equipped with a Peltier thermostat, using 1 cm path length black quartz cuvettes. Spectra of the copper(II)/peptide systems at different pH values were acquired in the spectral range 400-800 nm. Samples were prepared in doubly distilled water by mixing proper amounts of peptide and $CuCl_2$ stock solutions to obtain 0.24 mM to 0.70 mM final copper(II) concentrations. Copper(II):peptide ratios were 1:1.5 for R1, 1:3 for R3, and in the 1:1.2-2.1 range for R3A. The pH was adjusted to the desired values by addition of concentrated KOH or HCl.

Copper(I) binding to Tau fragments

Stock solutions. Concentrated (*ca.* 1-10 mM) solutions of R1 and R3 were prepared by weight in high-purity D_2O, and the peptide concentration was determined by NMR using an external standard. The titre was consistent with what obtained by potentiometric titrations. The stock solution of R3A was prepared at 4.5 mM concentration in Milli-Q water with the addition of 1 eq. of HCl. The titre of the solution was checked potentiometrically. Stock solutions (1.3-1.9 mM) of R3C were prepared in high-purity D_2O and the peptide concentration determined by NMR using an external standard. Ferrozine sodium salt (Fz) and sodium bicinchoninate (BCA) stock solutions (0.8-1.2 mM) were prepared by weight in aqueous HEPES buffer solution (100 mM, pH 7.4). Solutions of $[Cu(CH_3CN)_4]BF_4$ in acetonitrile (0.9-1.2 mM) were prepared by weight and their titre checked by spectrophotometric titrations with Fz. Aqueous HEPES buffer solutions (100 mM, pH 7.4) were prepared by suspending proper amounts of solid HEPES in doubly distilled water, correcting the pH to 7.4 using a concentrated NaOH solution. The buffer medium was passed through a 0.45 μM nylon filter, transferred into an air-free Schlenk vessel and deoxygenated by purge with a N2 stream before storage. The pH was checked every 2-3 days.

UV-visible spectroscopy. All spectral datasets were acquired at $T = 298.2$ K on a Thermo Scientific Evolution 260 Bio spectrophotometer provided with a Peltier thermostat. Spectrophotometric titrations were carried out in 100 mM aqueous HEPES buffer solutions at pH 7.4. To ensure the absence of atmospheric oxygen, buffer medium was degassed under vacuum and saturated with N_2 stream (5 min) before each sample preparation. The samples were prepared into gastight quartz cuvettes (1 cm path length) using gastight syringes, and all samples were nitrogen-bubbled beforehand. Typical samples contained the metallochromic indicator (either ferrozine or sodium bicinchoninate, *ca.* 80-120 μM), $[Cu(CH_3CN)_4]BF_4$ (40-55 μM) and sodium ascorbate (10 mM). The samples were titrated with up to 3.0, 5.0, 4.0 and 4.2 equivalents of R1, R3, R3A and R3C, respectively. Each titration run consisted of a minimum of 12 spectra, typically more than 20. Solutions of the metallochromic indicators were preliminarily evaluated by titrating an indicator solution (80-120 μM) containing sodium ascorbate with $[Cu(CH_3CN)_4]BF_4$ (0-1 equiv.). Similarly, $[Cu(CH_3CN)_4]BF_4$ solutions (55 μM) containing sodium ascorbate (10 mM) were titrated with metallochromic indicators up to 3-5 equivalents. These titrations allowed to determine the molar absorbance of the 1:1 and 1:2 adducts of copper(I) with the indicators under our experimental conditions. All titrations were performed in triplicate. Spectral data were treated using the software HypSpec2014[67,68].

REFERENCES

1. Fuster-Matanzo, A., Hernández, F. & Ávila, J. Tau Spreading Mechanisms; Implications for Dysfunctional Tauopathies. *Int. J. Mol. Sci.* ***19***, (2018).
2. Erkkinen, M. G., Kim, M. O. & Geschwind, M. D. Clinical Neurology and Epidemiology of the Major Neurodegenerative Diseases. *Cold Spring Harb. Perspect. Biol.* ***10***, a033118 (2018).
3. Elbaz, A., Carcaillon, L., Kab, S. & Moisan, F. Epidemiology of Parkinson's disease. *Rev. Neurol. (Paris).* ***172***, 14–26 (2016).
4. Glynn-Servedio, B. E. & Ranola, T. S. AChE Inhibitors and NMDA Receptor Antagonists in Advanced Alzheimer's Disease. *Consult. Pharm.* ***32***, 511–518 (2017).
5. Weingarten, M. D., Lockwood, A. H., Hwo, S. Y. & Kirschner, M. W. A protein factor essential for microtubule assembly. *Proc. Natl. Acad. Sci.* ***72***, 1858–1862 (1975).
6. Dotti, C. G., Banker, G. A. & Binder, L. I. The expression and distribution of the microtubule-associated proteins tau and microtubule-associated protein 2 in hippocampal neurons in the rat in situ and in cell culture. *Neuroscience* ***23***, 121–130 (1987).
7. Baas, P. W. & Qiang, L. Tau: It's Not What You Think. *Trends Cell Biol.* ***29***, 452–461 (2019).
8. Beharry, C. *et al.* Tau-induced neurodegeneration: mechanisms and targets. *Neurosci. Bull. 2014 302* ***30***, 346–358 (2014).
9. Sotiropoulos, I. *et al.* Atypical, non-standard functions of the microtubule associated Tau protein. *Acta Neuropathol. Commun. 2017 51* ***5***, 1–11 (2017).
10. Muralidar, S., Ambi, S. V., Sekaran, S., Thirumalai, D. & Palaniappan, B. Role of tau protein in Alzheimer's disease: The prime pathological player. *Int. J. Biol. Macromol.* ***163***, 1599–1617 (2020).
11. Gong, C.-X. & Iqbal, K. Hyperphosphorylation of Microtubule-Associated Protein Tau: A Promising Therapeutic Target for Alzheimer Disease. *Curr. Med. Chem.* ***15***, 2321 (2008).
12. Reynolds, M. R. *et al.* Tau Nitration Occurs at Tyrosine 29 in the Fibrillar Lesions of Alzheimer's Disease and Other Tauopathies. *J. Neurosci.* ***26***, 10636–10645 (2006).
13. Brister, M. A., Pandey, A. K., Bielska, A. A. & Zondlo, N. J. OGlcNAcylation and phosphorylation have opposing structural effects in tau: Phosphothreonine induces particular conformational order. *J. Am. Chem. Soc.* ***136***, 3803–3816 (2014).
14. Kontaxi, C., Piccardo, P. & Gill, A. C. Lysine-directed post-translational modifications of tau protein in Alzheimer's disease and related tauopathies. *Front. Mol. Biosci.* ***4***, 56 (2017).

15. Martin, L., Latypova, X. & Terro, F. Post-translational modifications of tau protein: Implications for Alzheimer's disease. *Neurochem. Int.* ***58***, 458–471 (2011).
16. Jeganathan, S., Von Bergen, M., Brutlach, H., Steinhoff, H. J. & Mandelkow, E. Global Hairpin Folding of Tau in Solution†. *Biochemistry* ***45***, 2283–2293 (2006).
17. Mukrasch, M. D. *et al.* Structural Polymorphism of 441-Residue Tau at Single Residue Resolution. *PLOS Biol.* **7**, e1000034 (2009).
18. Pîrşcoveanu, D. F. V. *et al.* Tau protein in neurodegenerative diseases - a review. *Rom. J. Morphol. Embryol.* ***58***, 1141–1150 (2017).
19. Von Bergen, M. *et al.* Assembly of τ protein into Alzheimer paired helical filaments depends on a local sequence motif (306VQIVYK311) forming β structure. *Proc. Natl. Acad. Sci.* ***97***, 5129–5134 (2000).
20. Bacchella, C. *et al.* Binding and Reactivity of Copper to R 1 and R 3 Fragments of tau Protein. *Inorg. Chem.* ***59***, 274–286 (2020).
21. Fitzpatrick, A. W. P. *et al.* Cryo-EM structures of tau filaments from Alzheimer's disease. *Nature* ***547***, 185–190 (2017).
22. Wang, Y. & Mandelkow, E. Tau in physiology and pathology. *Nat. Rev. Neurosci. 2015 171* **17**, 22–35 (2015).
23. Kanti Das, T., Wati, M. R. & Fatima-Shad, K. Oxidative Stress Gated by Fenton and Haber Weiss Reactions and Its Association With Alzheimer's Disease. *Arch. Neurosci.* ***2***, (2014).
24. Ischiropoulos, H. & Beckman, J. S. Oxidative stress and nitration in neurodegeneration: Cause, effect, or association? *J. Clin. Invest.* ***111***, 163–169 (2003).
25. Bhat, A. H. *et al.* Oxidative stress, mitochondrial dysfunction and neurodegenerative diseases; a mechanistic insight. *Biomed. Pharmacother.* **74**, 101–110 (2015).
26. Leal, S. S., Botelho, H. M. & Gomes, C. M. Metal ions as modulators of protein conformation and misfolding in neurodegeneration. *Coord. Chem. Rev.* ***256***, 2253–2270 (2012).
27. Benedictis, C. A. De, Vilella, A. & Grabrucker, A. M. The Role of Trace Metals in Alzheimer's Disease. *Exon Publ.* 85–106 (2019) doi:10.15586/ALZHEIMERSDISEASE.2019.CH6.
28. Dell'Acqua, S. *et al.* Prion Peptides Are Extremely Sensitive to Copper Induced Oxidative Stress. *Inorg. Chem.* ***56***, 11317–11325 (2017).
29. Dell'Acqua, S. *et al.* Copper(I) Forms a Redox-Stable 1:2 Complex with α-Synuclein N-Terminal Peptide in a Membrane-Like Environment. *Inorg. Chem.* ***55***, 6100–6106 (2016).
30. Carboni, E. & Lingor, P. Insights on the interaction of alpha-synuclein and metals in the pathophysiology of Parkinson's disease. *Metallomics* **7**, 395–404 (2015).

31. Valensin, D., Dell'Acqua, S., Kozlowski, H. & Casella, L. Coordination and redox properties of copper interaction with α-synuclein. *J. Inorg. Biochem.* ***163***, 292–300 (2016).
32. Dell'Acqua, S. *et al.* Reactivity of copper–α-synuclein peptide complexes relevant to Parkinson's disease. *Metallomics* **7**, 1091–1102 (2015).
33. De Ricco, R. *et al.* Copper(I/II), α/β-Synuclein and Amyloid-β: Menage à Trois? *ChemBioChem* ***16***, 2319–2328 (2015).
34. Ciregna, D., Monzani, E., Thiabaud, G. & Casella, L. Copper–β-amyloid peptides exhibit neither monooxygenase nor superoxide dismutase activities. *Chem. Commun.* ***49***, 4027–4029 (2013).
35. Pirota, V., Dell'Acqua, S., Monzani, E., Nicolis, S. & Casella, L. Copper-Aβ Peptides and Oxidation of Catecholic Substrates: Reactivity and Endogenous Peptide Damage. *Chem. – A Eur. J.* ***22***, 16964–16973 (2016).
36. Migliorini, C., Porciatti, E., Luczkowski, M. & Valensin, D. Structural characterization of Cu2+, Ni2+ and Zn2+ binding sites of model peptides associated with neurodegenerative diseases. *Coord. Chem. Rev.* ***256***, 352–368 (2012).
37. Kozlowski, H., Luczkowski, M., Remelli, M. & Valensin, D. Copper, zinc and iron in neurodegenerative diseases (Alzheimer's, Parkinson's and prion diseases). *Coord. Chem. Rev.* ***256***, 2129–2141 (2012).
38. Kepp, K. P. Alzheimer's disease: How metal ions define β-amyloid function. *Coord. Chem. Rev.* ***351***, 127–159 (2017).
39. La Mendola, D. & Rizzarelli, E. Evolutionary Implications of Metal Binding Features in Different Species' Prion Protein: An Inorganic Point of View. *Biomol. 2014, Vol. 4, Pages 546-565* ***4***, 546–565 (2014).
40. Du, X. *et al.* Inhibitory act of selenoprotein P on Cu+/Cu2+-induced tau aggregation and neurotoxicity. *Inorg. Chem.* ***53***, 11221–11230 (2014).
41. Shin, B. K. & Saxena, S. Insight into potential cu(II)-binding motifs in the four pseudorepeats of tau protein. *J. Phys. Chem. B* ***115***, 15067–15078 (2011).
42. Ma, Q. F. *et al.* Binding of copper (II) ion to an Alzheimer's tau peptide as revealed by MALDI-TOF MS, CD, and NMR. *Biopolymers* ***79***, 74–85 (2005).
43. Ma, Q. *et al.* Copper binding properties of a tau peptide associated with Alzheimer's disease studied by CD, NMR, and MALDI-TOF MS. *Peptides* ***27***, 841–849 (2006).
44. Zhou, L. X. *et al.* Copper (II) modulates in vitro aggregation of a tau peptide. *Peptides* ***28***, 2229–2234 (2007).
45. Soragni, A. *et al.* Structural characterization of binding of Cu(II) to Tau protein. *Biochemistry* ***47***, 10841–10851 (2008).
46. Pirota, V., Dell'Acqua, S., Monzani, E., Nicolis, S. & Casella, L. Copper-Aβ Peptides and Oxidation of Catecholic Substrates: Reactivity and Endogenous Peptide Damage. *Chem. - A Eur. J.* ***22***, 16964–16973 (2016).
47. Dell'Acqua, S. *et al.* Reactivity of copper–α-synuclein peptide complexes relevant to Parkinson's disease. *Metallomics* **7**, 1091–1102 (2015).

48. Dell'Acqua, S. *et al.* Prion Peptides Are Extremely Sensitive to Copper Induced Oxidative Stress. *Inorg. Chem.* ***56***, 11317–11325 (2017).
49. Billo, E. J. Copper(II) chromosomes and the rule of average environment. *Inorg. Nucl. Chem. Lett.* ***10***, 613–617 (1974).
50. Prenesti, E., Daniele, P. G., Prencipe, M. & Ostacoli, G. Spectrum-structure correlation for visible absorption spectra of copper(II) complexes in aqueous solution. *Polyhedron* ***18***, 3233–3241 (1999).
51. Sigel, H. & Martin, R. B. Coordinating properties of the amide bond. Stability and structure of metal ion complexes of peptides and related ligands. *Chem. Rev.* ***82***, 385–426 (1982).
52. Alderighi, L. *et al. Hyperquad simulation and speciation (HySS): a utility program for the investigation of equilibria involving soluble and partially soluble species*. *Coordination Chemistry Reviews* vol. 184 http://www.eawag.ch/soft/ (1999).
53. Soragni, A. *et al.* Structural characterization of binding of Cu(II) to Tau protein. *Biochemistry* ***47***, 10841–10851 (2008).
54. Young, T. R., Wedd, A. G. & Xiao, Z. Evaluation of Cu(i) binding to the E2 domain of the amyloid precursor protein-a lesson in quantification of metal binding to proteins via ligand competition. *Metallomics* ***10***, 108–119 (2018).
55. Quaretti, M. *et al.* Thermodynamic stability and structure in aqueous solution of the [Cu(PTA) 4] + complex (PTA = aminophosphine-1,3,5-triaza-7-phosphaadamantane). *J. Inorg. Biochem.* ***188***, 50–61 (2018).
56. Alies, B., Badei, B., Faller, P. & Hureau, C. Reevaluation of Copper(I) Affinity for Amyloid-β Peptides by Competition with Ferrozine-An Unusual Copper(I) Indicator. *Chem. - A Eur. J.* ***18***, 1161–1167 (2012).
57. Haas, K. L., Putterman, A. B., White, D. R., Thiele, D. J. & Franz, K. J. Model peptides provide new insights into the role of histidine residues as potential ligands in human cellular copper acquisition via Ctr1. *J. Am. Chem. Soc.* ***133***, 4427–4437 (2011).
58. Himes, R. A., Park, G. Y., Siluvai, G. S., Blackburn, N. J. & Karlin, K. D. Structural Studies of Copper(I) Complexes of Amyloid-β Peptide Fragments: Formation of Two-Coordinate Bis(histidine) Complexes. *Angew. Chemie Int. Ed.* ***47***, 9084–9087 (2008).
59. Conklin, S. E. *et al.* Specific Histidine Residues Confer Histatin Peptides with Copper-Dependent Activity against Candida albicans. *Biochemistry* ***56***, 4244–4255 (2017).
60. Bacchella, C. *et al.* Binding and Reactivity of Copper to R1 and R3 Fragments of tau Protein. *Inorg. Chem.* ***59***, 274–286 (2020).
61. Weast, R. *CRC Handbook of chemistry and physics.* (CRC Press, 1984).
62. Meites, L. *CRC Handbook Series in inorganic electrochemistry*. (CRC Press, 1980).
63. Martic, S., Rains, M. K. & Kraatz, H. B. Probing copper/tau protein interactions electrochemically. *Anal. Biochem.* ***442***, 130–137 (2013).

64. Nemeikaite-Čeniene, A., Imbrasaite, A., Sergediene, E. & Čenas, N. Quantitative structure–activity relationships in prooxidant cytotoxicity of polyphenols: Role of potential of phenoxyl radical/phenol redox couple. *Arch. Biochem. Biophys.* ***441***, 182–190 (2005).
65. Vogel, A. I. *Quantitative Inorganic Analysis Including Elementary Instrumental Analysis*. (Longmans, 1962).
66. Gran, G. Determination of the equivalence point in potentiometric titrations. Part II. *Analyst* **77**, 661 (1952).
67. Gans, P., Sabatini, A. & Vacca, A. Investigation of equilibria in solution. Determination of equilibrium constants with HYPERQUAD suite of programs. *Talanta* ***43***, 1739–1753 (1996).
68. Gans, P., Sabatini, A. & Vacca, A. To improve accuracy of the calculated pKa values,. *Annali di Chimica* vol. 89 45–49 (1999).

Chapter 2

Bioinspired copper peptides: role of His sites in oxidation reactions

INTRODUCTION

The binding of metal ions to peptides has been the target of systematic studies since the 1960s, and the coordinating properties of these bioligands have been gradually undisclosed throughout the years. Structural, thermodynamic, and kinetic features of metals bound to amino acids were initially investigated and characterised. This knowledge is critical to understand the interaction of metals with proteins, although the binding motifs of the latter only rarely could be precisely modelled by amino acids. It is However known that intrinsically disordered proteins (IDP) possess an ordered three-dimensional structure only when interacting with their biomolecular partners. For these proteins, oligopeptides represent models that are good enough to elucidate the structure and metal ions coordination properties of isolated portions of the protein sequence.

The chemistry of metallopeptides with different coordinating residues involved in metal coordination has been extensively explored in the past decades, and common features were identified among a great number of different peptides and the most biologically relevant metal ions.[1,2] This enormous amount of information allowed to elucidate and consequently to predict the role of specific residues in peptides and, more importantly, in proteins with respect of metal ions binding. On this ground the case of histidine residues is paradigmatic since the imidazole ring, for its chemical properties, is involved in Nature in the binding of numerous metal ions and metal cofactors.

Histidine-containing peptides are characterised by great affinities for numerous transition metal ions, which vary with the number and position of histidine residues in the peptide sequence, and the nature of the metal. Mono-histidine peptides are so far the best characterised in terms of interaction with metal ions, the latter being of remarkable stability when histidine is located close to or at the peptide N-terminus.[2–4] When multiple His residues are present in the sequence, the study of metal binding becomes more challenging, since small differences in the peptide sequences may result into large changes in the structure, stability and function of metal complexes. However, it was widely demonstrated that in these systems histidine residues act as anchor sites for metal ions, and several recurring features were identified.[4]

During the last two decades, metal complexes of multi-histidine peptides have drawn increasing attention due to their involvement in a remarkable number of biological processes. In particular, copper/histidine sites were found to be a major feature in a large number of biologically relevant metallopeptides and metalloenzymes. These systems provide a great variety of functions, from metal transport (*e.g.* ZIP[5] and Ctr[6] families) to antimicrobial,[7,8] monooxygenase,[9] SOD,[10] and catecholase[11] activities (among others). To evaluate the role of Cu/His sites in these biological systems, short peptide sequences containing or mimicking the active site of the desired enzyme were studied as an alternative to the whole enzyme sequence. This approach made easier the study of the portion of the

peptide that binds the metal, ideally isolated from the rest of the sequence. Interestingly, most of these model peptides contained two histidine residues arranged in a His-His (HH) motif, and possible additional His residues. Whether or not the HH site was located at the peptide N-terminus, it was shown without exception to coordinate copper(II) with high affinity.[6–11] However, the most interesting property of this site lies in its ability to bind copper not only in its oxidized +2 state, but also in its reduced +1 state.[6,8–11]

The peculiar capability of HH sites to bind copper in both oxidation states particularly emerged in the context of neurodegeneration, as His-His dyads represent a common motif among peptides related to neurodegenerative diseases. In this context, β-amyloid (Aβ) and tau protein are examples of these biomolecules. Indeed, the binding of both copper(II) and copper(I) to fragments of tau protein bearing either a single His or a His-His dyad was presented in *Chapter 1*.[12] As shown in that section, our results confirmed the high affinity of both mono-His and bis-His sites for copper(II) and, remarkably, the presence of the HH motif resulted essential for the binding of copper(I). Moreover, Cu/HH adducts showed a remarkable oxidase activity towards catechol substrates, which led to the production of reactive oxygen species (ROS).[12] This latter finding is in agreement with what reported in the literature for other peptides: the interaction of copper with amyloid-β and tau protein resulted to be involved in ROS production, oxidative stress, and pathologic protein aggregation.[13–22]

Cu-Aβ adducts are paradigmatic in the context of the study of copper/histidine sites in neuropeptides since Aβ sequences of both wild-type peptides and their models contain three His residues (*e.g.* $A\beta_{1\text{-}16}$: ^{1}DAEFR<u>H</u>DSGYEV<u>HH</u>NK^{16}). Their structural and thermodynamic properties have been largely characterised, and copper coordination modes involving His sites (***Figure 2.1***) resulted of outmost importance for the reactivity of Aβ species.[23–25] As also evidenced for tau protein, in the case of amyloid-β the pathologic activity of its copper complexes is related to the capability of the peptide to bind the metal in both oxidation states. Although copper(II) binding can occur at both the sites bearing one or two His residues (***Figure 2.1a***), the coordination of copper(I) is achieved exclusively through a linear interaction of two histidines (***Figure 2.1b***). Similarly, in Aβ this copper(II) binding at the His-His site occurs at the His13-His14 site. Also for this peptide the linear coordination of a bis-His motif was found responsible of promoting the reduction of copper(II) to copper(I) and of enabling the catalytic cycle leading to ROS production.[6,26–28] In this context, however, it is important to underline that recently a long-time elusive "in-between" coordination state of copper was found relevant in the catalytic formation of ROS species in copper/Aβ adducts.[22]

As previously mentioned, His-His sites are commonly found in metal binding peptides and in several of these peptides an additional histidine residue is present close to the HH dyad. For instance, fragments of histatin 5 containing mono-His and bis-His sites were reported to bind copper(II) mainly via the HH motif with a possible weak interaction of

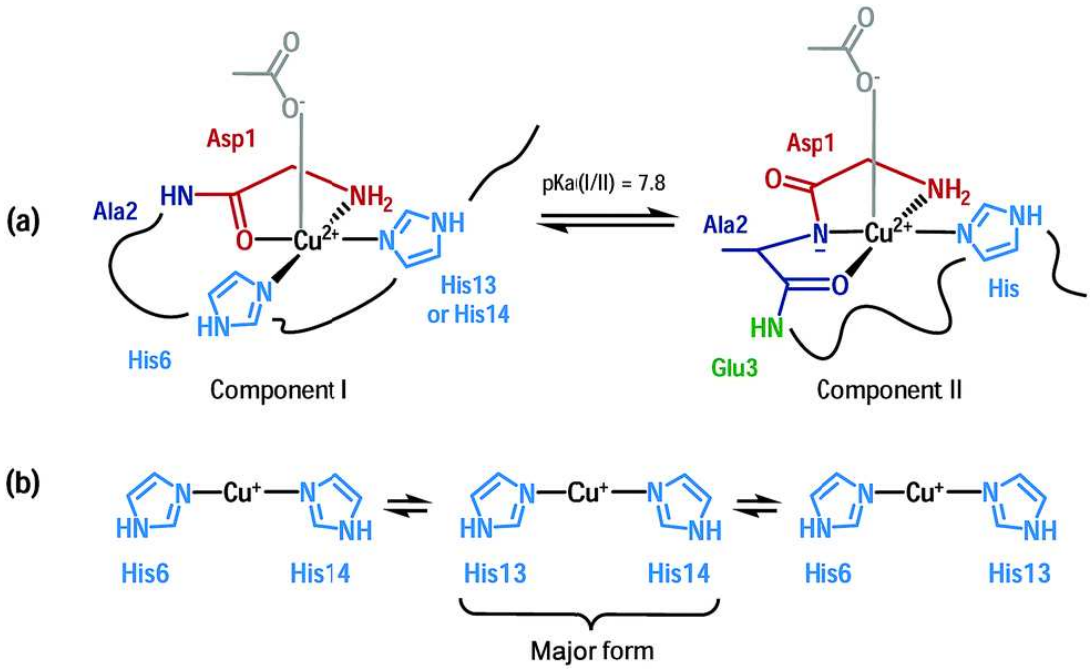


Figure 2.1 *Copper(II) coordination (a) and copper(I) coordination (b) to the Aβ peptide in the resting state. Reprinted from Ref. 22 with permission from the Royal Society of Chemistry.*

the single His residue.[7] Analogously, Cu(II) binds to the CuH site of peptidylglycine monooxygenase through a His-His-His coordination mode that involves two imidazole groups from a HH dyad and one imidazole donor from an additional histidine residue, the latter having however only a marginal role in metal binding.[9]

Conversely, in the field of model peptides of His/Pro-rich proteins, the Ac-HPHH-NH_2 sequence was reported to bind Cu(II) with high affinity and with the involvement of all three His residues.[11] Copper(II) complexes of this peptide were also explored in terms of oxidase-like activity. Interestingly, complex species characterised by a (3x N_{im}) coordination mode did not show any catalytic property,[11] as previously observed for a similar peptide (Ac-HHGH-OH).[29] This behaviour is perhaps expected: the presence of a third histidine that coordinates copper may destabilise the linear arrangement of two imidazoles, with formation of 3His complexes that in turn may stabilise copper in the +2 oxidation state. In this perspective, the third His residue has a detrimental effect on the catalytic properties of these metal adducts with respect oxidation reactions.[11] Finally, a histatin-like peptide bearing histidine residues at positions 3, 7 and 8 was reported to bind copper in both oxidation states. The essential role of the bis-His site (His7-His8) in copper(I) binding was again demonstrated, but in this case a possible Cu(I)-stabilising interaction of the mono-His site (His3) was proposed.[8]

Although the comprehension of how single-His sites may affect the structure, coordination properties, thermodynamics and reactivity of bis-His motifs is of outmost importance, it is evident that the role of an additional histidine residue in peptides containing a HH dyad is not yet fully understood, and more studies are advisable. On the other hand, it is critical to underline that, in this field of studies, difficulties may arise in trying to generalise the observations made on model systems. In fact and as an example, both Aβ and tau protein actually contain one or more His residues beyond those of the HH dyad. However, in tau the additional His residues are *ca.* 30 residues apart from the HH dyad, while in Aβ they are just 7 residues apart. For these reasons, short oligopeptides may work well as models for neuropeptides like Aβ, but results have to be taken with caution before generalising for larger proteins. Finally, the simultaneous involvement of mono- and bis-His sites in copper binding was already demonstrated for Aβ (***Figure 2.1***), by using model peptides which bear a free N-terminal amino group that can also coordinate copper(II). This structural aspect can complicate the elucidation of the sole role of the third His residue in affecting the properties of the His-His dyad. To the best of our knowledge, no systematic studies have been yet carried out in this perspective.

This work aims at clarifying how the presence and the position of an additional histidine residue affect the binding properties and the reactivity of peptides containing a His-His motif, in the presence of an acetylated N-terminal amino group. Taking inspiration from the previously studied 'R3' fragment of tau protein,[12] we *de novo* designed four peptides, which sequences are reported in ***Figure 2.2***. All peptides are characterised by the presence of a bis-His site, acetylated N-termini and amidated C-termini. In the case of PCtrl, the HH dyad is the only metal binding site and therefore this peptide could be used as a control for the reactivity of isolated bis-His sites. The other three peptides contain an additional mono-His site and differ for its position with respect to the HH dyad (indicated as '-2', '-4' and '-6').

PCtrl	Ac-GWKAASAGAI**HH**KPG-NH_2
PHis(-2)	Ac-GWKAASAG**H**I**HH**KPG-NH_2
PHis(-4)	Ac-GWKAAS**H**GAI**HH**KPG-NH_2
PHis(-6)	Ac-GWKA**H**SAGAI**HH**KPG-NH_2

Figure 2.2 *Amino acid sequences of the tau-inspired peptides studied in this work. The bis-His sites are highlighted in red, the mono-His sites in blue.*

The four peptides reported in ***Figure 2.2*** were synthesised at the Institut des Sciences Moléculaires de Marseille (iSm2, CNRS, University of Aix-Marseille) in collaboration with Prof. Thierry Tron and Prof. Olga Iranzo. The main features of these peptides will be addressed in the following sections, with particular focus on the binding of copper(II) and on the reactivity of Cu(II)-complexes in the oxidation of catechol substrates. Our findings will be discussed in terms of the possible influence of a mono-His site on the properties of His-His dyads.

BINDING OF COPPER(II) TO HIS-CONTAINING PEPTIDES

Peptide Protonation Equilibria

The acid-base behaviour of PCtrl, PHis(-2), PHis(-4) and PHis(-6) was explored through potentiometry. The PCtrl peptide contains four basic amino acids (2 His, 2 Lys) which are deprotonated in its neutral form (L). Therefore, PCtrl behaves as a tetraprotic acid when fully protonated (H_4L^{4+}). Analogously, the PHis(-2), PHis(-4) and PHis(-6) peptides have no acidic protons in their neutral form (L), but they can undergo protonation on the imidazole rings of the three His residues and on the amino group of the two lysine side chains. Hence, they all behave as pentaprotic acids in their fully protonated form (H_5L^{5+}). For all the peptides, the observed overall and stepwise protonation constants (***Table 2.1***) are fully consistent with the dissociation of the histidine imidazolium groups in the pH range 4-8.5, followed by the protonated lysine side chains at pH > 9. However, the proton dissociation steps are macroconstants, and the dissociation processes on individual sites are partially overlapped and only marginally sequential. It is interesting to note that the deprotonation of the imidazolium groups occurs at quite low pH, with the first step also being remarkably acidic for the PHis(-2), PHis(-4) and PHis(-6) peptides (pK_a *ca.* 5.4-5.5). Consequently, at pH 7.4 each peptide is mostly found in the H_2L^{2+} form, in which all of the His residues are deprotonated.

Table 2.1 *Overall protonation constants (log β_i^H) and proton dissociation constants (pK_a) of PCtrl, PHis(-2), PHis(-4) and PHis(-6) in aqueous solution (T = 298.2 K, I = 0.1 M in KNO3[a] or KCl[b]). L represents the completely deprotonated form of the peptides. Standard deviations on the last significant figure are given in parenthesis.*

Reaction		PCtrl[a]	PHis(-2)[b]	PHis(-4)[b]	PHis(-6)[b]
log β_i^H					
$L + H^+ \rightleftarrows HL^+$		10.23(2)	10.42(2)	10.50(1)	10.53(2)
$L + 2H^+ \rightleftarrows H_2L^{2+}$		20.17(1)	20.28(1)	20.38(1)	20.32(1)
$L + 3H^+ \rightleftarrows H_3L^{3+}$		26.79(2)	27.02(2)	27.16(1)	27.07(3)
$L + 4H^+ \rightleftarrows H_4L^{4+}$		32.60(2)	33.11(2)	33.32(1)	33.21(2)
$L + 5H^+ \rightleftarrows H_5L^{5+}$		-	38.49(2)	38.85(1)	38.74(2)
pK_a	*Diss. site*				
$HL^+ \rightleftarrows L + H^+$	NH_3^+ Lys	10.23(2)	10.42(2)	10.50(1)	10.53(2)
$H_2L^{2+} \rightleftarrows HL^+ + H^+$	NH_3^+ Lys	9.94(2)	9.85(2)	9.89(1)	9.79(2)
$H_3L^{3+} \rightleftarrows H_2L^{2+} + H^+$	$N_{im}H^+$ His	6.62(2)	6.75(2)	6.78(1)	6.75(2)
$H_4L^{4+} \rightleftarrows H_3L^{3+} + H^+$	$N_{im}H^+$ His	5.81(2)	6.08(2)	6.16(1)	6.14(3)
$H_5L^{5+} \rightleftarrows H_4L^{4+} + H^+$	$N_{im}H^+$ His	-	5.39(2)	5.53(1)	5.53(2)

Copper(II) Complexation Studies

The complexation of copper(II) with the PCtrl, PHis(-2), PHis(-4) and PHis(-6) peptides was studied by potentiometric titrations in aqueous solution. These studies were complemented by visible absorption and circular dichroism characterisation of Cu^{2+}/peptide systems at different pH values, to provide information on the coordination environment of copper in the complex species. The speciation models obtained by analysis of potentiometric data are reported in ***Table 2.2*** and ***Table 2.3***.

Before discussing formation equilibria of Cu^{2+}/peptide complexes in detail, an interesting observation can be made. Equilibrium constants on ***Table 2.2*** and ***Table 2.3*** show that all the four peptides have copper(II) complexation equilibria that parallel with each other. Not all species are present for all the peptides, for reasons that will be described further on. Nevertheless, the pK_a values suggest that the same stoichiometry corresponds to the same coordination mode. This has been indeed verified by UV-visible absorption studies, as shown below.

Table 2.2 *Logarithms of overall formation constants (log β) of the copper(II) complexes of PCtrl, PHis(-2), PHis(-4) and PHis(-6) in aqueous solution (T = 298.2 K, I = 0.1 M in KNO_3[a] or KCl[b]). L represents the completely deprotonated form of the peptides. Standard deviations on the last significant figure are given in parenthesis.*

Equilibrium reaction	$log\ \beta_{Cu_mH_hL_l}$			
	PCtrl[a]	PHis(-2)[b]	PHis(-4)[b]	PHis(-6)[b]
$Cu^{2+} + 3H^+ + L \rightleftarrows [CuH_3L]^{5+}$	-	32.72(1)	32.34(2)	32.25(3)
$Cu^{2+} + 2H^+ + L \rightleftarrows [CuH_2L]^{4+}$	25.83(1)	27.01(2)	26.90(1)	27.31(1)
$Cu^{2+} + H^+ + L \rightleftarrows [CuHL]^{3+}$	19.30(2)	20.60(2)	20.07(2)	19.94(3)
$Cu^{2+} + L \rightleftarrows [CuL]^{2+}$	12.67(1)	13.04(4)	12.59(2)	12.75(2)
$Cu^{2+} + L \rightleftarrows [CuLH_{-1}]^{+} + H^+$	4.21(2)	4.18(5)	3.53(3)	4.44(2)
$Cu^{2+} + L \rightleftarrows [CuLH_{-2}] + 2H^+$	-5.62(4)	-5.80(5)	-6.59(4)	-5.67(3)
$Cu^{2+} + L \rightleftarrows [CuLH_{-3}]^{-} + 3H^+$	-14.60(2)	-16.27(6)	-17.22(5)	-16.01(3)
$2Cu^{2+} + L \rightleftarrows [Cu_2L]^{4+}$	-	18.68(3)	18.23(4)	18.06(7)
$2Cu^{2+} + L \rightleftarrows [Cu_2LH_{-1}]^{3+} + H^+$	-	11.82(4)	11.86(3)	11.90(3)
$2Cu^{2+} + L \rightleftarrows [Cu_2LH_{-2}]^{2+} + 2H^+$	-	4.47(3)	4.51(4)	4.0(1)
$2Cu^{2+} + L \rightleftarrows [Cu_2LH_{-3}]^{+} + 3H^+$	-	-3.67(4)	-3.86(5)	-3.32(3)
$2Cu^{2+} + L \rightleftarrows [Cu_2LH_{-4}] + 4H^+$	-	-13.54(5)	-13.28(5)	-12.5(1)
$2Cu^{2+} + L \rightleftarrows [Cu_2LH_{-5}]^{-} + 5H^+$	-	-23.52(4)	-23.26(5)	-22.26(9)

Table 2.3 *Logarithms of proton dissociation constants (pK_a) of the copper(II) complexes of PCtrl, PHis(-2), PHis(-4) and PHis(-6) in aqueous solution (T = 298.2 K, I = 0.1 M in KNO_3[a] or KCl[b]). L represents the completely deprotonated form of the peptides. Standard deviations on the last significant figure are given in parenthesis.*

Equilibrium reaction	pK_a			
	PCtrl[a]	PHis(-2)[b]	PHis(-4)[b]	PHis(-6)[b]
$[CuH_3L]^{5+} \rightleftarrows [CuH_2L]^{4+} + H^+$	-	5.71(2)	5.44(2)	4.94(3)
$[CuH_2L]^{4+} \rightleftarrows [CuHL]^{3+} + H^+$	6.52(2)	6.41(3)	6.83(2)	7.36(3)
$[CuHL]^{3+} \rightleftarrows [CuL]^{2+} + H^+$	6.64(2)	7.56(3)	7.49(2)	7.20(3)
$[CuL]^{2+} \rightleftarrows [CuLH_{-1}]^{+} + H^+$	8.46(2)	8.87(4)	9.06(3)	8.30(2)
$[CuLH_{-1}]^{+} \rightleftarrows [CuLH_{-2}] + H^+$	9.84(5)	9.98(6)	10.12(5)	10.11(4)
$[CuLH_{-2}] \rightleftarrows [CuLH_{-3}]^{-} + H^+$	8.98(4)	10.47(8)	10.63(6)	10.34(4)
$[Cu_2L]^{4+} \rightleftarrows [Cu_2LH_{-1}]^{3+} + H^+$	-	6.86(5)	6.37(6)	6.16(9)
$[Cu_2LH_{-1}]^{3+} \rightleftarrows [Cu_2LH_{-2}]^{2+} + H^+$	-	7.34(5)	7.35(5)	8.0(1)
$[Cu_2LH_{-2}]^{2+} \rightleftarrows [Cu_2LH_{-3}]^{+} + H^+$	-	8.14(4)	8.38(5)	7.3(1)
$[Cu_2LH_{-3}]^{+} \rightleftarrows [Cu_2LH_{-4}] + H^+$	-	9.87(5)	9.41(6)	9.2(1)
$[Cu_2LH_{-4}] \rightleftarrows [Cu_2LH_{-5}]^{-} + H^+$	-	9.98(5)	9.98(5)	9.8(1)

Cu^{2+}/PCtrl complex-formation equilibria

In the presence of copper(II), the PCtrl peptide forms six complex species, all with a 1 to 1 copper/peptide stoichiometry and differing for their protonation states (***Table 2.2***, ***Table 2.3***). The analysis of the potentiometric curves suggests no formation of both 2:1 and 1:2 copper(II)/peptide species under the experimental conditions used. A representative speciation distribution diagram is reported in ***Figure 2.3***.

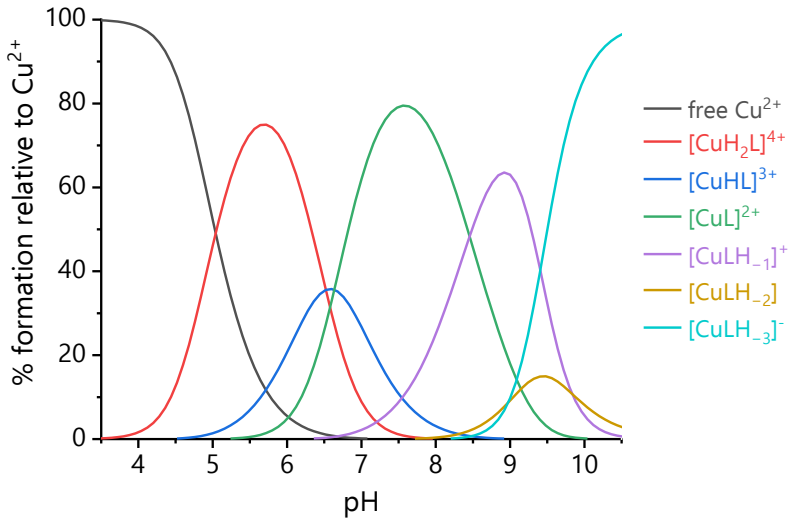


Figure 2.3 *Representative distribution diagram of the copper(II)/PCtrl (L) system in aqueous solution (Cu^{2+}:L = 1:1.2, C_{Cu} = 0.83 mM, T = 298.2 K, I = 0.1 M in KNO_3).*

The complexation of Cu^{2+} starts at pH 4 with the formation of $[CuH_2L]^{4+}$, which reaches *ca.* 75% of total copper at pH 5.8 and is predominant below pH 6.5. The broad band at *ca.* 660 nm observed in the visible absorption spectrum collected at pH 5.8 (***Figure 2.4***) is well accounted for a (N_{im}, N^-, 2x O_w) donor set, as calculated on the basis of the Billo's average environment theory (calc. λ_{max} = 661 nm).[30–32] This coordination environment may occur at both His11 or His12 residues, though the involvement of the latter would result in a stable 6-membered chelation ring (***Scheme 2.1a***), with respect to the 7-membered one formed in the case of His11.

A shift to *ca.* 610 nm of the d-d band is observed when increasing the pH to 6.6, where $[CuHL]^{3+}$ reaches its maximum (*ca.* 35% of total copper). The presence in this species of two imidazole donors, one deprotonated peptide nitrogen and a water molecule coordinated on the equatorial plane of copper(II) (***Scheme 2.1b***; calc. λ_{max} = 608 nm) explains quite well the observed λ_{max}. However, the absorption spectrum displays a broadened profile at both higher and lower wavelengths, that is ascribable to the presence in solution of *ca.* 30% of $[CuH_2L]^{4+}$ and $[CuL]^{2+}$, respectively. The latter species, which dominates the

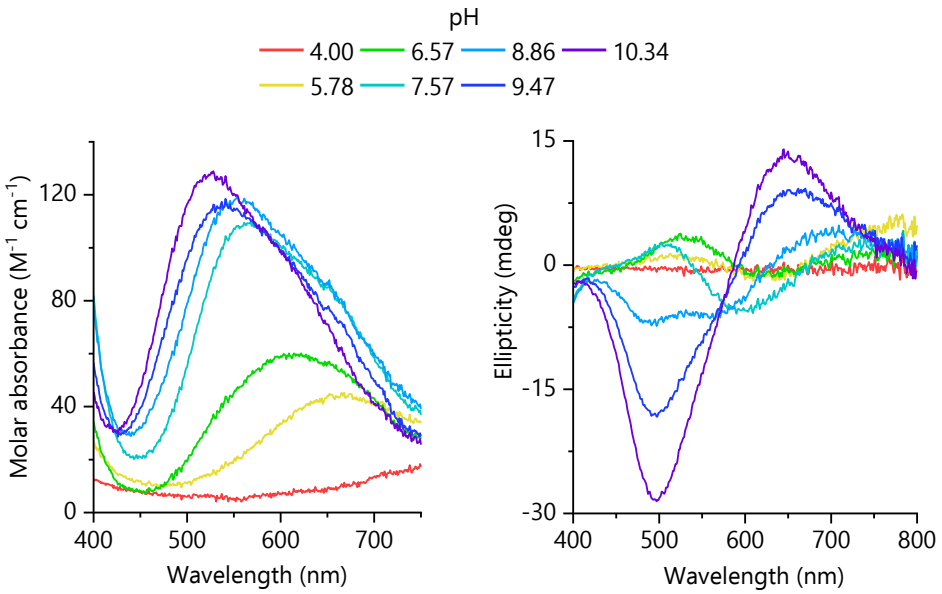
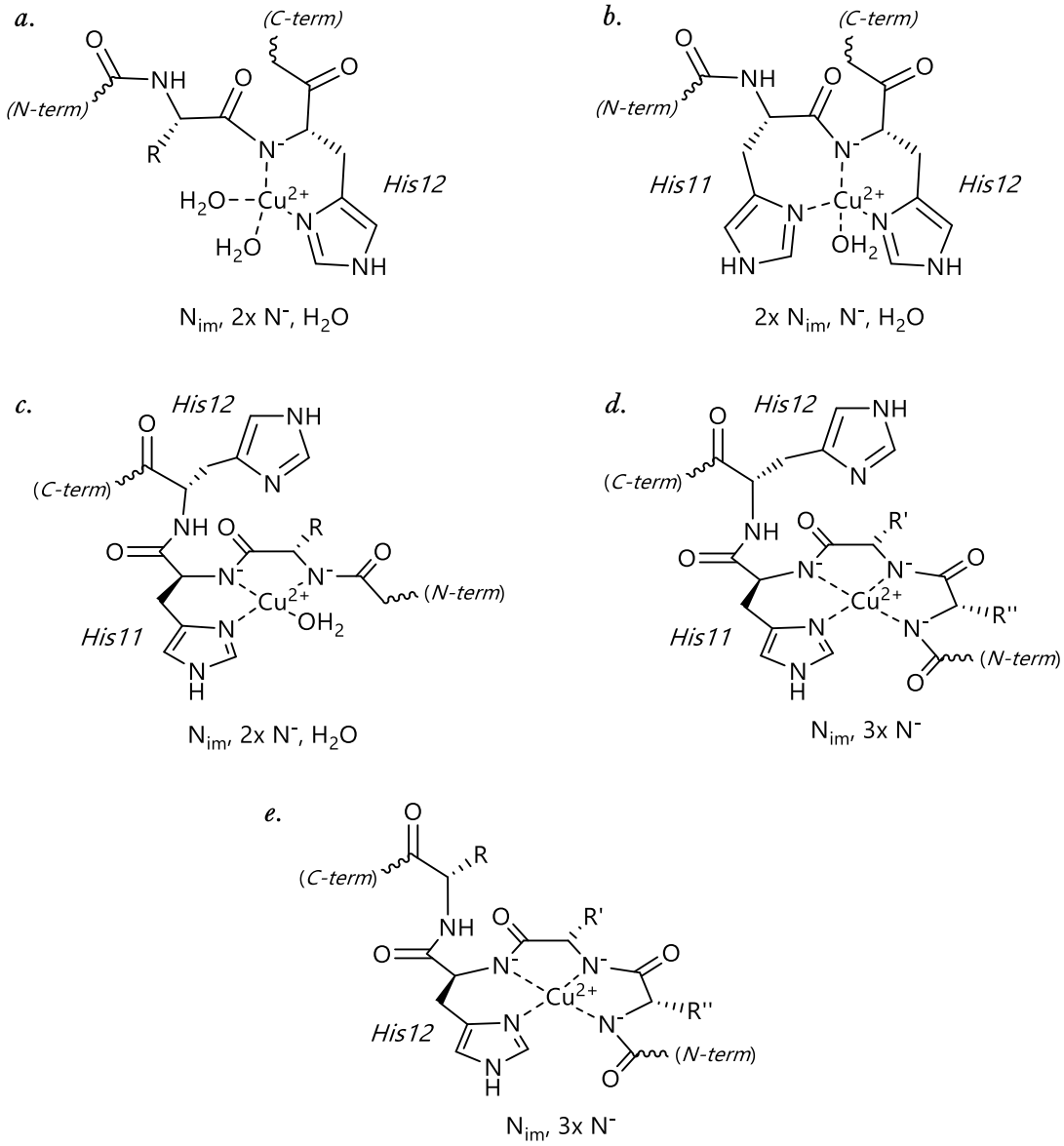


Figure 2.4 *Molar visible absorption spectra (absorbance/C_{Cu}; left) and visible CD spectra (right) of solutions of copper(II) and PCtrl (Cu^{2+}/PCtrl = 1:1.2, C_{Cu} = 0.83 mM, T = 298.2 K, I = 0.1 M in KNO_3).*

speciation in the pH range 6.7-8.4 and accounts for 80% of total copper at pH 7.6 (***Figure 2.3***), is likely to have Cu^{2+} bound in a (N_{im}, 2x N^-, H_2O) coordination environment. At pH 7.6, the observed absorption maximum at 569 nm (***Figure 2.4***) is lower than the expected one of the proposed coordination environment (calc. λ_{max} = 584 nm). This difference can be explained by considering the imidazole of His11 coordinated at an equatorial position of Cu^{2+}, while that of His12 possibly occupies one of the axial positions (***Scheme 2.1c***). This may likely result into a blue shift of the d-d band due to an increased ligand field with respect to the sole coordination of one equatorial His. Also, the shoulder peak that is observed at higher wavelengths is associated to the presence in solution of a significant amount of $[CuHL]^{3+}$.

$[CuLH_{-1}]^+$ starts forming at pH 7 and reaches its maximum at pH 8.8 (*ca.* 67% of total copper). For this species, the presence of a (N_{im}, 3x N^-) donor set (***Scheme 2.1d***; obs. λ_{max} = 552 nm, calc. λ_{max} = 523 nm) is likely. The discrepancy between the observed and calculated absorption maxima are attributed to the presence of *ca.* 20% copper in the form of $[CuL]^{2+}$ at pH 8.8. Finally, both absorption and dichroism data collected above pH 9 seem to suggest no further change in the coordination of copper(II), with the small variations observed in the spectroscopic profiles being related to the disappearance of the $[CuL]^{2+}$ species. Hence, it can be concluded that the two deprotonation steps leading from $[CuLH_{-1}]^+$ to $[CuLH_{-2}]$ and $[CuLH_{-3}]^-$ most likely involve the two Lys residues which remain uncoordinated in the entire pH range explored.



Scheme 2.1 *Schematic representation of the main coordination modes involving the His-His tandem site in copper(II) complexes of PCtrl, PHis(-2), PHis(-4) and PHis(-6). In the case of PCtrl, these coordination modes correspond to the species:* ***a.*** *$[CuH_2L]^{4+}$;* ***b.*** *$[CuHL]^{3+}$;* ***c.*** *$[CuL]^{2+}$;* ***d, e.*** *$[CuLH_{-1}]^{+}$, $[CuLH_{-2}]$ and $[CuLH_{-3}]^{-}$. Coordination modes* ***b****,* ***c*** *and* ***d*** *can be present only at the His-His site.*

Cu^{2+}/PHis(-2) complex-formation equilibria

In the presence of copper(II), the PHis(-2) peptide forms seven main species having 1 to 1 copper/peptide stoichiometry and differing for their protonation states (***Table 2.2***, ***Table 2.3***). The analysis of the potentiometric curves suggests no formation of 1:2 copper(II)/peptide species under the experimental conditions used. Instead, the formation of species with 2 to 1 copper/peptide stoichiometry seems to occur also in 1:1.2 copper/peptide ratio. However, using the latter copper/peptide ratio, these binuclear species never exceed 20% of total copper and their formation is negligible at physiological pH in 1:1 metal/peptide molar ratio. A representative distribution diagram for equimolar copper(II)/peptide ratio is shown in ***Figure 2.5***.

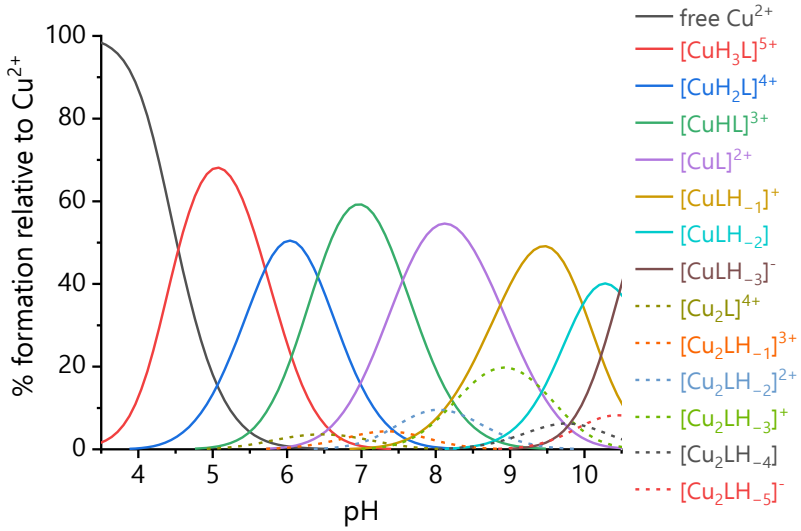


Figure 2.5 *Representative distribution diagram of the copper(II)/PHis(-2) (L) system in aqueous solution (Cu^{2+}:L = 1:1.2, C_{Cu} = 0.83 mM, T = 298.2 K, I = 0.1 M in KCl). Formation curves of bimetallic species are drawn as dashed lines.*

The first complex species is $[CuH_3L]^{5+}$, which starts forming at pH 3.5 and reaches almost 70% of total copper at pH 5.2. The absorption spectrum collected at this pH shows a maximum at *ca.* 690 nm (***Figure 2.6***). Exploiting Billo's method of the average environment,[30–32] this wavelength was ascribed to the presence in $[CuH_3L]^{5+}$ of two imidazole nitrogen atoms and two water molecules coordinated on the equatorial plane of copper(II) (calc. λ_{max} = 692 nm). The fact that this species is not formed with the PCtrl peptide suggests that the two imidazole groups coordinated to copper(II) are not those of the His-His dyad, but they belong to both His9 and either one among His11 and His12 instead.

A blue shift of the d-d band to *ca.* 650 nm is observed when increasing the pH to 6.0, where $[CuH_2L]^{4+}$ is predominant (*ca.* 50% of total copper). For this species it is reasonable

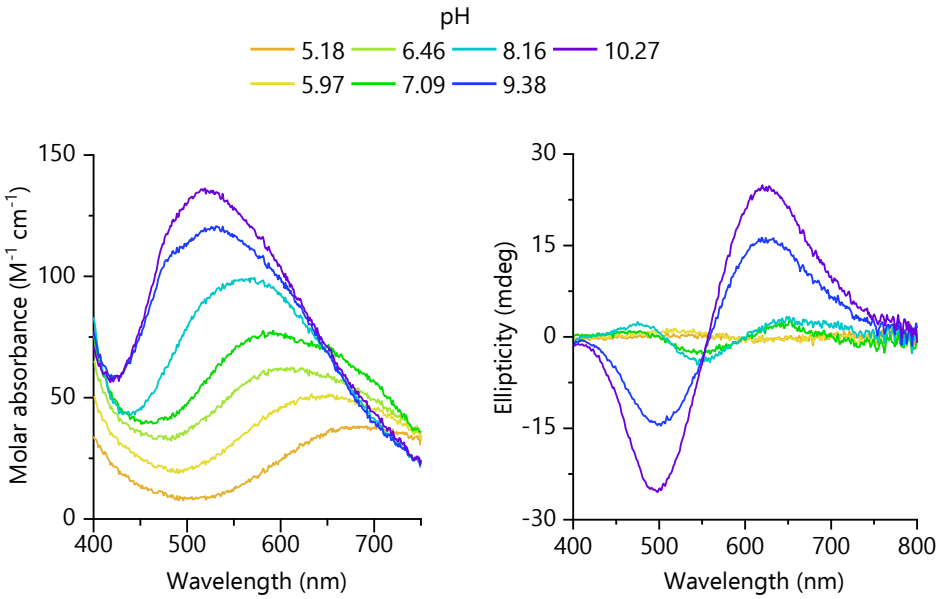


Figure 2.6 *Molar visible absorption spectra (absorbance/C_{Cu}; left) and visible CD spectra (right) of solutions of copper(II) and PHis(-2) (Cu^{2+}/PHis(-2) = 1:1.2, C_{Cu} = 0.83 mM, T = 298.2 K, I = 0.1 M in KCl).*

to propose a (N_{im}, N^-, 2x H_2O) equatorial coordination mode, for which the expected λ_{max} is 661 nm, where only one out of the two uncoordinated His residues is protonated. It is important to note that copper(II) binding with this coordination mode can occur at both the His9 and the His11/His12 sites.

Moving to pH 7, the deprotonation of the last imidazole group occurs and $[CuHL]^{3+}$ becomes the most abundant species, accounting for *ca.* 60% of total copper. Most likely, this species provides a (2x N_{im}, N^-, H_2O) donor set (***Scheme 2.1b***; calc. λ_{max} = 608 nm). Also, the absorption maximum at 590 nm at pH 7.1 (with a shoulder at 650 nm), can be reasonably ascribed to the presence in solution of significant amounts of $[CuL]^{2+}$, which absorbs at lower wavelengths (see below). The pK_a of the $[CuH_2L]$ species for PCtrl and PHis(-2) are very similar (6.52 and 6.41, respectively), perhaps confirming that copper binding in [CuHL] occurs at the His-His dyad in both peptides.

The $[CuL]^{2+}$ species is predominant in the pH range 7.6-8.9 (*ca.* 55% of total copper at pH 8.2). Two peptide nitrogen atoms and one imidazole group are coordinated to copper(II) resulting into a (N_{im}, 2x N^-, H_2O) coordination mode (***Scheme 2.1c***). The experimental λ_{max} at 565 nm at pH 8.2 is consistent with the presence of three nitrogen atoms and one oxygen in the copper(II) equatorial coordination plane, although slightly lower than that predicted using Billo's theory (calc. λ_{max} = 584 nm, ***Figure 2.6***). Further increase of the pH above 9 resulted into a large change in the circular dichroism profile (***Figure 2.6***). In this pH range the formation of $[CuLH_{-1}]^+$ occurs (*ca.* 50% of total copper at pH

9.4), with copper(II) being coordinated on the equatorial plane by one imidazole nitrogen and three deprotonated peptide nitrogen atoms (N_{im}, 3x N^-; ***Scheme 2.1***, ***d*** or ***e***), as suggested by the absorption band at 528 nm (calc. λ_{max} = 524 nm). It is worth mentioning that, for both the PCtrl and PHis(-2) peptides, the large increase in the CD signal occurs in correspondence of the formation of the $[CuLH_{-1}]^+$ species. Two successive deprotonations lead to $[CuLH_{-2}]$ (*ca.* 40% of total copper at pH 10.3) and $[CuLH_{-3}]^-$, the latter occurring above pH 9.5 where no shift of the absorption maximum and no significant change in the pattern of CD spectra were observed. These last deprotonation steps can therefore be ascribed to the two Lys residues which remain uncoordinated. It is important to point out that both the (N_{im}, 2x N^-, H_2O) and the (N_{im}, 3x N^-) coordination modes (species from $[CuL]^{2+}$ to $[CuLH_{-2}]$) can involve either the His9 or the His11/His12 site. Unfortunately, current data do not provide any indication on the preferential binding of copper(II) at either site. However, it should be noted that the (2x N_{im}, N^-, H_2O) mode in the $[CuHL]^{3+}$ species can only occur at the His11/His12 site.

Potentiometric data clearly show that bimetallic complexes form even in 1.2:1 PHis(-2)/Cu^{2+} ratio. The $[Cu_2L]^{4+}$ species starts forming at pH 5 followed by $[Cu_2LH_{-1}]^{3+}$, both not exceeding 5% of total copper (***Figure 2.5***). At pH 8, *ca.* 10% of total copper is accounted by the $[Cu_2LH_{-2}]^{2+}$ species, while $[Cu_2LH_{-3}]^+$ reaches *ca.* 20% of total copper at pH 8.9 representing the most abundant binuclear species. $[Cu_2LH_{-4}]$ and $[Cu_2LH_{-5}]^-$ successively appear above pH 9 and reach *ca.* 8-10% of total copper at their maxima. As it is reasonable to expect, in the excess of copper(II) (*ca.* 1.8:1 metal/peptide ratio) the bimetallic species dominate the speciation of the system, as shown by the distribution diagram in ***Figure 2.7***.

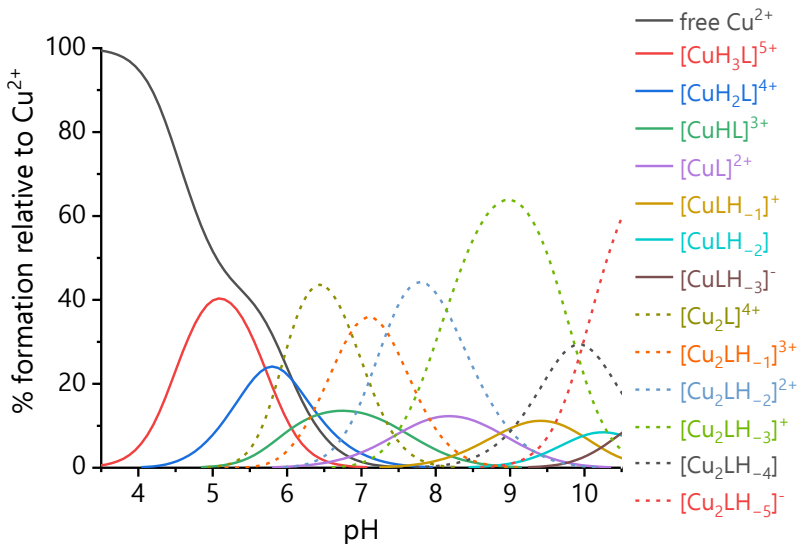


Figure 2.7 *Representative distribution diagram of the copper(II)/PHis(-2) (L) system in aqueous solution (Cu^{2+}:L = 1.7:1, C_{Cu} = 0.59 mM, T = 298.2 K, I = 0.1 M in KCl). Formation curves of bimetallic species are drawn as dashed lines.*

Visible absorption and CD spectra were collected in the excess of copper(II) (*ca.* 1.8:1 metal/peptide ratio) to gather insights on copper coordination environment of bimetallic species. Unfortunately, due to the complexity of the system, no precise interpretation as that in 1:1 metal/peptide ratio was possible (***Figure 2.8***). Nevertheless, in the initial approximation of considering the His9 and the His-His dyad as independent binding sites, it was possible to advance reasonable hypotheses on the copper(II) coordination modes by comparison with the data collected for monometallic species. ***Table 2.4*** reports the coordination modes of copper(II) in the bimetallic species at the two binding sites, which we label here ***H1*** and ***H2*** after the corresponding number of His residues. The coordination environments are proposed by comparing the observed λ_{max} in spectra collected at the pH where each species reaches a formation maximum with those calculated using the Billo's theory.

With regards to copper(II) coordination modes, some aspects need to be considered. First, both the His9 and the His-His dyad can theoretically be involved in forming the ***H1*** or ***H2*** sites. However, the (2x N_{im}, N^-, H_2O) coordination mode can evidently occur only at the ***H2*** site, as already stated. Secondly, for the PHis(-2) peptide, differently from PHis(-4) and PHis(-6), the (N_{im}, 3x N^-) coordination mode in the bimetallic $[Cu_2LH_{-3}]^+$, $[Cu_2LH_{-4}]$ and $[Cu_2LH_{-5}]^-$ species cannot involve the His11 residue at site ***H2***. In fact, if N_{im} of the (N_{im}, 3x N^-) macrochelate is His11, then the amidic nitrogen of His9 would be the third N^-. This is not possible since that donor atom would be involved in the formation of a (N_{im}, 2x N^-, H_2O) coordination environment at site ***H1***. In other words, the coordination mode depicted in ***Scheme 2.1d***, with His11 on the equatorial position and the second imidazole (His12) of the His-His dyad possibly interacting with copper(II) in axial position, cannot be obtained in bimetallic species at the ***H2*** site. However, the formation of the (N_{im}, 3x N^-) macrochelate can still form at the ***H2*** site by coordination of His12 in the equatorial position (***Scheme 2.1e***). Under these circumstances, however, His11 cannot interact in the axial position on copper(II) for geometric reasons.

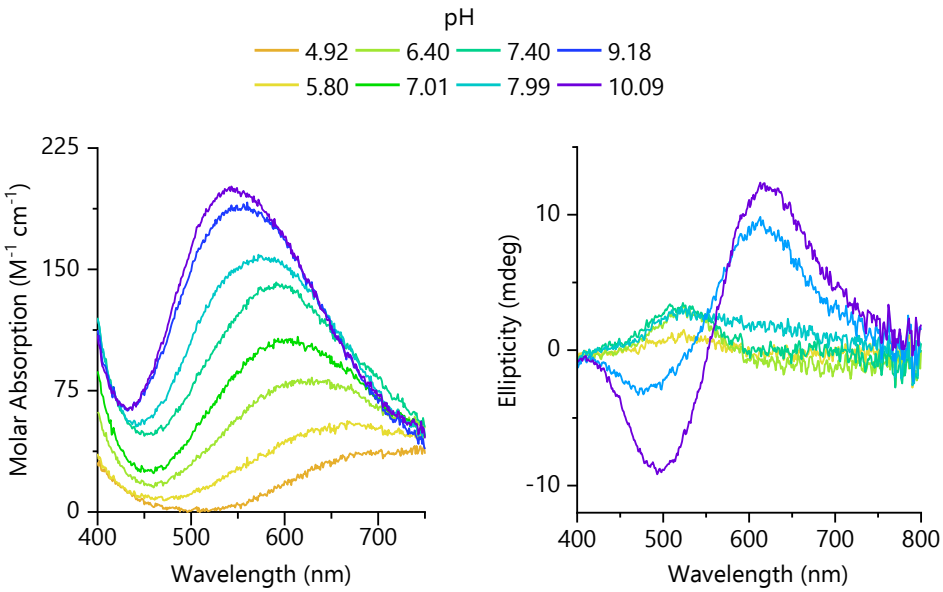


Figure 2.8 *Molar visible absorption spectra (absorbance/C_{Cu}; left) and visible CD spectra (right) of solutions of copper(II) and PHis(-2) (Cu^{2+}/PHis(-2) = 1.7:1, C_{Cu} = 0.59 mM, T = 298.2 K, I = 0.1 M in KCl).*

Table 2.4 *Proposed coordination environment of Cu^{2+} in binuclear copper(II)/PHis(-2) complex species. The λ_{max} of the various donor sets was calculated on the basis of Billo's method of the average environment.*[30–32] *For the coordination environment of the last three specie in the Table, see discussion in the text.*

Species	Cu(II) binding site ***H1*** (calc. λ_{max})	Cu(II) binding site ***H2*** (calc. λ_{max})
$[Cu_2L]^{4+}$	N_{im}, N^-, 2x H_2O (661 nm)	2x N_{im}, N^-, H_2O (608 nm)
$[Cu_2LH_{-1}]^{3+}$	N_{im}, 2x N^-, H_2O (584 nm)	2x N_{im}, N^-, H_2O (608 nm)
$[Cu_2LH_{-2}]^{2+}$	N_{im}, 2x N^-, H_2O (584 nm)	N_{im}, 2x N^-, H_2O (584 nm)
$[Cu_2LH_{-3}]^{+}$	N_{im}, 2x N^-, H_2O (584 nm) [a]	N_{im}, 3x N^- (523 nm) [a]
$[Cu_2LH_{-4}]$	N_{im}, 2x N^-, H_2O (584 nm) [a]	N_{im}, 3x N^- (523 nm) [a]
$[Cu_2LH_{-5}]^-$	N_{im}, 2x N^-, H_2O (584 nm)	N_{im}, 3x N^- (523 nm)

a In these species the two coordination modes can be distributed at both sites.

Cu^{2+}/PHis(-4) complex-formation equilibria

In the presence of copper(II), the PHis(-4) peptide forms ten complex species with both 1 to 1 and 2 to 1 copper/peptide stoichiometries, which differ for their protonation states (***Table 2.2***, ***Table 2.3***). The analysis of the potentiometric curves suggests no formation of 1:2 copper(II)/peptide species under the experimental conditions used. A representative distribution diagram is reported in ***Figure 2.9***.

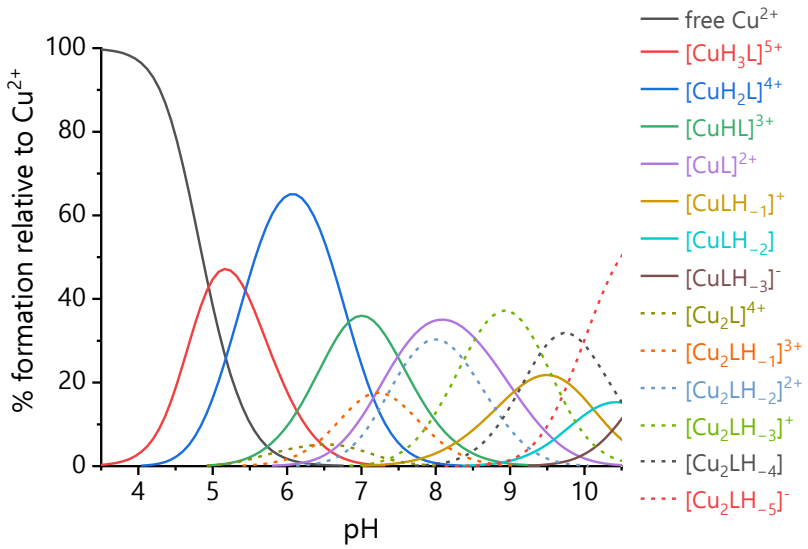


Figure 2.9 *Representative distribution diagram of the copper(II)/PHis(-4) (L) system in aqueous solution (Cu^{2+}:L = 1:1.2, C_{Cu} = 0.83 mM, T = 298.2 K, I = 0.1 M in KCl). Formation curves of bimetallic species are drawn as dashed lines.*

The complexation of Cu^{2+} starts at pH 3.5 with the formation of $[CuH_3L]^{5+}$ which accounts for 47% of total copper at pH 5.2, followed by $[CuH_2L]^{4+}$ that is the predominant species in the pH range 5.5-6.8 (*ca.* 65% of total copper at pH 6.1). Above pH 5, the $[CuHL]^{3+}$, $[Cu_2L]^{4+}$ and $[Cu_2LH_{-1}]^{3+}$ species appear, which at their maximum represent *ca.* 36%, 6% and 18% of total copper, respectively. At pH 6.5, both $[CuL]^{2+}$ and $[Cu_2LH_{-2}]^{2+}$ start forming and reach almost 35% and 30% of total copper at pH 8, respectively. At pH 7, the $[Cu_2LH_{-3}]^{+}$ and $[CuLH_{-1}]^{+}$ species appear, accounting for *ca.* 37% and 22% of total copper at pH 9 and 9.5 respectively, followed by $[Cu_2LH_{-4}]$ which represents 32% of total copper at pH 9.7. Above pH 10, the predominant species is $[Cu_2LH_{-5}]^{-}$, which starts forming at pH 9 together with $[CuLH_{-2}]$ and $[CuLH_{-3}]^{-}$ that however do not exceed 15% of total copper.

Visible absorption and CD spectra of the Cu^{2+}/PHis(-4) system at different pH values are reported in ***Figure 2.10*** (peptide/metal ratio = 1.2) and ***Figure 2.11*** (peptide/metal ratio = 0.6). Due to the great number of species that are simultaneously present in solution, it was not possible to assign each absorption λ_{max} to a single copper(II) coordination environment. Following the same approach used for the PCtrl and PHis(-2) peptides, copper(II) coordination mode at sites ***H1*** and ***H2*** were proposed on the basis of the different protonation states of the species and by comparing the experimental wavelength of the absorption maxima with those calculated using the Billo's method. These results are reported in ***Table 2.5***, while the involved coordination modes are those depicted in ***Scheme 2.1***.

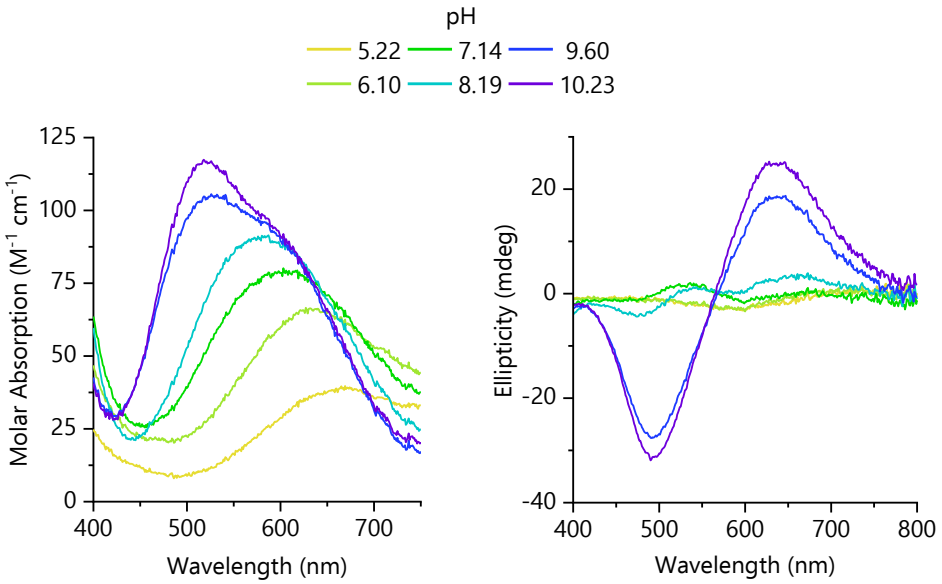


Figure 2.10 *Molar visible absorption spectra (absorbance/C_{Cu}; left) and visible CD spectra (right) of solutions of copper(II) and PHis(-4) (Cu^{2+}/PHis(-4) = 1:1.2, C_{Cu} = 0.83 mM, T = 298.2 K, I = 0.1 M in KCl).*

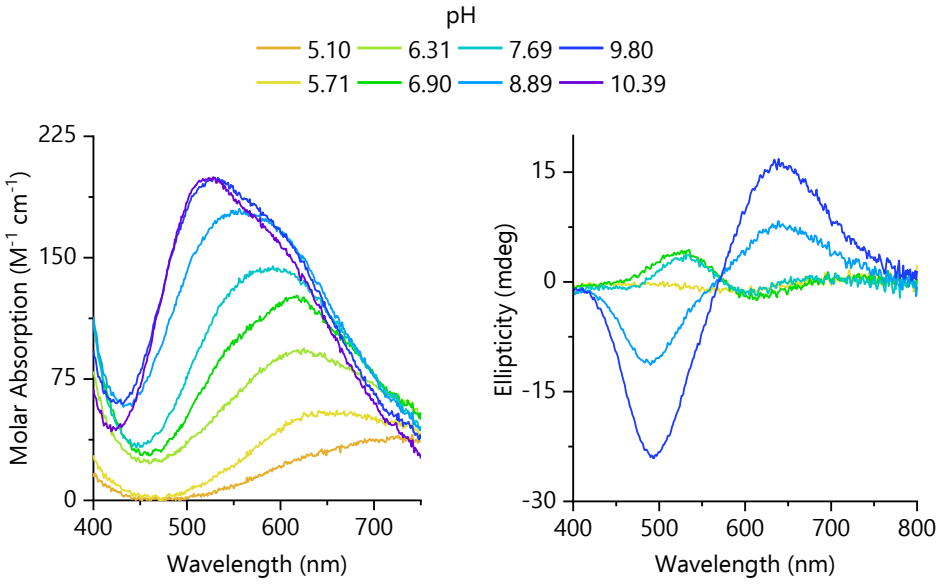


Figure 2.11 *Molar visible absorption spectra (absorbance/C_{Cu}; left) and visible CD spectra (right) of solutions of copper(II) and PHis(-4) (Cu^{2+}/PHis(-4) = 1.8:1, C_{Cu} = 0.68 mM, T = 298.2 K, I = 0.1 M in KCl).*

Table 2.5 *Proposed coordination environment of Cu^{2+} in each copper(II)/PHis(-4) complex species. The λ_{max} of the various donor sets was calculated on the basis of Billo's method of the average environment.*[30–32]

Species	Cu(II) binding site ***H1***	Cu(II) binding site ***H2***
$[CuH_3L]^{5+}$	-	2x N_{im}, 2x H_2O
$[CuH_2L]^{4+}$	-	N_{im}, N^-, 2x H_2O
$[CuHL]^{3+}$	-	2x N_{im}, N^-, H_2O
$[CuL]^{2+}$	-	N_{im}, 2x N^-, H_2O [a]
$[CuLH_{-1}]^{+}$	-	N_{im}, 3x N^- [a]
$[CuLH_{-2}]$	-	N_{im}, 3x N^- [a]
$[CuLH_{-3}]^{-}$	-	N_{im}, 3x N^- [a]
$[Cu_2L]^{4+}$	N_{im}, N^-, 2x H_2O	2x N_{im}, N^-, H_2O
$[Cu_2LH_{-1}]^{3+}$	N_{im}, 2x N^-, H_2O	2x N_{im}, N^-, H_2O
$[Cu_2LH_{-2}]^{2+}$	N_{im}, 2x N^-, H_2O	N_{im}, 2x N^-, H_2O
$[Cu_2LH_{-3}]^{+}$	N_{im}, 2x N^-, H_2O [b]	N_{im}, 3x N^- [b]
$[Cu_2LH_{-4}]$	N_{im}, 2x N^-, H_2O [b]	N_{im}, 3x N^- [b]
$[Cu_2LH_{-5}]^{-}$	N_{im}, 2x N^-, H_2O [b]	N_{im}, 3x N^- [b]

*a In these species this coordination environment can be present either at the **H1** or **H2** site. b In these species the two coordination modes can be distributed at both sites.*

Cu^{2+}/PHis(-6) complex-formation equilibria.

In the presence of copper(II), the PHis(-6) peptide forms ten complex species with both 1 to 1 and 2 to 1 copper/peptide stoichiometries and differing for their protonation states (***Table 2.2***, ***Table 2.3***). The analysis of the potentiometric curves suggests no formation of 1:2 copper(II)/peptide species under the experimental conditions used. A representative distribution diagram is reported in ***Figure 2.12***. A first evaluation of the speciation diagram reveals that in acidic pH (below 6.5) the speciation is dominated by monometallic species, while bimetallic species are mostly abundant above pH 8.

The complexation of Cu^{2+} starts at pH 3.5 with the sequential formation of $[CuH_3L]^{5+}$ (32% of total copper at pH 5) and $[CuH_2L]^{4+}$ that dominates the speciation below pH 7.3, accounting for *ca.* 90% of total copper at pH 6.1. In the pH range 6-8.5, several species are simultaneously present: $[CuHL]^{3+}$, $[CuL]^{2+}$, $[Cu_2L]^{4+}$, $[Cu_2LH_{-1}]^{3+}$, $[Cu_2LH_{-2}]^{2+}$ and $[Cu_2LH_{-3}]^{+}$. At their maxima, these species account for *ca.* 20%, 40%, 2%, 15%, 7% and 30% of total copper, respectively. At pH 9.2 $[CuLH_{-1}]^{+}$ is the most abundant species, representing *ca.* 50% of total copper, while $[Cu_2LH_{-4}]$ accounts for *ca.* 17% of total copper. The $[CuLH_{-2}]$, $[CuLH_{-3}]^{-}$ and $[Cu_2LH_{-5}]^{-}$ species appear above pH 9 and reach *ca.* 30% of total copper each at pH 10.3.

Visible absorption and CD spectra of the Cu^{2+}/PHis(-6) system at different pH values are reported in ***Figure 2.13*** (peptide/metal ratio = 1.2) and ***Figure 2.14*** (peptide/metal

ratio = 0.6). For each complex species, the copper(II) coordination modes at the ***H1*** and ***H2*** sites reported in ***Table 2.6*** and depicted in ***Scheme 2.1*** have been proposed using the same approach described for the previous three peptides.

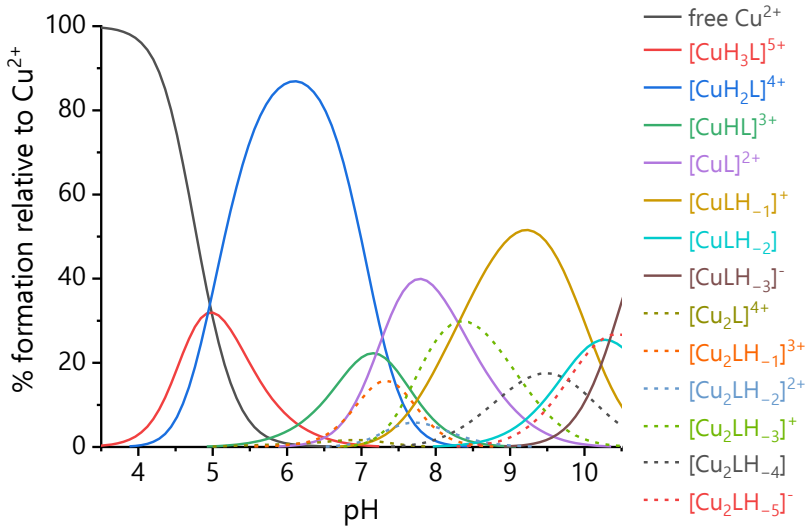


Figure 2.12 *Representative distribution diagram of the copper(II)/PHis(-6) (L) system in aqueous solution (Cu^{2+}:L = 1:1.2, C_{Cu} = 0.83 mM, T = 298.2 K, I = 0.1 M in KCl). Formation curves of bimetallic species are drawn as dashed lines.*

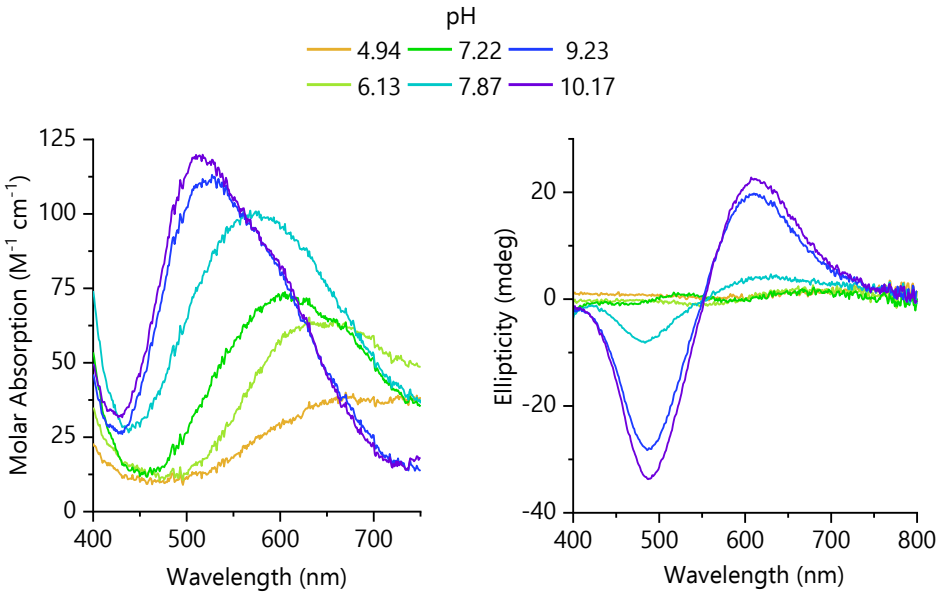


Figure 2.13 *Molar visible absorption spectra (absorbance/C_{Cu}; left) and visible CD spectra (right) of solutions of copper(II) and PHis(-6) (Cu^{2+}/PHis(-6) = 1:1.2, C_{Cu} = 0.83 mM, T = 298.2 K, I = 0.1 M in KCl).*

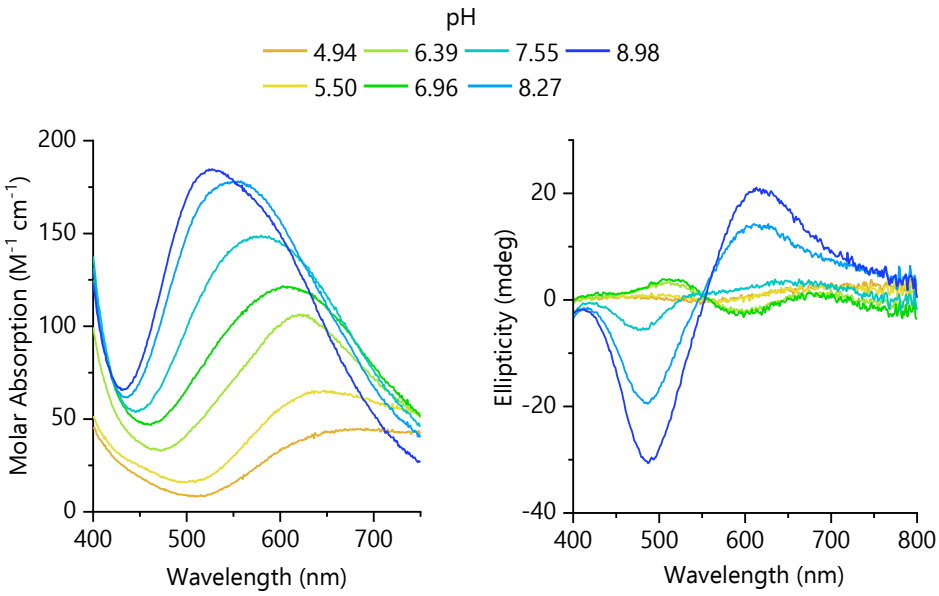


Figure 2.14 *Molar visible absorption spectra (absorbance/C_{Cu}; left) and visible CD spectra (right) of solutions of copper(II) and PHis(-6) (Cu^{2+}/PHis(-6) = 1.8:1, C_{Cu} = 1.69 mM, T = 298.2 K, I = 0.1 M in KCl).*

Table 2.6 *Proposed coordination environment of Cu^{2+} in each copper(II)/PHis(-6) complex species. The λ_{max} of the various donor sets was calculated on the basis of Billo's method of the average environment.*[30–32]

Species	Cu(II) binding site ***H1***	Cu(II) binding site ***H2***
$[CuH_3L]^{5+}$	-	2x N_{im}, 2x H_2O
$[CuH_2L]^{4+}$	-	N_{im}, N^-, 2x H_2O
$[CuHL]^{3+}$	-	2x N_{im}, N^-, H_2O
$[CuL]^{2+}$	-	N_{im}, 2x N^-, H_2O [a]
$[CuLH_{-1}]^{+}$	-	N_{im}, 3x N^- [a]
$[CuLH_{-2}]$	-	N_{im}, 3x N^- [a]
$[CuLH_{-3}]^{-}$	-	N_{im}, 3x N^- [a]
$[Cu_2L]^{4+}$	N_{im}, N^-, 2x H_2O	2x N_{im}, N^-, H_2O
$[Cu_2LH_{-1}]^{3+}$	N_{im}, 2x N^-, H_2O	2x N_{im}, N^-, H_2O
$[Cu_2LH_{-2}]^{2+}$	N_{im}, 2x N^-, H_2O	N_{im}, 2x N^-, H_2O
$[Cu_2LH_{-3}]^{+}$	N_{im}, 2x N^-, H_2O [b]	N_{im}, 3x N^- [b]
$[Cu_2LH_{-4}]$	N_{im}, 2x N^-, H_2O [b]	N_{im}, 3x N^- [b]
$[Cu_2LH_{-5}]^{-}$	N_{im}, 2x N^-, H_2O [b]	N_{im}, 3x N^- [b]

*a In these species this coordination environment can be present either at the **H1** or **H2** site. b In these species the two coordination modes can be distributed at both sites.*

Copper(II) binding affinities

The speciation models obtained through potentiometry allowed to calculate the conditional formation or *binding* constants (K_f^{II}) and affinities (K_d^{II}) at pH 7.4 for all Cu(II)/peptide systems (***Table 2.7***). A definition of both K_f^{II} and K_d^{II} was given in *Chapter 1*. Conditional constants depend, by definition, on the condition of the system. The values reported in ***Table 2.7*** were calculated by setting pH = 7.4 and equimolar copper(II) and peptide concentrations using the Hyss2009 software[33].

Table 2.7 *Logarithms conditional binding constants (log K_f^{II}) and conditional affinities (K_d^{II}) for Cu(II)/PCtrl, Cu(II)/PHis(-2), Cu(II)/PHis(-4) and Cu(II)/PHis(-6) systems at pH 7.4, calculated with the Hyss2009 software[33] using the speciation models reported in* ***Table 2.2****.*

Species	log K_f^{II}	K_d^{II}
[Cu^{II}PCtrl]	7.1(4)	79(7) nM
[Cu^{II}PHis(-2)]	7.7(4)	19(2) nM
[Cu^{II}PHis(-4)]	7.3(5)	50(6) nM
[Cu^{II}PHis(-6)]	7.4(8)	39(7) nM

The affinities of all the peptides are of the same order of magnitude, ranging from 19 nM to 79 nM. However, the higher capacity of PHis(-2), PHis(-4) and PHis(-6) to bind copper(II) is clearly evidenced by these K_d values. With respect to PCtrl, therefore, the presence of an additional copper(II) binding site other than the His-His tandem results in a higher affinity, which for PHis(-2) is *ca.* 4 times higher than that of PCtrl.

The effect of the additional His residue on the binding of copper(II) to the His-His tandem site was evaluated through the partial equilibrium constant for the dissociation of the $[CuHL]^{3+}$ species for the four peptides. As evidenced by the speciation models and distribution diagrams, all peptides form the $[CuHL]^{3+}$ species in which copper(II) is bound to the His-His tandem site (***Table 2.2***, ***Table 2.3***, and discussion). Therefore the log *K* of the dissociation equilibrium

$$[CuHL]^{3+} + H^{+} \rightleftarrows Cu^{2+} + H_2L^{2+}$$

contains the information related to the influence of the His in position 9, 7 and 5 respectively on the stability of copper(II) bound at the His11/His12 site. The *K* value of this latter equilibrium can be calculated as

$$K = \frac{[Cu^{2+}][H_2L^{2+}]}{[CuHL^{3+}][H^{+}]}$$

which corresponds to:

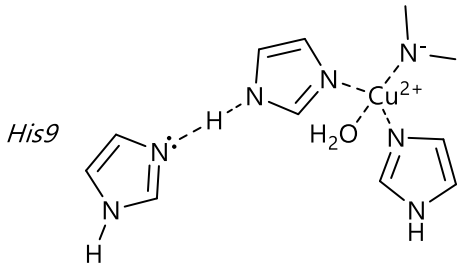
$$K = 10^{\log \beta_{H_2L} - \log \beta_{CuHL}}$$

For each peptide, log *K* values are reported in ***Table 2.8***.

Table 2.8 *Log K values for the $[CuHL]^{3+} + H^+ \rightleftarrows Cu^{2+} + H_2L^{2+}$ equilibrium. L represents the completely deprotonated form of the peptides. Log β_{H2L} and log β_{CuHL} for each peptide are also reported.*

Peptide (L)	log β_{H_2L}	log β_{CuHL}	log K
PCtrl	20.17(1)	19.30(2)	0.86(2)
PHis(-2)	20.28(1)	20.60(2)	-0.33(2)
PHis(-4)	20.38(1)	20.07(2)	0.31(2)
PHis(-6)	20.32(1)	19.94(3)	0.37(3)

With respect to PCtrl, the presence of an additional His residue in the peptide sequence results in a decrease in the log K values, indicating a stabilization of the binding of copper(II). The log K values span over more than 1 log unit, which corresponds to over 1 order of magnitude for K. This effect is larger when the His residue is in position 9, *i.e.* in proximity of the His-His tandem site, and it decreases with increasing distance between the His-His dyad and the third His in the sequence. Unfortunately, the nature of the interaction between the two binding sites that originate a stabilisation of the metal/peptide adduct cannot be established from current data. However, two hypotheses can be put forward at least for PHis(-2), for which the stabilisation effect is relevant. The first hypothesis is that the third His residue may interact directly with copper(II) bound at the His-His site. If this interaction occurs at the equatorial position by displacement of the coordinated water molecule, a blue shift from λ_{max} = 608 nm to 563 nm would be expected. However, a λ_{max} of *ca.* 590 nm was observed at the pH of maximum formation of $[CuHL]^{3+}$ of PHis(-2), which rules out a significant involvement of the third His in the coordination of copper(II) at the equatorial site. The second hypothesis is that the interaction of the third His with copper(II) occurs at the second-shell level, *e.g.* by formation of a hydrogen bond between the His in position 9, 7 or 5 and one of the two His of the dyad that are coordinated to copper. This interaction would increase the imidazolate character of the coordinated imidazole and should result into a stabler complex. Also, this interaction should result into a lower experimental wavelength of the absorption maximum compared to the expected one, which was indeed observed. However, although these hypotheses are particularly interesting, this aspect has not been further investigated and additional studies in the future are advisable. It is however possible to anticipate that data on catalysis confirm that the first hypothesis (the third His coordinated to copper(II) to explain the low log K values) is unlike, and that in the $[CuHL]^{3+}$ species of all peptides the equatorial coordination mode is reasonably (2x N_{im}, N^-, H_2O).



Scheme 2.2 *Representation of the possible second-shell (H-bond) interaction between His9 and the copper site at* ***H2*** *in* $[CuHL]^{3+}$ *of PHis(-2).*

In the attempt to summarize the data presented here, it can be stated that the analysis of both potentiometric and spectroscopic data points towards the presence in PHis(-2), PHis(-4) and PHis(-6) of two independent binding sites, namely the His-His (***H2***) site and the site represented by the third His residue (His9, 7 or 5, ***H1***). With the exception of the $[CuHL]^{3+}$ species where copper is bound to the His-His site and possibly interacting with the third histidine, in all other monometallic species the coordination modes proposed for the metal and reported in the discussion above can occur at either binding site without involvement of imidazole groups of the other site. Moreover, the capability of both sites to operate copper (II) binding also simultaneously is clearly indicated by the formation of binuclear copper(II)/peptide species. The formation of these species interestingly occurs not only in the excess of copper but also in Cu^{2+}/peptide equimolar conditions.

Oxidation of 4-methyl catechol by Copper(II)/peptide complexes

The potential pseudocatecholase catalytic properties of Cu/PCtrl, Cu/PHis(-2), Cu/PHis(-4) and Cu/PHis(-6) complexes was explored by measuring their oxidation kinetic profiles with 4-methylcatechol (MC) as the substrate, in analogy with the studies performed on the R1 and R3 tau fragments.[12] In these reactions, which mechanism is represented in ***Scheme 2.3***, MC is the reductant and dioxygen the oxidant. For each peptide, the oxidation of MC catalysed by copper(II) was studied in the presence of 3 mM substrate, 25 μM peptide and 0.25, 0.5, 0.75, 1 or 2 equivalents of Cu^{2+} *vs.* peptide. The reaction was monitored by UV-visible spectroscopy by measuring the absorption band of the product of the reaction (4-methylquinone, MQ) which absorbs at 401 nm.

The increase of absorbance at 401 nm during the oxidation of MC occurs with a biphasic behaviour (***Figure 2.15***). A fast initial step lasting *ca.* 200 s is followed by a linear

phase. This behaviour is consistent with the one previously reported for the oxidation of MC promoted by copper(II) adducts with the R3 fragment of tau. The reaction mechanism for catechol oxidation by copper complexes with His-containing peptides, including fragments of tau and prion proteins, is reported here.[12]

1. Cu^{2+} + peptide ⇆ [Cu^{2+}-peptide]
2. [Cu^{2+}-peptide] + catechol → [Cu^{+}-peptide] + semiquinone$^{\bullet+}$
3. [Cu^{+}-peptide] + catechol ⇆ [Cu^{+}-peptide-catechol]
4. [Cu^{+}-peptide-catechol] + O_2 → [Cu-peptide-catechol-O_2]
5. [Cu-peptide-catechol-O_2] →→ [Cu^{2+}-peptide] + quinone

Scheme 2.3 *Mechanism of catechol oxidation reaction by copper-complexes with His-containing peptides.*

The initial phase is ascribable to step 2, which occurs after the rapid complexation of copper(II) by the peptide. Step 2 actually consists in the reduction of copper(II) to copper(I) operated by MC with consequent release of semiquinone (sQ$^{\bullet+}$). This fast phase is named "stoichiometric phase" as it results into the stoichiometric reduction of Cu^{2+} into Cu^{+}. The successive reaction steps (3 and 4) involve the formation of a Cu^{+}/peptide/catechol adduct followed by its interaction with O_2. The latter (step 4) is the rate determining step of the MC oxidation reaction and it is related to the linear (slower) phase of the catalysis, corresponding to the increase of absorbance after the fast phase. After dioxygen is bound, copper(I) is quickly re-oxidised to copper(II) (step 5) which re-enters in the catalytic cycle, with release of MQ.

In the absence of copper(II), MC undergoes a very slow auto-oxidation process (black lines in ***Figure 2.15***). This process has a monophasic behaviour, and it is slightly faster in the presence of the peptides, possibly because it is assisted by some Brønsted acid-base catalysis (orange lines in ***Figure 2.15***). The rapid oxidation of MC is triggered by the presence of copper(II), and further increase in the rate is observed in the presence of copper(II) and the peptides. In particular, ***Figure 2.15*** shows that, at a given peptide concentration, the reaction rate progressively increases with increasing metal concentration, getting closer to a plateau upon reaching stoichiometric copper(II)/peptide conditions. This phenomenon is related to Cu + L = CuL equilibria, and it is consistent with the presence of progressively higher concentrations of Cu/peptide adducts in the reaction environment. The first important consideration on these kinetic data is therefore that the reaction rate increases for increasing concentration of copper(II)/peptide adducts. This is consistent with the mechanism, as step 4 is first-order in copper(II)/peptide concentration.

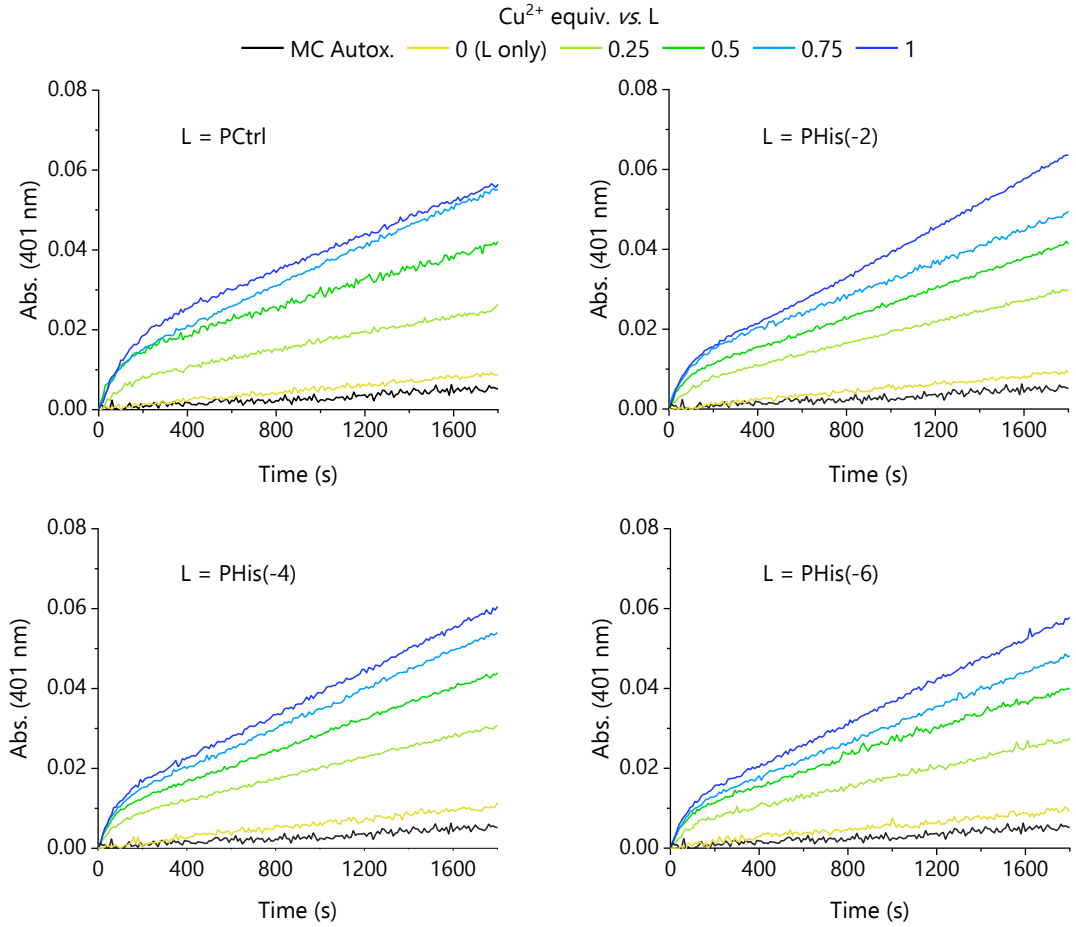


Figure 2.15 *Kinetic profiles (abs@401 nm vs. time) of 4-methyl catechol oxidation by Cu^{2+}/PCtrl (top left), Cu^{2+}/PHis(-2) (top right), Cu^{2+}/PHis(-4) (bottom left) and Cu^{2+}/PHis(-6) (bottom right) systems. For each data set: 3 mM MC, 25 μM peptide, 50 mM HEPES, pH 7.4, T = 293.2 K. MC Auto-oxidation (black curves) was evaluated in the absence of the peptides.*

In ***Figure 2.16***, the tracks of absorbance *vs.* time are grouped by Cu^{2+}/peptide ratio. Each plot groups together the kinetic profiles for the four different peptides at the same copper(II)/peptide stoichiometric ratio. Very interestingly, data in ***Figure 2.16*** show that for the same copper(II)/peptide ratio the plots of absorbance *vs.* time are very similar, irrespective of the nature of the peptide. Consequently, and unexpectedly, the rate of reaction does not depend on the presence in the peptide sequence of only the His-His dyad or of the dyad plus a third His residue. This effect is very pronounced for the curves collected in the presence of equimolar copper(II) and peptide amounts (***Figure 2.16*** upper left plot). Some differences were observed in the presence of a peptide excess, while the most significant differences, although not large, were observed in the excess of copper.

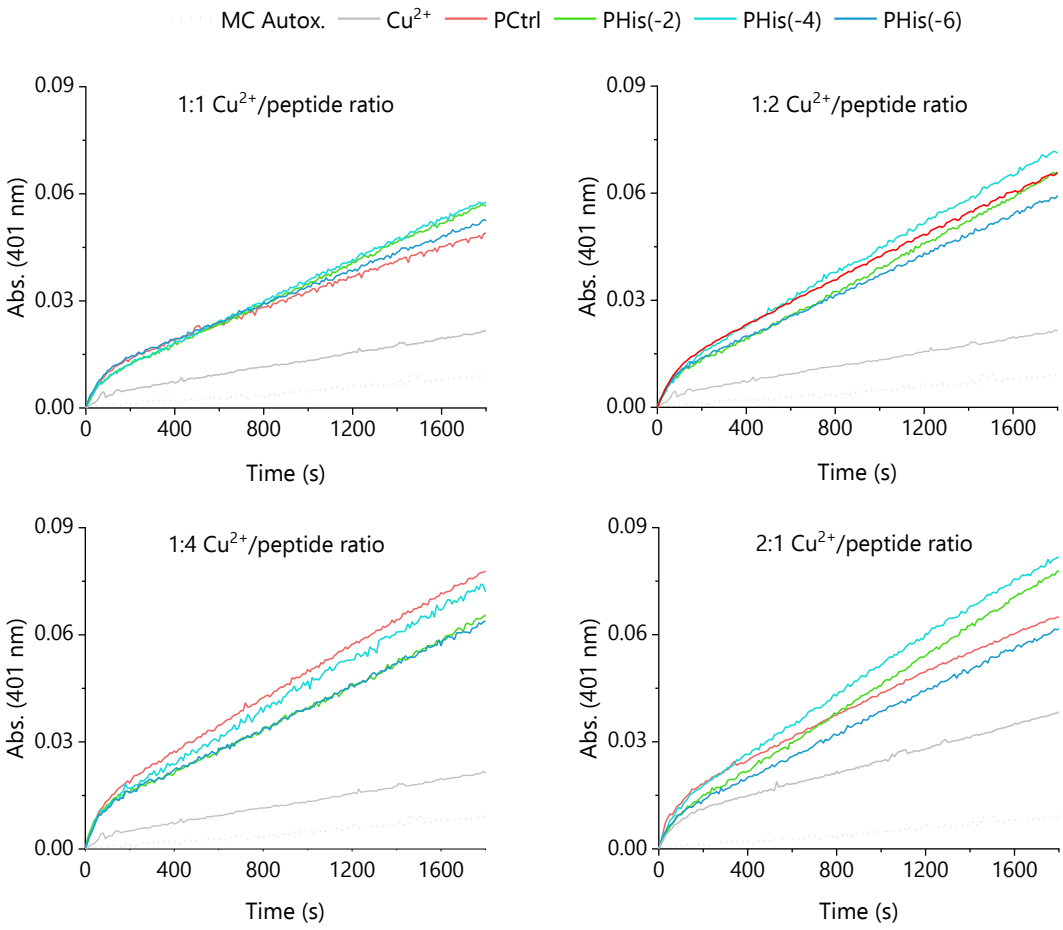


Figure 2.16 *Kinetic profiles (abs@401 nm vs. time) of 4-methyl catechol oxidation by Cu^{2+} in the presence of 1 (top left), 2 (top right), 4 (bottom left) and 0.5 (bottom right) equivalents of PCtrl (red), PHis(-2) (green), PHis(-4) (light blue) and PHis(-6) (blue) vs. Cu^{2+}. For each data set: 3 mM MC, 25 μM Cu^{2+}, 50 mM HEPES, pH 7.4, T = 293.2 K. MC Auto-oxidation (black curves) was evaluated in the absence of Cu^{2+}.*

Our work on the R1 and R3 fragments of tau (see *Chapter 1*) demonstrated that the catalytic behaviour pivots around the coordination of copper to the His-His site and in a (2x N_{im}, N^-, H_2O) coordination environment. In this perspective, data collected in equimolar copper(II)/peptide ratio are possibly most interesting. As it can be seen from the distribution diagrams in ***Figure 2.3***, ***Figure 2.5***, ***Figure 2.9*** and ***Figure 2.12***, the principal species that bears a copper(II) ion in a (2x N_{im}, N^-, H_2O) environment is $[CuHL]^{3+}$, which at pH 7.4 accounts for *ca.* 13%, 50%, 30% and 20% of total copper for PCtrl, PHis(-2), PHis(-4) and PHis(-6), respectively. For PHis(-4) and PHis(-6), the $[Cu_2LH_{-1}]^{3+}$ species, which bears the same copper site, is also present in significative amounts (*ca.* 16% of total copper for both peptides). The remaining copper(II) in all systems is basically bound to the

peptides in a (N_{im}, 2x N^-, H_2O) coordination mode, which is catalytically inactive. Based on these observations, the plots in ***Figure 2.16***for equimolar copper(II)/peptide ratio may appear difficult to rationalise only on the basis of the coordination environment of copper(II).

To try to get more information on the behaviour of these systems we have evaluated the oxidation of MC promoted by copper(II) in the presence of increasing amounts of the peptides. As a general trend, a small peptide effect (*i.e.* an increase in the rate upon increasing equivalents of peptide) was observed going from a peptide/copper(II) ratio of 1 to 4 (***Figure 2.16***). This effect may relate to a decrease in the amount of bimetallic species in favour of monometallic ones (in particular $[CuHL]^{3+}$) for PHis(-2), PHis(-4) and PHis(-6). In fact, the overall amount of bimetallic species for PHis(-4), for instance, decreases from 35% of total copper to less than 10%, with an increase of $[CuHL]^{3+}$ from 30 to 45%. However, this does not explain the peptide effect observed for PCtrl, which does not form bimetallic adducts. For this peptide, the increased reaction rate with increased equivalents of peptide is thought to relate with the reduction of free copper(II) in solution in favour of its complexed forms, overlapped with some peptide-assisted catalysis. Interestingly, a similar effect was observed also for the R3 fragment of tau (see *Chapter 1* and Ref. 12).

Overall, kinetic data show that the catalytic activity of copper adducts of PHis(-2), PHis(-4) and PHis(-6) is extremely similar, irrespective of the presence or the absence of a third His residue in the peptide sequence, and comparable to that of PCtrl. Moreover, kinetic traces in the excess of copper(II) (2:1 metal/peptide ratio) show that the reaction rate is similar to the one observed in equimolar condition, and only for PHis(-2) and PHis(-4) an increase of *ca.* 30% is observed. This behaviour can be explained by taking into account the binding of copper(I), rather than the speciation of copper(II), and the reaction mechanism. As demonstrated for the R1 and R3 fragments of tau, a single isolated His residue has an extremely low affinity for copper(I). Hence, the cycling of the copper ion between its +1 and +2 oxidation states is prevented at that site, resulting in a negligible catalytic activity. This provides a partial explanation of the catalytic behaviour: in the absence of a direct coordination interaction of the third His residue with the copper ion coordinated to the His-His site, the metal coordinated to the ***H2*** site operates the catalysis, as expected. However, PCtrl has only the His-His dyad as copper binding site, while for PHis(-2), PHis(-4) and PHis(-6) copper(II) is not coordinated exclusively by the His-His tandem site. Part of Cu^{2+}, in fact, is coordinated to the ***H1*** site also in equimolar metal/peptide ratio.

Our interpretation is therefore the following. During the stoichiometric reduction phase copper(I) gets progressively reduced and, since the process of re-oxidation by dioxygen is slow, it accumulates at the ***H2*** site. This occurs even in the case of copper not being all coordinated to the His-His dyad but rather distributed over the His11/His12 and the third His sites: the presence of free His-His dyad progressively bind copper upon its

reduction. Therefore, at the end of the stoichiometric phase copper(I) is all bound to the ***H2*** site and at that site it undergoes the catalytic cycles, likely without re-migration to the ***H1*** site (***Figure 2.17***). The ultimate consequence is that PCtrl, PHis(-2), PHis(-4) and PHis(-6) are all similar in terms of catalytical redox properties, in agreement with the experimental observations.

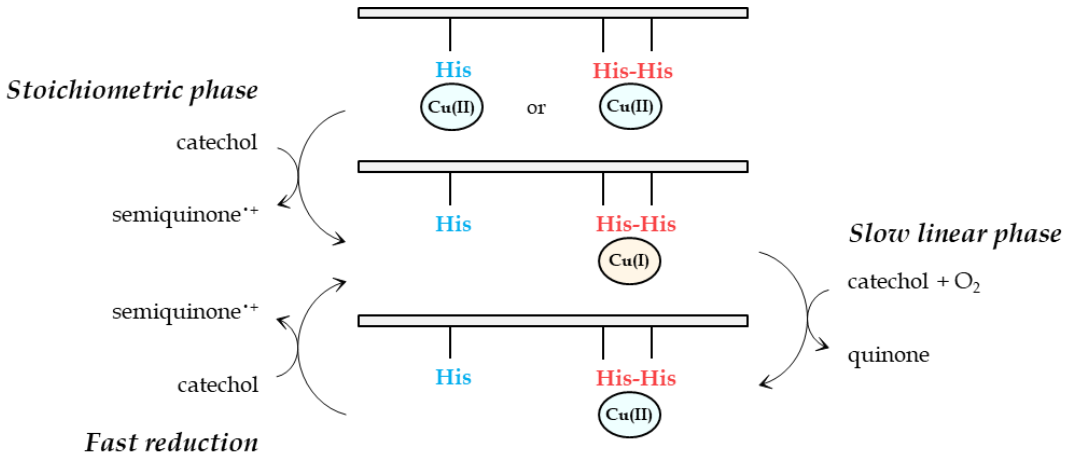


Figure 2.17 *Schematic representation of the proposed mechanism for copper reduction/oxidation by catechol in monometallic Cu/peptide species.*

Data collected in the excess of copper(II) (2 equiv. *vs.* peptide; ***Figure 2.16***, bottom right) show a limited increase in the reaction rate, largely less than doubled with respect to equimolar peptide/metal ratio. This behaviour can be rationalised as follows: with PHis(-2), PHis(-4) and PHis(-6) the first equivalent of copper(II) forces the His-His site to saturate, while the second equivalent is coordinated to the third His site. In the presence of catechol and dioxygen only the former ***H2*** site is catalytically active, and the behaviour of the three peptides parallel that of PCtrl (***Figure 2.18***). As for PCtrl, the excess of copper(II) cannot bind to the peptide. It rather remains uncomplexed copper(II), which is poorly active in promoting the oxidation reaction. Therefore, overall the four peptides operate pseudocatecholase activity in the same way, provided they are compared for the same copper/peptide ratio. Finally, it is important to underline that the kinetic data in the excess of copper(II) are those with the largest variability from peptide to peptide. The possible interaction between the two bound copper ions, and the effects of proximity (being them cooperative or not) are still not clear and their study is worth to be tackled in the future, along with the study of the binding affinities for copper(I).

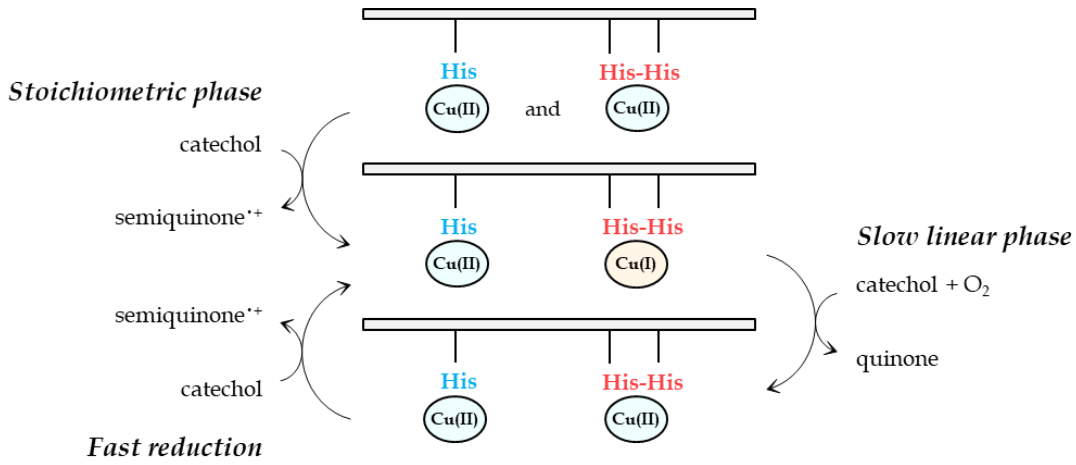


Figure 2.18 *Schematic representation of the proposed mechanism for copper reduction/oxidation by catechol in bimetallic Cu/peptide species.*

CONCLUSIONS

This work describes the systematic study of four tau-inspired peptides in terms of copper(II) binding and pseudocatecholase activity. The main feature of these peptides consists in the presence of a bis-His motif, which is known for promoting the reduction of copper(II) to copper(I) and for its involvement in ROS production. The presence of an additional histidine residue in the peptide sequence of PHis(-2), PHis(-4) and PHis(-6) was introduced with the aim of exploring the changes of the binding properties and the reactivity of their HH dyads, in the hypothesis that those depend on the relative position of the mono- and bis-His sites.

The analysis of both potentiometric and spectroscopic data suggests the presence in PHis(-2), PHis(-4) and PHis(-6) of two independent binding sites, namely the His-His dyad and the site represented by the third His residue (His9, 7 or 5). Moreover, these sites can bind copper(II) also simultaneously, as clearly indicated by the formation of binuclear copper(II)/peptide species which occurs not only in the excess of copper but also in Cu^{2+}/peptide equimolar conditions. For all PCtrl, PHis(-2), PHis(-4) and PHis(-6), the $[CuHL]^{3+}$ species is that in which copper(II) is bound to the bis-His site with a (2x N_{im}, N^-, H_2O) coordination mode, which was expected to be, among the mononuclear species, that with the catalytic activity.

The pseudocatecholase activity of the Cu(II)/peptide systems with 4-methyl catechol as the substrate resulted peculiar. Contrarily to our expectations, kinetic data show that the presence of a third His residue in the peptide sequence of PHis(-2), PHis(-4) and PHis(-6)

does not impair nor significantly accelerate the catalytic activity of copper adducts with all peptides, which remains comparable to those with PCtrl. This was explained on the basis of the extremely low binding affinity that a single isolated His residue has for copper(I), which prevents the cycling of copper between its +1 and +2 oxidation states. Rather, the copper(II) ions are reduced and progressively accumulate to the His-His site likely during the "stoichiometric phase" of the catalytic process. Beyond PCtrl, which has only the His-His dyad as copper binding site, all PHis(-2), PHis(-4) and PHis(-6) show therefore similar catalytic properties toward the oxidation of 4-methyl catechol.

In this work we provided insight into copper(II) binding to His-His motifs in the presence of mono-His sites, clarifying the effect that the latter have on the reactivity of these systems. These data will shed more light on the evaluation of the structure-stability-activity relationships of metallopeptides in the field of study of neurodegenerative diseases.

EXPERIMENTAL SECTION

Peptide synthesis and purification

PCtrl	Ac-GWKAASAGAIHHKPG-NH_2	mw 1528
PHis(-2)	Ac-GWKAASAGHIHHKPG-NH_2	mw 1594
PHis(-4)	Ac-GWKAASHGAIHHKPG-NH_2	mw 1594
PHis(-6)	Ac-GWKAHSAGAIHHKPG-NH_2	mw 1594

The four peptides were prepared using the standard fluorenyl methoxycarbonyl (Fmoc) solid-phase peptide synthesis in DMF[34–36]. Rink-amide resin MBHA (0.25 mmol scale, 0.59 mmol/g) was used as polymeric support, which yielded the peptide amidated at the C-terminus. In detail, the coupling reactions of the amino acids (4 equiv.) were carried out at 75°C under mixing for 5 min (50°C and 6.5-min mixing for histidine), using HBTU (3.9 equiv.) as coupling agent, and DIEA as base. Fmoc protecting groups were always removed with a 20% piperidine in DMF solution (room temperature, 13 min). All peptides were acetylated at the N-terminal using 10% acetic anhydride in DMF (room temperature, 10 min). Then, they were manually deprotected and cleaved from the resin by adding a TFA/TIS/water mixture (%v/v = 95 : 2.5 : 2.5) and stirring for 2h at room temperature. Crude peptides were precipitated by addition of cold diethyl ether and dissolved in a small amount of water. The crude peptides were purified by preparative reversed-phase HPLC (Phenomenex Jupiter Proteo column, 250 × 21.2 mm, 4 μm, 90 Å, C12) using solvent A (water/TFA %v/v = 99.9 : 0.1) and solvent B (CH_3CN/water/TFA %v/v = 90 : 9.9 : 0.1). The peptides were eluted from the column with a linear gradient from 20% to 25% B in 10 min at a flow rate of 10 mL/min. Their purity was checked by analytical reversed-phase HPLC (Phenomenex Jupiter Proteo column, 250 × 4.6 mm, 4 μm, 90 Å) and it was greater than 95%. Purified peptides were lyophilized and stored at -20°C until use.

Characterization of all the peptides was carried out at Spectropole, the Analytical Facility of Aix-Marseille University, by Electrospray Ionization-Mass spectrometry (ESI-MS) in positive mode using a Waters Synapt G2 HDMS (Manchester, UK) equipped with an ESI source employing the following parameters: ESI capillary voltage: +2.8 kV; extraction cone voltage: +20 V; desolvation gas (N_2) flow: 100 L/h; source temperature: 35 °C. Sample solutions were introduced in the ionization source at a 10 μL/min using a syringe pump.

PCtrl_data: m/z 764.9 ± 0.5 $[M + 2H]^{2+}$ and 510.3 ± 0.5 $[M + 3H]^{3+}$ (calculated for $[M + H]^+$: 1528.8 Da; $[M + 2H]^{2+}$: 764.9 Da; $[M + 3H]^{3+}$: 510.3 Da).

PHis(-2) data: m/z 1594.9 ± 0.5 $[M + H]^+$, 797.9 ± 0.5 $[M + 2H]^{2+}$, 532.3 ± 0.5 $[M + 3H]^{3+}$ and 399.5 ± 0.5 $[M + 4H]^{4+}$ (calculated for $[M + H]^+$: 1594.8 Da; $[M + 2H]^{2+}$: 797.9 Da; $[M + 3H]^{3+}$: 532.3 Da; $[M + 4H]^{4+}$: 399.5 Da).

PHis(-4) data: m/z 797.9 ± 0.5 $[M + 2H]^{2+}$, 532.3 ± 0.5 $[M + 3H]^{3+}$ and 399.5 ± 0.5 $[M + 4H]^{4+}$ (calculated for $[M + H]^{+}$: 1594.8 Da; $[M + 2H]^{2+}$: 797.9 Da; $[M + 3H]^{3+}$: 532.3 Da; $[M + 4H]^{4+}$: 399.5 Da).

PHis(-6) data: m/z 797.9 ± 0.5 $[M + 2H]^{2+}$, 532.3 ± 0.5 $[M + 3H]^{3+}$ and 399.5 ± 0.5 $[M + 4H]^{4+}$ (calculated for $[M + H]^{+}$: 1594.8 Da; $[M + 2H]^{2+}$: 797.9 Da; $[M + 3H]^{3+}$: 532.3 Da; $[M + 4H]^{4+}$: 399.5 Da).

Copper(II) binding to His-containing peptides

Formation equilibria studies. Concentrated (ca. 1-10 mM) solutions of PCtrl, PHis(-2), PHis(-4), and PHis(-6) were prepared by weight in Milli-Q water and the peptide concentration was determined by UV-Vis spectroscopy using the extinction coefficient of Trp (ε_{280nm} = 5690 $M^{-1}cm^{-1}$).[37] The titre was consistent with what obtained by potentiometric titrations. $CuCl_2$, and $Cu(NO_3)_2$ stock solutions were prepared by weight from analytical grade metal salts and standardized by complexometric titrations with EDTA following standard protocols.[38] Standard KOH solutions were prepared by diluting carbon dioxide-free concentrated solutions in freshly boiled, Milli-Q water under N_2 atmosphere. The title of KOH solutions (*ca.* 0.05 M) was determined by potentiometric titrations against potassium hydrogen phthalate. Alternatively, carbonate-free solutions of the titrant KOH were obtained at 0.100 M by using freshly prepared solutions from a Merck ampoule in 1000 mL of freshly boiled, Milli-Q water. These solutions were standardized by application of Gran's method.[39] Solutions of HCl in water were prepared by dilution of concentrated solutions of the acids and their titre determined by titrations with standard KOH using the Gran's method.

Potentiometric measurements. Potentiometric titrations were carried out in aqueous solution at T = 298.2 K and I = 0.1 M (KNO_3 or KCl) under a N_2 stream, using 25 to 28.4 mL sample volumes for PCtrl, and 1.5 to 2.1 mL sample volumes for PHis(-2), PHis(-4) and PHis(-6). The potentiometric setup has been described in the *Experimental section* of *Chapter 1*. A Metrohm semimicro electrode was used, which was calibrated in terms of $[H^+]$ by titrating HCl solutions with a KOH standardized solution. The electrode parameters were determined using the Gran's method, and the pK_w value resulted to be 13.76(1).[39] Protonation studies of the peptides were carried out by alkalimetric titrations in the pH range 2.8-10.2 ($C_{peptide}$ = 0.34-0.56 mM). Cu^{2+} complex formation constants were determined by alkalimetric titrations in the pH range 2.8-10.2 (Cu:peptide = 1:0.5/1.2/2.1, C_{Cu} = 0.19-0.81 mM). A minimum of two replicate titrations were performed for each system. Overall equilibrium constants $\beta_{M_mH_hL_l}$ (being $\beta_{M_mH_hL_l} = [M_mH_hL_l]/[M]^m[H]^h[L]^l$) were calculated by fitting the potentiometric data from protonation or complexation titrations with the HyperQuad 2013 software.[40] The calculated constants were used to draw the species distribution curves with the Hyss2009 software.[33]

UV-Visible spectroscopy. UV-Vis spectra were collected at T = 298.2 K and *I* = 0.1 M (KCl) on a Thermo Scientific Evolution 260 Bio spectrophotometer provided with a Peltier thermostat, using 1 cm path length black quartz cuvettes. Spectra of the copper(II)/peptide systems at different pH values were acquired in the spectral range 350-900 nm. Samples were prepared in doubly distilled water by mixing proper amounts of peptide and $CuCl_2$ stock solutions to obtain 0.83 mM final copper(II) concentration. Copper(II):peptide ratios were in the 1:1.2-2.1 range for all the peptides. The pH was adjusted to the desired values by addition of concentrated KOH or HCl.

Circular dichroism (CD) spectroscopy. CD spectra were collected at T = 298.2 K and *I* = 0.1 M (KCl) with a Jasco J715 spectropolarimeter equipped with a Peltier thermostat, using 1 cm path length black quartz cuvettes. Spectra of the copper(II)/peptide systems at different pH values were acquired in the spectral range 400-800 nm. Samples were prepared in doubly distilled water by mixing proper amounts of peptide and $CuCl_2$ stock solutions to obtain 0.83 mM final copper(II) concentration. Copper(II):peptide ratios were in the 1:1.2-2.1 range for all the peptides. The pH was adjusted to the desired values by addition of concentrated KOH or HCl.

Oxidation of 4-methyl catechol

The catalytic oxidation of MC by Cu^{2+} was studied at 20 °C in 50 mM aqueous HEPES buffer at pH 7.4, saturated with atmospheric oxygen. The reaction was monitored by UV-visible spectroscopy on a Photodiode Array Lambda 465 spectrophotometer provided with a Peltier thermostat, following the absorption band of MQ at 401 nm (ε = 1550 $M^{-1}cm^{-1}$) for a reaction time of 1800 sec. First, MC autoxidation experiment and kinetic profile of the reaction catalysed by Cu^{2+} were evaluated. The effect of PCtrl, PHis(-2), PHis(-4) and PHis(-6) peptides on the copper-catalysed MC oxidation was studied by adding the peptides in variable stoichiometries (0-4 equivalents *vs.* Cu^{2+}) to the reaction solution. These experiments were carried out by adding copper(II) chloride (25 μM) as the last reagent to the solution containing 3 mM MC and 0-100 μM PCtrl, PHis(-2), PHis(-4) or PHis(-6). Further measurements were performed on solutions containing 3 mM MC and 25 μM PCtrl, PHis(-2), PHis(-4) or PHis(-6) to which Cu^{2+} was added in variable stoichiometries (0.25, 0.5, 0.75 equivalents *vs.* peptide). All experiments were performed in triplicate.

REFERENCES

1. Kozłowski, H., Bal, W., Dyba, M. & Kowalik-Jankowska, T. Specific structure–stability relations in metallopeptides. *Coord. Chem. Rev.* ***184***, 319–346 (1999).
2. Sóvágó, I., Kállay, C. & Várnagy, K. Peptides as complexing agents: Factors influencing the structure and thermodynamic stability of peptide complexes. *Coord. Chem. Rev.* ***256***, 2225–2233 (2012).
3. Kozłowski, H., Kowalik-Jankowska, T. & Jezowska-Bojczuk, M. Chemical and biological aspects of Cu2+ interactions with peptides and aminoglycosides. *Coord. Chem. Rev.* ***249***, 2323–2334 (2005).
4. Sóvágó, I., Várnagy, K., Lihi, N. & Grenács, Á. Coordinating properties of peptides containing histidyl residues. *Coord. Chem. Rev.* ***327–328***, 43–54 (2016).
5. Potocki, S., Valensin, D. & Kozlowski, H. The specificity of interaction of Zn2+, Ni2+ and Cu2+ ions with the histidine-rich domain of the TjZNT1 ZIP family transporter. *Dalt. Trans.* ***43***, 10215–10223 (2014).
6. Pushie, M. J., Shaw, K., Franz, K. J., Shearer, J. & Haas, K. L. Model Peptide Studies Reveal a Mixed Histidine-Methionine Cu(I) Binding Site at the N-Terminus of Human Copper Transporter 1. *Inorg. Chem.* ***54***, 8544–8551 (2015).
7. Kulon, K. *et al.* The His-His sequence of the antimicrobial peptide demegen P-113 makes it very attractive ligand for Cu2+. *J. Inorg. Biochem.* ***102***, 960–972 (2008).
8. Conklin, S. E. *et al.* Specific Histidine Residues Confer Histatin Peptides with Copper-Dependent Activity against Candida albicans. *Biochemistry* ***56***, 4244–4255 (2017).
9. Chauhan, S., Hosseinzadeh, P., Lu, Y. & Blackburn, N. J. Stopped-Flow Studies of the Reduction of the Copper Centers Suggest a Bifurcated Electron Transfer Pathway in Peptidylglycine Monooxygenase. *Biochemistry* ***55***, 2008–2021 (2016).
10. Csire, G. *et al.* Coordination, redox properties and SOD activity of Cu(II) complexes of multihistidine peptides. *J. Inorg. Biochem.* ***177***, 198–210 (2017).
11. Jakab, N. I., Jancsó, A., Gajda, T., Gyurcsik, B. & Rockenbauer, A. Copper(II), nickel(II) and zinc(II) complexes of N-acetyl-His-Pro-His-His-NH2: Equilibria, solution structure and enzyme mimicking. *J. Inorg. Biochem.* ***102***, 1438–1448 (2008).
12. Bacchella, C. *et al.* Binding and Reactivity of Copper to R1 and R3 Fragments of tau Protein. *Inorg. Chem.* ***59***, 274–286 (2020).
13. Kozlowski, H., Luczkowski, M., Remelli, M. & Valensin, D. Copper, zinc and iron in neurodegenerative diseases (Alzheimer's, Parkinson's and prion diseases). *Coord. Chem. Rev.* ***256***, 2129–2141 (2012).

14. Pirota, V., Dell'Acqua, S., Monzani, E., Nicolis, S. & Casella, L. Copper-Aβ Peptides and Oxidation of Catecholic Substrates: Reactivity and Endogenous Peptide Damage. *Chem. – A Eur. J.* ***22***, 16964–16973 (2016).
15. Bagheri, S., Squitti, R., Haertlé, T., Siotto, M. & Saboury, A. A. Role of copper in the onset of Alzheimer's disease compared to other metals. *Front. Aging Neurosci.* ***9***, 1–15 (2018).
16. Faller, P., Hureau, C. & Berthoumieu, O. Role of Metal Ions in the Self-assembly of the Alzheimer's Amyloid-β Peptide. *Inorg. Chem.* ***52***, 12193–12206 (2013).
17. Hureau, C. Coordination of redox active metal ions to the amyloid precursor protein and to amyloid-β peptides involved in Alzheimer disease. Part 1: An overview. *Coord. Chem. Rev.* ***256***, 2164–2174 (2012).
18. Migliorini, C., Porciatti, E., Luczkowski, M. & Valensin, D. Structural characterization of Cu2+, Ni2+ and Zn2+ binding sites of model peptides associated with neurodegenerative diseases. *Coord. Chem. Rev.* ***256***, 352–368 (2012).
19. Ma, Q. F. *et al.* Binding of copper (II) ion to an Alzheimer's tau peptide as revealed by MALDI-TOF MS, CD, and NMR. *Biopolymers* ***79***, 74–85 (2005).
20. Shin, B. K. & Saxena, S. Insight into potential cu(II)-binding motifs in the four pseudorepeats of tau protein. *J. Phys. Chem. B* ***115***, 15067–15078 (2011).
21. Balogh, B. D. *et al.* Copper (II) binding properties of an octapeptide fragment from the R3 region of tau protein: A combined potentiometric, spectroscopic and mass spectrometric study. *J. Inorg. Biochem.* ***217***, 111358 (2021).
22. Cheignon, C. *et al.* Identification of key structural features of the elusive Cu–Aβ complex that generates ROS in Alzheimer's disease. *Chem. Sci.* ***8***, 5107–5118 (2017).
23. Hureau, C. Coordination of redox active metal ions to the amyloid precursor protein and to amyloid-β peptides involved in Alzheimer disease. Part 1: An overview. *Coord. Chem. Rev.* ***256***, 2164–2174 (2012).
24. Hureau, C. & Dorlet, P. Coordination of redox active metal ions to the amyloid precursor protein and to amyloid-β peptides involved in Alzheimer disease. Part 2: Dependence of Cu(II) binding sites with Aβ sequences. *Coord. Chem. Rev.* ***256***, 2175–2187 (2012).
25. De Gregorio, G. *et al.* Structural analysis of copper(I) interaction with amyloid β peptide. *J. Inorg. Biochem.* ***195***, 31–38 (2019).
26. Himes, R. A., Park, G. Y., Siluvai, G. S., Blackburn, N. J. & Karlin, K. D. Structural Studies of Copper(I) Complexes of Amyloid-β Peptide Fragments: Formation of Two-Coordinate Bis(histidine) Complexes. *Angew. Chemie Int. Ed.* ***47***, 9084–9087 (2008).
27. Shearer, J. & Szalai, V. A. The amyloid-β peptide of Alzheimer's disease binds cu 1 in a linear bis-his coordination environment: Insight into a possible neuroprotective mechanism for the amyloid-β peptide. *J. Am. Chem. Soc.* ***130***, 17826–17835 (2008).

28. Himes, R. A., Ga, Y. P., Barry, A. N., Blackburn, N. J. & Karlin, K. D. Synthesis and x-ray absorption spectroscopy structural studies of Cu(I) complexes of HistidylHistidine peptides: The predominance of linear 2-coordinate geometry. *J. Am. Chem. Soc.* ***129***, 5352–5353 (2007).
29. Jancsó, A. *et al.* Solution chemical properties and catecholase-like activity of the copper(II)–Ac-His-His-Gly-His-OH system, a relevant functional model for copper containing oxidases. *Dalt. Trans.* 3187–3194 (2005) doi:10.1039/B507655B.
30. Billo, E. J. Copper(II) chromosomes and the rule of average environment. *Inorg. Nucl. Chem. Lett.* ***10***, 613–617 (1974).
31. Prenesti, E., Daniele, P. G., Prencipe, M. & Ostacoli, G. Spectrum-structure correlation for visible absorption spectra of copper(II) complexes in aqueous solution. *Polyhedron* ***18***, 3233–3241 (1999).
32. Sigel, H. & Martin, R. B. Coordinating properties of the amide bond. Stability and structure of metal ion complexes of peptides and related ligands. *Chem. Rev.* ***82***, 385–426 (1982).
33. Alderighi, L. *et al. Hyperquad simulation and speciation (HySS): a utility program for the investigation of equilibria involving soluble and partially soluble species. Coordination Chemistry Reviews* vol. 184 http://www.eawag.ch/soft/ (1999).
34. Merrifield, R. B. Solid Phase Peptide Synthesis. I. The Synthesis of a Tetrapeptide. *J. Am. Chem. Soc.* ***85***, 2149–2154 (1963).
35. FIELDS, G. B. & NOBLE, R. L. Solid phase peptide synthesis utilizing 9-fluorenylmethoxycarbonyl amino acids. *Int. J. Pept. Protein Res.* ***35***, 161–214 (1990).
36. Chan, W. C.; White, P. D. *Fmoc Solid Phase Peptide Synthesis: A Practical Approach.* (Oxford University Press, 2000).
37. Pace, C. N. & Schmid, F. X. *Protein structure: a practical approach.* (Oxford University Press, 1997).
38. Vogel, A. I. *Quantitative Inorganic Analysis Including Elementary Instrumental Analysis.* (Longmans, 1962).
39. Gran, G. Determination of the equivalence point in potentiometric titrations. Part II. *Analyst* **77**, 661 (1952).
40. Gans, P., Sabatini, A. & Vacca, A. Investigation of equilibria in solution. Determination of equilibrium constants with HYPERQUAD suite of programs. *Talanta* ***43***, 1739–1753 (1996).

Chapter 3

Exploiting the SpyCatcher/SpyTag technology to design a new artificial metalloprotein

INTRODUCTION

"Protein design is no trivial task. It is an ultimate test of our understanding of structure–function relationships and requires a combined knowledge of physics, chemistry, and biology."

Yu et al., *Chem. Rev.* 2014

Metalloproteins promote several of the most complex biomolecular processes in Nature. Protein design has been often referred to as a powerful tool for the creation of tailored proteins that either do not exist or have different functions compared to those present in Nature.[1,2] As Yu and co-workers reported, designing new proteins or modifying already-existing ones it is not an easy task, and it comes with many challenges. Over the years, protein design has nonetheless become the target of a tremendous effort from the scientific community, as revealed by the exponential growth in the number of publications, that almost reached one million of systems in 2020.[3] In the specific field of metalloprotein design, this effort resulted into stories of extraordinary success.[4–9]

Protein design can be divided into two main categories: *protein redesign* and *de novo design*. 'Protein redesign' aims at conferring novel properties and functions to existing proteins of known structure. Conversely, '*de novo* design' or '*from scratch* design' involves the design of oligopeptide sequences that fold into a desired 3D structure and that are not strictly related to any other known natural construct. In the case of metalloproteins, for instance, protein redesign may consist into introducing mutations in the amino acid sequence with the aim of modifying either or both the first and second coordination sphere of a metal ion coordinated to a binding site (***Figure 3.1a***). On the other hand, *de novo* designed metalloproteins are based on peptide sequences that fold into minimalistic constructs that provide coordinating side chains that ultimately originate a metal binding site (***Figure 3.1b***).[1,2,10–14]

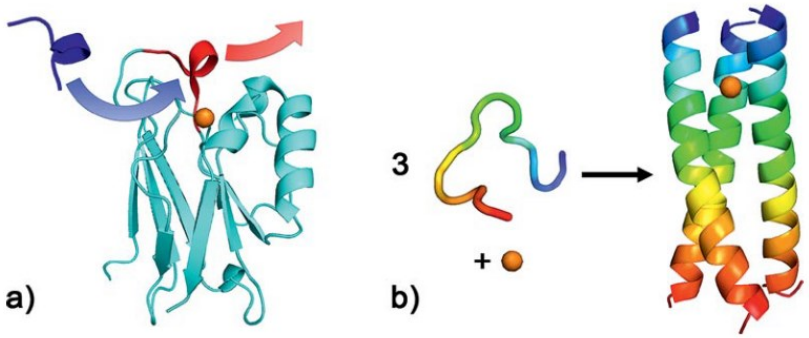


Figure 3.1 *Conceptual representation of protein redesign (a) and de novo design (b). Reprinted with permission from Ref. 1. Copyright © 2014 John Wiley and Sons*

Both protein redesign and *de novo* design can be achieved by exploiting two different strategies, namely 'combinatorial design' or 'rational design'. Both start from known protein scaffolds (especially within the protein redesign category), but they are very different in the approach. The combinatorial strategy consists in generating libraries of mutant sequences, which are then screened for desired structural features or functions.[15,16] For protein redesign this approach is powerful in the case of an incomplete knowledge (or no knowledge) of the structure of the starting protein. On the other hand, for protein *de novo* design this approach may allow to screen for the folding or assembly of a large number of peptide sequences, and possibly for their functional properties. Advancements in combinatorial techniques led to the development of *directed evolution* design strategies, for which the Nobel Prize was granted to Prof. Frances Arnold in 2018.[17] Without going into details, it is noteworthy that this approach mimics the natural selection process, and involves the evolution of a gene by iterative mutagenesis rounds, selection or screening for desired features, and amplification. These three steps are typically repeated using the best variant from a previous round as a template for the next, so that stepwise improvements are achieved.[18–21]

On the other side of protein design strategies stands the 'rational design'. This approach aims at tuning the properties of the final construct by the introduction of few mutations at specific positions of the protein scaffold. It is therefore evident that, in order to exploit this strategy, the knowledge of the structure and the understanding of reactivity features of the constructs to be evolved are required. Given the impact that slight modifications may have on protein structure, this design process is generally aided by computational methods for structure prediction and optimisation.[22–24]

Protein redesign: the Trojan horse approach

ARTZYMES (or ARTificial metalloenZYMES) have gained widespread interest in the recent years, as a consequence of an increasing demand for synthetic catalysts with enzyme-like performances. Due to their remarkable efficiency, enzymes are increasingly applied to a variety of industrial processes.[25,26] However, natural enzymes may not been evolved to perform many of the chemical conversions that are of interest in industry or bio(nano)technology. The need to develop artificial metalloenzymes capable to carry out non-natural catalysis therefore arises from the need of promoting (stereo)selective processes in aqueous medium that are not achievable using classic Werner-type or metallorganic metal complexes.[27]

Most of the redesigned Artzymes reported in the literature were obtained by introducing a catalytically active moiety (*i.e.* a metal cofactor) into biological chiral scaffolds such as proteins or DNA. Strategies went from the introduction of mutations of the peptide/nucleotide sequence to create a metal binding site, to the covalent grafting of a metal complex

to a biomolecule, to the use of non-covalent/supramolecular interactions.[25,28] Within the latter category, the 'Trojan horse' strategy in protein redesign is of particular interest for the introduction of a metal ion into a protein scaffold. Conceptually, the Trojan horse approach in metalloprotein redesign consists into taking a protein that does not bear any specific metal binding site, but that specifically interacts with a substrate (*e.g.* a small organic molecule) that can be functionalised with a metal-chelating ligand. If the functionalised cofactor is properly designed, it binds to the protein taking the metal cofactor in its (protein) proximity. As the Ulysses' wooden horse took the soldiers into the city of Troy, the substrate takes the metal close to the protein surface or inside a pocket or groove, providing it with a second-shell coordination environment. A conceptual graphical representation of this strategy is reported in ***Figure 3.2***.

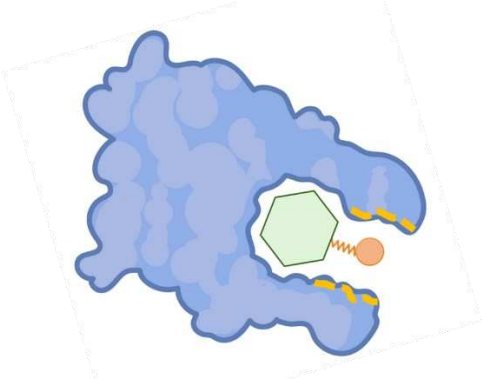


Figure 3.2 *Conceptual representation of a system designed with the Trojan horse approach. The protein (blue) binds the substrate/cofactor (green hexagon) and provides the metal site (orange) with a second-shell coordination environment (represented in yellow).*

Possibly, the most widely known example of this strategy for the design of artificial metalloenzymes is the *biotin-streptavidin* approach. This approach was first proposed by Wilson and Whitesides, and improved by Ward and colleagues.[29–31] By exploiting the remarkable affinity of (strept)avidin for biotin (K_a *ca.* 10^{15} M^{-1}), they managed to introduce transition metal complexes into the protein, obtaining artificial metalloenzymes for enantioselective catalysis. More in detail, a metal catalyst was covalently linked to a biotin anchor (*i.e.* the Trojan horse). The subsequent addition of a stoichiometric quantity of (strept)avidin resulted in the quantitative incorporation of the metal moiety into the protein scaffold, *de facto* obtaining an artificial metalloenzyme (***Figure 3.3***). By modifying the biotinylated metal catalyst, Ward and colleagues were also able to tune the catalytic properties of the biotin/(strept)avidin-based metalloenzymes to achieve desired reactivities.[31]

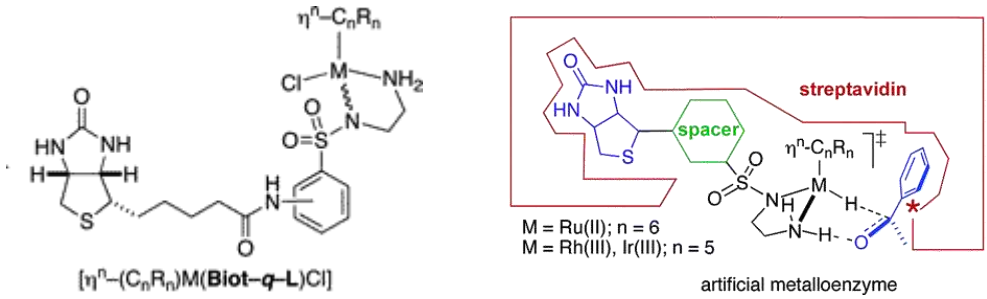


Figure 3.3 *A biotinylated Ru(II) complex (left) and a conceptual representation of a biotin/streptavidin-based metalloenzyme (right). Adapted with permission from Ref. 31. Copyright © 2010 American Chemical Society*

The SpyCatcher/SpyTag system

In this project, we had the aim to design and to develop a novel metalloenzyme by exploiting the Trojan horse approach, but in which the metal-binding cofactor that interacts with a protein is a peptide. The reason for this is that in perspective the introduction of an artificial metal-binding site on a peptide is in principle more straightforward than on a larger protein. Actually, oligopeptides can be prepared by chemical synthesis and easily purified even using non-natural amino acid. Also, peptides on which metal-binding moieties are introduced for instance through covalent grafting can be purified easily and easily characterized independently from the protein construct. If the oligopeptide acts as a cofactor of a protein, the final construct is entirely peptidic and through a Trojan horse approach an artificial metal site is introduced in a solely amino acidic protein.

Among the systems reported in the literature, the SpyCatcher/SpyTag peptide-tag[32] was chosen as a starting point for the design. The SpyCatcher/SpyTag construct was first developed by Howarth and co-workers exploiting the immunoglobulin-like collagen adhesion domain (CnaB2) that was found in the fibronectin binding protein of *Streptococcus pyogenes* (from which the 'Spy' acronym).[32] The CnaB2 domain is characterised by a β-barrel structure in which two β-strands (strand 1 in blue and strand 9 in green, ***Figure 3.4***) are linked through an isopeptide bond between Lys31 and Asp117, which confers an extraordinary stability to the protein.[33]

Howarth and colleagues found that by splitting the CnaB2 domain into a C-terminal peptide segment (13 amino acids) containing the Asp117 residue, and a N-terminal protein portion (138 amino acids) comprising Lys31, the formation of the specific isopeptide bond between the two fragments occurred spontaneously and irreversibly (***Figure 3.5***).

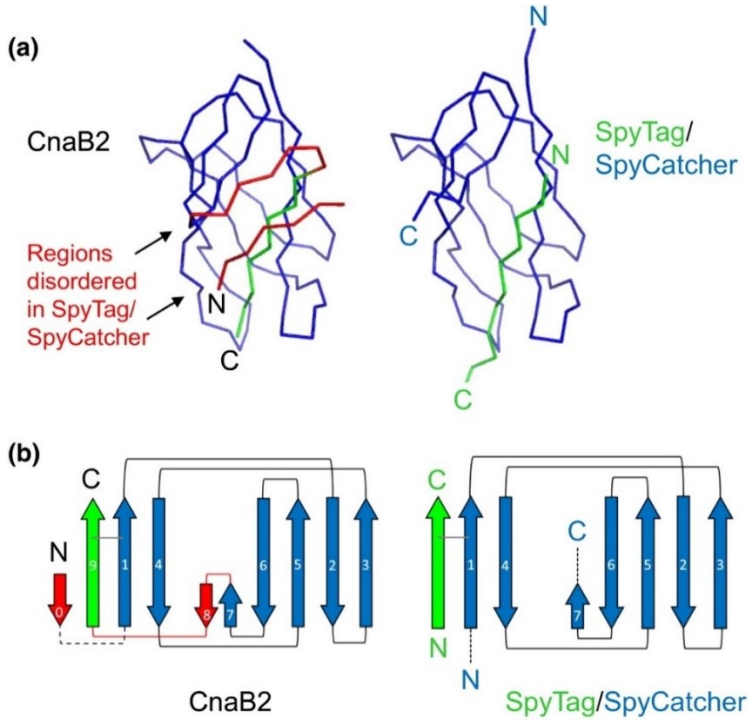


Figure 3.4 *Comparison between the 3D structures (a) and schematic representations (b) of CnaB2 domain and SpyTag/SpyCatcher. Blue: CnaB2 and SpyCatcher. Green: SpyTag. Red: disordered regions. Grey: isopeptide bond. Adapted from Ref. 34. Copyright © 2014, with permission from Elsevier*

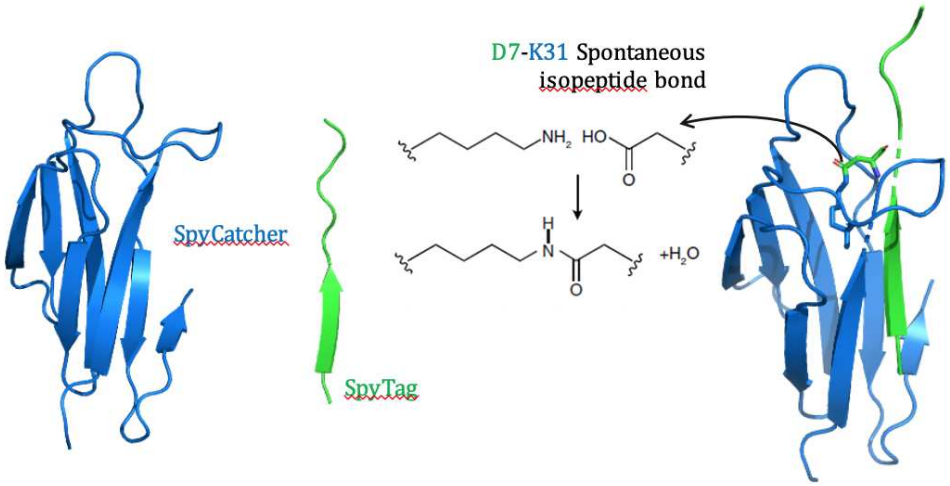


Figure 3.5 *Representation of the recombination between SpyCatcher and SpyTag and formation of isopeptide bond between Asp7 (SpyTag, green) and Lys31 (SpyCatcher, blue).*

A rational optimisation of the two sequences aimed at improving the recombination process led Howarth and coworkers to the first-generation SpyCatcher (SC) and SpyTag (ST) reported below (***Scheme 3.1***).[32] The interaction between the SpyCatcher construct and the SpyTag peptide is driven by the formation of a hydrogen bonds network and hydrophobic contacts between a conserved, consensus sequence of SpyTag (IVMVD) and a hydrophobic pocket of SpyCatcher that is formed by the Phe29, Ile27, Met44, Phe75, Ile90 and Phe92 residues. The SpyCatcher/SpyTag construct was found to be stable in a variety of conditions, such as pH (5-8), temperature (4-37 °C) and buffers, with high yields (> 97%) and recombination rates not exceeding 10 minutes.[32,34]

SpyCatcher001 (116 aa)
GAMVDTLSGLSSEQGQSGDMTIEEDSATHI**K**FSKRDEDGKELAGATMEL-
RDSSGKT-ISTWISDGQVKDFYLYPGKYTFVETAAPDGYEVA-
TAITFTVNEQGQVTVNGKATKG-DAHI

SpyTag001 (13 aa)
AHIVMV**D**AYKPTK

Scheme 3.1 *Amino acid sequences of the first-generation SpyCatcher protein and SpyTag peptide (PDB ID: 4MLI). The underlined residues in SpyTag represent the conserved sequence for the recombination with SpyCatcher, which occurs by the formation of an isopeptide bond between the residues highlighted in bold.*

Within few years from its first report, the use of the SpyCatcher/SpyTag system has greatly expanded and reported applications span from the production of biocompatible materials to vaccines development.[35] Among these, perhaps one of the most important is its use as a tagging system to obtain fusion proteins. If the SpyCatcher and SpyTag sequences are inserted at the N- or C-termini of two different proteins (tags), when incubated together these become fused into an artificial fusion protein made of two domains corresponding to the two individual proteins. This concept is represented in ***Figure 3.6***.

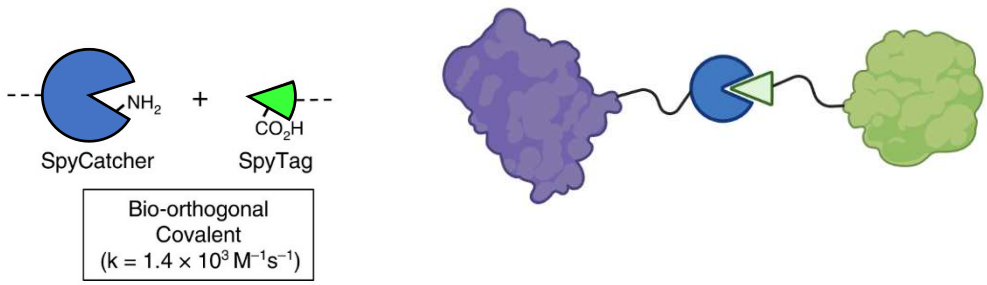


Figure 3.6 *Conceptual representation of the reaction between the SpyCatcher and SpyTag sequences to obtain a fusion protein. Adapted from Ref. 35, with permission from Elsevier*

This technology provides a unique covalent peptide-tagging system where a peptide tag (SpyTag) covalently binds its protein partner (SpyCatcher). However, although this tagging system has been extensively applied in a wide range of fields,[35] no attempt in the design of a SpyCatcher/SpyTag-based metalloenzyme is yet reported in the literature. The idea at the basis of the work presented herein is to explore the possibility of creating a metal moiety on the SpyTag peptide (by introducing a coordination site in its sequence) that can act as *Trojan horse* to introduce the metal into the SpyCatcher protein, exploiting the efficient and irreversible SpyCatcher/SpyTag recombination (see below). The obtained construct would actually represent a proof-of-concept for new metalloproteins in which the SpyTag peptide would provide the first metal coordination sphere, and the SpyCatcher protein may provide the second coordination sphere.

For the design of our system we used the second generation of the SpyCatcher/SpyTag construct.[36] Actually, after their first paper on the SpyCatcher/SpyTag construct, Howarth's group managed to improve the formation of the SpyCatcher/SpyTag adduct by selecting sequences with the aim of accelerating the recombination of the two units.[36] Phage display selections and site-directed mutagenesis strategies were exploited, and new sequences for both SpyCatcher and SpyTag were obtained. These sequences, those of the second-generation SpyCatcher/SpyTag, are reported in ***Scheme 3.2*** (the 002 label stands for second generation). In SpyTag002, the N-terminal VPT and the C-terminal RY residues were found to increase the peptide reaction rate. Similarly, a faster recombination was achieved by mutating the following residues of first-generation SpyCatcher: S12G, Q16P, I22T, K40R, Q65H, K108E and I116T. Interestingly, an unforeseen effect of these mutations was the formation of a covalent SpyCatcher/SpyCatcher adduct. In fact, the resemblance of the SpyCatcher N-terminal GAMVDT with SpyTag VMVDA (*i.e.* the conserved sequence for recombination) led to an unintended self-reactivity of SpyCatcher. To prevent this process, the N-terminal sequence was mutated to GAMVTT in SpyCatcher002 (***Scheme 3.2***).[36]

SpyCatcher002 (116 aa)
GAMVTTLSGLSGEQGPSGDMTTEEDSATHI**K**FSKRDEDGRELAGATMELRDSSG-
KTISTWISDGHVKDFYLYPGKYTFVETAAPDGYEVATAITFTVNEQGQVTVNGE-
ATKGDAHT

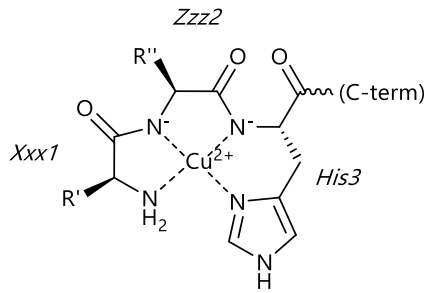
SpyTag002 (14 aa)
VPTIVMV**D**AYKRYK

Scheme 3.2 *Amino acid sequences of the second generation (002) of SpyCatcher and SpyTag. The underlined residues in SpyTag represent the conserved sequence for the recombination with SpyCatcher, which occurs by the formation of an isopeptide bond between the residues highlighted in bold. Mutated residues with respect to first-generation SpyCatcher and SpyTag are shown in red.*

The ATCUN metal binding site

Following the aforementioned approach, we would consider the SpyCatcher protein as the city of Troy that the SpyTag wooden horse would enter with an army (metal) inside. Therefore, the first step to achieve the invasion of Troy would be loading the horse with the army, *i.e.* binding a metal to the SpyTag peptide. With the aim of proving that this approach works, to choose wisely both the metal and the binding site is essential. For instance, a spectroscopically active metal would be simpler to characterise than a spectroscopically silent one. Also, it would be preferable to select a binding site with a high affinity for the chosen metal in order to avoid competition reactions with other possibly undisclosed binding sites on the construct. Finally, potentially catalytic properties of the metal/peptide adduct could be useful to confirm the occurring of the binding. As previously stated, in the perspective of a chemist the power of this strategy sounds quite evident. Introducing a metal within a folded protein without disrupting its structure is, especially for chemists, an extremely hard task. In contrast, the functionalisation of a short peptide (such as SpyTag) with a metal binding site is a much more achievable goal.

In this project, we decided to converge our attention on the ATCUN metal binding site and copper(II) as the metal of interest. The Amino Terminal Cu(II) and Ni(II) binding domain (ATCUN) is known for binding Cu^{2+} and Ni^{2+} with remarkable affinity (K_f = 10^{12} - 10^{15} M^{-1}).[37,38] First discovered in albumin, the ATCUN motif is characterised by a 'H_2N-Xxx-Zzz-His' (XZH) sequence, which represents the N-terminus of a peptide (or protein) having a His residue at the third position with respect to the free amino-terminal group.[39] When bound to an ATCUN site, the metal is found in a 4N coordination in which the free N-terminal amine, two deprotonated peptide nitrogen atoms and an imidazole nitrogen atom are present on the metal equatorial plane (***Scheme 3.3***). Several studies proved that the nature of the two amino acids that precede His is not relevant in defining the first coordination sphere of copper(II), and that provide only a small effect on the affinity for the metal ion.[40] Instead, the presence of a His residue at the third position is essential, as it allows the formation of two 5-membered and one 6-membered chelation rings, which determine a particularly stable copper(II) coordination that is represented in ***Scheme 3.3***.



Scheme 3.3 *Schematic representation of the coordination of copper(II) in an ATCUN (NH_2-Xxx-Zzz-His) motif.*

An important feature of the ATCUN motif is that, being located at the N-terminus, it can be easily added to most of the peptides or proteins without predicting dramatic changes in their 3D structure. Moreover, the presence of copper(II) bound to this site is easily detectable by UV-visible spectroscopies. A visible absorption at *ca.* 525 nm is indeed characteristic of a Cu^{2+}/ATCUN complex, along with a distinctive profile of circular dichroism spectra (*i.e.* a positive band at 490 nm and a negative one at *ca.* 560 nm), as depicted in ***Figure 3.7***. Together, these two spectroscopic features almost represent a fingerprint of copper(II) coordinated to an ATCUN site.

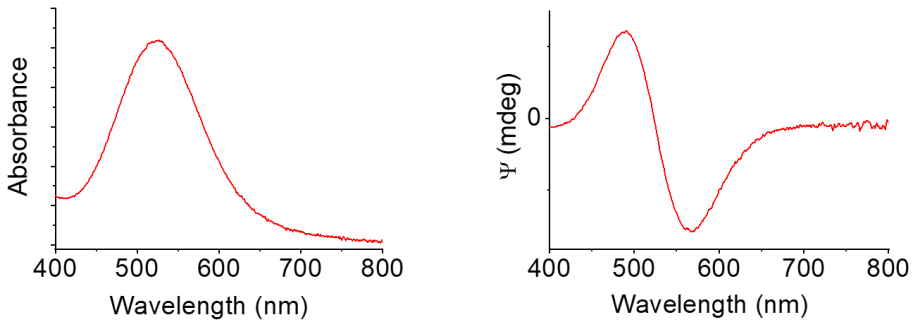


Figure 3.7 *Spectroscopic features of copper(II) bound to an ATCUN site. Left: visible absorption spectrum. Right: circular dichroism spectrum. Data available by the author, see Ref. 41*

Studies on the catalytic properties of copper(II)-ATCUN systems have highlighted, among others, an hydrolytic activity towards several biomolecules such as DNA, RNA, proteins and carbohydrates.[42] Applications of metal-ATCUN sites in chemistry, biology and medicine have been recently reviewed.[43] Overall, the great and versatile features of copper(II)/ATCUN systems (namely the strong and stable coordination of the metal, the easily recognisable spectroscopic fingerprint, and the catalytic properties) make this motif perfect to be used as a proof-of-concept for the design of a new SpyCatcher/SpyTag-based copper enzyme.

In the next sections, the design and structural evaluation of an artificial Cu(II)-ATCUN protein based on the SpyCatcher/SpyTag system will be discussed in detail. Then, spectroscopic characterisation of copper(II) binding to a SpyCatcher/SpyTag-ATCUN system will be extensively presented. Finally, the esterase properties of the Cu^{2+}/SpyCatcher/SpyTag-ATCUN adduct will be described. We anticipate here that the ATCUN site will be located on the SpyTag, obviously at the N-terminus. In ***Figure 3.8*** we therefore represent the SpyCatcher/SpyTag construct indicating the site on the SpyTag (in green) where the ATCUN site will be introduced.

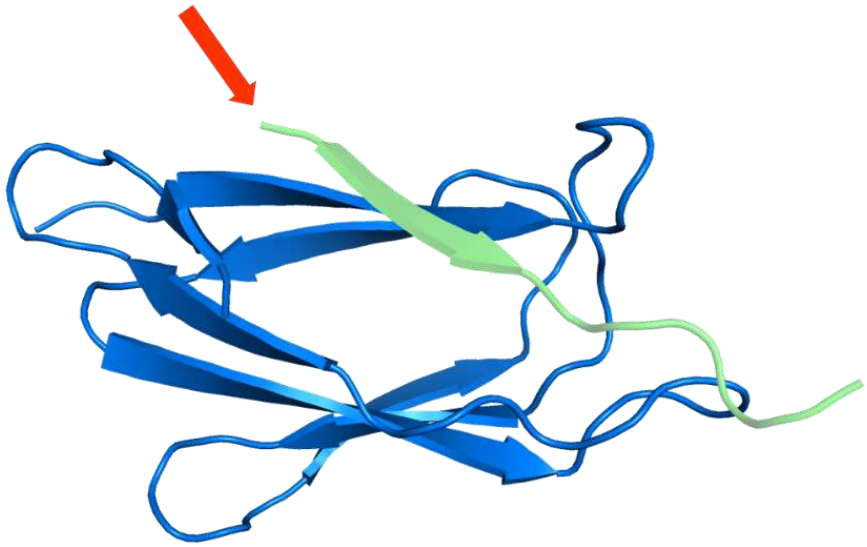


Figure 3.8 *Representation of the SpyCatcher/SpyTag construct (PDB 4MLI). The red arrow indicates the position where the ATCUN site was introduced, in our system.*

DESIGN OF THE NEW SPYCATCHER/SPYTAG-ATCUN SYSTEM

As previously mentioned, the second-generation SpyCatcher/SpyTag (SC/ST) system (***Scheme 3.4***) was chosen as a starting point for the design of the artificial metalloprotein, due to its enhanced SC/ST recombination rate with respect to the first-generation one.[36]

a. ***SpyCatcher002*** (SC002)
GAMVTTLSGLSGEQGPSGDMTTEE**DSATHIKFSKRDEDGRELAGATMELRDSS-GKTISTWISDGHVKDFYLYPGKYTFVETAAPDGYEVATAITFTVNEQGQVTVN**

b. ***SpyTag002*** (ST002)
VPTIVMVDAYKRYK

Scheme 3.4 *Amino acid sequences of SpyCatcher002 (a) and SpyTag002 (b). Bold: structured portion of the protein. Red: His residues. Green: acceleration sequence for SC/ST recombination. Underlined: conserved sequence for SC/ST interaction.*

Design of SpyCatcher

The first step to ensure an efficient copper coordination to the ATCUN site of SpyTag is to prevent the undesired binding of the metal to SpyCatcher to sites other than the unavoidable N-terminus. The second-generation SpyCatcher contains two His residues (His29 and His65), which are sites that may bind copper(II) with moderate to high affinity (***Scheme 3.4a***, red). As a consequence, SpyCatcher may compete with SpyTag for metal binding. A redesign of the SC protein was therefore necessary to remove these potential metal binding site and to guarantee the occurring of copper coordination at the sole SpyTag ATCUN site. However, changes in the sequence must not result in a disruption of the protein 3D structure or lack of folding, so mutations had to be chosen with caution.

In the tertiary structure of SC002, His29 and His65 are involved in hydrogen bonds in which the N7 of the imidazole ring acts as H-donor. While the residue in position 65 was actually a glutamate in generation one SpyCatcher, His29 is a conserved residue. Also, His29 forms an inter-strand interaction with Asp65. We therefore decided to mutate His29 into Gln29 in the attempt to preserve this interaction in the redesigned SC protein. The sequence is reported in ***Scheme 3.5***. To purify the protein by means of a HisTrap affinity column, a His-Tag was added to the N-terminus, together with a thrombin cleavage site (LVPRGS) to remove the HisTag after the purification step. The full sequence of the expressed protein is reported in ***Scheme 3.5a*** (HisTag-TEV-Thrombin-SpyCatcher002-mut (H29Q H65Q)). Thrombin performs proteolysis between Arg and Gly residues, and therefore GSGAM corresponds to the sequence at the N-terminus of the redesigned SC protein (***Scheme 3.5b***). It is worth to note that SC002 bears a Trp residue (Trp60) that we used for a spectrofluorimetric study of copper(II) binding, exploiting Trp fluorescence quenching caused by the metal as shown below. The sequence of the final construct after purification is reported in ***Scheme 3.5b*** and it will be hereafter referred to as 'SC-mut'. The study of the model of SC-mut and of his adduct with SpyTag is reported below.

a. HisTag-TEV-Thrombin-SpyCatcher002-mut (H29Q H65Q)

MGSSHHHHHHSENLYFQSGLVPRGSGAMVTTLSGLSGEQGPSGDMTTEE**DSAT-QIKFSKRDEDGRELAGATMELRDSSGKTISTWISDGQVKDFYLYPGKYTFVET-AAPDGYEVATAITFTVNEQGQVTVN**

b. SpyCatcher002-mut (H29Q H65Q) – 'SC-mut'

GSGAMVTTLSGLSGEQGPSGDMTTEE**DSATQIKFSKRDEDGRELAGATMELRD-SSGKTISTWISDGQVKDFYLYPGKYTFVETAAPDGYEVATAITFTVNEQGQVT-VN**

Scheme 3.5 *Amino acid sequence of the SC-mut construct prior (a) and after (b) purification. Green; His-Tag. Red: mutated residues. Underlined: thrombin cleavage site. Bold: structured portion of the protein.*

Design of SpyTag

The design of the SpyTag-ATCUN peptide started from the SpyTag002 sequence (***Scheme 3.4b***). The ATCUN site was introduced by adding a Gly-Ser-His triad at the N-terminus, and the sequence of a *E. coli* thioredoxin domain at the C-terminus to increase the dimension of the peptide and allow its characterization by SDS-PAGE. The full sequence of the expressed protein is reported in ***Scheme 3.6a*** (HisTag-TEV-Thrombin-SpyTag002 EcTrx-mut (H6V W28N W31Q)).

a. HisTag-TEV-Thrombin-SpyTag002 EcTrx-mut (H6V W28N W31Q)
MGSSHHHHHHSSGLVPRGSHVPTIVMVDAYKRYK*SDKIIVLTDDSFDTDVLKA-DGAILVDFNAEQCGPCKMIAPILDEIADEYQGKLTVAKLNIDQNPGTAPKYGI-RGIPTLLLFKNGEVAATKVGALSKGQLKEFLDANLA*

b. ATCUN-SpyTag002-EcTrx – 'ST-mut1'
GSHVPTIVMVDAYKRYK*SDKIIVLTDDSFDTDVLKADGAILVDFNAEQCGPCK-MIAPILDEIADEYQGKLTVAKLNIDQNPGTAPKYGIRGIPTLLLFKNGEVAAT-KVGALSKGQLKEFLDANLA*

Scheme 3.6 *SpyTag-ATCUN (ST-mut1) sequence prior (a) and after (b) removal of HisTag and purification. Bold: ATCUN site. Green: acceleration sequence for SC/ST recombination. Underlined: conserved sequence for SC/ST interaction (single line); thrombin cleavage site (double line). Italic: EcTrx domain (mutations in red).*

A thrombin cleavage site (LVPRGS) was introduced at the N-terminus of SpyTag, followed by a His residue (***Scheme 3.6a***). The role of this thrombin cleavage site is two-fold. On one hand, it allowed to take advantage of a His-Tag for protein purification. On the other hand, the proteolytic process carried out by thrombin at the Arg-Gly site resulted in the release of a Gly-Ser-His ATCUN site on ST-mut1, as evidenced in ***Scheme 3.6b***.

Similarly to SpyCatcher, also the *E. coli* thioredoxin (EcTrx) domain had to be mutated to prevent undesired copper binding at His sites. The sequence of this domain is that reported in the PDB as 2TRX (numbering of the sequence for Trx refer to the PDB deposited sequence). Fortunately, the EcTrx domain contains only one His (His6) which was mutated into valine due to its location on a β-strand and the absence of interactions involving this residue. Two Trp residues in the native Trx sequence were also mutated since they could bias the fluorescence studies involving the Trp residue on the SpyCatcher construct. Hence, Trp28 and Trp31 were mutated to Asn and Gln, respectively. While glutamine was introduced to preserve the H-bond interaction with Asp61 (W28-D61), asparagine was chosen to mutate Trp28 that is located in a β-strand segment of the protein.

An attempt to prepare a different ST construct bearing an ATCUN site was made by adding a Met-Ala-His triad at the N-terminus of ST002 (and the mutated EcTrx domain

at the C-terminus). The final sequence of this SpyTag-ATCUN, which was named 'ST-mut2', is reported in ***Scheme 3.7***.

ATCUN(MAH)-SpyTag002-EcTrx-mut (H6V W28N W31Q) – 'ST-mut2'
MAHVPTIVMVDAYKRYK*SDKIIVLTDDSFDTDVLKADGAILVDFNAEQCGPCK-MIAPILDEIADEYQGKLTVAKLNIDQNPGTAPKYGIRGIPTLLLFKNGEVAAT-KVGALSKGQLKEFLDANLA*

Scheme 3.7 *SpyTag-ATCUN(MAH) sequence (ST-mut2). Bold: ATCUN site. Green: acceleration sequence for SC/ST recombination. Underlined: conserved sequence for SC/ST interaction. Italic: EcTrx domain (mutations in red).*

It can be here anticipated that, unfortunately, the Met-Ala-His ATCUN site resulted unstable towards N-terminal Met Excision (NME), a process that cleaves the initial Met residue from mature proteins in the 55-70% of cases (see below).[44] For this reason, only the copper(II) binding to ST-mut1 was characterised in detail. A representation of the newly designed SC-mut/ST-mut1 system is reported in ***Figure 3.9***.

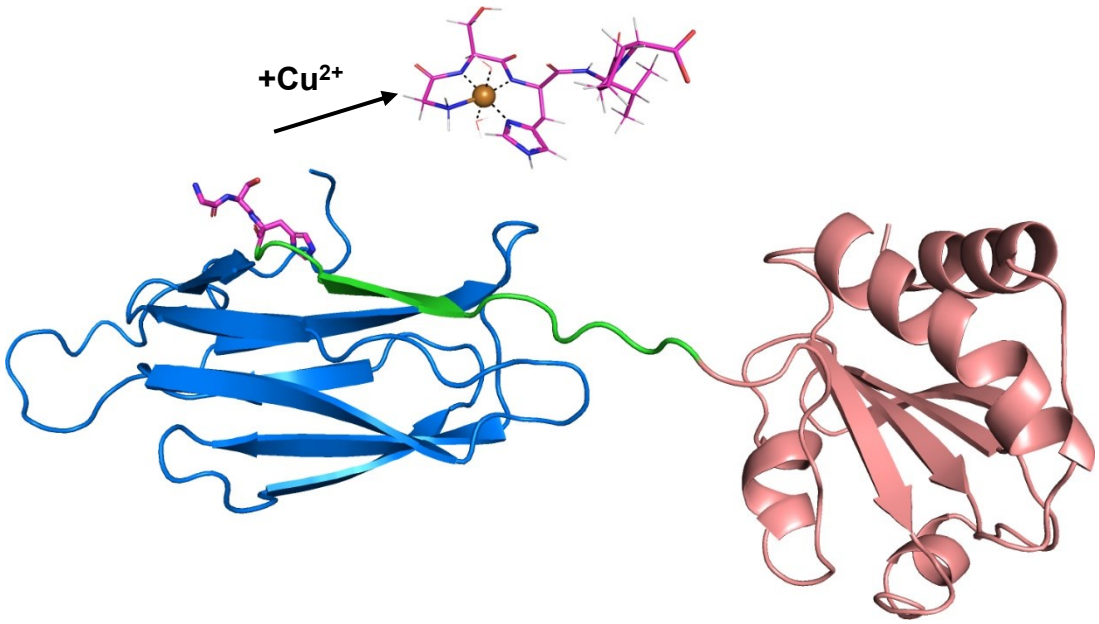


Figure 3.9 *Representation of the SpyCatcher/SpyTag-ATCUN (SC-mut/ST-mut1) system. Blue: SpyCatcher. Green: SpyTag. Magenta: Gly-Ser-His ATCUN site. Light pink: E. coli thioredoxin domain.*

In-silico study of the model of the designed SpyCatcher/SpyTag-ATCUN system (SC-mut/ST-mut1)

Prior to protein expression, the design of the new SpyCatcher/SpyTag-ATCUN system was studied by a computational analysis of its structural features, to investigate the effect of the newly introduced mutations.

The 3D structure of SpyCatcher SC-mut was first predicted using the SwissModel[45–49] software, then it was graphically compared to the PDB structure 4MLI (first generation SpyCatcher). No significant changes in the structured portion of the protein were observed upon sequence mutation with respect to the structure deposited in the PDB. The 26-amino acid sequence at the N-terminus (***Scheme 3.5b***) was modelled by SwissProt but it corresponds to a region not characterized in the X-ray structure of first generation SpyCatcher (PDB 4MLI). This region is likely unstructured in solution, as also attempts to predict a structure of the construct using Phyre2 (Protein Homology/analogy Recognition Engine V 2.0)[50] failed for that segment of the protein (while it supported the β-barrel fold of the remaining sequence). A structural optimisation of SC-mut1 was also performed using Foldit[51], starting from the model obtained using Swiss-Model. The resulting model of the structured portion of SC-mut (represented in ***Figure 3.10***) showed almost no differences with respect to the PDB structure of first generation SpyCatcher (PDB 4MLI).

Having obtained indications on the SpyCatcher robustness upon mutations (SC-mut), the model of SpyCatcher/SpyTag-ATCUN adduct (SC-mut/ST-mut1) was examined. Both the mutant SC and ST sequences (SC-mut and ST-mut1) were loaded in FoldIt, and the model of their adduct optimised (***Figure 3.9*** and ***Figure 3.11***). The representation of the SC-mut/ST-mut1 model overlaid that of the crystal structure PDB 4MLI (***Figure 3.12***).

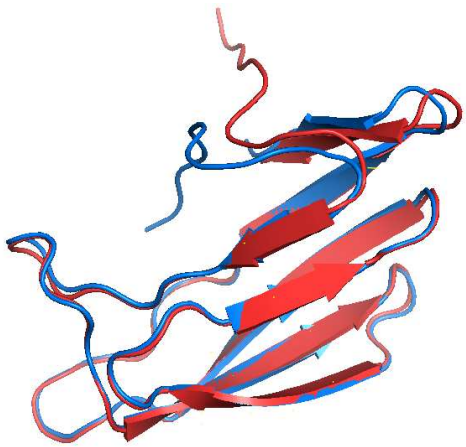


Figure 3.10 *Overlay of the SpyCatcher structure minimised by Foldit (red) and the SpyCatcher (blue; first generation, PDB 4MLI).*

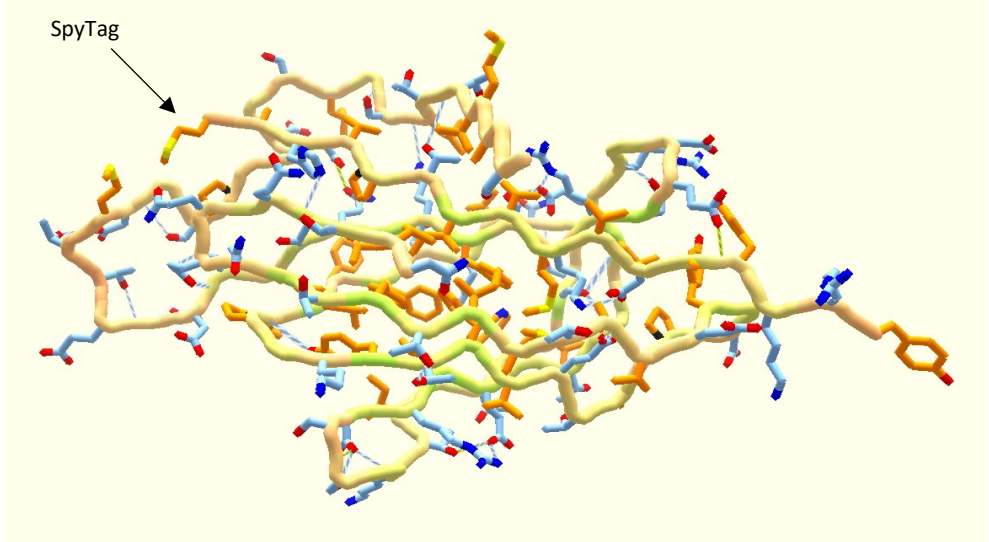


Figure 3.11 *Screenshot of the FoldIt modelisation of the SC-mut/ST-mut1 adduct.*

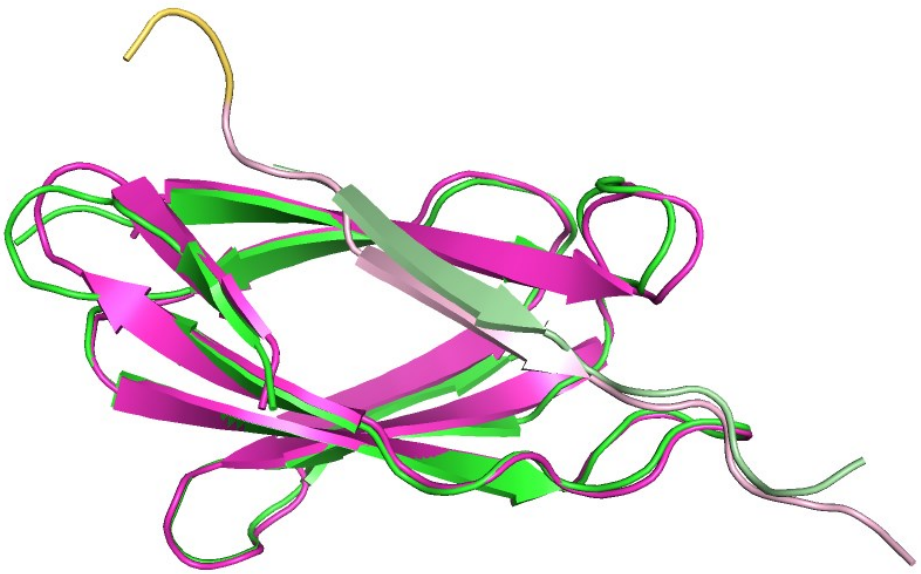


Figure 3.12 *Overlay of the FoldIt optimised SC-mut/ST-mut1 adduct (magenta) with the X-ray structure PDB 4MLI (green). Only structured regions of SC-mut/ST-mut1 are represented (unstructured regions: EcTrx domain and the 26–amino acid sequence at the N-Terminus). Pale magenta: ST in 4MLI; pale green: ST-mut1; yellow: ATCUN fragment.*

Comparison of the models clearly shows that the presence of the ATCUN peptide does not affect the overall SC-mut/ST-mut1 3D structure. However, several aspects need to be considered. First, the C-terminus sequence of ST-mut1 (YKRYK) that precedes the EcTrx domain extrudes out of the SpyCatcher construct enough to be considered a flexible spacer (***Figure 3.12***, tail on the bottom right). Under these circumstances it is reasonable, as suggested by the model from FoldIt, that the fold of the EcTrx domain is unaffected by its fusion to SpyTag and, in turn, to SpyCatcher. Second and possibly most important, also the ATCUN site extrudes from the SpyCatcher construct to a sufficient extent to provide it with a significant degree of flexibility and adaptability, which is considered important for copper(II) binding (magenta in ***Figure 3.9***, yellow in ***Figure 3.12***). Moreover, the presence of a carboxylate-rich region within the SpyCatcher protein needs to be highlighted (***Figure 3.13***). The evaluation of the electrostatic potential of the protein construct using PyMol (***Figure 3.13***) revealed that this latter region is associated to a negatively charged region, as possibly expected, which may act as a copper(II) coordination site, as we will show below. Finally, it is interesting to note that in SC-mut Gln31 corresponds to His26 in the parent SC/ST adduct (PDB 4MLI), the latter one of the two His that were mutated to avoid anchoring residues for copper(II). Differently from His26 in 4MLI which forms a H-bond with Asp65, Gln31 forms a H-bond with Tyr72, as represented in ***Figure 3.14***.

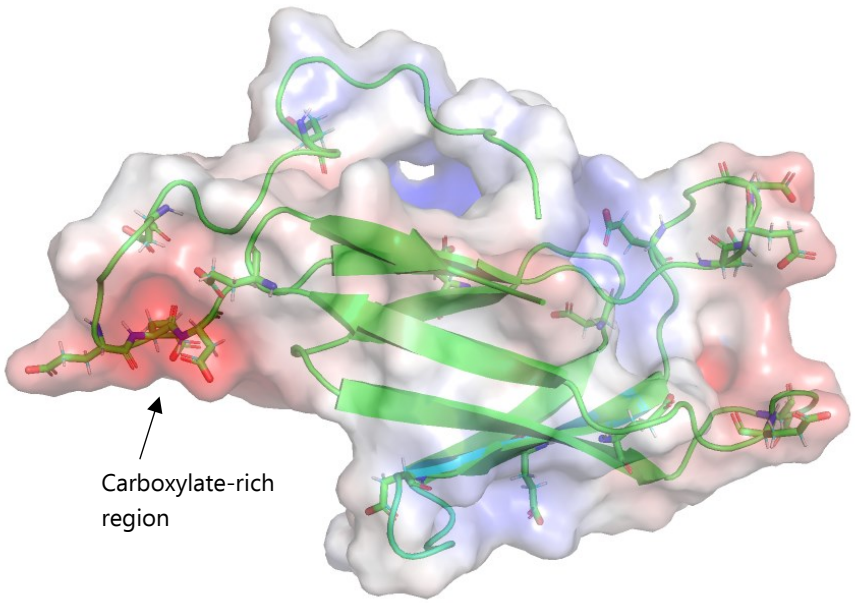


Figure 3.13 *Graphic representation of the SpyCatcher protein (model obtained from FoldIt) overlapped to its electrostatic potential surface (PyMol). Red: negatively charged regions; blue: positively charged regions. Asp and Glu residues are represented as sticks.*

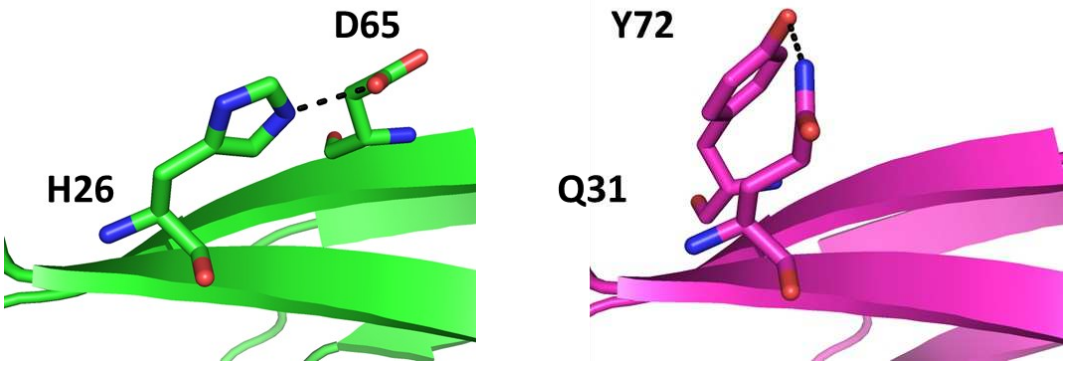


Figure 3.14 *Representation of the interaction between His26 and Asp65 in the SC/ST adduct studied by X-ray diffraction (PDB 4MLI, green), and the FoldIt model of SC-mut/ST-mut1 (magenta). Q31 corresponds to H26Q using the numbering in PDB 4MLI.*

SpyTag-ATCUN sequencing and recombination studies with SpyCatcher

The sequence of both the ST-mut1 and ST-mut2 SpyTag were studied by digesting the proteins with trypsin prior to infusion in ESI-MS of the digested samples at 2 μM concentration. The experiments were carried out in collaboration with the group of Prof. Maria Careri and Prof. Monica Mattarozzi (University of Parma). The obtained mass spectra are reported in ***Figure 3.15*** and ***Figure 3.16***. The deconvoluted mass spectrum of ST-mut1 (***Figure 3.15***, bottom) shows one signal at 13455.0 Da, which is fully consistent with the full-length sequence reported in ***Scheme 3.6b***. On the contrary, the deconvoluted mass spectrum of ST-mut2 (***Figure 3.16***, bottom) displays four main peaks, indicating that the sample is a mixture of four proteins. The analysis of the signals allowed to assign the signal at m/z 13380 Da to the sequence of ST-mut2 (***Scheme 3.7***) devoid of the Met residue at the N-terminus. The signal at 13394 Da is the most abundant and it is likely related to the previous one: since it corresponds to +14 Da, it is possibly the result of a post-expression methylation on one residue.

As anticipated, the post-expression removal of the initial Met in a protein sequence is not an uncommon process, as it involves *ca.* 55% of the known proteins.[44] It is known that the translation of mRNA for protein production involves a start codon that always codes for methionine in eukaryotes (generally the AUG codon). As a consequence, aside from few rare exceptions, every new protein is expressed starting from a Met residue. However,

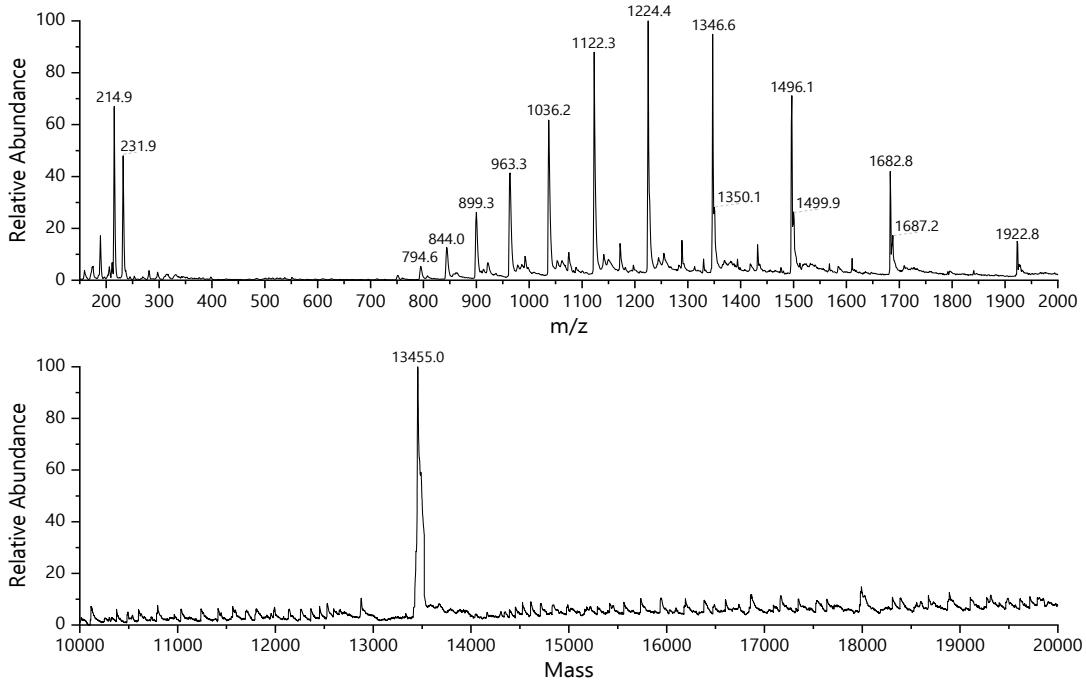


Figure 3.15 *Raw (top) and deconvoluted (bottom) mass spectra of ST-mut1 digested with trypsin.*

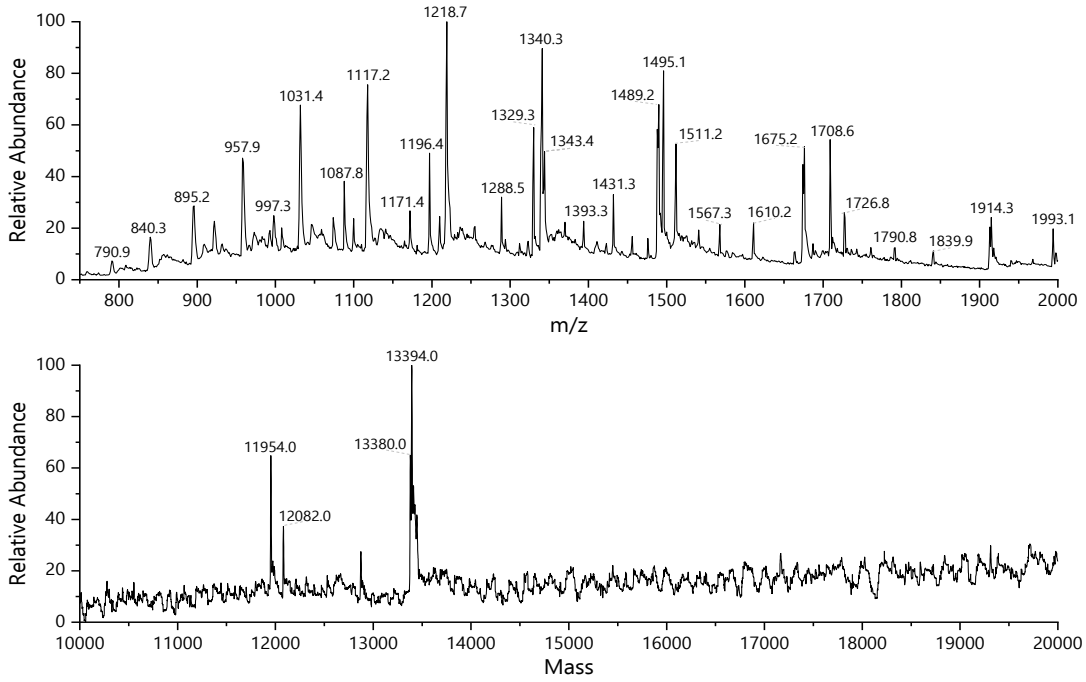


Figure 3.16 *Raw (top) and deconvoluted (bottom) mass spectra of ST-mut2 digested with trypsin.*

this residue is frequently (55-70%) removed from mature proteins in favour of a less hindering one (*e.g.* Ala, Cys, Gly, Pro, Ser, Thr or Val) through a process called N-terminal Met Excision, which likely occurred also in the case of the ST-mut2 protein.[44]

The signals at 11954 and 12082 Da correspond to the expected ST-mut2 sequence lacking the initial MAHVPTIVMVDAY and MAHVPTIVMVDAYK fragments, respectively. These fragments basically correspond to the entire SpyTag peptide sequence that precedes that of EcTrx, and that were therefore cleaved from EcTrx. It can be thus hypothesised that the removal of those segments is possibly the result of the action of some proteases that remove that portion, leaving the thioredoxin domain unaffected. Also, these data do not show the presence of +14 gemini signals on these two latter peaks, therefore suggesting that the methylation process proposed above occurs on one of the residues of the SpyTag sequence.

Overall, these data reveal that the expression and purification of ST-mut1 was successful, while ST-mut2 was obtained as a mixture of at least four constructs. Among the latter, none is expected to contain an ATCUN site, an observation that has been confirmed by spectroscopic analyses (see below). ESI-MS seems to suggest that in the ST-mut2 sample, despite the cleavage of the SpyTag sequence needed to bind to the SpyCatcher construct, *ca.* two thirds of the protein still contain the conserved IVMVD fragment. We decided therefore to use the ST-mut2 samples for the recombination studies with SpyCatcher (see below). It is however clear that of the different design, expression, and purification strategies used to obtain ST-mut1 and ST-mut2, only those of the former are effective to generate an ATCUN SpyTag construct.

Studies of recombination of SC-mut with ST-mut1 were carried out by incubation of the two components in 25 mM TRIS buffer (added with 1 M of KP) at pH 6 for 5 min (***Figure 3.17***). Increasing amounts of SpyTag *vs.* SpyCatcher up to 5 equivalents resulted in the selective appearance of a band corresponding to the 1:1 adduct. Since SC and ST are known to recombine in 1:1 ratio, the complete disappearance of the band of SC-mut observed for the addition of 2 equivalents of ST-mut1 was explained with an underestimation of the ST-mut1 nominal concentration.

Further studies aimed at clarifying the possible role of copper(II) in preventing the recombination of SC-mut with ST-mut1 were carried out. SC-mut and an equimolar amount of copper(II) sulphate were incubated together prior to the addition of ST-mut1. Recombination reactions were carried out for 5 min, then the gel was loaded. SDS-PAGE analysis (***Figure 3.18***) showed that the two proteins have the same recombination features irrespective of the presence or absence of copper(II). It can be therefore concluded that copper(II) does not interfere in the recombination of the two constructs.

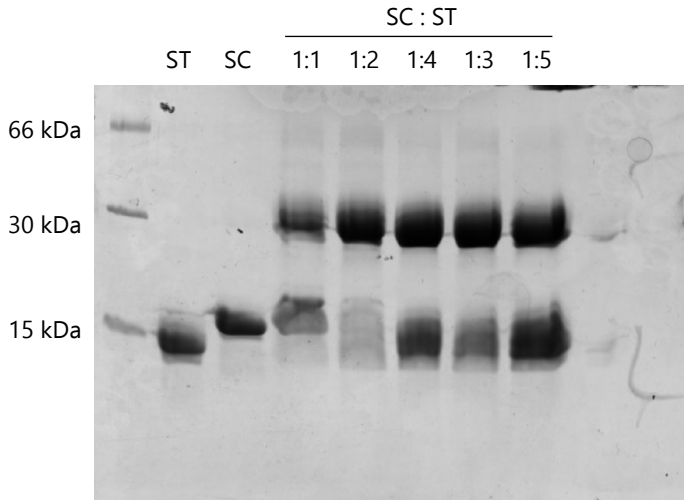


Figure 3.17 *SDS-PAGE analysis of samples of SC-mut (SC) and ST-mut1 (ST). Incubation was carried out using 12 μg of SC-mut.*

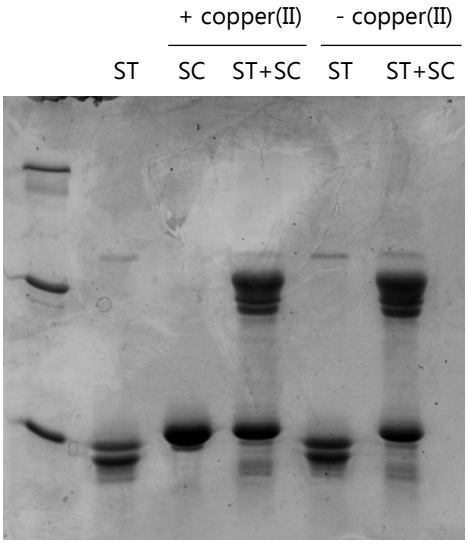


Figure 3.18 *SDS-PAGE analysis of samples of SC-mut (SC) and ST-mut1 (ST) in the presence (+) or absence (-) of equimolar copper(II) sulfate in the SC-mut sample prior to the addition of ST-mut1.*

COPPER(II) BINDING TO THE SPYCATCHER/SPYTAG-ATCUN SYSTEM

A further step towards the characterisation of the SpyCatcher/SpyTag-ATCUN artificial metalloenzyme was the study of the binding of copper(II). Adducts of the proteins with Cu(II) were spectroscopically characterised by means of UV-visible absorption spectroscopy, circular dichroism, and fluorescence spectroscopy. Prior to the study of Cu^{2+}/SpyCatcher/SpyTag-ATCUN interaction, binding of copper(II) to SC-mut and ST-mut1 was examined.

Cu^{2+}/SpyCatcher (SC-mut)

The binding of copper(II) to the SpyCatcher protein (SC-mut, referred to as 'SC' from now on) was studied at physiological pH (7.4 if not otherwise stated) in buffer media. Visible absorption and circular dichroism techniques were first employed to provide information on the coordination of copper in the Cu^{2+}/SC adduct. These studies were complemented by experiments of quenching of fluorescence, which aimed at determining the K_f value of the copper(II)/SpyCatcher adduct.

Figure 3.19 reports the spectrophotometric and CD titrations of SpyCatcher solutions with Cu^{2+} in aqueous HEPES buffer. Upon the addition of 1.0 metal equivalents, absorption spectra exhibit a λ_{max} of *ca.* 580 nm, which shifted to 700 nm upon addition of further equivalents of the metal.

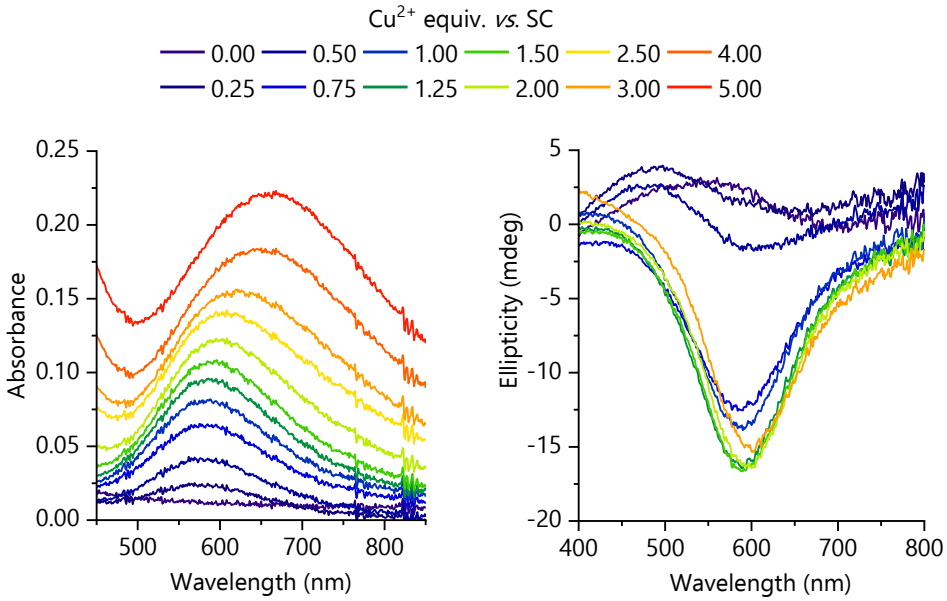
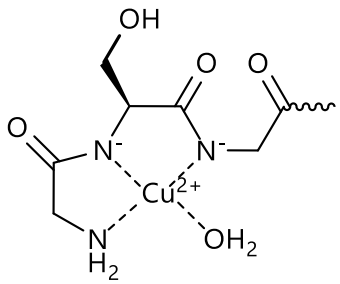


Figure 3.19 *Visible absorption (left) and CD (right) spectra for the titration of SC solutions with Cu^{2+} (C_{SC} = 948 μM, Cu:SC = 0-5:1, 20 mM aqueous HEPES buffer + 0.1 M NaCl solution, pH 7.4).*

Exploiting Billo's method of the average environment,[52–54] a λ_{max} of *ca.* 580 nm can be ascribed to the presence of one amino nitrogen, two deprotonated peptide nitrogen atoms, and a water molecule coordinated on the equatorial plane of copper(II) (***Scheme 3.8***; calc. λ_{max} = 576 nm).[52–54] This (N_{am}, 2x N^-, H_2O) donor set was associated to the coordination of copper(II) to the N-terminal Gly-Ser-Gly residues of the SpyCatcher sequence, which indeed is an unavoidable copper(II) binding site in peptides and proteins that bear a free NH_2 group at the N-terminus.



Scheme 3.8 *Schematic representation of Cu^{2+} coordination at the N-terminus (Gly-Ser-Gly) of the SpyCatcher protein.*

Further copper(II) additions to SC led to a red-shift and broadening of the visible absorption band (***Figure 3.19***). Difference spectra (***Figure 3.20***) allowed to identify the presence of a second absorption band at *ca.* 700 nm that appears in the excess of copper(II). This suggests the presence of an additional coordination site, and addition of copper(II) beyond equimolar ratio results in the formation of species with putative 2:1 metal/protein ratio.

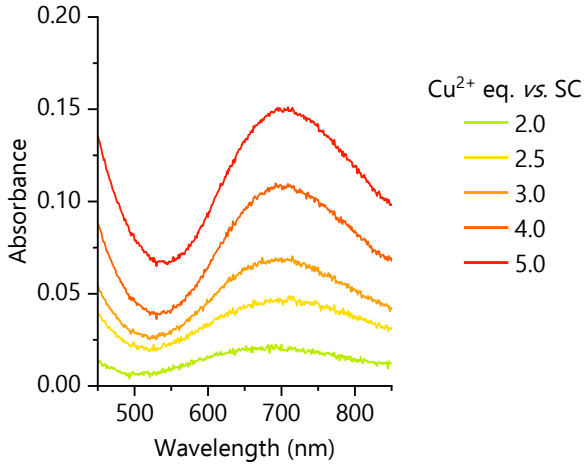


Figure 3.20 *Difference spectra of Cu^{2+} additions to SC (C_{SC} = 948 μM, Cu:SC = 2-5:1, 20 mM aqueous HEPES buffer + 0.1 M NaCl solution, pH 7.4) obtained subtracting the spectrum recorded for 1.5 Cu^{2+} eq. to those recorded for the equivalents added as reported in the legend.*

Further proof of the existence of two binding sites for copper(II) was obtained by performing a UV-Vis titration analogous to that reported in ***Figure 3.19***, but collecting more spectra (smaller titrant aliquots, ***Figure 3.21***). The absorbance values at 580 and 700 nm *vs.* the equivalents of copper(II) added were then plotted. This plot, represented in ***Figure 3.21***, shows a marked increase of the absorbance at 580 nm until 1 equivalent of Cu^{2+}. Beyond this copper(II):protein ratio, the absorbance at 580 nm slightly flattens. When 2 equivalents of Cu^{2+} are reached, the increase of absorbance is significant at 700 nm. The hypothesis of the presence of this unexpected copper(II) binding site, other than that at the N-terminus, is likewise supported by the significant change observed in the CD spectra, which are dominated by the appearance of a negative signal at 590 nm that becomes intense only after the addition of 1 equivalent of metal ion (***Figure 3.19***).

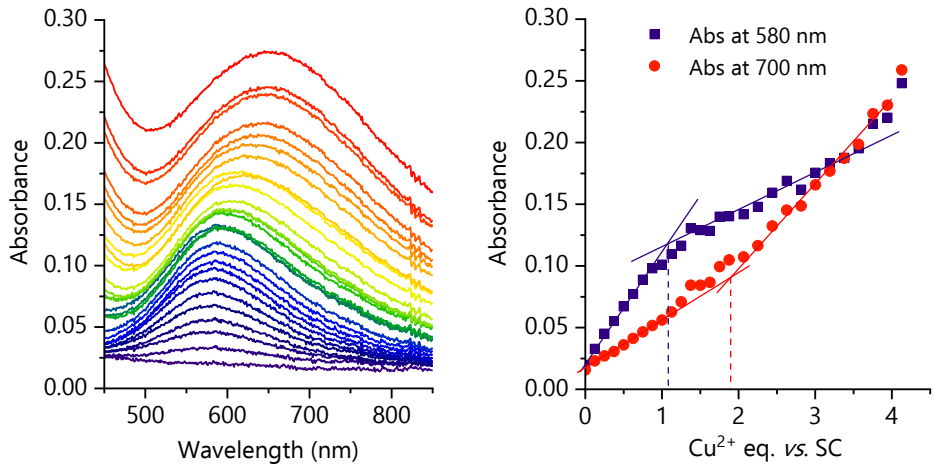


Figure 3.21 *Left: Visible absorption spectra for the titration of SC solutions with* Cu^{2+} *(*C_{SC} *= 948 μM, Cu:SC = 0-4.4:1, 20 mM aqueous HEPES buffer + 0.1 M NaCl solution, pH 7.4). Left: Visible absorption spectra. Right: absorbance at 580 nm (blue) and 700 nm (red) as a function of copper(II) equivalents.*

A λ_{max} of *ca.* 700 nm, following the Billo's theory, is likely associated to the coordination of 3 to 4 carboxylate groups. It is therefore reasonable to hypothesise that the carboxylate-rich region that was identified by the structural evaluation of SpyCatcher (***Figure 3.13***) actually acts as a metal binding site. Indeed, a (4x O_{COO^-}) coordination mode would result in an absorption expected at λ_{max} of 708 nm,[52–54] which is consistent with the experimental data (observed λ_{max} *ca.* 700 nm).

We explored the possibility to take advantage of the presence of a Trp residue in the sequence of SpyCatcher for the study of the affinity of the protein for copper(II) through fluorescence spectroscopy. Spectrofluorimetric titrations of SC solutions with Cu^{2+} were performed in aqueous HEPES buffer. A representative spectral data set is shown in ***Figure 3.22***.

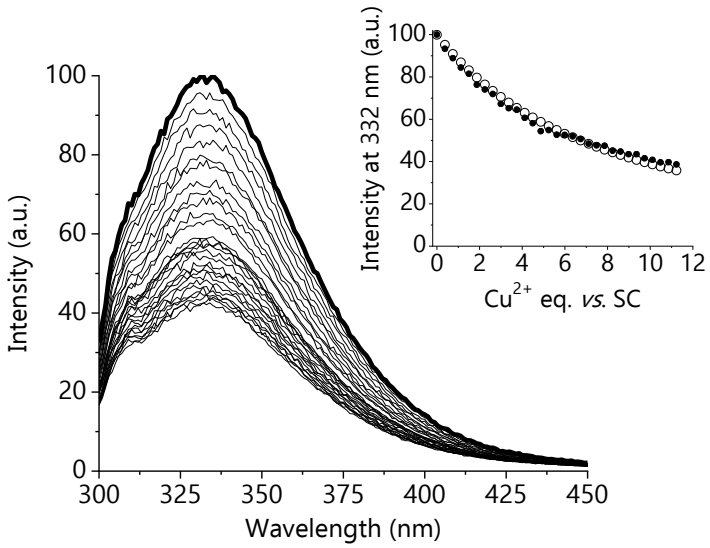


Figure 3.22 *Fluorescence emission spectra of the titration of a SpyCatcher solution with up to 11.25 eq. of* Cu^{2+} *(*C_{SC} *= 10 μM, 20 mM aqueous HEPES buffer + 0.1 M NaCl solution, pH 7.4). Inset: observed (filled circles) and calculated (open circles) intensities at 332 nm vs. copper(II) equivalents.*

The fluorescence spectrum of SpyCatcher acquired in the absence of copper(II) (***Figure 3.22***, bold line) exhibits the characteristic emission profile of tryptophan, *i.e.* a broad band at *ca.* 335 nm. The addition of 2 equivalents of Cu^{2+} to the solution resulted in a limited (*ca.* 20%) although significant fluorescence quenching (***Figure 3.22***, inset). A quenching of Trp emission of *ca.* 60% was observed in the presence of a large metal excess (*ca.* 11 equiv. *vs.* SC). The large distance between the primary metal coordination site of the protein (*i.e.* the N-terminus in the unstructured portion of the protein) and the Trp residue could explain the low-efficiency quenching upon copper addition.

Treatment of the data in ***Figure 3.22*** allows to calculate the affinity of SC for copper(II). The treatment of spectrofluorimetric data with the HypSpec software[55,56] yielded a log K_f of 5.6(1). This value is consistent with (just slightly lower than) the conditional binding constant of Cu^{2+} to pentaglycine N-terminus (6.17)[57]. The log K_f value therefore suggests that the primary metal binding site of copper(II) is the N-terminus H_2N-GSG- site. These considerations are valid in the approximation that the coordination of copper(II) to the secondary binding site is negligible at the concentrations used in spectrofluorimetric titrations. Simulations using Hyss[58] showed indeed that a stepwise log K_f = 3 for the binding of the second equivalent of copper(II) has a negligible impact on the speciation of this system, a behaviour that justifies the aforementioned approximation.

Cu^{2+}/SpyTag-ATCUN (ST-mut1)

The SpyTag-ATCUN (ST-mut1, or 'ST' in short) construct was preliminarily characterised in terms of copper(II) binding by visible absorption spectroscopy and circular dichroism. During the purification of ST, some issues connected with the relatively low solubility of the protein (less than 500 μM) were experienced. Therefore, it was decided to collect the very first preliminary spectra at very low concentration (*ca.* 0.25 mM of Cu^{2+} and ST) to verify the presence of an ATCUN site. Despite low protein concentration, which resulted into small absorbance values, the absorption band at *ca.* 530 nm is fully consistent with ATCUN coordination, as proven by the comparison of CD spectra of an ATCUN peptide (AAHAWG-NH_2)[41] and the SpyTag protein (***Figure 3.23***).

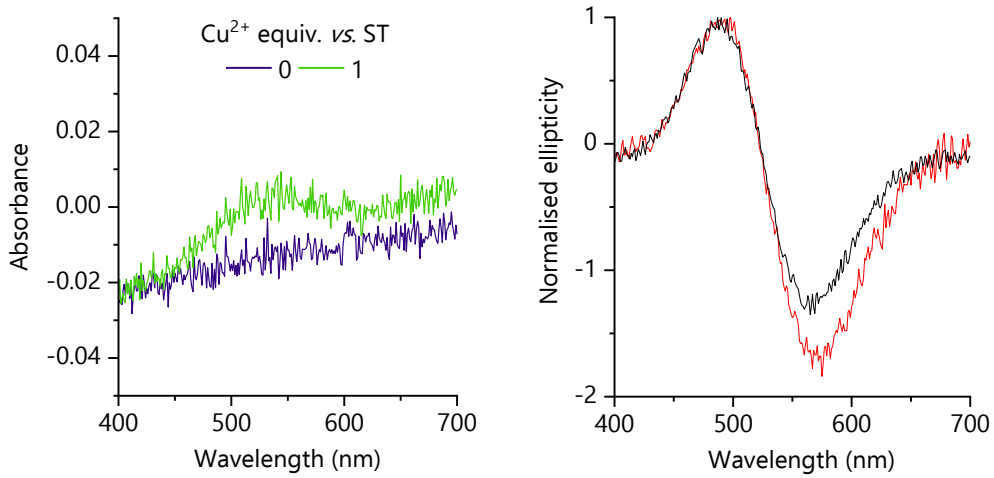


Figure 3.23 *Left: absorption spectra of the ST protein before (purple) and after (green) the addition of 1 eq. of Cu^{2+} equiv. (C_{ST} = 246 μM, 20 mM aqueous HEPES buffer + 0.1 M NaCl solution, pH 7.4). Right: normalised CD spectra of the ST protein (red; C_{ST} = 246 μM, 20 mM aqueous HEPES buffer + 0.1 M NaCl solution, pH 7.4) and of the AAHAWG-NH_2 peptide (black) after the addition of 1 Cu^{2+} equiv. Normalisation was carried out rescaling both spectra at 1 mdeg at λ = 490 nm.*

Once the presence of the ATCUN site was confirmed, spectrophotometric and CD titrations with Cu^{2+} of ST solutions in higher concentration were carried out in aqueous HEPES. Spectral data sets are reported in ***Figure 3.24***. Upon Cu^{2+} additions, an absorption band at ca. 530 nm is observed. Exploiting Billo's average environment theory, the experimental λ_{max} is ascribed to the presence of one amino nitrogen, two deprotonated peptide nitrogen atoms and an imidazole nitrogen on the equatorial plane of Cu^{2+} (***Scheme 3.9***; calc. λ_{max} = 536 nm)[52–54] which is indeed characteristic of an ATCUN binding site. Furthermore, CD spectra show a positive signal at 490 nm and a negative one at 570 nm. These spectral features are consistent with those in ***Figure 3.23*** and they confirm that the binding of copper(II) occurs at the ATCUN motif.

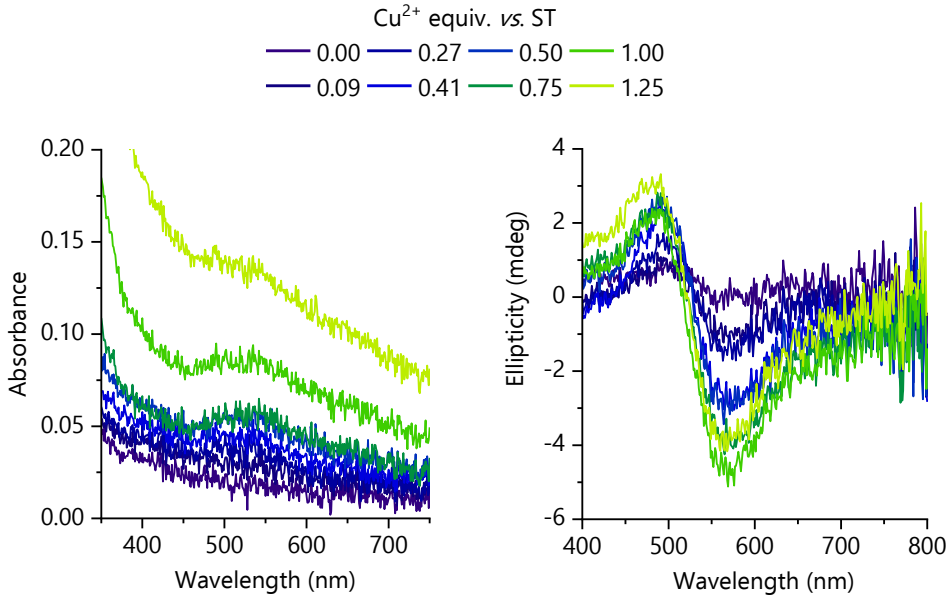
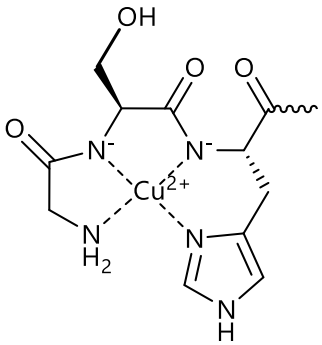


Figure 3.24 *Visible absorption (left) and CD (right) spectral data sets for the titration of ST solutions with up to 1.25 eq. of Cu^{2+} (C_{ST} = 401 μM, 20 mM aqueous HEPES buffer + 0.1 M NaCl solution, pH 7.4).*



Scheme 3.9 *Schematic representation of the ATCUN coordination site in ST.*

Unfortunately, the baseline of the spectra in ***Figure 3.24*** moves to higher absorbance values. This phenomenon was ascribed to the relatively low solubility of ST, which possibly becomes even less soluble in the presence of copper(II). This interpretation is supported by the appearance of turbidity in this protein sample for further addition of copper(II).

Although the data presented above confirm the presence of an ATCUN binding site in ST, it became evident that further characterization had to be made at lower protein concentration. In accordance with this, the study of the affinity of ST for copper(II) was tackled using fluorescence spectroscopy. A first approach consisted into taking advantage of the

presence of Tyr residues in the SpyTag sequence. Spectrofluorimetric titrations of ST solutions with Cu^{2+} were therefore performed in aqueous HEPES buffer, with the aim of observing possible Tyr luminescence quenching phenomena promoted by the presence of the metal ion. Also, the affinity of copper(II)/ATCUN complexes is typically in the pico-/femto-molar range, and therefore its determination cannot be achieved by direct metal/protein titrations at concentrations that are appropriate for UV-vis absorption. A representative spectral data set is shown in ***Figure 3.25***.

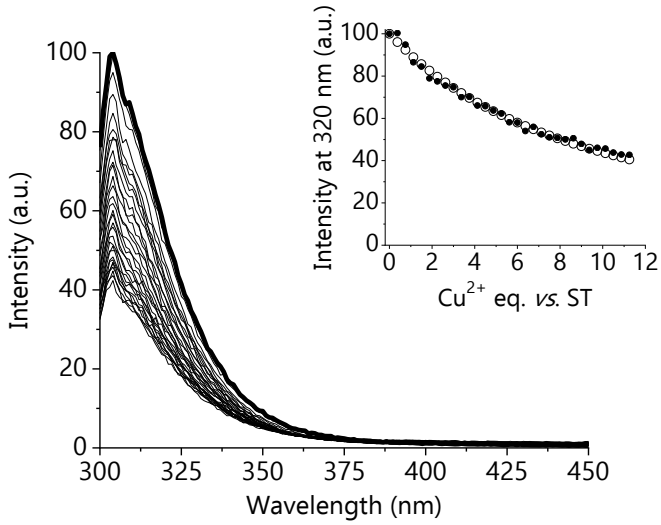


Figure 3.25 *Fluorescence emission spectra for the titration of ST with Cu^{2+} (C_{ST} = 10 μM, Cu^{2+}:ST = 0-11.25:1, 20 mM aqueous HEPES buffer + 0.1 M NaCl solution, pH 7.4). Inset: observed (filled circles) and calculated (open circles) intensities at 320 nm vs. copper(II) equivalents.*

The fluorescence spectra of SpyTag acquired in the absence of copper(II) (***Figure 3.25***, bold line) displays the characteristic emission band of tyrosine at *ca.* 305 nm. The addition of 2 equivalents of Cu^{2+} to the solution resulted in a limited (*ca.* 20%) fluorescence quenching (***Figure 3.25***, inset). The emission intensity then decreased to *ca.* 45% only in the presence of a large metal excess (*ca.* 11 equiv. *vs.* ST). Data treatment of these spectrofluorimetric data yielded a log K_f for the adduct of Cu^{2+} with ST of 5.5(1). This value is surprisingly lower than the one reported in literature for an ATCUN site (log K = 12.0-14.7)[43], and substantially inconsistent with the nature of the binding site that was ascertained to be present in ST. A reasonable interpretation is that the Tyr residues are too distant from the ATCUN site to result into a detectable efficiency in quenching the aromatic emission upon binding of the first equivalent of copper(II). Rather, the observed reduction of luminescence of Tyr along the titration may be ascribed to unselective binding of copper(II) to the construct, which is likely at very high copper(II) concentration.

To solve this issue we decided to go through a competition titration, by titrating with copper(II) a solution of the ATCUN peptide AAHAWG-NH_2 in the absence and in the presence of ST. It was previously demonstrated that the AAHAWG-NH_2 peptide, when titrated with copper(II), undergoes a steep reduction of luminescence of Trp at 350 nm to *ca.* 20% of that of the apo-peptide in correspondence of 1 equivalent of copper(II). The emission remains constant for further metal additions (***Figure 3.26***).[41] Also, the formation constant of the Cu^{2+}/AAHAWG-NH_2 peptide is known (log K_f = 13.5 at pH 7.4, calculated from the speciation obtained by potentiometry).[41]

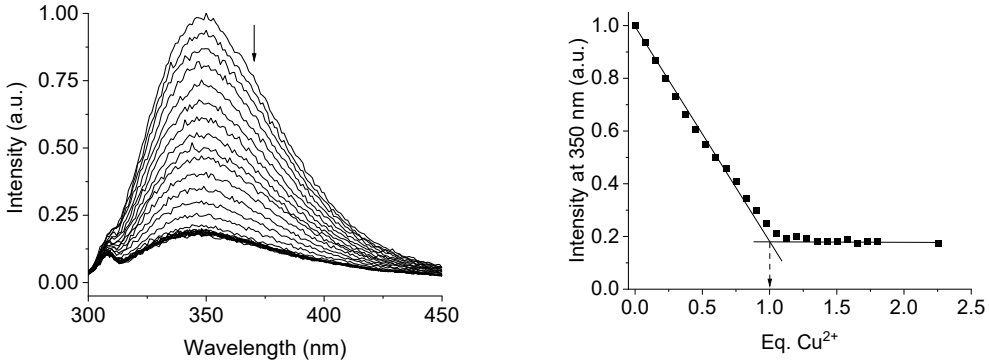


Figure 3.26 *Fluorescence emission spectra (left) and intensity of emission at 350 nm (right) for the titration of a solution of AAHAWG-NH_2 with Cu^{2+} ($C_{peptide}$ = 48 μM, Cu^{2+}/peptide = 0-2.5, 25 mM HEPES, pH 7.4). Data taken from original experiments, see Ref. 41*

Overall, this information suggested that the stability of Cu^{2+}/ST adduct could be calculated from luminescence data obtained by titration of a mixture of ST and AAHAWG-NH_2. Actually, a Trp residue is present solely in the oligopeptide, so that ST operates as a competing dye for AAHAWG-NH_2. Under these circumstances, it is expected that the luminescence of the mixture does not reaches a plateau for the addition of 1 equivalent of copper(II) *vs.* the oligopeptide, as part of the metal ion is sequestered by ST.

Figure 3.27 reports the spectral data set for the titration with copper(II) of a solution containing both ST and AAHAWG-NH_2. The initial spectrum acquired in the absence of copper(II) clearly shows the presence of the emission bands of both SpyTag and AAHAWG-NH_2 at 305 and *ca.* 350 nm, respectively. Upon the addition of 1 Cu^{2+} equivalent, fluorescence intensity is quenched only to 75% of the initial intensity, in contrast to what observed for the AAHAWG-NH_2 peptide alone (***Figure 3.26***). This behaviour proves that a competition for metal binding between the two peptides takes place, as expected. Also, it suggests that the log K_f value is much higher than that previously estimated by direct titration. Indeed, treatment of the spectrofluorimetric data carried out considering a log K_f of 13.5 for the Cu^{2+}/AAHAWG-NH_2 adduct resulted in a log K_f of 13.8(1) for the copper(II)/ST, in complete agreement with values for Cu^{2+}/ATCUN complexes.

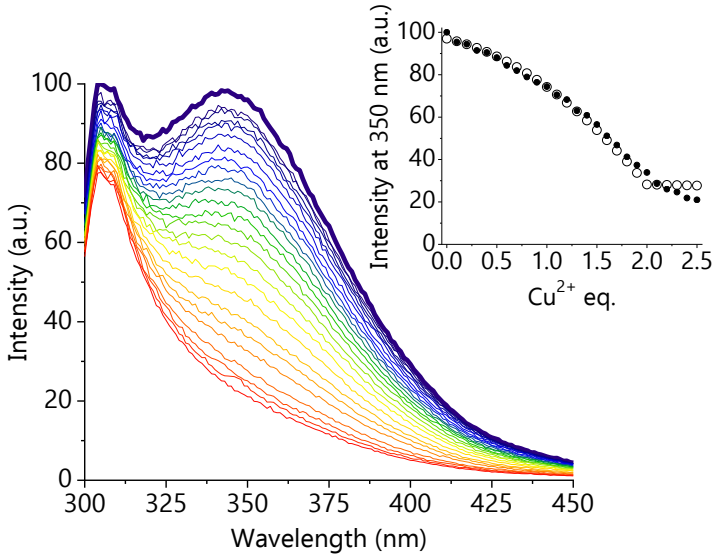


Figure 3.27 *Fluorescence emission spectra for the titration of a solution of ST and AAHAWG-NH₂ with up to 2.5 eq. of Cu^{2+} (C_{ST} = C_{AAH} = 10.6 μM, 20 mM HEPES + 0.1 M NaCl, pH 7.4). Inset: observed (filled circles) and calculated (open circles) intensities at 350 nm vs. copper(II) equivalents.*

Cu^{2+}/SpyTag (ST-mut2)

For the sake of completeness in the characterisation of the SpyTag construct ST-mut2, the binding of copper(II) was studied also for this protein. Although ESI-MS studies revealed that this protein has been obtained as a mixture of four different sequences, the same data suggest that constructs bearing an H_2N-Ala-His- N-terminus may prevail over those in which the SpyTag sequence was cleaved. Therefore, analogously to SpyCatcher the study of copper(II) coordination to ST-mut2 at pH 7.4 in buffer media was carried out. The Cu^{2+}/ST-mut2 adduct was characterised by visible absorption spectroscopy and circular dichroism, which provided information on copper(II) coordination environment.

Absorption and CD spectra for titrations with Cu^{2+} of solutions of ST-mut2 in aqueous HEPES buffer are reported in ***Figure 3.28***. Given the results of the sequencing studies, it was not expected to observe the typical ATCUN spectroscopic fingerprint upon copper(II) binding (*i.e.* an absorption λ_{max} of *ca.* 530 nm and a peculiar CD spectrum presenting a positive band at 490 nm and a negative one at 570 nm, see ***Figure 3.24*** left). Indeed, Cu^{2+} addition to the protein solution resulted in the appearance of an absorption band at 600 nm (***Figure 3.28***, left). Moreover, CD spectra displayed a positive signal at 600 nm and a negative one at 510 nm (***Figure 3.28***, right).

The spectroscopic features of the copper(II)/ST-mut2 are similar to those showed by the coordination of Cu^{2+} to a GHTD-NH_2 peptide which bears a His residue in position

2 (***Figure 3.29***). For that system, an absorption λ_{max} of *ca.* 600 nm and a CD profile that is very similar to that of our Cu^{2+}/ST-mut2 system were observed.[59] In agreement with that previewed by Billo's theory, Cu^{2+} is coordinated to ST-mut2 in a (NH_2, N^-, N_{im}, H_2O) environment (***Scheme 3.10***). On the equatorial plane one amino nitrogen, a deprotonated peptide nitrogen atom, an imidazole nitrogen and a water molecule are coordinated (***Scheme 3.10***; calc. λ_{max} = 599 nm)[52–54]. This model is fully consistent with the experimental data (observed λ_{max} = 600 nm).

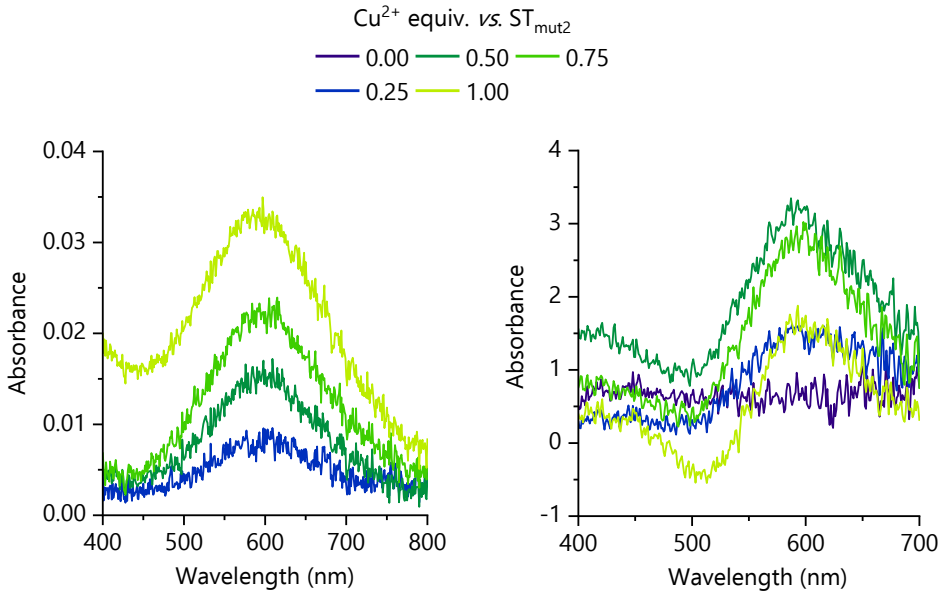


Figure 3.28 *Difference absorption (left) and CD (right) spectra for the titration of a ST-mut2 solution with up to 1 eq. of Cu^{2+} (C_{ST} = 401 μM, 20 mM aqueous HEPES buffer + 0.1 M NaCl solution, pH 7.4). Difference absorption spectra were obtained by subtracting to the spectra that at 0 eq. of copper(II).*

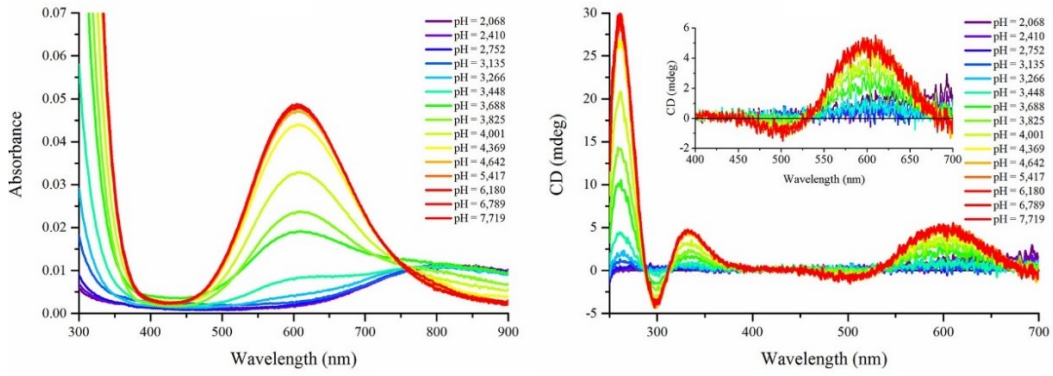
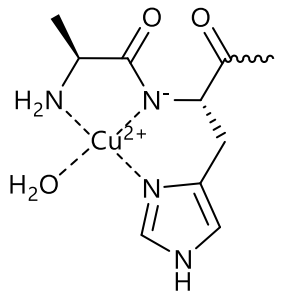


Figure 3.29 *UV–vis and CD pH-metric titrations of 1 mM GHTD-NH_2 in the presence of 0.8 mM $CuCl_2$. The order of the spectra is rainbow-coded from purple to red for increasing pH. Adapted with permission from Ref. 59. Copyright © 2018 American Chemical Society*



Scheme 3.10 *Schematic representation of Cu^{2+} coordination at the N-terminus of ST-mut2.*

Cu^{2+}/SpyCatcher/SpyTag-ATCUN (SC-mut/ST-mut1)

The SpyCatcher/SpyTag construct obtained by recombination of SC-mut and ST-mut1 (hereafter named ST⊂SC in short) is a peculiar, redesigned protein which bears two different N-termini that act as copper(II) binding sites with affinities for Cu^{2+} that are orders of magnitude apart. The ATCUN motif on the SpyTag sequence has an affinity for copper(II) of the order of the high femtomolar ($K_f = 10^{13.8}$), while the SpyCatcher N-terminus, lacking His residues, has an affinity of the order of the micromolar ($K_f = 10^{5.6}$). We therefore intended to prove that in the fusion SpyCatcher/SpyTag construct protein the binding of copper(II) occurs sequentially on the ATCUN site followed by the Spy-Catcher N-terminus, in agreement with the affinities determined for the two individual components. Also for this system the coordination of copper(II) was studied by means of spectrophotometric and CD titrations, which were carried out by adding the metal ion to solutions of SC-mut and ST-mut1 incubated in buffered solution for 10 min prior to metal addition.

The visible absorption spectra for the titration with Cu^{2+} of the ST⊂SC protein are reported in ***Figure 3.30***. A band at *ca.* 525 nm is clearly observed for additions of Cu^{2+} up to *ca.* 1 equivalent (see difference spectra in ***Figure 3.30***), whereas for higher metal amounts the absorption band undergoes a *red shift*. The λ_{max} of 525 nm is the same observed for the coordination of 1 equivalent of copper(II) to SC-mut, and therefore this value suggests that the binding of the metal initially occurs on the ST-mut1 sequence through an AT-CUN coordination mode. Beyond the addition of 0.75 equivalents of metal, the λ_{max} value shifts to 562 nm. Although slightly lower than that observed for the copper(II)/SC-mut adduct (580 nm), these spectroscopic data suggest that the binding of the second equivalent of Cu^{2+} occurs in a (3N, O) coordination environment. Finally, starting from the addition of 1.5 equivalents of copper(II), the absorption spectra become biased by baseline shift due to the formation of insoluble material which became clearly visible in correspondence of 2 equivalents of metal ion added. It is therefore evident that the insol-ubility of the SpyTag construct, although less marked than for the individual component,

is still significant when recombined with the SpyCatcher construct. Finally, this insolubility prevented the detection of the possible binding of copper(II) to the carboxylate-rich region, which was expected to become significant for the addition of a third equivalent of metal ion.

The circular dichroism spectra reported in ***Figure 3.31*** (left) exhibit the characteristic CD profile of ATCUN and they confirm the involvement of this coordination mode for the addition of 1 equivalent of copper(II). The comparison between the CD spectra of the Cu^{2+}/ST and Cu^{2+}/ST⊂SC systems supports this hypothesis (***Figure 3.31***, right). Interestingly and perhaps unexpectedly, a limited change in the CD spectrum was observed when copper(II) is added beyond 1 equivalent.

Overall, these data show that our design was successful to obtain a construct in which the metal cofactor (copper(II)) binds almost selectively to the *Trojan horse* component (ST-mut1). For the ST⊂SC system, the two primary binding sites correspond to the two N-termini of the recombined peptide sequences, which are separated by *ca.* 8 orders of magnitude in their affinity for copper(II). Therefore, for ST⊂SC the condition of having Cu^{2+} selectively bound to the ATCUN site on SpyTag is to have an excess of protein or at maximum an equimolar metal/protein stoichiometric condition. Studies of the activity of this Cu^{2+}/ATCUN construct in promoting the hydrolysis of the activated phosphoester 4-nitrophenyl phosphate are presented in the following section.

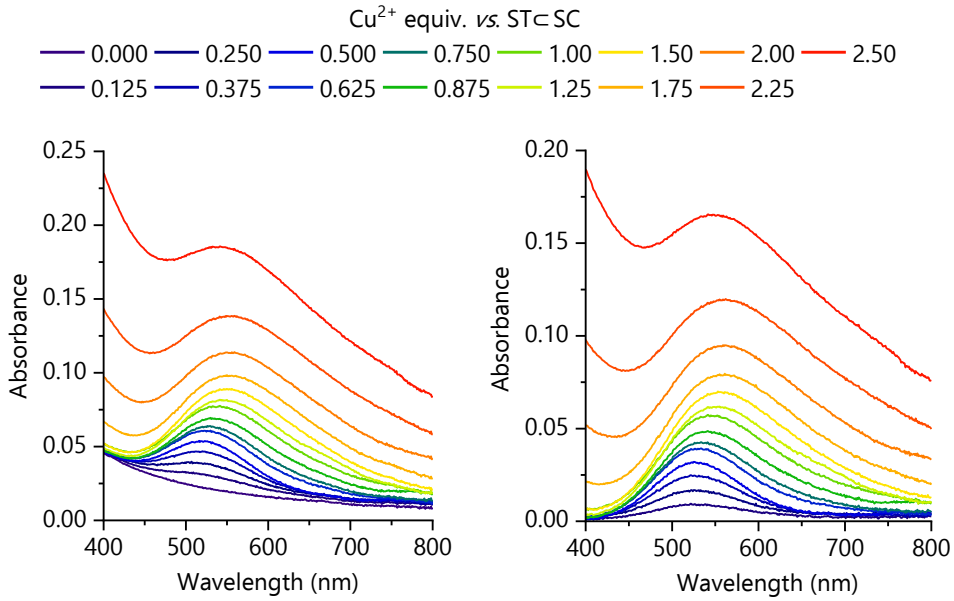


Figure 3.30 *Raw (left) and difference (right) absorption spectra for the titration with Cu^{2+} of a SC-mut /ST-mut1 solution (ST⊂SC) (C_{SC} = C_{ST} = 500 μM, Cu/ST⊂SC = 0-2.5, 20 mM aqueous HEPES buffer + 0.1 M NaCl solution, pH 7.4). Difference absorption spectra were obtained by subtracting to the spectra that at 0 eq. of copper(II).*

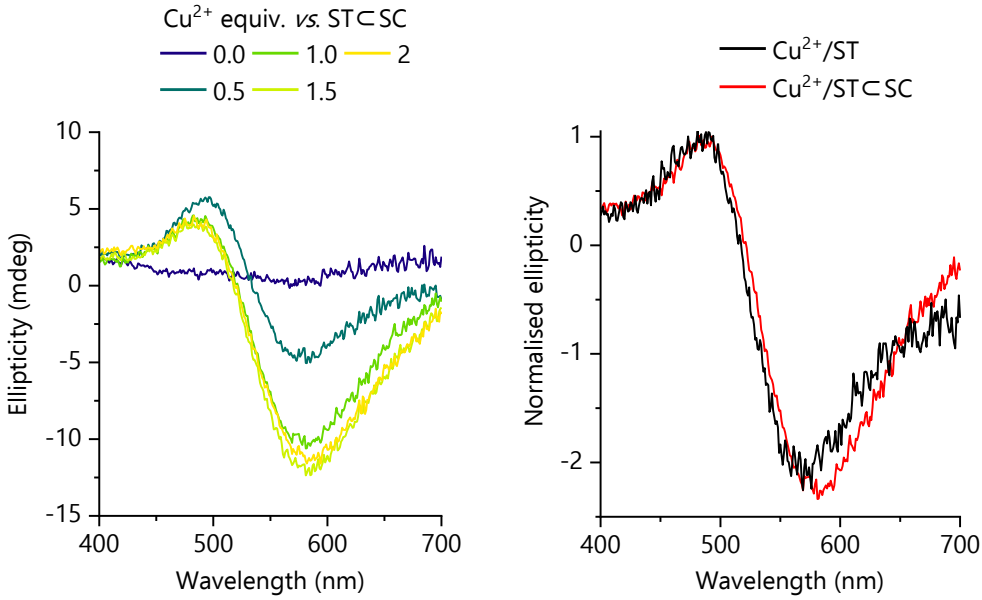
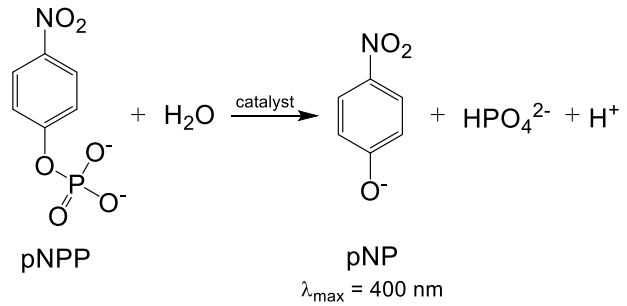


Figure 3.31 *Left: CD spectra for the titration with Cu^{2+} of a SC-mut /ST-mut1 solution (ST⊂SC) ($C_{SC} = C_{ST} = 500\ \mu M$, Cu/ST⊂SC = 0-2, 20 mM aqueous HEPES buffer + 0.1 M NaCl solution, pH 7.4). Right: normalised CD spectra of Cu^{2+}/ST-mut1 (black) and Cu^{2+}/SC-mut /ST-mut1 (red). Normalisation was carried out rescaling both spectra at 1 mdeg at λ = 490 nm.*

HYDROLYSIS OF P-NITROPHENYL PHOSPHATE BY CU(II)/SPYCATCHER/SPYTAG-ATCUN (SC-MUT/ST-MUT1)

In 2008 Singh and co-workers presented a study of a copper(II)/ATCUN tripeptide that exhibited a hydrolytic behaviour toward the activated phosphoric ester *p*-nitrophenyl phosphate (pNPP).[60] In line with that work, the potential esterase activity of the copper(II)/ST⊂SC system was explored by measuring its hydrolytic activity against pNPP as the substrate. The reaction of hydrolysis of pNPP to hydrogen phosphate and the yellow-coloured *p*-nitrophenolate (pNP) was monitored by UV-visible spectroscopy by measuring the absorbance change at 405 nm (***Scheme 3.11***).

Preliminary experiments were carried out in cuvette in the presence of 1 mM substrate, 0.21 mM SpyTag (ST-mut1) and 0.17 mM Cu^{2+}. The plot of absorbance at 400 nm *vs.* time is reported in ***Figure 3.32***.



Scheme 3.11 *Scheme of the reaction of hydrolysis of pNPP.*

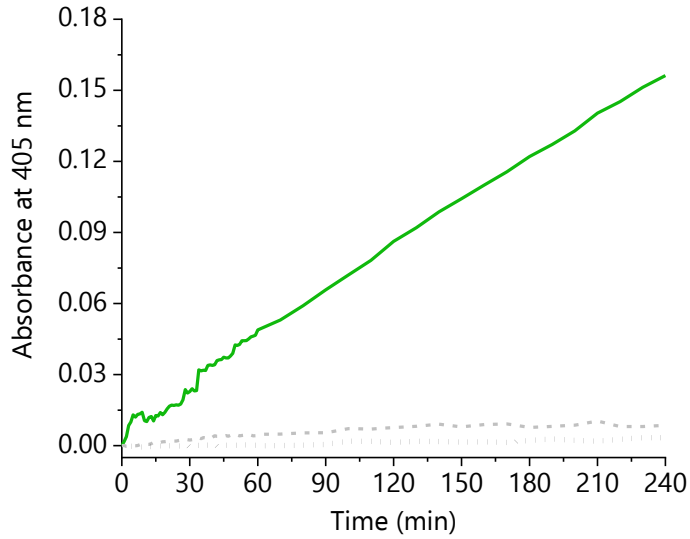


Figure 3.32 *Kinetic profile (abs@405 nm vs. time) for the hydrolysis of p-nitrophenyl phosphate in the presence of Cu^{2+}/ST-mut1 (green), Cu^{2+} (dashed grey), and pNPP self-hydrolysis (dotted black) (1 mM pNPP, 0.21 mM ST-mut1, 0.17 mM Cu^{2+}, 20 mM HEPES, 0.1 M NaCl, pH 7.4, T = 298.2 K).*

The Cu^{2+}/SpyTag (ST-mut1) system resulted to have a quite low catalytic efficiency, with the hydrolytic reaction being extremely slow given the quite high concentration for an enzyme. Actually, given a molar absorbance for pNP of *ca.* 12190 M^{-1} cm^{-1} (from spectrophotometric titrations of pNP, unpublished results), the observed variation in the absorbance of 0.16 over 4 hours corresponds to *ca.* 0.06 turnovers. *Turnover* here is defined as the concentration of product over the concentration of catalyst. It was therefore clear that the study of this system using this experimental setup (cuvette with 1 path length, volume *ca.* 350 μL) required high concentrations of protein and it was therefore excessively protein-consuming to perform all the controls and replicates.

An experimental setup that took advantage of a microplate reader and 96 well plates as reaction vessels was therefore developed. This approach required small quantities of protein

though maintaining high concentrations of the constructs. Actually, it has been verified during the optimisation of the method that the minimum sample volume is 50 μL in wells against 350 μL for microcuvettes, therefore needing lower volumes of protein stock solutions. The system was also optimised in terms of sample volumes and quantities of protein. Round-bottom plates were chosen over flat-bottom ones as they allowed to obtain reproducible absorbance values for pNP, which was used as the coloured standard by using smaller sample volumes. Then, the system was calibrated in terms of registering the absorbance values for increasing pNP concentration, obtaining the calibration curve reported in ***Figure 3.33***. In this way it was also possible to correlate the absorbance of the samples at 405 nm with the concentration of the reaction product pNP through the least-square fitting equation $A^{405\ nm} = 2400 \cdot C_{pNP} + 0.06097$ (C_{pNP} = molar concentration of pNP).

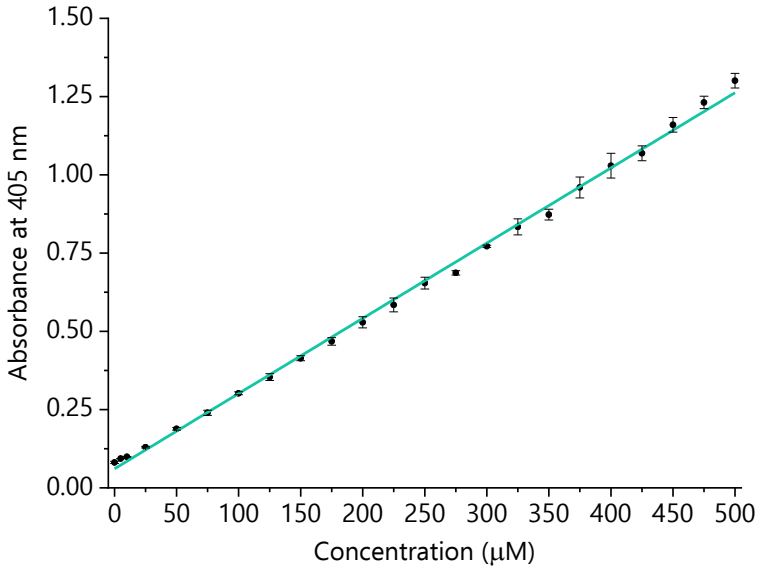


Figure 3.33 *Calibration curve (abs@405 nm vs. concentration) for pNP solutions in wells of 96 wells plate reader (C_{pNP} = 0-500 μM, V_{sample} = 50 μL, 20 mM HEPES, 0.1 M NaCl, pH 7.4).*

Finally, pNPP hydrolysis experiments were performed on the Cu^{2+}/SC, Cu^{2+}/ST and Cu^{2+}/ST⊂SC systems. The reactivity of the substrate (alone and in the presence of Cu^{2+}) and of the apo-proteins (*i.e.* in the absence of the metal) were evaluated as control samples. All the acquired absorbance values of the kinetic profiles for each well were subtracted with those of the blank wells (matrix only). This allowed to obtain the kinetic profiles reported in ***Figure 3.34***.

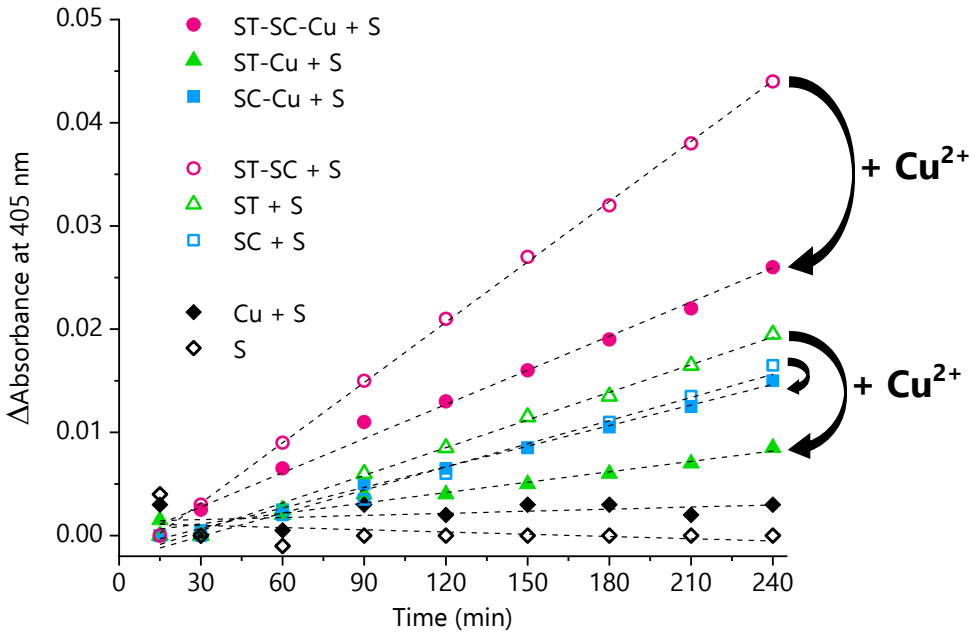


Figure 3.34 *Kinetic profiles (Δabs@405 nm vs. time) of p-nitrophenyl phosphate (S) hydrolysis by the Cu^{2+}/ST⊂SC system (1 mM pNPP, 0.23 mM SC, 0.21 mM ST, 0.17 mM Cu^{2+}, 20 mM HEPES, 0.1 M NaCl, pH 7.4, T = 298.2 K). Δabs was calculated as the difference between the absorbance of the sample and the absorbance of the blank wells.*

The pNPP substrate shows a negligible self-hydrolysis reaction even in the presence of copper(II). In contrast, the constructs containing coordinated Cu^{2+} (to either SpyCatcher or SpyTag) display and enhanced esterase activity compared to aqueous copper(II). It is however most interesting that the apo-proteins (SC, ST, or the ST⊂SC) are more active in assisting the hydrolytic reaction of pNPP than the corresponding metalated constructs. It is also interesting that for apo-ST⊂SC, which is the most active construct, after 240 min the reaction underwent just 0.09 turnovers. Therefore, it can be stated that it was not observed a strictly *catalytic* esterase activity of these constructs, but rather an activity in *promoting* the hydrolysis of the substrate.

The slope of the kinetic profiles of apoproteins is higher than those of the corresponding metalated ones. Also, the slope of the profile of apo-ST⊂SC is slightly higher than the sum of the slopes of the profiles of SC and ST. For samples also containing 1 equivalent of Cu^{2+}, the slope diminishes significantly for the constructs that contain the ATCUN sequence, while this effect is very small for SpyCatcher. These data overall seem to suggest that the capacity of these constructs to promote the hydrolysis of pNPP is mostly construct-based rather than copper(II)-based. Instead, copper(II) seems to act as an inhibitor for this esterase activity.

The nature of the interaction between the substrate and the proteins that are at the basis of the observed reactivity cannot be disclosed from current data. It is likely that some Brønsted acid-base mechanisms operate in promoting the hydrolytic process. At the same time, the fact that copper(II) inhibits more ST than SC seems to suggest that, at least for SpyTag and the fusion protein ST⊂SC, the ATCUN site is involved in the interaction with pNPP. As for the role of copper(II), its role in inhibiting effects seem to be more significant than the reaction promotion effects.

Conclusions

This work describes the design and isolation of the first metalloenzyme based on the SpyCatcher/SpyTag technology. The second-generation SC/ST tagging system was chosen as starting point for the design, and a Trojan horse strategy was used. The metalloprotein consisted of two redesigned moieties: a mutated SpyCatcher protein (SC-mut), and a SpyTag-ATCUN peptide (ST-mut1) that was intended to bind copper(II). The ATCUN motif was chosen as a proof-of-concept due to its high affinity Cu(II)-complexes, which in turn are easily detectable by absorption spectroscopy and circular dichroism and were reported to present some catalytic activity towards phosphoester hydrolysis.

Extensive characterisation of the Cu(II)/SC-mut, Cu(II)/ST-mut1 and Cu(II)/SC-mut/ST-mut1 systems proved the success of the aforementioned design strategy for binding 1 equivalent of copper(II) solely at the SpyTag fragment. While absorption and CD spectroscopies evidenced the formation of a Cu(II)-ATCUN site, the binding affinity was determined by competition fluorescence titrations ($K_f = 10^{13.8}$). On the other hand, also SpyCatcher binds copper(II) at its N-terminus, but with a micromolar affinity. It was finally demonstrated that the coordination of copper(II) to the SC-mut/ST-mut1 construct occurred sequentially on the ATCUN site followed by the SpyCatcher N-terminus, in agreement with the affinities determined for the two individual components.

In terms of esterase activity, the Cu(II)/SC-mut/ST-mut1 system revealed a limited reactivity towards the hydrolysis of the phosphoester p-nitrophenyl phosphate. Interestingly, a higher reactivity was observed for the SC-mut/ST-mut1 construct in the absence of Cu(II), the latter appearing to inhibit the reaction. Moreover, apo-SpyCatcher/SpyTag protein showed a turnover of 0.06-0.09 and hence it showed to be able to promote the hydrolysis of pNPP, although we cannot yet demonstrate a full catalytic activity.

Overall, in this project we demonstrated that the SpyCatcher/SpyTag system can be proficiently used to design a metalloprotein using a Trojan horse strategy. A relatively simple metallopeptide can be introduced into a protein scaffold, to obtain a construct that can bind a metal ion on the SpyTag fragment. A future perspective may be the study of

His-His tandem site present in the SpyTag sequence (which we know are catalytic in oxidation reactions). Metals and cofactors different from copper(II) (*e.g.* heme-peptides) can also be introduced in the SpyTag moiety. Finally, aside the modification of the metal-SpyTag construct, SpyCatcher variants will also be explored in view of tuning the properties of the second coordination sphere that SC provides the metal with. The power of approaching new artificial metalloproteins using this construct relies, in perspective, in the fact that dealing with a peptide (*e.g.* synthesis, characterization, use of non-natural amino acids, *etc.*) is more straightforward than to introduce a metal cofactor into an already folded protein, or to mutate a protein with non-natural residues. This work hence paves the way for a rather straightforward design of innovative metalloenzymes.

EXPERIMENTAL SECTION

Structural and computational studies

The study of protein models was carried out using *PyMOL* (The PyMOL Molecular Graphics System, Version 2.3.4, Schrödinger, LLC). The structural model of SpyCatcher (SC-mut) was obtained from *SWISS-MODEL*[45–49]. A portion of the 3D structure of SpyTag-ATCUN (ST-mut1) was modelled using the software *Phyre²* (Protein Homology/analogy Recognition Engine V 2.0)[50]. All structural models were optimised using *FoldIt*[51] software.

Protein expression and purification

SpyCatcher (SC-mut), SpyTag-ATCUN (ST-mut1) and SpyTag (ST-mut2) were provided by the research group of Prof. A. Bolchi and Prof. S. Ottonello of the University of Parma. Both proteins were expressed in *E. coli* and purified to homogeneity by affinity and ion exchange chromatography. The SpyCatcher (SC-mut) present in the soluble fraction of the bacterial lysate was enriched by an his-tag affinity chromatographic step; after his-tag removal by thrombin digestion, the SpyCatcher protein was isolated in a pure form with an anion exchange column refinement step. The SpyTag-ATCUN (ST-mut1) protein was isolated from the *E. coli* insoluble fraction. The protein was dissolved in 6M urea and eluted from an his-tag affinity column by applying an in-column renaturation protocol. After his-tag thrombin cleavage, the SpyTag-ATCUN protein was purified to homogeneity by anion exchange chromatography. The protein purity was assessed by SDS-PAGE . The correct size of the isolated proteins was verified by electrospray ionisation mass spectrometry analysis (ESI-MS).

Protein sequencing and recombination studies

SpyTag sequencing studies. The sequencing analyses of the ST-mut1 and ST-mut2 proteins were carried out using a liquid chromatograph coupled to a linear ion trap mass spectrometer by an electrospray interface. In particular, the peptide mixture deriving from the tryptic digestion of the protein was subjected to analysis. The reducing agent DL-dithiothreitol (DTT) was added to the initial solution of 2 μM protein in ammonium acetate (pH 8) at a final concentration of 10 mM, for a reaction time of 40 minutes. Subsequently, the alkylation was carried out using iodoacetamide at a final concentration of 40 mM, for 40 minutes in the dark. After reduction and alkylation, trypsin from bovine pancreas was added to the solution in enzyme:substrate ratio of 1:50 (w/w), and the sample was left to incubate at 37 °C overnight. The acquisition of the signal was carried out in data dependent analysis (DDA), using a full scan as a survey scan and two dependent events (determination of the state of charge and acquisition of the MS/MS spectrum). The raw

data was analyzed using BioWorks bioinformatics software, using the FASTA sequence of the protein.

SpyCatcher/SpyTag recombination studies. The studies of recombination were carried out by incubation of SC-mut (12 μg) with proper amounts of ST-mut1 in aqueous 25 mM TRIS buffer and 1 M KP at pH 6 for 5 minutes. SDS-PAGE analyses were carried out in aqueous 25 mM TRIS buffer and 0.15 M NaCl at pH 7.5. In the experiments aimed at clarifying the role of copper(II) in influencing the recombination, an equimolar amount of $CuSO_4$ (*vs.* SC-mut) was added to the protein before the addition of ST-mut1.

Spectroscopic studies

Stock solutions. Concentrated (*ca.* 0.4-2 mM) solutions of SC-mut, ST-mut1 and ST-mut2 were obtained after protein expression and purification, in 20 mM HEPES buffer + 0.1 M NaCl at pH 7.4. Concentrated (*ca.* 3 mM) solutions of AAHAWG-NH_2 were prepared by weight in 20 mM HEPES + 0.1 M NaCl solution at pH 7.4, and the peptide concentration was determined by UV-Vis spectroscopy using the extinction coefficient of Trp (ε_{280nm} = 5690 M^{-1} cm^{-1}).[61] $CuCl_2$, and $Cu(NO_3)_2$ stock solutions were prepared by weight from analytical grade metal salts and standardised by complexometric titrations with EDTA following standard protocols.[62] Aqueous HEPES buffer solutions (20 mM, pH 7.4) containing 0.1 M NaCl were prepared by suspending proper amounts of solid HEPES and NaCl in doubly distilled water, correcting the pH to 7.4 using a concentrated NaOH solution. The buffer medium was passed through a 0.45 μM nylon filter, and the pH was checked every 2-3 days.

UV-Visible spectroscopy. UV-Vis spectra were collected in the spectral range 300-900 nm at T = 298.2 K on a Thermo Scientific Evolution 260 Bio spectrophotometer provided with a Peltier thermostat, using 1 cm path length black quartz cuvettes. Samples were prepared in 20 mM HEPES + 0.1 M NaCl solution at pH 7.4 by adding proper volumes of SC-mut, ST-mut1 or ST-mut2 stock solutions to obtain 0.4-1 mM final protein concentration. Spectrophotometric titrations were carried out by adding $CuCl_2$ to sample solutions up to 4-6, 1.25, and 1 metal equiv. with respect to SC-mut, ST-mut1, and ST-mut2 respectively. Spectrophotometric titrations of SC-mut/ST-mut1 (1:1 SC-mut/ST-mut1 ratio, $C_{protein}$ = 0.25-0.5 mM) were performed by adding $CuCl_2$ up to 4-6 metal equiv. All titrations were performed in duplicate. Spectral data were treated using the software HypSpec2014[55,56].

Circular dichroism (CD) spectroscopy. CD spectra were collected in the visible range (400-800 nm) at T = 298.2 K with a Jasco J715 spectropolarimeter equipped with a Peltier thermostat, using 1 cm path length black quartz cuvettes. Samples were prepared in 20 mM HEPES + 0.1 M NaCl solution at pH 7.4 by adding proper volumes of SC-mut, ST-mut1 or ST-mut2 stock solutions to obtain 0.4-1 mM final protein concentration. Spectrophotometric titrations were carried out by adding Cu^{2+} to sample solutions, up to 4-6

equiv., 1.25 equiv., and 1 equiv. with respect to SC-mut, ST-mut1, and ST-mut2 respectively. Spectrophotometric titrations of SC-mut/ST-mut1 (1:1 SC-mut/ST-mut1 ratio, $C_{protein}$ = 0.25-0.5 mM) were performed by adding Cu^{2+} up to 4-6 equiv. All titrations were performed in duplicate.

Fluorescence spectroscopy. Fluorescence emission spectra were recorded at T = 298.2 K with a spectrofluorometer Horiba Jobin Yvon Fluoromax 3, using 1cm x 1cm path length quartz cuvette. The following acquisition parameters were set: λ_{EM} = 300-450 nm (slit width = 3 nm) and $\lambda_{EX(Trp)}$ = 280 nm (slit width = 2 nm) for all samples containing SC-mut; λ_{EM} = 300-450 nm (slit width = 3 nm) and $\lambda_{EX(Tyr)}$ = 275 nm (slit width = 2 nm) for samples containing only ST-mut1. For the Cu^{2+}/SC-mut, Cu^{2+}/ST-mut1 and Cu^{2+}/SC-mut /ST-mut1 systems, spectrofluorimetric titrations were carried out in aqueous buffer (20 mM HEPES, 0.1 M NaCl, pH 7.4) containing protein(s) at 10 μM concentration. Each solution was titrated with $CuCl_2$ up to 11.25 metal equiv. Spectrofluorimetric competition experiments were carried out in aqueous buffer (20 mM HEPES, 0.1 M NaCl, pH 7.4). Samples containing 10.58 μM ST-mut1 and an equimolar amount of the AA-HAWG-NH_2 peptide were titrated with $CuCl_2$ up to 2.5 metal equiv. All titrations were performed in duplicate. Spectral data were treated using the HypSpec2014[55,56] and Hyss[58] software.

Hydrolysis of p-nitrophenyl phosphate

The catalytic hydrolysis of pNPP by Cu^{2+} adducts with SC-mut, ST-mut1, and SC-mut/ST-mut1 was studied at 20 °C in 20 mM aqueous HEPES buffer containing 0.1 M NaCl at pH 7.4, saturated with atmospheric oxygen. The reaction was monitored by UV-visible spectroscopy on a Photodiode Array Lambda 465 spectrophotometer provided with a Peltier thermostat, following the absorption band of p-nitrophenolate anion at 400 nm (ε = 18500 $M^{-1}cm^{-1}$)[63] for a reaction time of 240 min.

First, pNPP autohydrolysis and kinetic profile of the reaction catalysed by Cu^{2+} were evaluated. The effect of the SC-mut and ST-mut1 proteins, and the SC-mut/ST-mut1 adduct on the copper-catalysed pNPP hydrolysis was studied by adding the proteins to the reaction solution. These experiments were carried out by adding pNPP (1 mM) as the last reagent to solutions containing Cu^{2+} adducts with SC-mut, ST-mut1, or SC-mut/ST-mut1 at 0.17 mM concentration (with 18% protein excess to ensure complete copper complexation). All experiments were performed in duplicate.

Lastly, kinetic experiments were replicated on a round bottom 96-wells plate using a Biochrom EZ Read 400 microplate reader. The system was calibrated with 20 mM HEPES + 0.1 M NaCl solutions (pH 7.4) containing pNP at different concentrations (0-500 μM).

REFERENCES

1. Tegoni, M. De Novo Designed Copper α-Helical Peptides: From Design to Function. *Eur. J. Inorg. Chem.* ***2014***, 2177–2193 (2014).
2. Yu, F. *et al.* Protein Design: Toward Functional Metalloenzymes. (2014) doi:10.1021/cr400458x.
3. Pereira, J. M., Vieira, M. & Santos, S. M. Step-by-step design of proteins for small molecule interaction: A review on recent milestones. *Protein Sci.* ***30***, 1502–1520 (2021).
4. Pinter, T. B. J., Koebke, K. J. & Pecoraro, V. L. Catalysis and Electron Transfer in De Novo Designed Helical Scaffolds. *Angew. Chemie Int. Ed.* ***59***, 7678–7699 (2020).
5. Almhjell, P. J. & Mills, J. H. Metal-chelating non-canonical amino acids in metalloprotein engineering and design. *Curr. Opin. Struct. Biol.* ***51***, 170–176 (2018).
6. Akcapinar, G. B. & Sezerman, O. U. Computational approaches for de novo design and redesign of metal-binding sites on proteins. *Biosci. Rep.* ***37***, 20160179 (2017).
7. Nastri, F. *et al.* Engineering Metalloprotein Functions in Designed and Native Scaffolds. *Trends Biochem. Sci.* ***44***, 1022–1040 (2019).
8. 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.* ***33***, 745–754 (2000).
9. Yu, F. *et al.* Protein Design: Toward Functional Metalloenzymes. *Chem. Rev.* ***114***, 3495–3578 (2014).
10. Lu, Y., Berry, S. M. & Pfister, T. D. Engineering Novel Metalloproteins: Design of Metal-Binding Sites into Native Protein Scaffolds. (2001) doi:10.1021/cr0000574.
11. Lu, Y. Metalloprotein and Metallo-DNA/RNAzyme Design: Current Approaches, Success Measures, and Future Challenges. *Inorg. Chem.* ***45***, 9930–9940 (2006).
12. Hellinga, H. W. The construction of metal centers in proteins by rational design. *Fold. Des.* ***3***, R1–R8 (1998).
13. Lu, Y., Yeung, N., Sieracki, N. & Marshall, N. M. Design of functional metalloproteins. *Nature* ***460***, 855–862 (2009).
14. Woolfson, D. N. A Brief History of De Novo Protein Design: Minimal, Rational, and Computational. *J. Mol. Biol.* ***433***, 167160 (2021).
15. Zhao, H. & Arnold, F. H. Combinatorial protein design: strategies for screening protein libraries. *Curr. Opin. Struct. Biol.* **7**, 480–485 (1997).
16. Saven, J. G. Combinatorial protein design. *Curr. Opin. Struct. Biol.* ***12***, 453–458 (2002).
17. The Nobel Prize in Chemistry 2018. https://www.nobelprize.org/prizes/chemistry/2018/summary/.

18. Voigt, C. A., Kauffman, S. & Wang, Z. G. Rational evolutionary design: The theory of in vitro protein evolution. *Adv. Protein Chem.* ***55***, 79–160 (2001).
19. Lutz, S. Beyond directed evolution—semi-rational protein engineering and design. *Curr. Opin. Biotechnol.* ***21***, 734–743 (2010).
20. Goldsmith, M. & Tawfik, D. S. Directed enzyme evolution: beyond the low-hanging fruit. *Curr. Opin. Struct. Biol.* ***22***, 406–412 (2012).
21. Zeymer, C. & Hilvert, D. Directed Evolution of Protein Catalysts. *Annu. Rev. Biochem.* ***87***, 131–157 (2018).
22. Hellinga, H. W. Commentary Rational protein design: Combining theory and experiment. ***94***, 10015–10017 (1997).
23. Zastrow, M. L. & Pecoraro, V. L. Influence of active site location on catalytic activity in de novo -designed zinc metalloenzymes. *J. Am. Chem. Soc.* ***135***, 5895–5903 (2013).
24. Tiwari, M. K., Singh, R., Singh, R. K., Kim, I. W. & Lee, J. K. COMPUTATIONAL APPROACHES FOR RATIONAL DESIGN OF PROTEINS WITH NOVEL FUNCTIONALITIES. *Comput. Struct. Biotechnol. J.* ***2***, e201204002 (2012).
25. Schmid, A. *et al.* Industrial biocatalysis today and tomorrow. *Nat. 2001 4096817* ***409***, 258–268 (2001).
26. Aehle, W. *Enzymes in Indurstry*. (Wiley-VCH, 2007).
27. Deuss, P. J., Denheeten, R., Laan, W. & Kamer, P. C. J. Bioinspired Catalyst Design and Artificial Metalloenzymes. *Chem. – A Eur. J.* ***17***, 4680–4698 (2011).
28. Lin, Y. W. Rational Design of Artificial Metalloproteins and Metalloenzymes with Metal Clusters. *Mol. 2019, Vol. 24, Page 2743* ***24***, 2743 (2019).
29. Wilson, M. E. & Whitesides, G. M. Conversion of a protein to a homogeneous asymmetric hydrogenation catalyst by site-specific modification with a diphosphinerhodium(I) moiety. *J. Am. Chem. Soc.* ***100***, 306–307 (1978).
30. Collot, J. *et al.* Artificial metalloenzymes for enantioselective catalysis based on biotin-avidin. *J. Am. Chem. Soc.* ***125***, 9030–9031 (2003).
31. Ward, T. R. Artificial Metalloenzymes Based on the Biotin−Avidin Technology: Enantioselective Catalysis and Beyond. *Acc. Chem. Res.* ***44***, 47–57 (2010).
32. Zakeri, B. *et al.* Peptide tag forming a rapid covalent bond to a protein, through engineering a bacterial adhesin. *Proc. Natl. Acad. Sci. U. S. A.* ***109***, E690–E697 (2012).
33. Hagan, R. M. *et al.* NMR spectroscopic and theoretical analysis of a spontaneously formed Lys-Asp isopeptide bond. *Angew. Chem. Int. Ed. Engl.* ***49***, 8421–8425 (2010).
34. Li, L., Fierer, J. O., Rapoport, T. A. & Howarth, M. Structural Analysis and Optimization of the Covalent Association between SpyCatcher and a Peptide Tag. *J. Mol. Biol.* ***426***, 309–317 (2014).
35. Reddington, S. C. & Howarth, M. Secrets of a covalent interaction for biomaterials and biotechnology: SpyTag and SpyCatcher. *Curr. Opin. Chem. Biol.* ***29***, 94–99 (2015).

36. Keeble, A. H. *et al.* Evolving Accelerated Amidation by SpyTag/SpyCatcher to Analyze Membrane Dynamics. *Angew. Chemie Int. Ed.* ***56***, 16521–16525 (2017).
37. Rózga, M., Sokołowska, M., Protas, A. M. & Bal, W. Human serum albumin coordinates Cu(II) at its N-terminal binding site with 1 pM affinity. *JBIC J. Biol. Inorg. Chem.* ***12***, 913–918 (2007).
38. Płonka, D. & Bal, W. The N-terminus of hepcidin is a strong and potentially biologically relevant Cu(II) chelator. *Inorganica Chim. Acta* ***472***, 76–81 (2018).
39. Harford, C. & Sarkar, B. Amino Terminal Cu(II)- and Ni(II)-Binding (ATCUN) Motif of Proteins and Peptides: Metal Binding, DNA Cleavage, and Other Properties†. *Acc. Chem. Res.* ***30***, 123–130 (1997).
40. Mlynarz, P. *et al.* Impact of the peptide sequence on the coordination abilities of albumin-like tripeptides towards Cu 2+ , Ni 2+ and Zn 2+ ions. Potential albumin-like peptide chelatorsy. *New J. Chem* ***26***, 264–268 (2002).
41. Perinelli, M. *et al.* Cu(II) coordination to His-containing linear peptides and related branched ones: Equalities and diversities. *J. Inorg. Biochem.* ***205***, 110980 (2020).
42. Gonzalez, P. *et al.* N-Terminal Cu-Binding Motifs (Xxx-Zzz-His, Xxx-His) and Their Derivatives: Chemistry, Biology and Medicinal Applications. *Chem. - A Eur. J.* ***24***, 8029–8041 (2018).
43. Maiti, B. K., Govil, N., Kundu, T. & Moura, J. J. G. Designed Metal-ATCUN Derivatives: Redox- and Non-redox-Based Applications Relevant for Chemistry, Biology, and Medicine. *iScience* vol. 23 101792 (2020).
44. Giglione, C., Boularot, A. & Meinnel, T. Protein N-terminal methionine excision. *Cellular and Molecular Life Sciences* vol. 61 1455–1474 (2004).
45. Bienert, S. *et al.* The SWISS-MODEL Repository—new features and functionality. *Nucleic Acids Res.* ***45***, D313–D319 (2017).
46. Bertoni, M., Kiefer, F., Biasini, M., Bordoli, L. & Schwede, T. Modeling protein quaternary structure of homo- and hetero-oligomers beyond binary interactions by homology. *Sci. Rep.* **7**, 10480 (2017).
47. Guex, N., Peitsch, M. C. & Schwede, T. Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: A historical perspective. *Electrophoresis* ***30***, S162–S173 (2009).
48. Waterhouse, A. *et al.* SWISS-MODEL: Homology modelling of protein structures and complexes. *Nucleic Acids Res.* ***46***, W296–W303 (2018).
49. Studer, G. *et al.* QMEANDisCo-distance constraints applied on model quality estimation. *Bioinformatics* ***36***, 1765–1771 (2020).
50. Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N. & Sternberg, M. J. E. The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.* ***10***, 845–858 (2015).
51. Kleffner, R. *et al.* Foldit Standalone: a video game-derived protein structure manipulation interface using Rosetta. *Bioinformatics* ***33***, 2765–2767 (2017).
52. Billo, E. J. Copper(II) chromosomes and the rule of average environment. *Inorg. Nucl. Chem. Lett.* ***10***, 613–617 (1974).

53. Prenesti, E., Daniele, P. G., Prencipe, M. & Ostacoli, G. Spectrum-structure correlation for visible absorption spectra of copper(II) complexes in aqueous solution. *Polyhedron* ***18***, 3233–3241 (1999).
54. Sigel, H. & Martin, R. B. Coordinating properties of the amide bond. Stability and structure of metal ion complexes of peptides and related ligands. *Chem. Rev.* ***82***, 385–426 (1982).
55. Gans, P., Sabatini, A. & Vacca, A. Investigation of equilibria in solution. Determination of equilibrium constants with HYPERQUAD suite of programs. *Talanta* ***43***, 1739–1753 (1996).
56. Gans, P., Sabatini, A. & Vacca, A. To improve accuracy of the calculated pKa values,. *Annali di Chimica* vol. 89 45–49 (1999).
57. Sóvágó, I., Sanna, D., Dessí, A., Várnagy, K. & Micera, G. EPR and potentiometric reinvestigation of copper(II) complexation with simple oligopeptides and related compounds. *J. Inorg. Biochem.* ***63***, 99–117 (1996).
58. Alderighi, L. *et al. Hyperquad simulation and speciation (HySS): a utility program for the investigation of equilibria involving soluble and partially soluble species. Coordination Chemistry Reviews* vol. 184 http://www.eawag.ch/soft/ (1999).
59. Kotuniak, R., Fraczyk, T., Skrobecki, P., Płonka, D. & Bal, W. Gly-His-Thr-Asp-Amide, an Insulin-Activating Peptide from the Human Pancreas Is a Strong Cu(II) but a Weak Zn(II) Chelator. *Inorg. Chem.* ***57***, 15507–15516 (2018).
60. Singh, U., Singh, R., Sharma, N. & Prasad, R. DNA Cleavage Study Using Copper (II)-GlyAibHis: A Tripeptide Complex Based on ATCUN Peptide Motifs. *Protein Pept. Lett.* ***15***, 13–19 (2008).
61. Pace, C. N. & Schmid, F. X. *Protein structure: a practical approach*. (Oxford University Press, 1997).
62. Vogel, A. I. *Quantitative Inorganic Analysis Including Elementary Instrumental Analysis*. (Longmans, 1962).
63. Selmeczi, K., Giorgi, M., Speier, G., Farkas, E. & Réglier, M. Mono- versus Binuclear Copper(II) Complexes in Phosphodiester Hydrolysis. *Eur. J. Inorg. Chem.* ***2006***, 1022–1031 (2006).

Concluding remarks

The work reported in this thesis has tackled the investigation of copper/histidine sites in designed peptides and artificial proteins. Our research efforts were directed at disclosing the relationship between the nature of these metal centres and the reactivity of the systems in which they are contained. Our studies involved peptide systems and approaches progressively more complex, and each chapter of this dissertation represents a stepwise increase in the research complexity.

The work presented in *Chapter 1* relates to the study of copper binding to fragments of tau protein that are characterised by a single metal binding motif, being either a mono- or a bis-His site. These systems have a great biological relevance since Cu/tau adducts are involved in neurodegenerative processes. The intriguing aspect of this project is the rationalisation of Cu/tau pathological activity in terms of metal coordinating features of mono- and bis-His sites. In this perspective, our study provides a solid description of the structural and thermodynamic features that may trigger pathological Cu/tau-mediated processes. Our findings on progressively longer fragments of tau also represent an important step in view of a future evaluation of the full-length protein.

Given the encouraging results obtained for tau fragments, we decided to explore the reactivity of peptide systems containing two copper/histidine sites. This project was more ambitious than the previous one, not only for the presence of an additional metal binding site that may complicate the study, but also because new peptide sequences had to be designed and synthesised. The investigation of these systems, described in *Chapter 2*, led to an unexpected but most fascinating outcome. Indeed, we observed that the presence of an isolated histidine residue did not influence the reactivity of bis-His motifs and that the two sites behaved as independent metal binding sites, irrespective of their distance. This is in contrast with what reported in the literature for β-amyloid peptides, for instance, in which a His-His dyad interacts with a third His residue in what was defined as an 'in-between' state. Our findings may raise questions and perhaps lead to a comprehensive re-evaluation of the coordination properties of these systems.

Finally, our ultimate challenge consisted in the realisation of an artificial metalloprotein (*Chapter 3*). Through protein design, we aimed at conferring to a peptide tagging system an entirely novel function, that is, the capacity of binding a metal ion and possibly display a catalytic activity. The success of our approach is extremely relevant as it paves the way for the rather straightforward creation of metalloproteins in which the complexity of the design is moved from the protein to the peptide. The resulting proteins can therefore bear a large number of artificial catalytic sites and still be fully peptidic in their components, with potential in promoting not only biological reactions but also chemical conversions of interest in industry or bio(nano)technology.