UNIVERSITY OF NAPLES FEDERICO II

Polytechnic and Basic Sciences School

Ph.D. School in Chemical Sciences

De Novo Designed Copper-Containing Metalloenzymes for Oxidative Chemistry

SALVATORE LA GATTA

Advisor:

Prof. Angela Lombardi
Prof. Flavia Nastri

Examiner:

Prof. Lucia Panzella

XXXV Cycle 2019-2022

Coordinator: *Prof.* Angela Lombardi

This page is intentionally left blank.

Siamo chimici, cioè cacciatori: nostre sono "le due esperienze della vita adulta" di cui parlava Pavese, il successo e l'insuccesso, uccidere la balena bianca o sfasciare la nave; non ci si deve arrendere alla materia incomprensibile, non ci si deve sedere. Siamo qui per questo, per sbagliare e correggerci, per incassare colpi e renderli. Non ci si deve mai sentire disarmati: la natura è immensa e complessa, ma non è impermeabile all'intelligenza; devi girarle intorno, pungere, sondare, cercare il varco o fartelo.

Primo Levi, *Il Sistema Periodico*, Nichel, 1975.

This page is intentionally left blank.

Table of Contents

SUMMARY ..... - 1 -

1. INTRODUCTION ..... - 4 -

1.1 COPPER IN BIOLOGICAL SYSTEMS ..... - 4 -

1.2 DIOXYGEN ACTIVATION IN COPPER METALLOENZYMES ..... - 8 -

1.3 NATURAL COPPER OXIDOREDUCTASE ..... - 9 -

1.3.1 Type 2 Copper Proteins: the elusive nature of the histidine brace copper-binding site ..... *- 10 -*

1.3.2 Type 3 Copper Proteins: altering the enzymatic activity by tuning outer shell interactions ..... *- 18 -*

1.3.3 Multicopper Oxidases and Particulate Methane Monooxygenases: different copper centers inside the same protein ..... *- 26 -*

1.4 NESTING COPPER ACTIVE SITES INTO DE NOVO DESIGNED PROTEINS ..... - 31 -

1.4.1 Protein design of helical bundles ..... *- 31 -*

1.4.2 De novo design of copper proteins ..... *- 37 -*

1.4.3 De novo design of catalytic copper proteins ..... *- 39 -*

1.5 AIM OF THE THESIS ..... - 45 -

1.6 REFERENCES ..... - 47 -

2. ENGINEERING A TYPE 3 COPPER SITE INTO A DE NOVO DESIGNED SCAFFOLD ..... **- 58 -**

2.1 RESULTS AND DISCUSSION ..... - 58 -

2.1.1 DR1 Design ..... *- 58 -*

2.1.2 Shaping active site accessibility and design validation ..... *- 61 -*

2.1.3 DR1 Folding, Copper Binding and Thermodynamic Stability ..... *- 69 -*

2.1.4 Spectroscopic Characterization of the Di-Copper Center ..... *- 72 -*

2.1.5 Di-Copper-DR1 Diphenolase Activity: DTBC oxidation ..... *- 74 -*

2.1.6 DTBQ interaction within the active site ..... *- 77 -*

Table of Contents

2.1.7 Di-Cu^{2+}-DR1 kinetic parameters and substrate recognition by Hammett analysis ..... - 79 -
2.1.8 di-Cu^{2+}-DR1 substrate recognition by DFT calculations ..... - 82 -
2.1.9 Mechanistic studies of dioxygen activation ..... - 86 -
2.2 EXPERIMENTAL SECTION ..... - 94 -
2.2.1 Design ..... - 94 -
2.2.2 Molecular Dynamics Simulation ..... - 95 -
2.2.3 DFT calculations ..... - 96 -
2.2.4 Synthesis ..... - 97 -
2.3.5 Gel Filtration Chromatography ..... - 101 -
2.2.6 CD Spectroscopy ..... - 102 -
2.2.7 Isothermal Titration Calorimetry ..... - 103 -
2.2.8 UV-Vis absorption spectroscopy ..... - 104 -
2.2.9 EPR spectroscopy ..... - 105 -
2.2.10 Products identification ..... - 106 -
2.2.11 Catalytic assays ..... - 108 -
2.2.12 Nuclear Magnetic Resonance spectroscopy ..... - 110 -
2.2.13 Amplex Red assay ..... - 112 -
2.2.14 Kinetic simulation of progress curves ..... - 114 -
2.3 CONCLUSIONS ..... - 117 -
2.4 REFERENCES ..... - 118 -

3. MIMICKING THE HISTIDINE BRACE MOTIF IN DE NOVO DESIGNED PROTEINS ..... - 121 -

3A.1 RESULTS AND DISCUSSION ..... - 122 -
3A.1.1 Design strategy of miniLPMO ..... - 122 -
3A.1.2 miniLPMO folding and oligomerization state ..... - 126 -
3A.1.3 miniLPMO copper binding ..... - 127 -
3A.1.4 ^{14}N ENDOR and HYSCORE spectroscopy ..... - 132 -

Table of Contents

3A.1.5 Hypothetical coordination geometries of the species across the pH ..... - 135 -
3A.1.6 Hydrogen peroxide activation by Cu-miniLPMO ..... - 136 -
3A.2 EXPERIMENTAL SECTION ..... - 138 -
3A.2.1 Design ..... - 138 -
3A.2.2 Molecular Dynamics simulations ..... - 139 -
3A.2.3 Synthesis ..... - 140 -
3A.2.4 CD Spectroscopy ..... - 141 -
3A.2.5 UV-Vis absorption spectroscopy ..... - 142 -
3A.2.6 CW EPR spectroscopy ..... - 143 -
3A.2.7 Davies ^{14}N ENDOR spectroscopy ..... - 145 -
3A.2.8 HYSCORE spectroscopy ..... - 146 -
3A.2.9 Catalytic assays ..... - 146 -
3A.3 REFERENCES ..... - 148 -
3B.1 RESULTS AND DISCUSSION ..... - 150 -
3B.1.1 Design strategy of dHisB ..... - 150 -
3B.1.2 DHisB folding and oligomerization state ..... - 156 -
3B.1.3 DHisB copper binding and spectroscopic analysis ..... - 159 -
3B.1.4 Role of N-Terminal Amine in the copper binding ..... - 168 -
3B.1.5 dHisB reduction and Azide Binding ..... - 169 -
3B.1.6 Dioxygen activation by Cu-dHisB ..... - 172 -
3B.2 EXPERIMENTAL SECTION ..... - 177 -
3B.2.1 Design ..... - 177 -
3B.2.2 Molecular Dynamics simulations ..... - 177 -
3B.2.3 Synthesis ..... - 177 -
3B.2.4 Size-exclusion chromatography ..... - 180 -
3B.2.5 CD spectroscopy ..... - 182 -
3B.2.6 Copper reconstitution ..... - 183 -
3B.2.7 CW-EPR spectroscopy ..... - 183 -
3B.2.8 UV-Vis absorption spectroscopy ..... - 188 -

Table of Contents

3B.2.9 Amplex Red assay ..... *- 188 -*

3.3. CONCLUSIONS ..... - 190 -

3B.3 REFERENCES ..... - 192 -

4. CONCLUSIONS AND PERSPECTIVES ..... - 195 -

5. LIST OF ABBREVIATIONS ..... - 197 -

Summary

Copper proteins are involved in several key biological processes, such as electron-transfer, dioxygen binding and activation, pigmentation, methane monooxygenation, and denitrification. The copper's ubiquity is mostly related to its abundance on the earth crust, the two accessible oxidation states, and its coordinative promiscuity. It is, in fact, found in a range of coordination environments, which are classified by their spectroscopic properties. Moreover, the protein matrix strictly controls selectivity and catalytic efficiency of the copper center by primary and outer sphere interactions. The most abundant copper sites are the so-called Type 1, 2, and 3. While the subtle interplay of bonding and non-bonding interactions has been exhaustively explored for simpler and closed-shell Type 1 proteins, Type 2 and 3 remain more elusive in their structure-function correlations.

Among several approaches to study natural metalloenzymes, such as protein engineering and site-directed mutagenesis, de novo protein design represents a valuable strategy to transplant the target metal-binding-site in a simpler and possibly more stable small-sized scaffolds. This generally allows for direct evaluation of metal-protein interactions by a bottom-up approach, in which structural determinants for metal-binding are engineered from scratch. In this thesis, three copper-binding designed proteins are showcased: (i) DR1, a four-helical bundle hosting a Type 3 copper site; (ii) miniLPMO, a quite unnatural homodimeric four-helical bundle hosting a Type 2 copper site; (iii) dHisB, a heterodimeric three-helical bundle hosting a Type 2 copper site.

Summary

DR1 (Due Rame, in Italian) is a newly designed protein that mimics polyphenol oxidases and contains a di-copper site. The first and second di-metal coordination spheres were hierarchically engineered in order to nest the di-copper site into a simpler scaffold made of a four-helix bundle. DR1 recapitulates the Type 3 copper site, supporting several copper redox states and being active in the O_2-dependent oxidation of catechols to o-quinones, according to spectroscopic, thermodynamic, and functional analysis. Most importantly, DR1 is endowed with substrate recognition thanks to the careful design of the binding pocket residues, as confirmed by Hammett analysis and computational studies on substituted catechols. Moreover, deeper spectroscopic characterization prompted us to propose a kinetic model for dioxygen activation for this synthetic enzyme, which involves hydrogen peroxide as an intermediate.

In parallel, the de novo design of a peptide-based model inspired by lytic polysaccharide monooxygenase (LPMO) proteins was undertaken. Motivated by the elusive nature of the histidine brace copper-binding site, as found in LPMO proteins, the design process was performed on two different scaffolds of increasing complexity. MiniLPMO is a C_2-symmetric de novo designed miniprotein that homodimerizes to form a bisecting four-helix bundle, as confirmed by CD analysis. MiniLPMO binds copper with a histidine brace-like motif, as assessed by UV-Visible absorption and continuous wave/pulsed EPR spectroscopy and shows oxidase activity towards the model substrate 4-nitropheyl-β-D-glucopyranoside. In a second approach, symmetry was released and a heterodimeric construct was designed, featuring a helix-loop-helix (α_2 motif) that specifically recognizes a helical peptide (α chain), generating a heterodimeric three-helix bundle, called dHisB (designed Histidine

Brace). The design process led to a single histidine brace site hosted on the N-terminal side of the α chain, facing toward the loop of the α_2 motif that exerts secondary shell interactions. The desired folding and heterodimerization were assessed by size-exclusion chromatography, CD spectroscopy and direct quantification of peptide content by reversed-phase HPLC. Using a combination of continuous wave EPR and UV-Visible absorption spectroscopy over a range of pH values, it has been shown that the active site of dHisB can exist in different protonation states. These small models recapitulate the spectroscopic fingerprint of natural LPMOs, confirming the power of the design strategy. Interestingly, dHisB is able to cycle between both copper oxidation states under mild conditions and bind exogenous ligands like sodium azide. Finally, O_2-activation has been shown by the Amplex Red assay with 3-fold apparent enhancement with respect to free copper in aqueous solution.

Remarkably, the three models allowed a very fine spectroscopic characterization, not always attainable with natural proteins. The Type 2 copper-containing miniproteins allowed unraveling for the first time, to the best of our knowledge, the specific contribution to the EPR spectrum of the different protonation states of the histidine brace center over a wide range of pH (from pH 2 to 13). Noteworthy, all the designed models activate O_2 or H_2O_2 in a similarly way to their natural counterparts. These simple models represent a milestone in the development of synthetic metalloenzymes for the degradation and conversion of biomass into second-generation fuels.

1. Introduction

Keywords: Copper • Metalloenzymes • Oxidative Chemistry • Protein Design • Biomimetics

1.1 Copper in biological systems

Metals are omnipresent in biology, from microorganisms to superior animals and plants.[1] They are essential for numerous biological processes, which include enzymatic catalysis, oxygen transport and activation, and electron transfer. Life relies on accumulation, transport and use of many different metal ions: iron, zinc, copper, molybdenum, cobalt, chromium, vanadium and nickel, in order of decreasing abundance.[2] Such d-block elements are commonly bound to cofactors, such as porphyrins and cobalamines, or to proteins in which residues with N, O and S-containing sidechains are arranged to form metal-binding sites. Copper, investigated within this work, has a rich bioinorganic chemistry, second only to iron in its complexity, and can be found in metalloproteins with widely different functions from all biological domains.[3] The emergence of copper metalloproteins seems to be related to the great oxidation event that accompanied the evolution and proliferation of early cyanobacteria, and the subsequent rapid increase of oxygen concentration in the atmosphere.[4] Whereas early redox biochemistry was performed mostly through iron cofactors, the accumulation of oxygen led to increased availability of copper ions, among others, and consequently to the development of entirely new bioinorganic pathways. Nowadays, copper is an essential cofactor for

several protein activities, including electron transfer (ET) and catalysis. The first coordination sphere and the second coordination sphere/interactions are tuned inside the complex protein matrix to control these functions. In this context, copper proteins perform various functions. The active sites of most representative natural copper-binding proteins are depicted in figure 1.1.

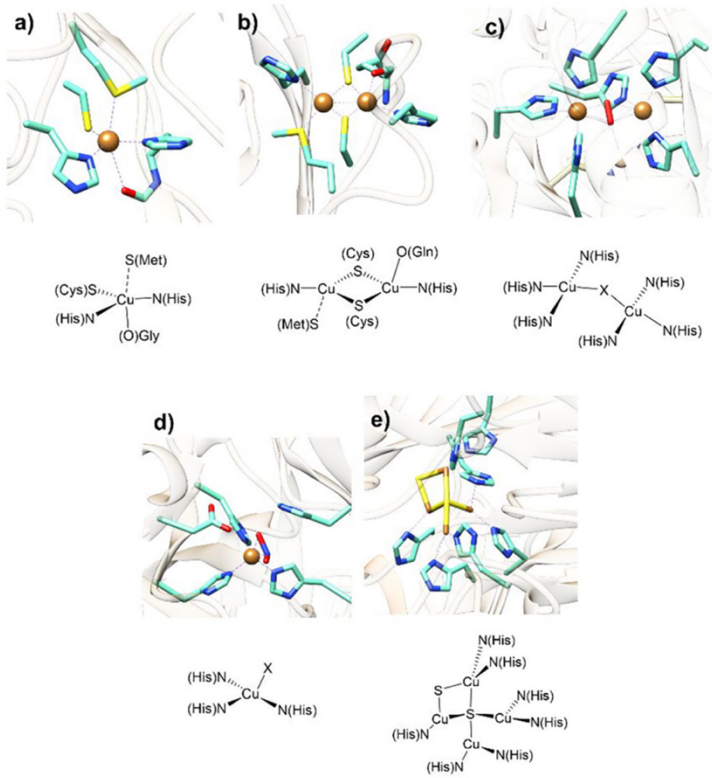


Figure 1.1. Structures of active sites from a variety of typical naturally occurring copper-binding proteins. a) T1Cu of azurin from *Pseudomonas aeruginosa* (PDB ID: 4azu). b) Cu_A site of the cupredoxin domain from *Thermus thermophilus* cytochrome *ba3* oxidase (PDB ID: 2cua). c) Substrate-bound (O_2) T3Cu site of laccase from *Melanocarpus albomyces* (PDB ID: 1gw0). d) Substrate-bound (nitrite) T2Cu site of *Alcaligenes xylosoxidans* nitrite reductase (PDB ID: 2xwz). e) Cu_z site of the *Pseudomonas stutzeri* N_2OR (PDB ID: 3sbr). Atom coloring: N (blue); C (cyan); O (red); S (yellow); Cu (copper). Adapted with permission from reference.[5] Copyright 2022 American Chemical Society.

1. Introduction

Cu proteins may act as electron transfer proteins using either type 1 Cu (T1Cu) center (Figure 1.1a) or copper A (Cu_A) center (Figure 1.1b). T1Cu proteins are also known as blue copper proteins because of their intense blue color due to a strong ligand-to-metal charge transfer (LMCT) absorption band around 600 nm. In all these proteins, a central Cu^{2+} is coordinated by three equatorial residues (two histidine (His) and one cysteine (Cys)), with a trigonal geometry (Figure 1.1).[6] Additional weakly bound axial ligands are observed. In plastocyanin and stellacyanin, the fourth ligands are a methionine (Met) sulfur atom and a glutamine (Gln) side chain amide oxygen atom, respectively. In azurin, the fourth ligand is a glycine backbone carbonyl oxygen, which coordinates Cu^{2+}. A Met residue also exists in the Cu^{2+} coordination environment in azurin, but at a longer distance.

The involvement of T1Cu proteins in electron transfer processes entails reducing Cu^{2+} to Cu^{+}, which prefers the tetrahedral coordination (when four-coordinate). From a functional point of view, the trigonal coordination observed for the Cu^{2+} ion represents a compromise between the stereochemical and electronic needs of Cu^{+} and Cu^{2+}. T1Cu proteins are a unique example where the protein structure determines the coordination geometry of the metal ion, forcing it toward unusual geometries. This entatic (or strained) state enhances the proteins' reactivity.[7]

In contrast, the Cu_A center is more complex,[8,9] consisting of a spatially near, electronically delocalized dicopper core that is connected by two Cys residues, one His, and one variable axial ligand on each side (Figure 1.1b). The geometry of Cu_A is more rigid compared to T1Cu, and its reorganization energy for the $Cu^{1.5+}$-$Cu^{1.5+}$/Cu^{2+}-Cu^{2+} redox pair is

comparable to that of T1Cu. Cu_A sites are often associated with metalloenzymes having catalytically active sites, such as heme-copper oxidase (HCO), nitrous oxide reductase (N_2OR), and non-heme nitric oxide reductase.

Other than electron transfer proteins, copper-proteins are also enzymes involved in oxidative chemistry, harboring Type 2 Copper (T2Cu), Type 3 Copper (T3Cu) and Cu_z centers.

In T2Cu proteins, copper is usually coordinated by nitrogen and oxygen-based ligands (mostly His and tyrosine (Tyr) residues) in a distorted square planar or square pyramidal geometry (Figure 1.1d) with an empty site allowing for the binding of small oxidants molecules such as molecular dioxygen or superoxide.[10] This copper center is commonly found in oxidoreductases such as amine oxidase, galactose oxidase, Cu-Zn superoxide dismutase, nitrite reductase, and lytic polysaccharide monooxygenases (LPMOs). Among these enzymes, LPMOs host a peculiar T2Cu center, and it will be discussed in more detail later. Beyond the first coordination sphere, the secondary coordination sphere and long-range interactions play an essential role in T2Cu enzymes, as they not only tune the copper center, but also stabilize and/or activate substrates via hydrophobic interactions or hydrogen bonds, or they directly contribute to the stabilization of highly reactive species in the catalytic mechanism.[5]

Type 3 Copper (T3Cu) proteins (Figure 1.1c) contain coupled binuclear sites and they are found in oxygen carrier proteins such as hemocyanin (Hc) and oxidase and oxygenase enzymes such as catechol oxidase (CO) and tyrosinase (TYR), respectively.[11] In T3Cu proteins, each copper ion is bound in a nearly planar trigonal geometry by three

His residues. T3Cu centers are EPR (Electron Paramagnetic Resonance) silent because of strong antiferromagnetic coupling between the two Cu^{2+} atoms with S = 1/2 caused by different bridging ligands.

The Cu_Z copper site (Figure 1.1e) is only found in N_2O reductase and is a tetranuclear copper sulfide cluster in which the copper ions are coordinated by seven His and a bridging inorganic sulfur atom.[12] The Cu ions form a distorted tetrahedron with one copper bonded to a single His and to an undetermined oxygen species.

1.2 Dioxygen activation in copper metalloenzymes

One of the fundamental processes in biological, synthetic, and industrial systems is the reductive activation of molecular dioxygen (O_2).

The ground state of O_2 is a spin triplet (S = 1). The spin selection rule prohibits its reaction with an organic substrate, which is almost universally assumed to be in a spin singlet ground state, to produce spin singlet products. In order to enable the triplet-state O_2 to react with a singlet-state substrate and provide singlet-state products, oxidoreductase manage the change in the spin state (intersystem crossing), thanks to cofactors, capable of O_2 activation.[13,14] Among such cofactors, transition metals (and particularly copper) are prevalent because spin-orbit coupling at the metal ion considerably enhances the intersystem crossing between distinct spin states, allowing for the simple and quick activation of O_2. In addition to their ability to allow intersystem crossing, first-row transition metals can also be involved in one or more electron redox processes involving the substrate and/or O_2 as redox partners.

Copper oxidoreductase enzymes catalyze the oxidation of substrates by O_2 (or its reduced forms, e.g., superoxide). These enzymes use the

oxidizing power of O_2 in order to accept electrons and protons as part of a process that is frequently coupled to the incorporation of either one or both of the O atoms from O_2 into an oxidized product.

These enzymes are involved in several key biological processes, such as electron-transfer, dioxygen binding and activation, pigmentation, methane monooxygenation, and denitrification. They are able to generate a wide spectrum of O_2-derived species (Figure 1.2). The formation of these so-called reactive oxygen species (ROS) is the mechanism underlying the enzyme promoted oxidative reaction.

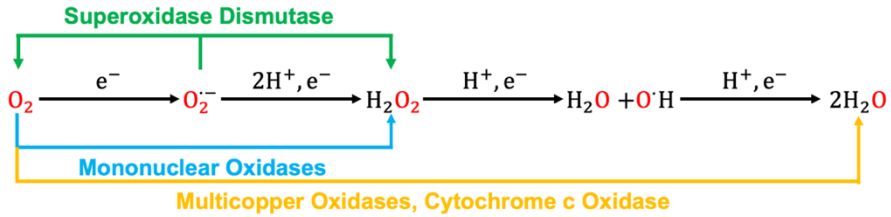


Figure 1.2. Aqueous reduction of O_2 and pertinent intermediates. Copper-containing enzymes that perform parts of this reduction are shown.

Given the limited set of ligands in the active site of these enzymes and the limited number of O_2-derived species, the environment beyond the first coordination sphere must play a key part in conferring and adjusting their remarkable catalytic features, allowing active sites with otherwise comparable primary coordination sphere to accomplish a broad array of biological tasks.

1.3 Natural Copper Oxidoreductase

Among natural copper oxidoreductase, LPMOs and polyphenol oxidases (PPOs) are of considerable importance for their unique network of interactions beyond the first coordination sphere, which finely tunes

their activity. LPMOs and PPOs contain a peculiar type 2 copper center and a type 3 copper site, respectively.

Often these sites are found within the same protein matrix, making their isolation and consequent structural and functional investigation quite difficult. These proteins include the multicopper oxidase (MCO) proteins and the particulate methane monooxygenase (pMMO) proteins. These classes of copper oxidoreductases will be discussed in the following paragraphs, as they are of particular interest for the purpose of this Ph.D. thesis.

1.3.1 Type 2 Copper Proteins: the elusive nature of the histidine brace copper-binding site

A class of recently discovered copper oxidoreductases, LPMOs, has attracted growing attention thanks to their ability to degrade woody biomass, a crucial preliminary step in the production of biofuels.[15] LPMOs are a superfamily of copper-dependent enzymes, whose putative homologs are found in all domains of life except for archaea, responsible for oxidative cleavage of glycosidic bonds in recalcitrant polysaccharides, such as cellulose, hemicellulose, chitin, amylose, amylopectin. The first LPMO was discovered in 2010 as CBP21 (Carbohydrate Binding Protein 21) in *Serratia marcescens* and was identified as a protein boosting the degradation of chitin and promoting the activity of chitinases.[16] Since then, numerous other members of the LPMO family have been identified primarily in fungi and bacteria. Currently, the most widely used classification is the one adopted by the CAZy (Carbohydrate Active Enzymes) database, in which LPMOs are divided into seven AA (Auxiliary Activity) families based on phylogenetic relationships and

substrate specificity: AA9, AA10, AA11, AA13, AA14, AA15, AA16 and AA17.[17] Five of these families are only found in fungi (AA9, AA11, AA13, AA14 and AA16), AA10 are found mainly in bacteria while AA15 and AA17 have been identified in eukaryotes, and other organisms. Interestingly, some putative chitin-active LPMOs belonging to family AA10 have been discovered in viruses, where they seem to facilitate insect infection, by degrading the chitin-rich exoskeleton of these animals;[18] recently, the AA17 family was discovered to be involved in plant infection.[19] LPMOs belonging to different families have specificity towards preferred polysaccharide substrates and can perform oxidations with different regioselectivity, for example type I and type II AA9 oxidize only C-1 and C-4 in cellulose, respectively, while type III can oxidize both positions.

All LPMOs share a common structure formed by a core β-sandwich (Figure 1.4a), like immunoglobulins, that contains the two copper-binding His, one of which is the N-terminal residue, and a Tyr in axial position (figure 1.4b). In addition to the copper-binding catalytic domain, various carbohydrate or polysaccharide binding modules (CBM, WSC and GbpA) can be present, appended to the C-terminus, which are involved in binding and specificity towards the substrate.[20–22]

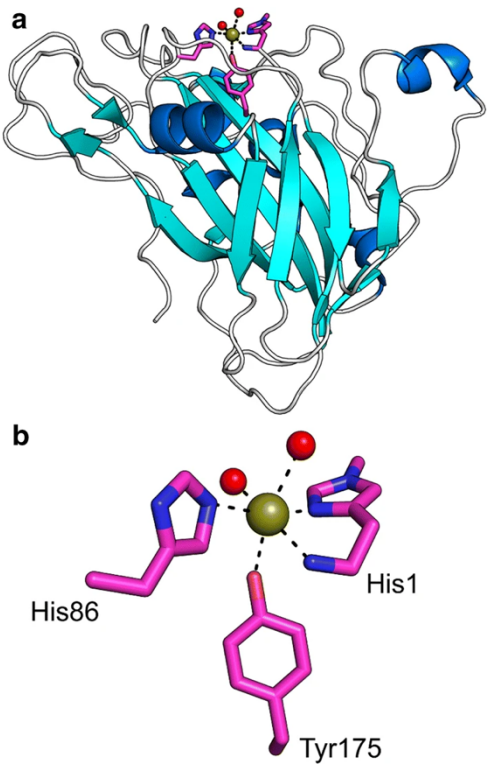


Figure 1.3. X-ray structure of a typical LPMO and its active site. a) The crystal structure and b) zoom of the active site of AA9 LPMO from the fungus *Thermoascus aurantiacus* (PDB ID: 2yet) The crystal structure is displayed in cartoon representation. The active site residues are shown as sticks with pink colored carbon atoms. The copper atom is shown as a golden sphere and water molecules coordinated by the copper atom are shown as red colored spheres. Adapted with permission from reference.[23] Copyright 2019, The Authors.

The active site present in LPMOs, formally a T2Cu center, is unique among copper proteins and it is the hearth of the displayed activity. Two His residues (a His terminal and a His internal) coordinate the copper through three nitrogen atoms (two from imidazole side chains and one from the amino terminus) creating a T-shaped geometry and forming a metal site named by Walton P.H. *Histidine Brace* (HB).[24,25] Due to their high susceptibility to photoreduction, several crystal structures feature a combination of Cu^{+} and Cu^{2+}.[26,27] The 3-coordinate geometry of the HB

appears to be typical for the enzyme in its reduced form, while the Cu^{2+} species tends to attract more exogenous ligands (typically water or chloride ions).[28,29] The aromatic residue below the copper, phenylalanine (Phe) or Tyr (Tyr175 in Figure 1.3), is highly conserved among the LPMO families. The precise function of these residues is unknown, although it is supposed to play a role in preserving the active site T-shaped geometry.[30] Further, such aromatic residue may represent a component of a charge-transfer process, known as "hole-hopping",[31] comprising a pathway of conserved Tyr and tryptophan (Trp) residues, that can prevent inactivation of the protein active site during uncoupled turnover.[32]

Fungal LPMOs exhibit a post-translational modification, where the Nε of His1 is methylated. The reason for this modification was unclear until a few weeks ago. In fact, LPMOs function without such modification, and theoretical calculations showed that it does not affect the oxidative mechanism.[33] Only recently, a preprint article appeared online, where the authors demonstrated that the methylation on His1 is necessary to stabilize a putative histidyl radical, involved in an oxidation repair mechanism.[34]

The HB active site, however simple in its coordination, still has unsolved questions. The different activity shown by the various LPMOs is certainly dictated by other interactions, involving the second coordination sphere.

The AA10s is unusual in showing a conserved alanine (Ala) residue occupying a position above the HB in the second coordination sphere. The role of this residue has been attributed to blocking ligation in the axial

position, thus orientating the binding and reactivity of dioxygen or hydrogen peroxide in the empty equatorial position of Cu^{+}.[26]

The LPMO crystal structures have revealed that secondary sphere residues establish H-bonding networks with buffer and solvent molecules in the active site.[24,35,36] In the *Mt*PMO3*, a glutamine residue in the second coordination sphere (Gln167) is important for the stabilization through H-bond of high oxidation state copper-oxygen species during the catalytic cycle and it increases the ligand donor strength of Tyr169 to the copper via a hydrogen-bonding interaction (Figure 1.4). Further, a His161 is involved in a proton transfer process,[37] and Thr74 forms a H-bond with the N-terminal amino group of His1. This last interaction makes the N-terminal amino ligand more nucleophilic and could stabilize highly reactive copper-oxygen species.[38]

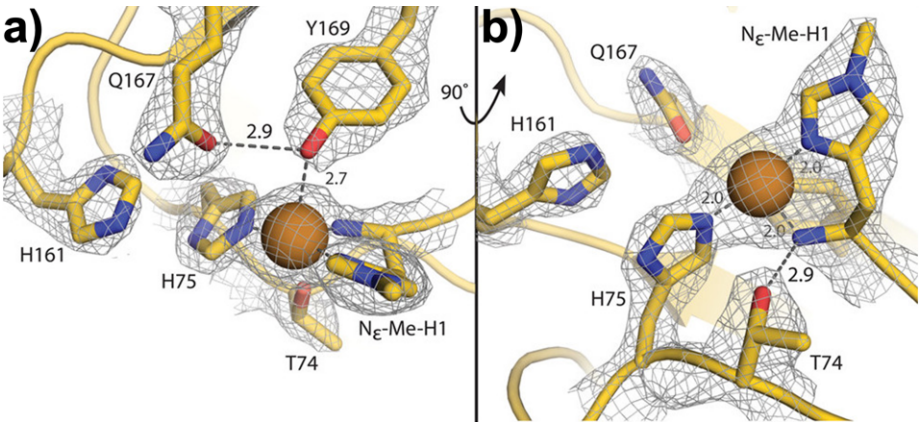
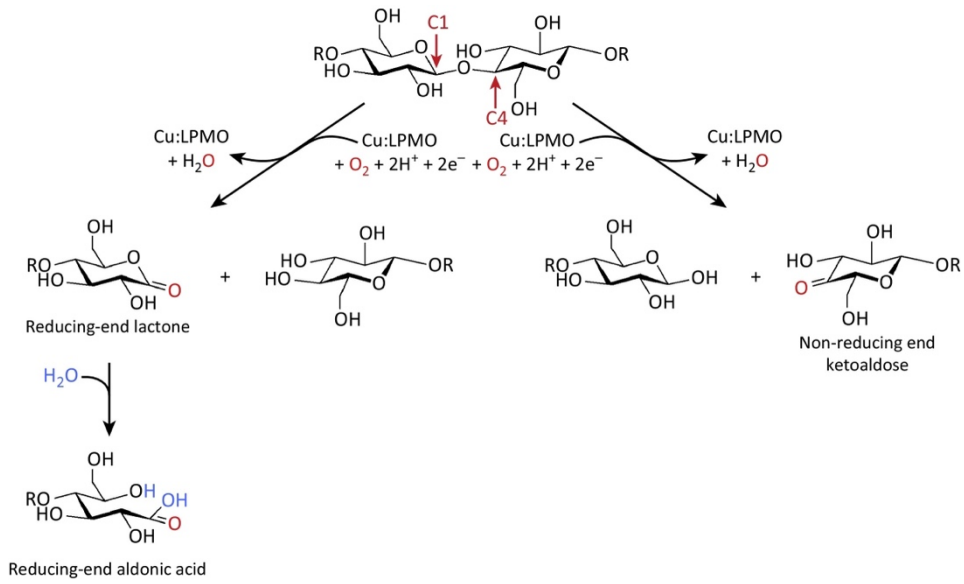


Figure 1.4. *Mt*PMO3* active site displaying primary (N_{ε}-Me-H1, H75, Y169) and secondary (T74, H161, Q167) sphere residues with mesh electron density map contoured to 1.5 σ. H161 and Q167 are positioned to H-bond with ligands in the solvent-facing equatorial position. Q167 H-bonds with the axial Tyr ligand. T74 H-bonds with the N-terminal amino group of H1 and forms a bond with an axial solvent ligand when present. Hydrogen-bonding distances are shown in Ångstroms. Adapted with permission from reference.[38] Copyright 2017 American Chemical Society.

Recently, Vu and coworkers inferred modulation of the redox state of copper and stabilization of copper-oxygen species by the distal hydrophobic loop in AA13.[39]

The copper atom is present as Cu^{2+} in the resting state, and the enzyme is activated by reduction to Cu^{+}; however, several aspects of the catalytic cycle are still unclear. It is known that LPMOs activate O_2 or H_2O_2 to break the glycosidic bonds of polysaccharides by hydroxylation of either the C1-H or C4-H bonds, followed by the removal of the scissile glycosidic link and production of aldonic acids or 4-keto sugars at oxidized chain ends, respectively (Scheme 1).

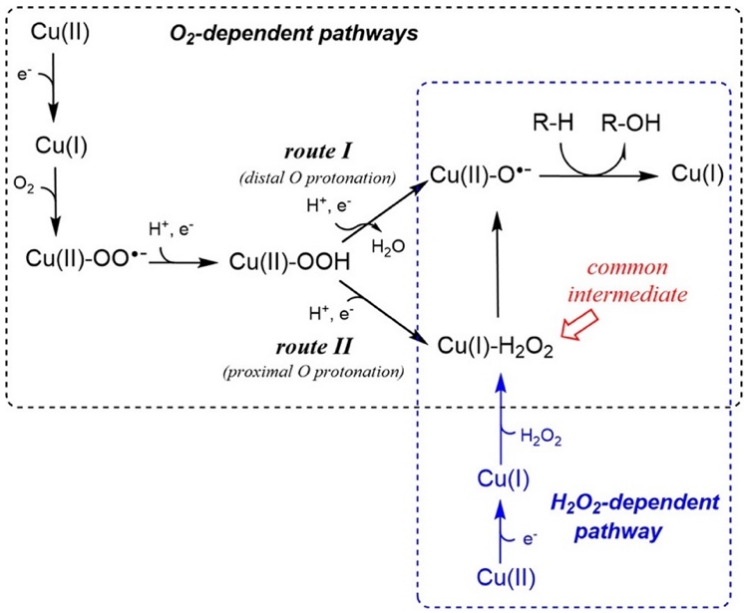
Scheme 1. General reaction for O_2-dependent activity of LPMOs toward oxidation of a polysaccharide substrate at C1 and C4. The formation of a lactone happens as a result of oxidation at C1, and it then becomes hydrated to turn into a reducing-end aldonic acid. On the other hand, oxidation at C4 leads to the formation of a ketoaldose at the non-reducing end. Adapted with permission from reference.[40] Copyright 2015 Elsevier Ltd. All rights reserved.



The nature of the co-substrate (O_2 or H_2O_2) used by LPMOs is still quite controversial. Based on the higher rates of substrate oxidation by H_2O_2 as compared to the identical reaction carried out using O_2 as the oxidant, several authors support the idea that H_2O_2 is the main LPMO co-substrate.[41–46] The H_2O_2 activation is performed similarly to heme peroxidases[47] or to the mononuclear nonheme iron enzyme HppE.[48] However, the LPMO pathway seems to be different from these canonical peroxygenases because they need an external reducing agent to get activated. Other authors reported that O_2 is the main LPMO co-substrate because they found a similar rate of oligosaccharide release, when fungal secretomes were used to break down polysaccharides with both H_2O_2 and O_2.[49] Paradisi et al.[32] and Bissaro et al.[41] showed that H_2O_2 can cause a significant amount of damage to proteins. In some cases, the production of the oxidized oligosaccharide is higher over time when the reaction is fueled by O_2 instead of H_2O_2, even though the initial rate is lower. Ambiguities regarding multiple aspects of LPMO functioning also extend to the mechanistic details of the catalytic cycle. A particularly debated point is the identity of the key intermediates, especially the reactive intermediate responsible for hydrogen atom abstraction (HAA) from the substrate, a difficult step that requires C-H bond cleavage. Initially, the intermediate involved in HAA was thought to be a Cu^{2+}-O_2^{-}, but subsequently DFT computations seem to suggest that this species is converted to the hydroperoxo-species Cu^{2+}-OOH, which is itself incapable of HAA.[50] QM/MM calculations showed that the O_2 and H_2O_2 oxidative pathways are linked and share a common intermediate Cu^{+}-H_2O_2 (Scheme 2). This reactive intermediate produces a copper-oxyl (Cu^{2+}-$O^{\cdot}$), which reacts with the substrate to oxidize it.[51,52] Most of the recent

computational work done on LPMOs also points toward a copper-oxyl as key reactive intermediate among both catalytic cycles that rely on O_2 and H_2O_2; however, the details of possible electron transfer mechanisms are still being understood.[53,54] There is also evidence supporting a possible mechanism of oxygen activation coupled to substrate binding.[55] In the absence of the substrate, the O_2-pathway produces the Cu^{+}-H_2O_2 intermediate. This intermediate breaks apart to give the free enzyme and H_2O_2 in solution, which explains why H_2O_2 is produced when the substrate is absent.[56,57]

Scheme 2. Different possible pathways for LPMO catalytic mechanism depending on cosubstrate identity and protonation events. Adapted with permission from reference.[40] Copyright 2020 American Chemical Society.



As mentioned, enzyme activation requires the reduction of Cu^{2+} to Cu^{+} and many possible electron donors could be involved in a biological context. LPMOs have reduction potentials that range between 155 and 370 mV.[26,30,36,58–60] This relatively low potential could explain why

different chemical, biological, and light-based reductants can reduce the LPMO active site.

In vitro studies usually employ reducing agents such as ascorbic acid (Asc), Cys and gallic acid (GA), however, these are unlikely candidates for *in vivo* reduction.[40,52] Studies performed on some fungal LPMOs seem to suggest cellobiose dehydrogenase (CDH) as a biologically relevant electron-donor, however not all fungi expressing LPMO also possess genes encoding for CDH.[61] A protein that could be involved is pyrroloquinoline-quinone (PQQ)-dependent pyranose dehydrogenases,[62,63] another sugar-oxidizing metalloenzyme, but in both cases the detailed electron transfer mechanism to LPMOs has yet to be determined. Moreover, light-based natural reductants seem to be involved in electron transfer reactions.[64–66] From this brief introduction on LPMOs, it is clear that many aspects remain to be clarified, such as the role of the first coordination sphere residues, identity of co-substrate and biological reducing agent, exact catalytic mechanism, function of His methylation, outer sphere interactions and structural basis for substrate selectivity.

1.3.2 Type 3 Copper Proteins: altering the enzymatic activity by tuning outer shell interactions

T3Cu copper proteins contain a binuclear copper site, which binds and, eventually, activates molecular oxygen. The main components of this class are Hcs,[67] COs[68] and TYRs.[69] COs and TYRs, together with Aurone Synthase (AUS),[70] are classified as PPOs.[71] In the active site, each copper ion is bound in a nearly planar trigonal geometry by three His residues, provided by an antiparallel α-helix pair with the

consequent formation of a not common crossed four-helix bundle (Figure 1.5).[72–74]

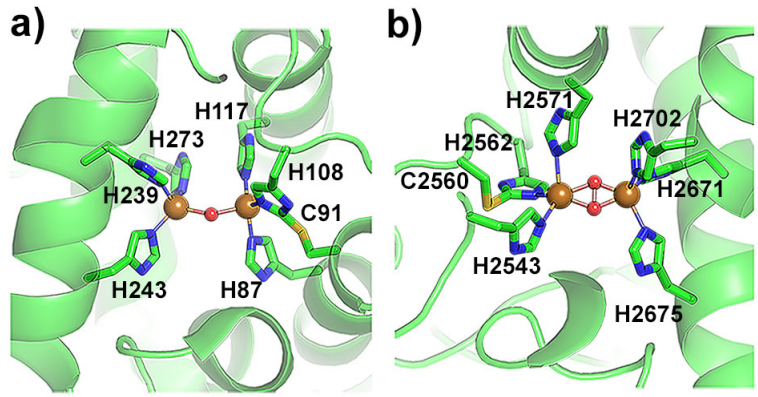
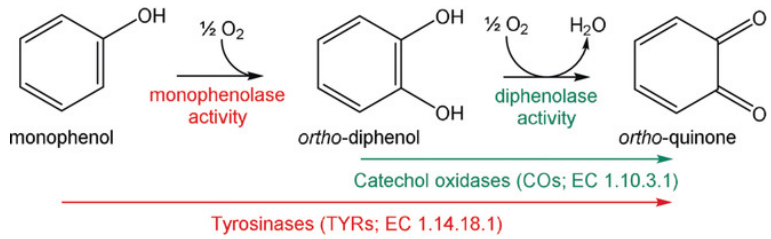


Figure 1.5. Active site structures of a) *met*-TYR from walnut leaves (Juglans regia, PDB ID: 5ce9)[69] with the two copper ions bridged by a hydroxide anion; b) *oxy*-Hc from Octopus (PDB ID 1js8)[75] with the peroxide bridging the two copper ions in a μη2 : η2 coordination mode. The copper ions are depicted as brown spheres, the coordinating His residues as sticks and the bridging ligands as red spheres. The protein is shown as green cartoon. The thioether cross links between His108 and Cys91, and His2562 and Cys2560, are also shown. Adapted with permission from reference.[76] Copyright 2022 Wiley-VCH GmbH.

Hcs are large multimeric proteins freely dissolved in the hemolymph of many invertebrates, such as arthropods and mollusks.[77] While Hc reversibly binds and transports the molecular oxygen, TYR and CO perform monophenolase/diphenolase and diphenolase activity, respectively (Scheme 3).

Scheme 3. Reaction catalyzed by CO and TYR. Adapted with permission from reference.[78] Copyright 2019 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim.



1. Introduction

The metal cofactor can be found in different oxidation states, which have been extensively characterized both in structure[73,79] and spectroscopic features,[80–86] and they are involved in the dioxygen binding/activation. The dicupreous deoxy form consists of two Cu^{+} ions with no bridging ligands, the dicupric met form has two Cu^{2+} bridged by one or two hydroxylates, whereas the oxy form is bridged by the characteristic side-on (μ-η^2:η^2) peroxo anion. The Cu···Cu distance ranges between 3.5 – 4.6 Å[75,69] and in both met and oxy form copper ions are antiferromagnetically coupled (S = 0) via the superexchange pathway offered by the bridging ligands. Despite being EPR silent, this peroxo core has been characterized by several techniques. Indeed, this core gives rise to the peculiar spectroscopic features of T3Cu proteins:

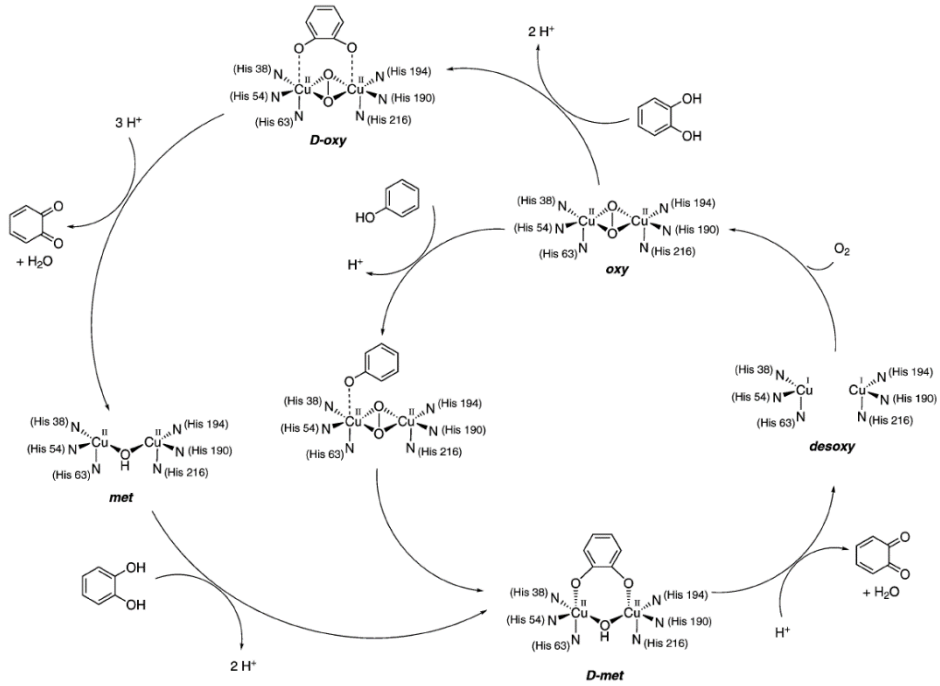
- The absorption electronic spectra display LMCT transitions at about 340 nm ($\varepsilon \sim 20$ $mM^{-1}cm^{-1}$) and around 600 nm ($\varepsilon \sim 1$ $mM^{-1}cm^{-1}$)[81,87] even though some variability on these values are observed among different family of T3Cu proteins;[84,88,89]
- The Resonance Raman spectra, upon irradiation of the visible band, result in the enhancement of a sharp band around 750 cm^{-1} assignable to the symmetric stretching of the O-O bond;[90–92]
- EXAFS analysis highlights the presence of a rather unique outer shell feature at ~3.7 Å corresponding to the Cu···Cu interaction with a low Debye–Waller factor.[93]

In HCs, the peroxo core serves as dioxygen transporter, and these proteins control the deoxy/oxy form equilibrium by subtle allosteric regulations[94,95] that lead to controlled dioxygen-release, similarly to hemoglobins.[96] The side-on Cu_2O_2 core is also able to perform very demanding hydroxylations as in TYRs, by activating aromatic C-H

bonds, or just to oxidize aromatic substrates as in catechol oxidases, without the release of appreciable hydrogen peroxide as a byproduct.

The mechanism proposed for the monophenolase and diphenolase activity of TYR is shown in Scheme 4. Both monophenolase and diphenolase activity (the latter being valid also for CO), with the key intermediates, are shown. Native TYR occurs mainly in the met-form,[97] in which a hydroxyl ion is bound to the copper ions, and is able to bind and to oxidize only catechols, with the concomitant reduction of copper ions and the formation of the deoxy state. The *deoxy*-TYR binds the molecular oxygen to generate the $Cu_2(\mu\text{-}\eta^2\text{:}\eta^2)$ peroxo core and, therefore, the *oxy*-form. The *oxy*-TYR can bind both monophenols and catechols: if monophenol is bound, it is first converted to o-diphenol and, subsequently, to the corresponding quinone; otherwise, the *oxy*-TYR oxidizes two equivalents of catechol to the o-quinone; in both cases, the deoxy form is reformed.

Scheme 4. Monophenolase and diphenolase catalytic cycle of TYR. Adapted with permission from reference.[98] Copyright 2011 The Royal Society of Chemistry.



Even though Hc, CO and TYR host the same cofactor, the tuning of their activity is a shiny example of how nature has evolved these enzymes to perform different activities. The main differences rely on the residues beyond the first coordination sphere.

It is generally accepted that the inaccessibility of the T3Cu center to potential substrates is one of the primary causes of the lack of enzymatic activity in Hc.[74]

A more intriguing scenario concerns the difference in activity between CO and TYR. Initially, a bulky amino acid that limits substrate access to the active site was identified in the pioneering crystal structure of the CO from *Ipomoea batatas*[99]. This residue was referred to as a "blocker residue" since it was thought to be essential for regulating mono-

diphenolase selectivity.[89,100–103] However, the first crystal structure of a plant TYR (from *Juglans regia*)[69] reveled the presence of the blocker residue in a similar position observed in plant COs (*ib*CO and *vv*CO), bringing the idea that the differentiation between mono- and diphenolase activity is independent on the degree of restriction at the active site. This suggested that other factors, such as second shell interactions, play an important role in enzymatic specificity. Rompel and coworkers[69] demonstrated, through the X-ray structures of *jr*TYR and of a bacterial TYR (from *Bacillus megaterium*) with Tyr bound to the active site, the importance of residues lining the active site for the tuning of enzymatic activity. Analysis of the X-ray structure of *jr*TYR revealed a hydrophobic leucine (Leu) residue lining the substrate access (Figure 1.6a), whereas plant COs possesses, at the same position, positively charged arginine or lysine residues, both capable of stabilizing acidic substrate functionalities. Kinetic studies demonstrated faster turnover for monophenols lacking a carboxylate group. This was attributed to the Leu244 residue, which stabilizes hydrophobic moieties. In contrast, the *bm*TYR structure has an arginine residue at the corresponding position, able to make hydrogen bonds with substrates that contain a carboxylic group, as revealed in the Tyr- and L-DOPA-bound structures.[103] Furthermore, it was shown that the bulky residue Phe260, in conjunction with His243, forms a gate leading into the active site through which the substrate can only pass in the proper orientation.

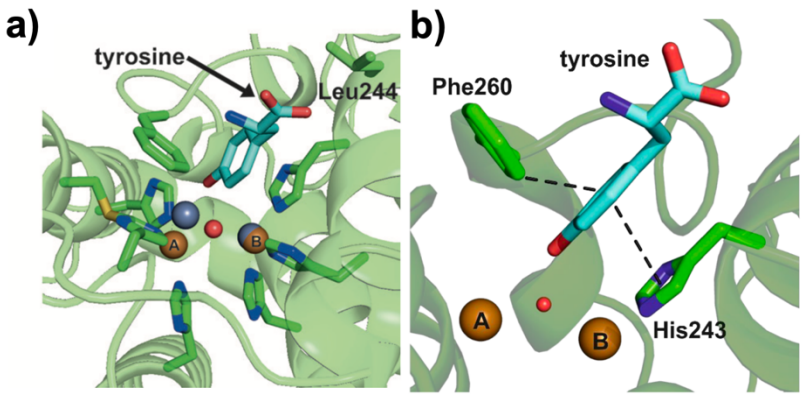
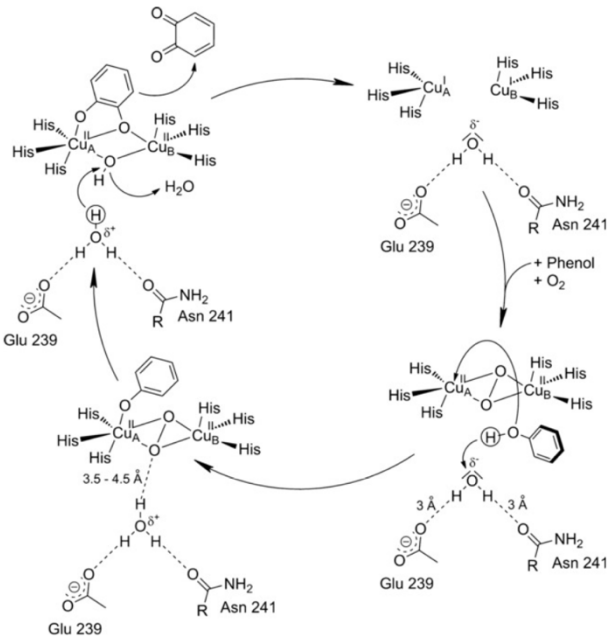


Figure 1.6. a) Superimposition of Tyr from the *bm*TYR + Tyr structure (PDB ID: 4p6r) with *jr*TYR. The copper ions from *jr*TYR are depicted as brown spheres and the superimposed zinc ions from *bm*TYR + Tyr are depicted as silver spheres. Leu244 is located within the second shell of the active site, and it is in close proximity to the carboxylic tail of Tyr. Substrates possessing a hydrophobic tail instead of a carboxylic moiety are able to interact (hydrophobically) with Leu244 and are thus stabilized (as indicated by the kinetic data in the ref. [69]). b) Gate formed by His243 and Phe260. Adapted with permission from reference.[69] Copyright 2015 The Authors. Published by Wiley-VCH Verlag GmbH & Co. KGaA.

Furthermore, substrate deprotonation is a required step for monophenolase reaction. In TYR, a glutamate and asparagine (Asn) residues bind a conserved water molecule (Scheme 5).[104] As a result, the conserved water molecule becomes basic enough to extract the proton from the phenolic substrates upon coordination to the active site. In CO, only the glutamate is located at the binding site, and it is unable to lower the pKa of the water molecule. In this case only diphenolase activity is possible, which does not depend on a deprotonation step.

Scheme 5. Deprotonation of a phenolic substrate in the TYR cycle. Adapted with permission from reference.[104] Copyright 2016 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim.



It is evident from these findings that Nature, during evolution, altered outer sphere interactions to modulate substrate access and tune enzymatic activity. Therefore, it is hypothesized that the electrostatic environment, conformation, and nature of second-shell residues at the entrance of the mono-/diphenolase active site might be the determinant of mono-/diphenolase activity. However, it is not completely clear which molecular constraints specify the T3 copper core toward mono- and/or diphenolase activity and a specific selectivity.

1.3.3 Multicopper Oxidases and Particulate Methane Monooxygenases: different copper centers inside the same protein

The MCOs are a family of enzymes that combines the oxidation of four substrate equivalents by one electron with the reduction of oxygen to water by four electrons.[105] At least four Cu atoms, split across T1Cu, T2Cu, and T3Cu centers (Figure 1.7), make up the catalytic motif in these proteins. The trinuclear center (TNC) responsible for oxygen reduction is formed by the combination of T2Cu T3Cu sites.[106]

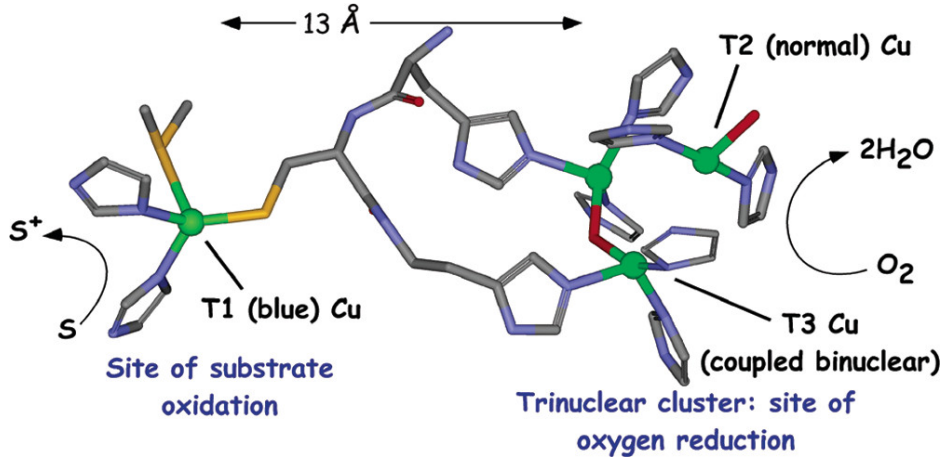


Figure 1.7. The catalytic core of MCOs. Green spheres represent copper ions. Figure generated from the crystal structure of ascorbate oxidase (PDB ID: 1aoz). Adapted with permission from reference.[107] Copyright 2007 American Chemical Society.

MCOs are found in a variety of living organisms. They are involved in the formation of lignin in plants (plant laccases),[108] pigment formation, lignin degradation,[109] detoxification processes in fungi (fungal laccases),[110] iron metabolism in yeast (Fet3p)[111] and mammals (hCp and hephaestin), copper homeostasis in bacteria (CueO),[112] and manganese oxidation by bacterial spores.[113]

In the reaction mechanism of MCOs, the main electron transfer (ET) steps are the (i) substrate reducing the T1Cu site, (ii) ET from the T1Cu site to the trinuclear cluster, and (iii) the trinuclear cluster reducing O_2.[114] Thus, the T1Cu site's main task is to transfer electrons from the substrate to the trinuclear Cu cluster site, which is about 13 Å away. The coordination sphere of the T1Cu site always has two His ligands and one Cys, providing MCOs their typical blue color. Interestingly, although having essentially similar copper active sites for catalysis, MCOs have a wide range of substrates and catalytic rates. Moreover, it is still unclear the specific function played by each of the three copper ions of the trinuclear cluster.

Given the coordinative promiscuity of copper in these enzymes, studies to explain the differences in activity have been conducted at individual copper sites.

The investigations have mostly focused on type 2 depleted (T2D) form of laccase,[115] the influence of second coordination sphere residues near the trinuclear cluster,[116,117] and computational analyses of the dioxygen reduction mechanism. Like other binuclear T3Cu enzymes, T2D laccase only retains a single binuclear T3Cu site at the trinuclear cluster. Although it has been demonstrated that the T3Cu site in Hc can reversibly bind di-oxygen in the form of peroxide,[118] the reduced site in laccase does not interact with di-oxygen.[115] Studies have also focused on the role of the D94 (located on the edge of the trinuclear cluster) and E487 residues (located near the T3Cu site) in the second coordination sphere. It has been shown that D94 is necessary for the trinuclear cluster to react with O_2, and E487 donates a proton during the reductive cleavage of the O-O bond.

Solomon and coworkers performed an experimental evaluation of the unique functions of each copper center in O_2 reactivity by deliberately perturbing the trinuclear cluster in MCOs using mutagenesis.[119] The three copper ions in the cluster are coordinated by eight His residues. To produce an altered, but stable and intact trinuclear cluster, each His residue was mutated to Gln. Three of the eight mutants contain the entire trinuclear cluster and were isolated in their type 1 depleted (T1D) forms to remove the strong spectroscopical features of the T1Cu center. This represents a valuable study to isolate the different types of copper centers within MCOs, without spectral overlap between the 3 different copper centers.

The catalytic promiscuity of MCOs is also given by the modulation of the reduction potential of the T1Cu site. All MCOs require effective electron transport from T1Cu to the active site, which is tightly coupled to the T1Cu redox potential. A fine-tuning of the redox potential of the T1Cu center can alter the catalytic activity of MCOs.[120,121]

In *Bacillus subtilis* laccase, changing the axial T1Cu ligand from Met to Leu or Phe resulted in a 100 mV increase in reduction potential, which lowered laccase activity by 24-fold in the Leu mutant and nearly completely in the Phe mutant.[122] Changing the native axial Phe of *Thapsia villosa* laccase to Met decreased the T1Cu potential by 100 mV.[123] This mutation raised k_{cat} and K_m while shifting the optimum pH value to a more basic value. Kamitaka and coworkers reported that the M467Q mutation in *Myrothecium verrucaria* bilirubin oxidase reduced the T1Cu formal potential by 230 mV while enhancing the kinetics of the electrocatalytic oxygen reduction process.[124]

Another class of proteins that host different copper active sites are the pMMOs. Soluble methane monooxygenase (sMMO), which requires iron, and pMMO, which requires copper, catalyze the conversion of methane to methanol.[125,126] While the structure and spectroscopy of the diiron metal active site in sMMO are well-established, the copper sites in pMMO are not completely characterized.

pMMO is an integral-membrane enzyme, whose purification is a challenging task, being a membrane protein. [127–129] It is composed of three subunits, PmoA, PmoB, and PmoC, arranged in a 300 kDa $\alpha_3\beta_3\gamma_3$ complex. Based on the crystal structures of various pMMO, it has been determined that two copper centers, one situated in PmoB (Cu_B) and one located in PmoC (Cu_C), have the potential to serve as active sites (Figure 1.8).

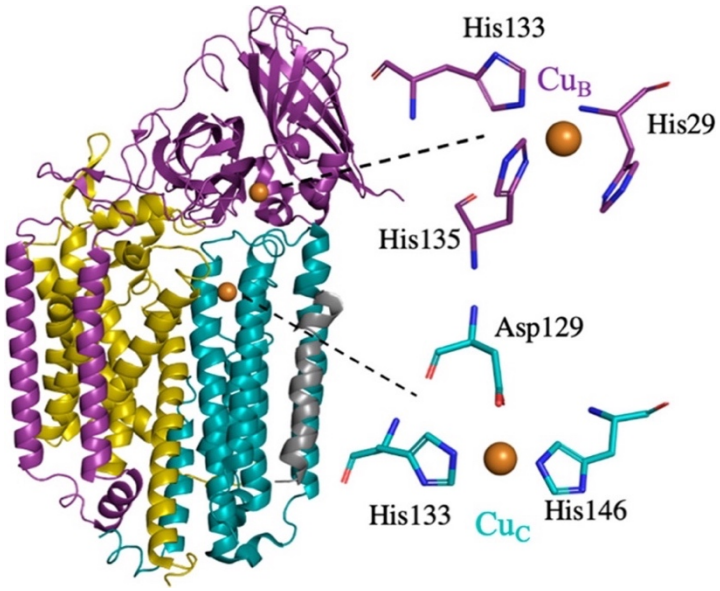


Figure 1.8. A single protomer from the crystal structure of *Mc.* sp. str. Rockwell pMMO (PDB ID: 4phz). PmoA is depicted in gold, PmoB in purple, PmoC in teal, and an unidentified helix is gray. The two copper ions are depicted as spheres and to the right are expanded to show the putative coordinating residues. Adapted with permission from reference.[130] Copyright 2021 American Chemical Society.

The Cu_B first coordination sphere is composed of three imidazole nitrogens from His29, His133 and His135 and one amino terminal nitrogen from His29.

Although in the past this site has been modeled as both mononuclear and binuclear,[131–134] it has now been confirmed to be a mononuclear site.[135,136] Cu_B nuclearity, coordination environment and redox state (2+) are also supported by Continuous Wave-Electron Paramagnetic Resonance (CW-EPR), ENDOR (Electron Nuclear Double Resonance) spectroscopy and XAS (X-ray Absorption Spectroscopy).[137,138] However, given the no-conserved His residues in all pMMOs,[139] the saturated equatorial coordination, and the oxidation state (2+) present even *in vivo*, it is not plausible that Cu_B is the binding site of O_2 and methane.

Crystallography has also shown that the enzyme has a Cu_C metal-binding site in the subunit PmoC whose ligand set, includes Asp129. His133 and His146, are the same in all types of methanotrophs.[139] The copper in Cu_C is present as Cu(I) *in vivo*[140] and in the presence of Asc as reductant,[137] making it reasonable to assume that it represents the active site for dioxygen activation and methane binding.

Over the year, bioinorganic chemists performed a thorough spectroscopic investigation of the Cu sites, in order to dissect the main features responsible for pMMO activity. The catalytic mechanism is still difficult to interpret, nevertheless, it is supposed that the species responsible for the oxidation of methane is copper-oxyl species, as hypothesized for LPMOs.[141]

1.4 Nesting copper active sites into de novo designed proteins

As discussed in the previous paragraphs, numerous natural copper oxygenases/oxidases have been structurally characterized. Consequently, most of our understanding of enzyme mechanisms is derived from these studies. Significant advances have been made in this respect. However, the different modes through which copper enzymes catalyze specific reactions with high efficiency remain still elusive. Dissecting the specific contribution of several elements, such as first and outer coordination sphere ligands, through a fine structural, spectroscopic, and functional characterization of multi-site metalloenzymes represent a complicated and challenging task.

The design of artificial metalloenzymes, made up of robust scaffolds, easy to produce in good yield, may help in understanding the catalytic properties of the parent natural enzymes.

Three different strategies are usually applied for developing artificial metalloenzymes: (i) providing a metal-binding site to a native and stable protein, lacking a metal center; (ii) changing an existing metal-binding site in a native metalloenzyme to generate a non-native reactivity; and (iii) building a metal-binding site into a de novo designed protein scaffold. This last strategy is used in this Ph.D. thesis, and it will be discussed in more detail.

1.4.1 Protein design of helical bundles

Proteins are versatile biomolecules adapted to perform countless functions indispensable to life in all its forms. The structure of a protein, encoded in its amino acid sequence, has been carefully sculpted by

evolution and is the basis for its ability to perform catalysis, signaling, transport, scaffolding or any other required function. Because the three-dimensional structure of a protein is completely specified by its sequence, gaining a sufficiently detailed understanding of the sequence/structure relationship would enable predicting the native structure, thereby solving what is known as the "protein folding problem".[142] Protein design[143] is the opposite endeavor, namely given a three-dimensional structure required for a desired function, to find the sequence encoding for it, and is therefore also called "inverse folding problem".[144] De novo design refers to the construction of new proteins from first principles, unrelated to any sequence found in nature, and encompasses both computational design methodologies and sequence-based approaches, like binary patterning.[145,146] The totality of all proteins of all organisms that exist and have existed is known as the "protein universe", and it represents only a very small fraction of all possible sequences, which is enormous even for relatively short polypeptides.[147] The still unexplored regions of the sequence space can be sampled through de novo design, to obtain the optimal sequences for entirely new structures. Since proteins adopt the structure that corresponds to the energy minimum, an accurate model to calculate the energy associated with each conformation is necessary for both design and structure prediction, together with a conformation sampling method. Possible strategies in de novo design have evolved during the last decades parallel to both technological advancements and to the understanding of the biophysical principles underlying protein folding. The first examples of de novo design, from the 1970s to 1980s, were done manually using model-building and physical models.[148,149] Later efforts, between the late 1980s and early 2000s, were

guided by physicochemical principles and were based on equations defining backbone geometry and sidechain repacking algorithms.[150,151] Starting from the 2000s, progress in bioinformatics and rapid increase in PDB entry number have allowed the rise of fragment-based and bioinformatically informed computational protein design. These advancements have led to the first computationally designed structure (α3D)[152] and fold (Top7),[153] in 1999 and 2003 respectively, demonstrating the feasibility of protein design from scratch. Almost half of all existing proteins bear metal-binding sites with either a structural or functional role, such as transport, catalysis, electron transfer and dioxygen binding. Metalloprotein design[154] is a natural extension of protein design methodologies and can be implemented by grafting a metal site onto an existing scaffold or by designing de novo a new structure satisfying precise spatial requirements for the coordinating residues (commonly His, Glu, Asp, Met and Cys). Second-shell hydrogen bonds can be introduced with first-shell residues to further increase the rigidity and binding affinity of the newly formed metal site. Among possible scaffolds for metalloprotein de novo design, coiled coil helical bundles play a prominent role. α-Helical bundles have been extensively studied and parametric equations can be used to design both right-handed and left-hended coiled coils.[155] The sequence requirements for a α-helical coiled coil can be stated in terms of heptad repeats "*abcdefg*",[156,157] as a consequence of the structural features of the α-helix and of the relative position of the helices forming the bundle (Figure 1.9). Positions *a* and *d* form the interior of the bundle and are occupied by hydrophobic residues, positions *b*, *c* and *f* point towards the exterior and are generally occupied by polar residues, while positions *e* and *g* are

located at the interface between subunits and are usually polar, but their identity depends on the context.[158] It was shown that it is possible to design self-assembling α-helical coiled coils from scratch simply by designing the sequence according to the aforementioned rules.[156] However, the stoichiometry and topology of the bundle can be highly dependent on specific packing and polar interaction between subunits.[159,160]

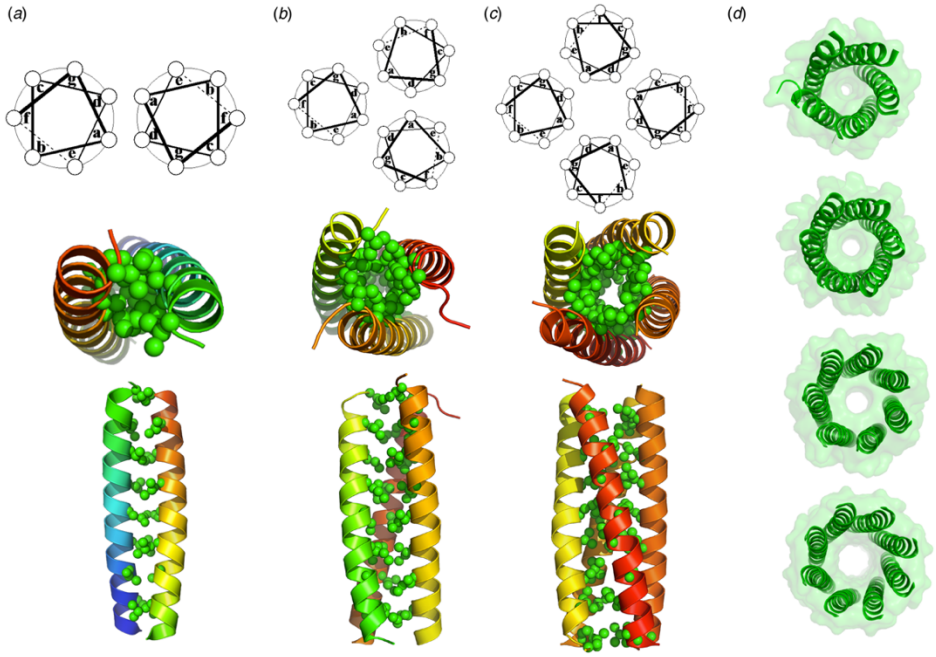


Figure 1.9. Crystal structures and helical wheel diagrams for: a) two-helix bundle protein GCN4 (PDB ID: 2zta); b) three-helix bundle (PDB ID: 1gcm); c) four-helix bundle (PDB ID: 1gcl); d) end on views of de novo designed penta-, hexa-, hepta-, and octameric bundles (PDB ID: 4pnd, 4h8o, 5ez8, 6g67). Adapted with permission from reference.[143] Copyright Cambridge University Press 2020.

The reliable parametrization and simple design rules available for α-helical coiled coils have made them popular scaffolds for de novo metalloprotein design. Pecoraro and coworkers, starting form coil-Ser peptide,[161] designed a homotrimeric three-helix bundle bearing a

trigonal 3-Cys site that binds Hg^{2+} ions with high affinity and thermodynamic stability.[162,163] By substituting the apolar *a* position in different points of the bundle, it was also possible to engineer an appropriate coordination environment for the binding of various metal ions such as Zn^{2+}, Cd^{2+}, Pb^{2+}, As^{3+} and Bi^{3+}.[164–174] The structural stability offered by the formation of a 3-Cys mercury site was later exploited to build a three-helix bundle bearing both the $(Cys)_3Hg^{2+}$ site and a catalytically active $(His)_3Zn^{2+}$ capable of CO_2 hydration.[175] The zinc-based active site that was designed has also been relocated into α3D (a 73-residue single-chain three-helical bundle anti-parallel protein discussed earlier) by changing three of the core Leu residues into His residues.[176,177] The same scaffold has been modified to nest a tetrathiolate environment able to bind iron[178] and, recently, molybdenum.[179] The trimeric coiled coils can be used not only for catalysis. Peacock and coworkers have used them successfully to make magnetic resonance imaging probes with good relaxivity[180,181] and Pecoraro, Ivancich and coworkers to bind heme cofactor.[182]

Four helix bundles are another topology that has been successfully adapted to the introduction of metal-binding sites. Lombardi, DeGrado and coworkers designed the DF (Due Ferri, in Italian) de novo family (Figure 1.10), a series of minimal proteins possessing a Glu_4His_2-diiron site within a four-helix bundle, by using parametric equations to generate the folded structure and algorithmic protocols for packing.[183] The first model, DF1, showed remarkable agreement between the designed and experimentally determined structure and the position of the coordinating residues was largely preorganized as demonstrated by the almost identical *apo* and *holo* form structures.[184] Whereas the

original DF1 was a C_2-symmetric homodimer, subsequent modifications led to related structures formed by a single chain (DFsc)[185] and four chains (DFtet).[186] Construction of the tetrameric bundle DFtet was the first example of a Monte Carlo algorithm that implemented both positive and negative design computations and enabled the creation of two-component A_2B_2 and three-component $A_AA_BB_2$ heterotetrameric bundles[187] that assembled with high selectivity and affinity to form the desired structure. In DF protein scaffolds, a range of catalytic and binding activities have been generated By altering the surrounding environment of the diiron site, it was feasible to construct DF analogs that catalyze the O_2-dependent oxidation of dihydro-quinones[188] and amino phenols[189] at comparable rates to the alternative oxidase enzyme. Additionally, by asymmetrically incorporating an extra His ligand (as well as additional second- and third-shell hydrogen bonding groups), the DF protein was modified to catalyze aniline hydroxylase, resembling a family of non-heme enzymes.[190–192] Moreover, the radical semiquinone anion, which is ordinarily unstable in aqueous solution, was stabilized using the DFsc scaffold.[193]

A tetranuclear Zn^{2+} cluster was also introduced into a DF protein by placing four bridging Asp and four terminal His inside the bundle, as well as replacing some apolar residues inside the bundle with polar ones to build a well-defined hydrogen bonded network around the metal site.[194,195]

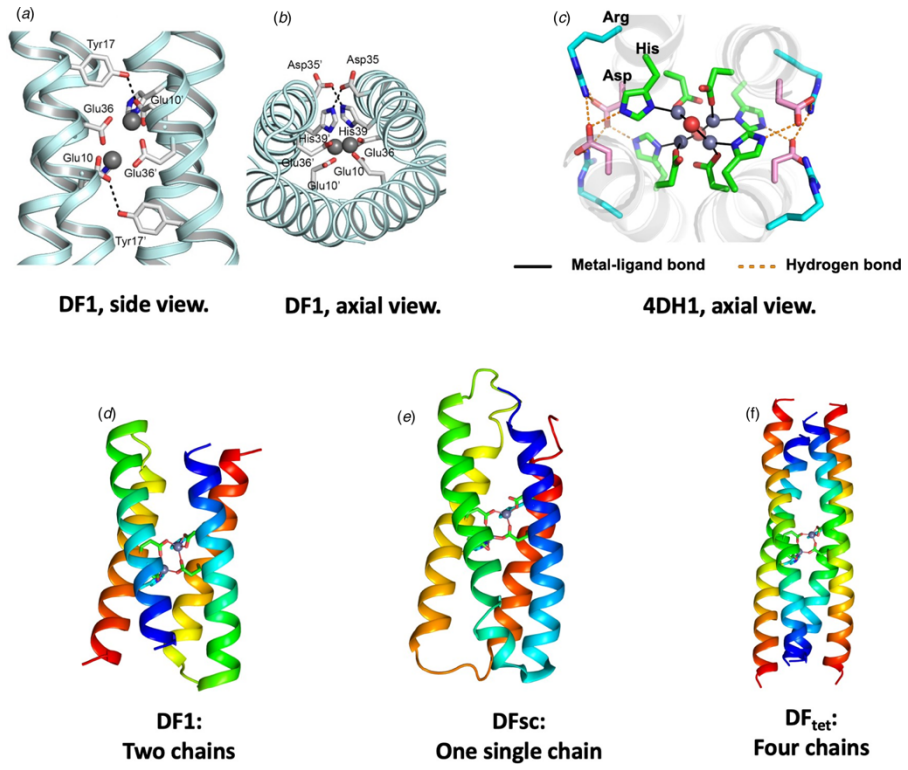


Figure 1.10. DF protein family. a) and b) DF1 protein metal-binding site and d) crystal structure (PDB ID: 1ec5); c) 4DH1 tetrazinc cluster site (PDB ID: 5wll); e) single chain and f) tetrameric variants created from DF1. Adapted with permission from reference.[143] Copyright Cambridge University Press 2020.

Over the years the four-helix bundles have been engineered to bind small molecules, such as metalloporphyrins,[196–198] or to perform natural functions such as allostery.[199]

1.4.2 De novo design of copper proteins

Copper proteins have been chosen as candidates for protein design to test our understanding of metalloprotein structure and function. However, it is worth mentioning that some first examples of de novo designed copper proteins are the copper-dependent electron transfer

protein. An early example of de novo copper proteins was reported by Tanaka and coworkers. They inserted a T1Cu and a Cu_A center within a native-like single-stranded four-helix bundle, called AM2C, to take benefit from the intrinsic asymmetry of the construct.[200–202] Later, Pecoraro and coworkers designed cupredoxins using rational design inside a three-helix bundle and were able to replicate all mononuclear cupredoxin classes within the same scaffold.[203–206]

More intriguing is the case of de novo designed T2Cu proteins. The first examples of a structural T2Cu center in a de novo designed three-stranded coiled coil containing a $(His)_3$ binding site was reported by Ghadiri and Case.[207] A 19-residues peptide sequence with a 2,2'-bipyridyl (bpy) moiety connected at the N terminus and a His residue at the C terminus was used to build the construct. The bpy group was treated with Ru^{2+} to produce a *fac*-$[Ru(bpyPep)_3]^{2-}$ unit that connects the N termini of three helices. This construct led to the formation of a completely folded homotrimeric three-helix bundle with a $(His)_3$ coordination environment available for copper-binding.

Other examples of structural T2Cu sites inside a de novo designed scaffold are offered by Tanaka and coworkers inside the peptide family of isoleucine zipper (IZ)[208] and inside a single chain four-helix bundle.[209]

Although these systems recapitulate the T2Cu center spectroscopic fingerprints, they were not studied for redox catalytic behavior (e.g., oxidation of organic substrate). In the design of catalytic active T2Cu proteins, the scaffold should: (i) be flexible to change the coordination environment according to changes in the oxidation state of the copper ion; (ii) enable the activation of O_2 (or its derivative) without ROS

formation and copper-binding deactivation; (iii) enable the access for the substrate to the metal site for the catalysis.

1.4.3 De novo design of catalytic copper proteins

As it was discussed earlier, copper oxidoreductases contain mainly mononuclear T2Cu or binuclear T3Cu (or both). Differently from mononuclear T2Cu sites, which have been widely constructed in a variety of de novo scaffolds, the T3 di-nuclear sites have been mainly studied in natural proteins and small inorganic complexes.[210,211] Very recently, Song and coworkers installed a di-copper site at the C_2-symmetric interface of the hexameric acetyltransferase, using the unnatural bipyridyl-alanine as a tight chelating residue.[212] The first example of a de novo designed T3Cu protein, published at beginning of 2023,[76] has been developed within this thesis, and will be discussed in the following chapter. The following paragraph describes some outstanding examples of catalytic active T2Cu de novo designed copper proteins, in order to highlight the progress made in the field.

As a model for the catalytic center of nitrite reductase (NiR), Pecoraro, Tegoni and coworkers reported the insertion of a catalytic T2Cu site within a de novo-designed three-helix bundle.[213] In algae, fungi and bacteria, the nitrite reductase (NiR) enzyme facilitates the one-electron reduction of nitrite (NO_2^-) to nitric oxide (NO). This protein hosts a T1Cu electron transfer site and a T2Cu center, where nitrite is reduced.[214] As discussed earlier, de novo metalloprotein design allows scientists to dissect a complicated system like Cu-NiR and analyze its cofactor individually. In the crystal structure of Hg^{2+}-Zn^{2+}-TRIL9CL23H (a de novo designed model for Carbonic Anhydrase),[175] the Zn^{2+}-$(His)_3$

binding site superimposes very well with the T2Cu center of native CuNiR. For this reason, to characterize Cu^{+} and Cu^{2+} binding, a simpler scaffold (TRIL23H), missing the heavy metal-binding Cys3 site, was developed (Figure 1.11).

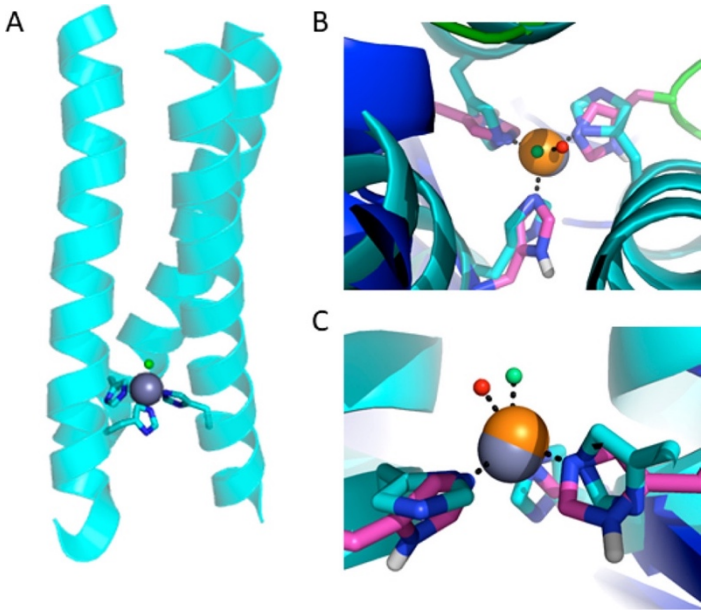


Figure 1.11. a) Model representation of de novo copper protein Cu-$(TRIL23H)_3$ based on the structure of $Hg^{2+}{}_S[Zn^{2+}{}_N(H_2O)](\mathbf{CS}L9CL23H)_3^{+}$ (PDB ID: 3pbj). b) View of the Zn^{2+}-$(H_2O)(His)_3$ site along the pseudo-three-fold axis of $Hg^{2+}{}_S[Zn^{2+}{}_N(H_2O)](\mathbf{CS}L9CL23H)_3^{+}$ (cyan), overlapped on the $T2Cu^{2+}(H_2O)(His)_3$ site in *R. sphaeroides* NiR (PDB ID: 2dy2, backbone depicted as green cartoon and side chain as pink sticks). Zinc is depicted as grey sphere, copper as orange sphere, water-bond on zinc as green sphere, water-bond on copper as red sphere. c) Side view of the same superposition in b). Adapted with permission from reference.[215] Copyright 2018 American Chemical Society.

TRIL23H bounds Cu^{+} and Cu^{2+} with affinity in the range of pM and nM, respectively. The X-ray absorption near edge structure (XANES) analysis of the 1s to 4p transition study of Cu^{+}-TRIL23H was compatible with a three-coordinate Cu^{+} complex, and the extended X-ray absorption fine structure (EXAFS) analysis verified His binding through outer-shell scattering consistent with imidazole ligation. Cu^{2+}-TRIL23H was studied

by UV-Vis and EPR spectroscopy, confirming a Cu^{2+}-His_3 site with one or two bound water molecules.

Interestingly, when Cu^{+}-TRIL23H is mixed with nitrite, NO gas production is observed (trapped as $[Fe(NO)(EDTA)]^{2-}$). Moreover, in the presence of excess Asc, the system acts as a catalyst for nitrite reduction for at least 5 turnovers. Lastly, FT-IR spectral analysis did not detect N_2O, a common side product observed with small molecule models of CuNiR.

This scaffold was the first de novo T2Cu protein that was stable, functional, and water-soluble, but with catalytic constants still far from those reported for the natural CuNiR.[213]

Making functional a T2Cu center inside a de novo designed scaffold, the modulation of the catalytic activity was studied by changing the identity of the residues beyond the first coordination sphere. The modulation of the negative charge around the T2Cu center leads to: (i) a decrease of 2 orders in Cu^{+} affinity, (ii) a 100 mV reduction in the redox potential and (iii) a 3.5-fold increase in the CuNiR activity.[216] Cu-TRIL23H includes Leu residues both above and below the copper-binding site; hence, water access is restricted. By altering these Leu packing layers to smaller Ala, larger Ile or H-bonding Asp residues, the importance of second sphere residues was proved to be a key in the modulation of the catalytic activity.[217]

Next, methylated His residues were used to enforce His binding by either the δ- or ε-nitrogen (using N_ε or N_δ methylated His, respectively). It was found that N_δ methylated His residues together with a Leu to Ala mutations in the second coordination sphere boost the CuNiR activity.[218] This work is an outstanding illustration of design methods made possible

by de novo scaffolds that are more difficult to implement into native systems.

Pecoraro, Policar and coworkers designed binding sites for Cu-only Superoxide Dismutase (SOD) into a single chain three-helical bundle scaffold. The asymmetry offered by this scaffold allowed to investigate how changing the identity of the fourth ligand affected the enzyme reactivity and the redox potential. They tested three-His (H_3), four-His (H_4), His_2Asp binding plane with a third His outside this plane (H_2DH) or a His_3 binding plane with an Asp outside the plane (H_3D). From this study, it appears that a fourth ligand from the scaffold is not required for Cu-only SOD activity, but may be employed to modulate the copper redox potential.[219]

Copper de novo designed proteins have recently been investigated for their potential to activate O-O and O-H bonds. Chakraborty and coworkers recently investigated this idea by including a $Cu(His)_3$ site within a homotrimeric three-helix bundle,[220] very similar to the design of the CuNiR described by the Pecoraro group,[213] albeit with a different peptide sequence and the $Cu(His)_3$ located closer to the N-terminus of the scaffold. Adding H_2O_2 to ArCuP (Articial Copper Peptide), a putative copper oxygen species is formed, even if several residues within 5–28 Å of the copper active site were oxidized as hydroxyl radicals were generated in the process (Figure 1.12).[220]

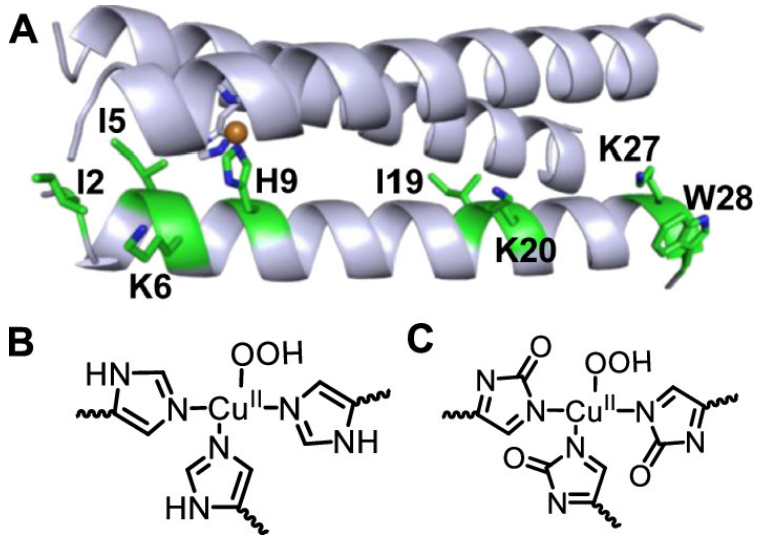


Figure 1.12. a) Structure of ArCuP (PDB ID: 7l33) highlighting the oxidizable residues on one helix as detected by LC-MS/MS. Proposed copper-oxygen species showing b) unoxidized His residues or c) proposed oxidative damage to the first coordination sphere during catalysis. Adapted with permission from reference.[220] Copyright 2021 American Chemical Society.

ArCuP is electrochemically active in the C-H bond activation of various organic substrates, using H_2O_2 as a co-substrate. Furthermore, modifications of the residues of the second coordination sphere lead to a modulation of the catalytic activity.[221]

Finally, in addition to α helical bundle, Korendovych and coworkers used catalytic Cu-bond amyloid peptides to study the O_2 activation by a T2Cu center. EPR spectroscopy confirmed the presence of a T2Cu center and one peptide showed a 1-fold increase in activity toward 2,6-dimethoxyphenol (DMP) oxidation with respect to blank reaction (Figure 1.13).[222]

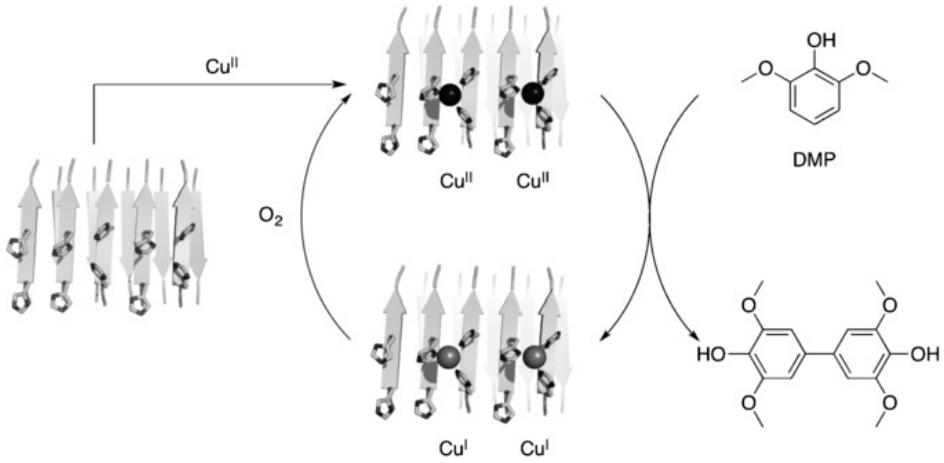


Figure 1.13. Oxidation of DMP catalyzed by copper-bound amyloid peptides. Adapted with permission from reference.[222] Copyright 2016 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim.

Catalytic amyloids have the potential to act as heterogeneous catalysts, which is one of their many advantages. A Cu-bond amyloid peptide was studied for the hydrolysis of paraoxon, a common organophosphate insecticide that persists in the environment for long periods of time due to its poor background hydrolysis. The authors used a 0.22 μm syringe filter to capture Cu-bond amyloid peptide. Despite the reduced activity after being deposited on the filter, multiple passes of the substrate solution can be done to promote paraoxon hydrolysis.[223]

Aim of the thesis

Copper metalloproteins play important roles in a variety of biological processes. Although there have been significant advancements in various fields, the mechanisms by which copper-enzymes efficiently and specifically activate dioxygen are still not fully understood. The first coordination sphere of many natural copper-enzymes, as well as the outer shell interactions, are located within complex and intricate protein structures, adding complexity to their study. As previously highlighted, many copper metalloproteins contain multiple copper-cofactors that may produce overlapping spectroscopic signals, making it challenging to specifically examine the different contribute of each cofactor.

The goal of this Ph.D. thesis is to develop artificial copper-containing metalloenzymes by using de novo design strategies. This represents an outstanding challenge, as this task requires a suitable protein scaffold apt to fulfill three essential functions: (i) adapt to changes in the coordination environment as a result of the copper ion oxidation state; (ii) facilitate the activation of O_2/H_2O_2 without the formation of ROS and (iii) enable the substrate to access the active site.

Along these lines, Type 3 and Type 2 copper sites were engineered into helical bundles. Indeed, as outlined in the previous chapter, three- and four-stranded helical bundles are particularly suitable to host metal centers. For the Type 3 copper center, the aim of this project was to prove the feasibility of reshaping the first and second coordination sphere of the previously developed di-iron containing DF1 model, to obtain a functional protein (DR1) housing a 6-His di-copper binding site, endowed with substrate selectivity, a key feature of natural enzymes.

For T2Cu, the goal was to assess if the coordination behaviors and catalytic features of the LPMO active site, namely the HB, naturally housed on the surface of β-sandwich structures, can be implanted in completely different folds, as simple helical bundle scaffolds. To this end, the de novo design of four-stranded (miniLPMO) and three- stranded (dHisB) helical bundles was exploited.

All the results highlight the power of de novo design to afford protein scaffolds suitable for the development of artificial copper metalloenzymes, which could have significant implications in different fields, by providing new tools for the synthesis of complex molecules and the production of high-value chemicals in a sustainable and environmentally friendly way.

References

[1] R. R. Crichton, in *Pract. Approaches Biol. Inorg. Chem. Second Ed.* (Eds.: R.R. Crichton, R.O. Louro), Elsevier, **2020**, pp. 1–16.

[2] S. J. Lippard, J. M. Berg, *Principles of Bioinorganic Chemistry*, University Science Books, **1994**.

[3] E. I. Solomon, D. E. Heppner, E. M. Johnston, J. W. Ginsbach, J. Cirera, M. Qayyum, M. T. Kieber-Emmons, C. H. Kjaergaard, R. G. Hadt, L. Tian, *Chem. Rev.* **2014**, *114*, 3659–3853.

[4] T. W. Lyons, C. T. Reinhard, N. J. Planavsky, *Nature* **2014**, *506*, 307–315.

[5] C. Van Stappen, Y. Deng, Y. Liu, H. Heidari, J.-X. Wang, Y. Zhou, A. P. Ledray, Y. Lu, *Chem. Rev.* **2022**, *122*, 11974–12045.

[6] H. B. Gray, B. G. Malmström, R. J. P. Williams, *JBIC J. Biol. Inorg. Chem.* **2000**, *5*, 551–559.

[7] O. Maglio, F. Nastri, A. Lombardi, in *Ion. Interact. Nat. Synth. Macromol.* (Eds.: A. Ciferri, A. Perico), John Wiley & Sons, Inc., Hoboken, NJ, USA, **2012**, pp. 361–450.

[8] B. G. Malmström, R. Aasa, *FEBS Lett.* **1993**, *325*, 49–52.

[9] M. G. Savelieff, Y. Lu, *JBIC J. Biol. Inorg. Chem.* **2010**, *15*, 461–483.

[10] M. A. McGuirl, D. M. Dooley, in *Encycl. Inorg. Bioinorg. Chem.*, John Wiley & Sons, Ltd, **2011**.

[11] C. Kaintz, S. G. Mauracher, A. Rompel, in *Adv. Protein Chem. Struct. Biol.* (Ed.: C.Z. Christov), Academic Press, **2014**, pp. 1–35.

[12] T. Rasmussen, B. C. Berks, J. Sanders-Loehr, D. M. Dooley, W. G. Zumft, A. J. Thomson, *Biochemistry* **2000**, *39*, 12753–12756.

[13] T. Goto, *Pure Appl. Chem.* **1968**, *17*, 421–442.

[14] L. Gabison, C. Chopard, N. Colloc'h, F. Peyrot, B. Castro, M. E. Hajji, M. Altarsha, G. Monard, M. Chiadmi, T. Prangé, *Proteins Struct. Funct. Bioinforma.* **2011**, *79*, 1964–1976.

[15] K. S. Johansen, *Biochem. Soc. Trans.* **2016**, *44*, 143–149.

[16] G. Vaaje-Kolstad, B. Westereng, S. J. Horn, Z. Liu, H. Zhai, M. Sørlie, V. G. H. Eijsink, *Science* **2010**, *330*, 219–222.

[17] E. Drula, M.-L. Garron, S. Dogan, V. Lombard, B. Henrissat, N. Terrapon, *Nucleic Acids Res.* **2022**, *50*, D571–D577.

[18] W. Mitsuhashi, in *Trends Insect Mol. Biol. Biotechnol.* (Eds.: D. Kumar, C. Gong), Springer International Publishing, Cham, **2018**, pp. 129–141.

[19] F. Sabbadin, S. Urresti, B. Henrissat, A. O. Avrova, L. R. J. Welsh, P. J. Lindley, M. Csukai, J. N. Squires, P. H. Walton, G. J. Davies, N. C. Bruce, S. C. Whisson, S. J. McQueen-Mason, *Science* **2021**, *373*, 774–779.

[20] B. Danneels, M. Tanghe, T. Desmet, *Biotechnol. J.* **2019**, *14*, 1800211.

[21] Z. Forsberg, G. Courtade, *Essays Biochem.* **2022**, EBC20220162.

References – 1. Introduction

[22] A. Srivastava, P. Nagar, S. Rathore, N. Adlakha, *Microbiol. Spectr.* **2022**, *10*, e02697-21.

[23] V. G. H. Eijsink, D. Petrovic, Z. Forsberg, S. Mekasha, Å. K. Røhr, A. Várnai, B. Bissaro, G. Vaaje-Kolstad, *Biotechnol. Biofuels* **2019**, *12*, 58.

[24] R. J. Quinlan, M. D. Sweeney, L. Lo Leggio, H. Otten, J.-C. N. Poulsen, K. S. Johansen, K. B. R. M. Krogh, C. I. Jørgensen, M. Tovborg, A. Anthonsen, T. Tryfona, C. P. Walter, P. Dupree, F. Xu, G. J. Davies, P. H. Walton, *Proc. Natl. Acad. Sci.* **2011**, *108*, 15079–15084.

[25] L. Ciano, G. J. Davies, W. B. Tolman, P. H. Walton, *Nat. Catal.* **2018**, *1*, 571–577.

[26] G. R. Hemsworth, E. J. Taylor, R. Q. Kim, R. C. Gregory, S. J. Lewis, J. P. Turkenburg, A. Parkin, G. J. Davies, P. H. Walton, *J. Am. Chem. Soc.* **2013**, *135*, 6069–6077.

[27] C. H. Kjaergaard, M. F. Qayyum, S. D. Wong, F. Xu, G. R. Hemsworth, D. J. Walton, N. A. Young, G. J. Davies, P. H. Walton, K. S. Johansen, K. O. Hodgson, B. Hedman, E. I. Solomon, *Proc. Natl. Acad. Sci.* **2014**, *111*, 8797–8802.

[28] G. R. Hemsworth, G. J. Davies, P. H. Walton, *Curr. Opin. Struct. Biol.* **2013**, *23*, 660–668.

[29] G. R. Hemsworth, B. Henrissat, G. J. Davies, P. H. Walton, *Nat. Chem. Biol.* **2014**, *10*, 122–126.

[30] Z. Forsberg, A. K. Mackenzie, M. Sørlie, Å. K. Røhr, R. Helland, A. S. Arvai, G. Vaaje-Kolstad, V. G. H. Eijsink, *Proc. Natl. Acad. Sci.* **2014**, *111*, 8446–8451.

[31] J. R. Winkler, H. B. Gray, *Q. Rev. Biophys.* **2015**, *48*, 411–420.

[32] A. Paradisi, E. M. Johnston, M. Tovborg, C. R. Nicoll, L. Ciano, A. Dowle, J. McMaster, Y. Hancock, G. J. Davies, P. H. Walton, *J. Am. Chem. Soc.* **2019**, *141*, 18585–18599.

[33] S. Kim, J. Ståhlberg, M. Sandgren, R. S. Paton, G. T. Beckham, *Proc. Natl. Acad. Sci.* **2014**, *111*, 149–154.

[34] J. Zhao, Y. Zhuo, D. Diaz, M. Shanmugam, A. Telfer, P. Lindley, D. Kracher, T. Hayashi, L. Seibt, F. Hardy, O. Manners, T. Hendison, K. Hollywood, S. Diaz-Moreno, N. Scrutton, M. Tovborg, G. Davies, P. Walton, D. Heyes, A. Green, *The Histidine-Brace Active Site of a Copper Monooxygenase Is Redox Active and Forms Part of an in-Built Enzyme Repair Mechanism*, In Review, **2022**.

[35] M. Wu, G. T. Beckham, A. M. Larsson, T. Ishida, S. Kim, C. M. Payne, M. E. Himmel, M. F. Crowley, S. J. Horn, B. Westereng, K. Igarashi, M. Samejima, J. Ståhlberg, V. G. H. Eijsink, M. Sandgren, *J. Biol. Chem.* **2013**, *288*, 12828–12839.

[36] A. S. Borisova, T. Isaksen, M. Dimarogona, A. A. Kognole, G. Mathiesen, A. Várnai, Å. K. Røhr, C. M. Payne, M. Sørlie, M. Sandgren, V. G. H. Eijsink, *J. Biol. Chem.* **2015**, *290*, 22955–22969.

[37] S. Banerjee, S. J. Muderspach, T. Tandrup, K. E. H. Frandsen, R. K. Singh, J. Ø. Ipsen, C. Hernández-Rollán, M. H. H. Nørholm, M. J. Bjerrum, K. S. Johansen, L. Lo Leggio, *Biomolecules* **2022**, *12*, 194.
[38] E. A. Span, D. L. M. Suess, M. C. Deller, R. D. Britt, M. A. Marletta, *ACS Chem. Biol.* **2017**, *12*, 1095–1103.
[39] S. T. Ngo, H. N. Phan, C. X. Luu, C. N. Le, G. T. Ho, N. T. C. Ngo, L. Q. Le, B. K. Mai, H. T. T. Phung, H.-D. Nguyen, K. B. Vu, V. V. Vu, *J. Phys. Chem. B* **2022**, *126*, 7567–7578.
[40] B. Wang, Z. Wang, G. J. Davies, P. H. Walton, C. Rovira, *ACS Catal.* **2020**, *10*, 12760–12769.
[41] B. Bissaro, Å. K. Røhr, G. Müller, P. Chylenski, M. Skaugen, Z. Forsberg, S. J. Horn, G. Vaaje-Kolstad, V. G. H. Eijsink, *Nat. Chem. Biol.* **2017**, *13*, 1123–1128.
[42] S. Kuusk, B. Bissaro, P. Kuusk, Z. Forsberg, V. G. H. Eijsink, M. Sørlie, P. Väljamäe, *J. Biol. Chem.* **2018**, *293*, 523–531.
[43] Z. Forsberg, M. Sørlie, D. Petrović, G. Courtade, F. L. Aachmann, G. Vaaje-Kolstad, B. Bissaro, Å. K. Røhr, V. G. Eijsink, *Curr. Opin. Struct. Biol.* **2019**, *59*, 54–64.
[44] B. Bissaro, B. Streit, I. Isaksen, V. G. H. Eijsink, G. T. Beckham, J. L. DuBois, Å. K. Røhr, *Proc. Natl. Acad. Sci.* **2020**, *117*, 1504–1513.
[45] F. Filandr, P. Man, P. Halada, H. Chang, R. Ludwig, D. Kracher, *Biotechnol. Biofuels* **2020**, *13*, 37.
[46] S. M. Jones, W. J. Transue, K. K. Meier, B. Kelemen, E. I. Solomon, *Proc. Natl. Acad. Sci.* **2020**, *117*, 11916–11922.
[47] T. L. Poulos, *Chem. Rev.* **2014**, *114*, 3919–3962.
[48] B. Wang, J. Lu, K. D. Dubey, G. Dong, W. Lai, S. Shaik, *J. Am. Chem. Soc.* **2016**, *138*, 8489–8496.
[49] J. A. Hangasky, A. T. Iavarone, M. A. Marletta, *Proc. Natl. Acad. Sci.* **2018**, *115*, 4915–4920.
[50] L. Bertini, R. Breglia, M. Lambrughi, P. Fantucci, L. De Gioia, M. Borsari, M. Sola, C. A. Bortolotti, M. Bruschi, *Inorg. Chem.* **2018**, *57*, 86–97.
[51] B. Wang, E. M. Johnston, P. Li, S. Shaik, G. J. Davies, P. H. Walton, C. Rovira, *ACS Catal.* **2018**, *8*, 1346–1351.
[52] B. Wang, P. H. Walton, C. Rovira, *ACS Catal.* **2019**, *9*, 4958–4969.
[53] E. D. Hedegård, U. Ryde, *Chem. Sci.* **2018**, *9*, 3866–3880.
[54] E. D. Larsson, G. Dong, V. Veryazov, U. Ryde, E. D. Hedegård, *Dalton Trans.* **2020**, *49*, 1501–1512.
[55] G. Courtade, L. Ciano, A. Paradisi, P. J. Lindley, Z. Forsberg, M. Sørlie, R. Wimmer, G. J. Davies, V. G. H. Eijsink, P. H. Walton, F. L. Aachmann, *Proc. Natl. Acad. Sci.* **2020**, *117*, 19178–19189.
[56] R. Kittl, D. Kracher, D. Burgstaller, D. Haltrich, R. Ludwig, *Biotechnol. Biofuels* **2012**, *5*, 79.
[57] A. A. Stepnov, V. G. H. Eijsink, Z. Forsberg, *Sci. Rep.* **2022**, *12*, 6129.

[58] F. L. Aachmann, M. Sørlie, G. Skjåk-Bræk, V. G. H. Eijsink, G. Vaaje-Kolstad, *Proc. Natl. Acad. Sci.* **2012**, *109*, 18779–18784.
[59] D. Zouraris, M. Dimarogona, A. Karnaouri, E. Topakas, A. Karantonis, *Bioelectrochemistry* **2018**, *124*, 149–155.
[60] C. M. Cordas, G. N. Valério, A. Stepnov, E. Kommedal, Å. R. Kjendseth, Z. Forsberg, V. G. H. Eijsink, J. J. G. Moura, *J. Inorg. Biochem.* **2023**, *238*, 112056.
[61] C. M. Phillips, W. T. I. Beeson, J. H. Cate, M. A. Marletta, *ACS Chem. Biol.* **2011**, *6*, 1399–1406.
[62] K. Takeda, H. Matsumura, T. Ishida, M. Samejima, H. Ohno, M. Yoshida, K. Igarashi, N. Nakamura, *PLOS ONE* **2015**, *10*, e0115722.
[63] A. Várnai, K. Umezawa, M. Yoshida, V. G. H. Eijsink, *Appl. Environ. Microbiol.* **2018**, *84*, e00156-18.
[64] D. Cannella, K. B. Möllers, N.-U. Frigaard, P. E. Jensen, M. J. Bjerrum, K. S. Johansen, C. Felby, *Nat. Commun.* **2016**, *7*, 11134.
[65] B. Bissaro, E. Kommedal, Å. K. Røhr, V. G. H. Eijsink, *Nat. Commun.* **2020**, *11*, 890.
[66] E. G. Kommedal, F. Sæther, T. Hahn, V. G. H. Eijsink, *Proc. Natl. Acad. Sci.* **2022**, *119*, e2204510119.
[67] S. Kato, T. Matsui, C. Gatsogiannis, Y. Tanaka, *Biophys. Rev.* **2018**, *10*, 191–202.
[68] C. Eicken, B. Krebs, J. C. Sacchettini, *Curr. Opin. Struct. Biol.* **1999**, *9*, 677–683.
[69] A. Bijelic, M. Pretzler, C. Molitor, F. Zekiri, A. Rompel, *Angew. Chem. Int. Ed.* **2015**, *54*, 14677–14680.
[70] C. Molitor, S. G. Mauracher, A. Rompel, *Proc. Natl. Acad. Sci.* **2016**, *113*, E1806–E1815.
[71] A. M. Mayer, *Phytochemistry* **2006**, *67*, 2318–2331.
[72] A. Volbeda, W. G. J. Hol, *J. Mol. Biol.* **1989**, *209*, 249–279.
[73] B. Hazes, K. H. Kalk, W. G. j. Hol, K. A. Magnus, C. Bonaventura, J. Bonaventura, Z. Dauter, *Protein Sci.* **1993**, *2*, 597–619.
[74] K. A. Magnus, B. Hazes, H. Ton-That, C. Bonaventura, J. Bonaventura, W. G. Hol, *Proteins* **1994**, *19*, 302–309.
[75] M. E. Cuff, K. I. Miller, K. E. van Holde, W. A. Hendrickson, *J. Mol. Biol.* **1998**, *278*, 855–870.
[76] F. Pirro, S. La Gatta, F. Arrigoni, A. Famulari, O. Maglio, P. Del Vecchio, M. Chiesa, L. De Gioia, L. Bertini, M. Chino, F. Nastri, A. Lombardi, *Angew. Chem. Int. Ed.* **2023**, *62*, e202211552.
[77] K. Kusche, H. Ruhberg, T. Burmester, *Proc. Natl. Acad. Sci.* **2002**, *99*, 10545–10548.
[78] S. M. Prexler, M. Frassek, B. M. Moerschbacher, M. E. Dirks-Hofmeister, *Angew. Chem. Int. Ed.* **2019**, *58*, 8757–8761.
[79] M. Perbandt, E. W. Guthöhrlein, W. Rypniewski, K. Idakieva, S. Stoeva, W. Voelter, N. Genov, C. Betzel, *Biochemistry* **2003**, *42*, 6341–6346.

[80] T. Nakamura, H. S. Mason, *Biochem. Biophys. Res. Commun.* **1960**, *3*, 297–299.
[81] R. S. Himmelwright, N. C. Eickman, C. D. LuBien, E. I. Solomon, *J. Am. Chem. Soc.* **1980**, *102*, 5378–5388.
[82] G. L. Woolery, L. Powers, M. Winkler, E. I. Solomon, K. Lerch, T. G. Spiro, *Biochim. Biophys. Acta BBA - Protein Struct. Mol. Enzymol.* **1984**, *788*, 155–161.
[83] S. D. Longa, I. Ascone, A. Bianconi, A. Bonfigli, A. C. Castellano, O. Zarivi, M. Miranda, *J. Biol. Chem.* **1996**, *271*, 21025–21030.
[84] A. Rompel, H. Fischer, D. Meiwes, K. Büldt-Karentzopoulos, R. Dillinger, F. Tuczek, H. Witzel, B. Krebs, *JBIC J. Biol. Inorg. Chem.* **1999**, *4*, 56–63.
[85] L. Bubacco, J. Salgado, A. W. J. W. Tepper, E. Vijgenboom, G. W. Canters, *FEBS Lett.* **1999**, *442*, 215–220.
[86] S. Hirota, N. Tanaka, I. Mičetić, P. D. Muro, S. Nagao, H. Kitagishi, K. Kano, R. S. Magliozzo, J. Peisach, M. Beltramini, L. Bubacco, *J. Biol. Chem.* **2010**, *285*, 19338–19345.
[87] N. C. Eickman, R. S. Himmelwright, E. I. Solomon, *Proc. Natl. Acad. Sci.* **1979**, *76*, 2094–2098.
[88] C. Eicken, F. Zippel, K. Büldt-Karentzopoulos, B. Krebs, *FEBS Lett.* **1998**, *436*, 293–299.
[89] C. Gerdemann, C. Eicken, B. Krebs, *Acc. Chem. Res.* **2002**, *35*, 183–191.
[90] J. S. Loehr, T. B. Freedman, T. M. Loehr, *Biochem. Biophys. Res. Commun.* **1974**, *56*, 510–515.
[91] T. B. Freedman, J. S. Loehr, T. M. Loehr, *J. Am. Chem. Soc.* **1976**, *98*, 2809–2815.
[92] N. C. Eickman, E. I. Solomon, J. A. Larrabee, T. G. Spiro, K. Lerch, *J. Am. Chem. Soc.* **1978**, *100*, 6529–6531.
[93] G. L. Woolery, L. Powers, M. Winkler, E. I. Solomon, T. G. Spiro, *J. Am. Chem. Soc.* **1984**, *106*, 86–92.
[94] M. Brenowitz, C. Bonaventura, J. Bonaventura, *Arch. Biochem. Biophys.* **1984**, *230*, 238–249.
[95] S. Hirota, T. Kawahara, M. Beltramini, P. D. Muro, R. S. Magliozzo, J. Peisach, L. S. Powers, N. Tanaka, S. Nagao, L. Bubacco, *J. Biol. Chem.* **2008**, *283*, 31941–31948.
[96] W. A. Eaton, E. R. Henry, J. Hofrichter, S. Bettati, C. Viappiani, A. Mozzarelli, *IUBMB Life* **2007**, *59*, 586–599.
[97] S. Hirota, T. Kawahara, E. Lonardi, E. de Waal, N. Funasaki, G. W. Canters, *J. Am. Chem. Soc.* **2005**, *127*, 17966–17967.
[98] M. Rolff, J. Schottenheim, H. Decker, F. Tuczek, *Chem. Soc. Rev.* **2011**, *40*, 4077.
[99] T. Klabunde, C. Eicken, J. C. Sacchettini, B. Krebs, *Nat. Struct. Biol.* **1998**, *5*, 1084–1090.
[100] C. Olivares, J. C. García-Borrón, F. Solano, *Biochemistry* **2002**, *41*, 679–686.

[101] H. Decker, T. Schweikardt, F. Tuczek, *Angew. Chem. Int. Ed.* **2006**, *45*, 4546–4550.
[102] M. Sendovski, M. Kanteev, V. S. Ben-Yosef, N. Adir, A. Fishman, *J. Mol. Biol.* **2011**, *405*, 227–237.
[103] M. Goldfeder, M. Kanteev, S. Isaschar-Ovdat, N. Adir, A. Fishman, *Nat. Commun.* **2014**, *5*, 1–5.
[104] E. Solem, F. Tuczek, H. Decker, *Angew. Chem. Int. Ed.* **2016**, *55*, 2884–2888.
[105] E. I. Solomon, U. M. Sundaram, T. E. Machonkin, *Chem. Rev.* **1996**, *96*, 2563–2606.
[106] M. D. Allendorf, D. J. Spira, E. I. Solomon, *Proc. Natl. Acad. Sci.* **1985**, *82*, 3063–3067.
[107] L. Quintanar, C. Stoj, A. B. Taylor, P. J. Hart, D. J. Kosman, E. I. Solomon, *Acc. Chem. Res.* **2007**, *40*, 445–452.
[108] Y. Tobimatsu, M. Schuetz, *Curr. Opin. Biotechnol.* **2019**, *56*, 75–81.
[109] R. S. Granja-Travez, T. D. H. Bugg, *Arch. Biochem. Biophys.* **2018**, *660*, 97–107.
[110] T. Sonoki, S. Kajita, S. Ikeda, M. Uesugi, K. Tatsumi, Y. Katayama, Y. Iimura, *Appl. Microbiol. Biotechnol.* **2005**, *67*, 138–142.
[111] Ö. Eyice, N. Myronova, A. Pol, O. Carrión, J. D. Todd, T. J. Smith, S. J. Gurman, A. Cuthbertson, S. Mazard, M. A. Mennink-Kersten, T. D. Bugg, K. K. Andersson, A. W. Johnston, H. J. Op den Camp, H. Schäfer, *ISME J.* **2018**, *12*, 145–160.
[112] C. Rensing, G. Grass, *FEMS Microbiol. Rev.* **2003**, *27*, 197–213.
[113] C. A. Francis, K. L. Casciotti, B. M. Tebo, *Arch. Microbiol.* **2002**, *178*, 450–456.
[114] E. I. Solomon, P. Chen, M. Metz, S.-K. Lee, A. E. Palmer, *Angew. Chem. Int. Ed.* **2001**, *40*, 4570–4590.
[115] C. D. LuBien, M. E. Winkler, T. J. Thamann, R. A. Scott, M. S. Co, K. O. Hodgson, E. I. Solomon, *J. Am. Chem. Soc.* **1981**, *103*, 7014–7016.
[116] L. Quintanar, C. Stoj, T.-P. Wang, D. J. Kosman, E. I. Solomon, *Biochemistry* **2005**, *44*, 6081–6091.
[117] A. J. Augustine, L. Quintanar, C. S. Stoj, D. J. Kosman, E. I. Solomon, *J. Am. Chem. Soc.* **2007**, *129*, 13118–13126.
[118] H. Decker, N. Hellmann, E. Jaenicke, B. Lieb, U. Meissner, J. Markl, *Integr. Comp. Biol.* **2007**, *47*, 631–644.
[119] A. J. Augustine, C. Kjaergaard, M. Qayyum, L. Ziegler, D. J. Kosman, K. O. Hodgson, B. Hedman, E. I. Solomon, *J. Am. Chem. Soc.* **2010**, *132*, 6057–6067.
[120] S. M. Jones, E. I. Solomon, *Cell. Mol. Life Sci.* **2015**, *72*, 869–883.
[121] T. Sakurai, K. Kataoka, *Cell. Mol. Life Sci.* **2007**, *64*, 2642.
[122] P. Durão, I. Bento, A. T. Fernandes, E. P. Melo, P. F. Lindley, L. O. Martins, *JBIC J. Biol. Inorg. Chem.* **2006**, *11*, 514.

[123] F. Xu, A. E. Palmer, D. S. Yaver, R. M. Berka, G. A. Gambetta, S. H. Brown, E. I. Solomon, *J. Biol. Chem.* **1999**, *274*, 12372–12375.
[124] Y. Kamitaka, S. Tsujimura, K. Kataoka, T. Sakurai, T. Ikeda, K. Kano, *J. Electroanal. Chem.* **2007**, *601*, 119–124.
[125] C. E. Tinberg, S. J. Lippard, *Acc. Chem. Res.* **2011**, *44*, 280–288.
[126] S. I. Chan, W.-H. Chang, S.-H. Huang, H.-H. Lin, S. S.-F. Yu, *J. Inorg. Biochem.* **2021**, *225*, 111602.
[127] D. D. S. Smith, H. Dalton, *Eur. J. Biochem.* **1989**, *182*, 667–671.
[128] H.-H. T. Nguyen, S. J. Elliott, J. H.-K. Yip, S. I. Chan, *J. Biol. Chem.* **1998**, *273*, 7957–7966.
[129] M. Takeguchi, M. Ohashi, I. Okura*, *Biometals* **1999**, *12*, 123–129.
[130] R. J. Jodts, M. O. Ross, C. W. Koo, P. E. Doan, A. C. Rosenzweig, B. M. Hoffman, *J. Am. Chem. Soc.* **2021**, *143*, 15358–15368.
[131] R. L. Lieberman, A. C. Rosenzweig, *Nature* **2005**, *434*, 177–182.
[132] S. M. Smith, S. Rawat, J. Telser, B. M. Hoffman, T. L. Stemmler, A. C. Rosenzweig, *Biochemistry* **2011**, *50*, 10231–10240.
[133] S. Sirajuddin, D. Barupala, S. Helling, K. Marcus, T. L. Stemmler, A. C. Rosenzweig, *J. Biol. Chem.* **2014**, *289*, 21782–21794.
[134] S. Y. Ro, M. O. Ross, Y. W. Deng, S. Batelu, T. J. Lawton, J. D. Hurley, T. L. Stemmler, B. M. Hoffman, A. C. Rosenzweig, *J. Biol. Chem.* **2018**, *293*, 10457–10465.
[135] L. Cao, O. Caldararu, A. C. Rosenzweig, U. Ryde, *Angew. Chem. Int. Ed.* **2018**, *57*, 162–166.
[136] S. Y. Ro, L. F. Schachner, C. W. Koo, R. Purohit, J. P. Remis, G. E. Kenney, B. W. Liauw, P. M. Thomas, S. M. Patrie, N. L. Kelleher, A. C. Rosenzweig, *Nat. Commun.* **2019**, *10*, 2675.
[137] M. O. Ross, F. MacMillan, J. Wang, A. Nisthal, T. J. Lawton, B. D. Olafson, S. L. Mayo, A. C. Rosenzweig, B. M. Hoffman, *Science* **2019**, *364*, 566–570.
[138] G. E. Cutsail, M. O. Ross, A. C. Rosenzweig, S. DeBeer, *Chem. Sci.* **2021**, *12*, 6194–6209.
[139] C. W. Koo, A. C. Rosenzweig, in *Encycl. Inorg. Bioinorg. Chem.*, John Wiley & Sons, Ltd, **2020**, pp. 1–8.
[140] D. W. Choi, W. E. Antholine, Y. S. Do, J. D. Semrau, C. J. Kisting, R. C. Kunz, D. Campbell, V. Rao, S. C. Hartsel, A. A. Y. 2005 DiSpirito, *Microbiology* **n.d.**, *151*, 3417–3426.
[141] W. Peng, X. Qu, S. Shaik, B. Wang, *Nat. Catal.* **2021**, *4*, 266–273.
[142] K. A. Dill, J. L. MacCallum, *Science* **2012**, *338*, 1042–1046.
[143] I. V. Korendovych, W. F. DeGrado, *Q. Rev. Biophys.* **2020**, *53*, DOI 10.1017/S0033583519000131.
[144] P.-S. Huang, S. E. Boyken, D. Baker, *Nature* **2016**, *537*, 320–327.
[145] M. H. Hecht, A. Das, A. Go, L. H. Bradley, Y. Wei, *Protein Sci.* **2004**, *13*, 1711–1723.

[146] M. H. Hecht, S. Zarzhitsky, C. Karas, S. Chari, *Curr. Opin. Struct. Biol.* **2018**, *48*, 124–132.
[147] R. Kolodny, L. Pereyaslavets, A. O. Samson, M. Levitt, *Annu. Rev. Biophys.* **2013**, *42*, 559–582.
[148] B. Gutte, M. Däumigen, E. Wittschieber, *Nature* **1979**, *281*, 650–655.
[149] R. Moser, R. M. Thomas, B. Gutte, *FEBS Lett.* **1983**, *157*, 247–251.
[150] W. F. DeGrado, L. Regan, S. P. Ho, *Cold Spring Harb. Symp. Quant. Biol.* **1987**, *52*, 521–526.
[151] B. I. Dahiyat, S. L. Mayo, *Protein Sci.* **1996**, *5*, 895–903.
[152] S. T. R. Walsh, H. Cheng, J. W. Bryson, H. Roder, W. F. DeGrado, *Proc. Natl. Acad. Sci.* **1999**, *96*, 5486–5491.
[153] B. Kuhlman, G. Dantas, G. C. Ireton, G. Varani, B. L. Stoddard, D. Baker, *Science* **2003**, *302*, 1364–1368.
[154] M. J. Chalkley, S. I. Mann, W. F. DeGrado, *Nat. Rev. Chem.* **2022**, *6*, 31–50.
[155] P. B. Harbury, J. J. Plecs, B. Tidor, T. Alber, P. S. Kim, *Science* **1998**, *282*, 1462–1467.
[156] S. Y. Lau, A. K. Taneja, R. S. Hodges, *J. Biol. Chem.* **1984**, *259*, 13253–13261.
[157] K. T. O'Neil, W. F. DeGrado, *Science* **1990**, *250*, 646–651.
[158] F. H. C. Crick, *Acta Crystallogr.* **1953**, *6*, 685–689.
[159] P. B. Harbury, T. Zhang, P. S. Kim, T. Alber, *Science* **1993**, *262*, 1401–1407.
[160] P. B. Harbury, P. S. Kim, T. Alber, *Nature* **1994**, *371*, 80–83.
[161] B. Lovejoy, S. Choe, D. Cascio, D. K. McRorie, W. F. DeGrado, D. Eisenberg, *Science* **1993**, *259*, 1288–1293.
[162] G. R. Dieckmann, D. K. McRorie, D. L. Tierney, L. M. Utschig, C. P. Singer, T. V. O'Halloran, J. E. Penner-Hahn, W. F. DeGrado, V. L. Pecoraro, *J. Am. Chem. Soc.* **1997**, *119*, 6195–6196.
[163] G. R. Dieckmann, D. K. McRorie, J. D. Lear, K. A. Sharp, W. F. DeGrado, V. L. Pecoraro, *J. Mol. Biol.* **1998**, *280*, 897–912.
[164] B. T. Farrer, N. P. Harris, K. E. Balchus, V. L. Pecoraro, *Biochemistry* **2001**, *40*, 14696–14705.
[165] M. Matzapetakis, B. T. Farrer, T.-C. Weng, L. Hemmingsen, J. E. Penner-Hahn, V. L. Pecoraro, *J. Am. Chem. Soc.* **2002**, *124*, 8042–8054.
[166] B. T. Farrer, V. L. Pecoraro, *Proc. Natl. Acad. Sci.* **2003**, *100*, 3760–3765.
[167] K.-H. Lee, M. Matzapetakis, S. Mitra, E. N. G. Marsh, V. L. Pecoraro, *J. Am. Chem. Soc.* **2004**, *126*, 9178–9179.
[168] M. Matzapetakis, D. Ghosh, T.-C. Weng, J. E. Penner-Hahn, V. L. Pecoraro, *JBIC J. Biol. Inorg. Chem.* **2006**, *11*, 876–890.
[169] D. S. Touw, C. E. Nordman, J. A. Stuckey, V. L. Pecoraro, *Proc. Natl. Acad. Sci.* **2007**, *104*, 11969–11974.
[170] O. Iranzo, C. Cabello, V. L. Pecoraro, *Angew. Chem.* **2007**, *119*, 6808–6811.

[171] S. Chakraborty, D. S. Touw, A. F. A. Peacock, J. Stuckey, V. L. Pecoraro, *J. Am. Chem. Soc.* **2010**, *132*, 13240–13250.
[172] S. Chakraborty, J. Yudenfreund Kravitz, P. W. Thulstrup, L. Hemmingsen, W. F. DeGrado, V. L. Pecoraro, *Angew. Chem. Int. Ed.* **2011**, *50*, 2049–2053.
[173] O. Iranzo, S. Chakraborty, L. Hemmingsen, V. L. Pecoraro, *J. Am. Chem. Soc.* **2011**, *133*, 239–251.
[174] S. Chakraborty, O. Iranzo, E. R. P. Zuiderweg, V. L. Pecoraro, *J. Am. Chem. Soc.* **2012**, *134*, 6191–6203.
[175] M. L. Zastrow, A. F. A. Peacock, J. A. Stuckey, V. L. Pecoraro, *Nat. Chem.* **2012**, *4*, 118–123.
[176] M. L. Zastrow, V. L. Pecoraro, *J. Am. Chem. Soc.* **2013**, *135*, 5895–5903.
[177] V. M. Cangelosi, A. Deb, J. E. Penner-Hahn, V. L. Pecoraro, *Angew. Chem. Int. Ed.* **2014**, *53*, 7900–7903.
[178] A. G. Tebo, T. B. J. Pinter, R. García-Serres, A. L. Speelman, C. Tard, O. Sénéque, G. Blondin, J.-M. Latour, J. Penner-Hahn, N. Lehnert, V. L. Pecoraro, *Biochemistry* **2018**, *57*, 2308–2316.
[179] P. M. S. Bragança, M. S. P. Carepo, S. R. Pauleta, T. B. J. Pinter, M. Elia, C. M. Cordas, I. Moura, V. L. Pecoraro, J. J. G. Moura, *J. Inorg. Biochem.* **2023**, *240*, 112096.
[180] M. R. Berwick, D. J. Lewis, A. W. Jones, R. A. Parslow, T. R. Dafforn, H. J. Cooper, J. Wilkie, Z. Pikramenou, M. M. Britton, A. F. A. Peacock, *J. Am. Chem. Soc.* **2014**, *136*, 1166–1169.
[181] M. R. Berwick, L. N. Slope, C. F. Smith, S. M. King, S. L. Newton, R. B. Gillis, G. G. Adams, A. J. Rowe, S. E. Harding, M. M. Britton, A. F. A. Peacock, *Chem. Sci.* **2016**, *7*, 2207–2216.
[182] K. J. Koebke, T. Kühl, E. Lojou, B. Demeler, B. Schoepp-Cothenet, O. Iranzo, V. L. Pecoraro, A. Ivancich, *Angew. Chem. Int. Ed.* **2021**, *60*, 3974–3978.
[183] A. Lombardi, C. M. Summa, S. Geremia, L. Randaccio, V. Pavone, W. F. DeGrado, *Proc. Natl. Acad. Sci.* **2000**, *97*, 6298–6305.
[184] O. Maglio, F. Nastri, V. Pavone, A. Lombardi, W. F. DeGrado, *Proc. Natl. Acad. Sci.* **2003**, *100*, 3772–3777.
[185] J. R. Calhoun, H. Kono, S. Lahr, W. Wang, W. F. DeGrado, J. G. Saven, *J. Mol. Biol.* **2003**, *334*, 1101–1115.
[186] C. M. Summa, M. M. Rosenblatt, J.-K. Hong, J. D. Lear, W. F. DeGrado, *J. Mol. Biol.* **2002**, *321*, 923–938.
[187] E. N. G. Marsh, W. F. DeGrado, *Proc. Natl. Acad. Sci.* **2002**, *99*, 5150–5154.
[188] M. Faiella, C. Andreozzi, R. T. M. de Rosales, V. Pavone, O. Maglio, F. Nastri, W. F. DeGrado, A. Lombardi, *Nat. Chem. Biol.* **2009**, *5*, 882–884.
[189] J. Kaplan, W. F. DeGrado, *Proc. Natl. Acad. Sci.* **2004**, *101*, 11566–11570.

[190] A. J. Reig, M. M. Pires, R. A. Snyder, Y. Wu, H. Jo, D. W. Kulp, S. E. Butch, J. R. Calhoun, T. Szyperski, E. I. Solomon, W. F. DeGrado, *Nat. Chem.* **2012**, *4*, 900–906.
[191] R. A. Snyder, J. Betzu, S. E. Butch, A. J. Reig, W. F. DeGrado, E. I. Solomon, *Biochemistry* **2015**, *54*, 4637–4651.
[192] R. A. Snyder, S. E. Butch, A. J. Reig, W. F. DeGrado, E. I. Solomon, *J. Am. Chem. Soc.* **2015**, *137*, 9302–9314.
[193] G. Ulas, T. Lemmin, Y. Wu, G. T. Gassner, W. F. DeGrado, *Nat. Chem.* **2016**, *8*, 354–359.
[194] S.-Q. Zhang, M. Chino, L. Liu, Y. Tang, X. Hu, W. F. DeGrado, A. Lombardi, *J. Am. Chem. Soc.* **2018**, *140*, 1294–1304.
[195] M. Chino, S.-Q. Zhang, F. Pirro, L. Leone, O. Maglio, A. Lombardi, W. F. DeGrado, *Biopolymers* **2018**, *109*, e23339.
[196] N. F. Polizzi, Y. Wu, T. Lemmin, A. M. Maxwell, S.-Q. Zhang, J. Rawson, D. N. Beratan, M. J. Therien, W. F. DeGrado, *Nat. Chem.* **2017**, *9*, 1157–1164.
[197] N. F. Polizzi, W. F. DeGrado, *Science* **2020**, *369*, 1227–1233.
[198] S. I. Mann, A. Nayak, G. T. Gassner, M. J. Therien, W. F. DeGrado, *J. Am. Chem. Soc.* **2021**, *143*, 252–259.
[199] F. Pirro, N. Schmidt, J. Lincoff, Z. X. Widel, N. F. Polizzi, L. Liu, M. J. Therien, M. Grabe, M. Chino, A. Lombardi, W. F. DeGrado, *Proc. Natl. Acad. Sci.* **2020**, *117*, 33246–33253.
[200] D. Shiga, D. Nakane, T. Inomata, Y. Funahashi, H. Masuda, A. Kikuchi, M. Oda, M. Noda, S. Uchiyama, K. Fukui, K. Kanaori, K. Tajima, Y. Takano, H. Nakamura, T. Tanaka, *J. Am. Chem. Soc.* **2010**, *132*, 18191–18198.
[201] D. Shiga, Y. Hamano, M. Kamei, Y. Funahashi, H. Masuda, M. Sakaguchi, T. Ogura, T. Tanaka, *JBIC J. Biol. Inorg. Chem.* **2012**, *17*, 1025–1031.
[202] D. Shiga, Y. Funahashi, H. Masuda, A. Kikuchi, M. Noda, S. Uchiyama, K. Fukui, K. Kanaori, K. Tajima, Y. Takano, H. Nakamura, M. Kamei, T. Tanaka, *Biochemistry* **2012**, *51*, 7901–7907.
[203] J. S. Plegaria, M. Duca, C. Tard, T. J. Friedlander, A. Deb, J. E. Penner-Hahn, V. L. Pecoraro, *Inorg. Chem.* **2015**, *54*, 9470–9482.
[204] J. S. Plegaria, C. Herrero, A. Quaranta, V. L. Pecoraro, *Biochim. Biophys. Acta BBA - Bioenerg.* **2016**, *1857*, 522–530.
[205] K. J. Koebke, L. Ruckthong, J. L. Meagher, E. Mathieu, J. Harland, A. Deb, N. Lehnert, C. Policar, C. Tard, J. E. Penner-Hahn, J. A. Stuckey, V. L. Pecoraro, *Inorg. Chem.* **2018**, *57*, 12291–12302.
[206] K. J. Koebke, V. S. Alfaro, T. B. J. Pinter, A. Deb, N. Lehnert, C. Tard, J. E. Penner-Hahn, V. L. Pecoraro, *J. Am. Chem. Soc.* **2020**, *142*, 15282–15294.
[207] M. R. Ghadiri, M. A. Case, *Angew. Chem. Int. Ed. Engl.* **1993**, *32*, 1594–1597.

[208] T. Kiyokawa, K. Kanaori, K. Tajima, M. Koike, T. Mizuno, J.-I. Oku, T. Tanaka, *J. Pept. Res.* **2004**, *63*, 347–353.
[209] D. Shiga, D. Nakane, T. Inomata, H. Masuda, M. Oda, M. Noda, S. Uchiyama, K. Fukui, Y. Takano, H. Nakamura, T. Mizuno, T. Tanaka, *Biopolymers* **2009**, *91*, 907–916.
[210] C. Citek, S. Herres-Pawlis, T. D. P. Stack, *Acc. Chem. Res.* **2015**, *48*, 2424–2433.
[211] A. Brinkmeier, R. A. Schulz, M. Buchhorn, C.-J. Spyra, S. Dechert, S. Demeshko, V. Krewald, F. Meyer, *J. Am. Chem. Soc.* **2021**, *143*, 10361–10366.
[212] S.-M. Jung, M. Yang, W. J. Song, *Inorg. Chem.* **2022**, *61*, 12433–12441.
[213] M. Tegoni, F. Yu, M. Bersellini, J. E. Penner-Hahn, V. L. Pecoraro, *Proc. Natl. Acad. Sci.* **2012**, *109*, 21234–21239.
[214] A. C. Merkle, N. Lehnert, *Dalton Trans.* **2012**, *41*, 3355–3368.
[215] K. J. Koebke, V. L. Pecoraro, *ACS Catal.* **2018**, *8*, 8046–8057.
[216] F. Yu, J. E. Penner-Hahn, V. L. Pecoraro, *J. Am. Chem. Soc.* **2013**, *135*, 18096–18107.
[217] K. J. Koebke, F. Yu, E. Salerno, C. Van Stappen, A. G. Tebo, J. E. Penner-Hahn, V. L. Pecoraro, *Angew. Chem. Int. Ed.* **2018**, *57*, 3954–3957.
[218] K. J. Koebke, F. Yu, C. Van Stappen, T. B. J. Pinter, A. Deb, J. E. Penner-Hahn, V. L. Pecoraro, *J. Am. Chem. Soc.* **2019**, *141*, 7765–7775.
[219] E. Mathieu, A. E. Tolbert, K. J. Koebke, C. Tard, O. Iranzo, J. E. Penner-Hahn, C. Policar, V. Pecoraro, *Chem. – Eur. J.* **2020**, *26*, 249–258.
[220] S. Mitra, D. Prakash, K. Rajabimoghadam, Z. Wawrzak, P. Prasad, T. Wu, S. K. Misra, J. S. Sharp, I. Garcia-Bosch, S. Chakraborty, *ACS Catal.* **2021**, *11*, 10267–10278.
[221] D. Prakash, S. Mitra, M. Murphy, S. Chakraborty, *ACS Catal.* **2022**, *12*, 8341–8351.
[222] O. V. Makhlynets, P. M. Gosavi, I. V. Korendovych, *Angew. Chem. Int. Ed.* **2016**, *55*, 9017–9020.
[223] Z. Lengyel, C. M. Rufo, Y. S. Moroz, O. V. Makhlynets, I. V. Korendovych, *ACS Catal.* **2018**, *8*, 59–62.

2. Engineering a Type 3 Copper site into a De Novo Designed Scaffold

Key words: Artificial Metalloenzymes • Polyphenol Oxidases • Protein Design • Substrate Selectivity • T3 di-Copper site

2.1 Results and Discussion

2.1.1 DR1 Design

To build the T3 site into the four-helix bundle scaffold, the symmetric crystal structure of di-Zn^{2+}-DF1 (PDB ID: 1ec5) was used.[1] The DF1 scaffold consists of two non-covalently linked helix-loop-helix motifs (α_2), with the first coordination sphere ligands of the di-nuclear Glu_4His_2 site buried within the protein (Figure 2.1a). The entire design procedure was carried out while maintaining the C_2 symmetry of DF1. The first step was aimed at identifying suitable positions to host the coordinating His triad for each Cu ion. Potential locations were the *a* and *d* positions, with the Cα atoms placed within 7 Å of the Zn^{2+} ions and facing inwards the bundle.[2] In addition to locations 10, 36, and 39, which correspond to the coordinating residues of the DF1 di-nuclear site, position 32' (numbers with primes indicate symmetry-related residues in the symmetric dimer) appeared to be the only appropriate candidate (Figure 2.1b). Next, the X-

ray structure of di-Zn^{2+}-DF1 was superimposed with the structures of several T3 copper proteins (Figure 2.1b and Table 2.1). To limit the number of potential permutations, His39 was maintained as the initial member of the His triad at the original di-metal location. Then, the optimal place for His mutation was established among Glu10, Glu36, and Ile32' by repeatedly superimposing His39 onto each His residue in the natural proteins. His10, His32', and His39 were determined to be the optimal triad in terms of ligand distance and coordination geometry. Glu10/10' and Ile32/32' were therefore changed to His. Notably, in 3His-DF_{SC}, a single-chain variant of DFs, Ile was replaced with His in a similar position in order to reprogram the reactivity from oxidation of activated quinols to N-oxygenation of anilines.[3] In order to prevent competition for Cu binding, Glu36/36' were changed to Ala, creating a 6His di-copper-binding site in the homodimeric structure (Figure 2.1c).

Subsequently, hydrogen-bonded networks were introduced, as they are crucial for stabilizing metal sites in de novo designed proteins.[4,5] As already exploited in DF1, His39 was successfully stabilized by its N-terminal interaction with Asp35' (Figure 2.1a). Additionally, this Asp interacts with Lys38 via a salt bridge (Figure 2.1d). Ile14Glu and Ile40Lys mutations were required for His10 second shell interaction. By interacting with His Nε, Lys and Glu sidechains form a salt-bridge that may prevent the formation of unwanted side chain/main chain H-bonds (Figure 2.1e). Finally, similar to 3His-DF_{sc}, by swapping Leu43His, Leu3Asn, and Leu47Glu, His32' was stabilized by an H-bond network spanning nearly the whole length of the bundle (Figure 2.1f).

2. Engineering a Type 3 Copper site into a De Novo Designed Scaffold

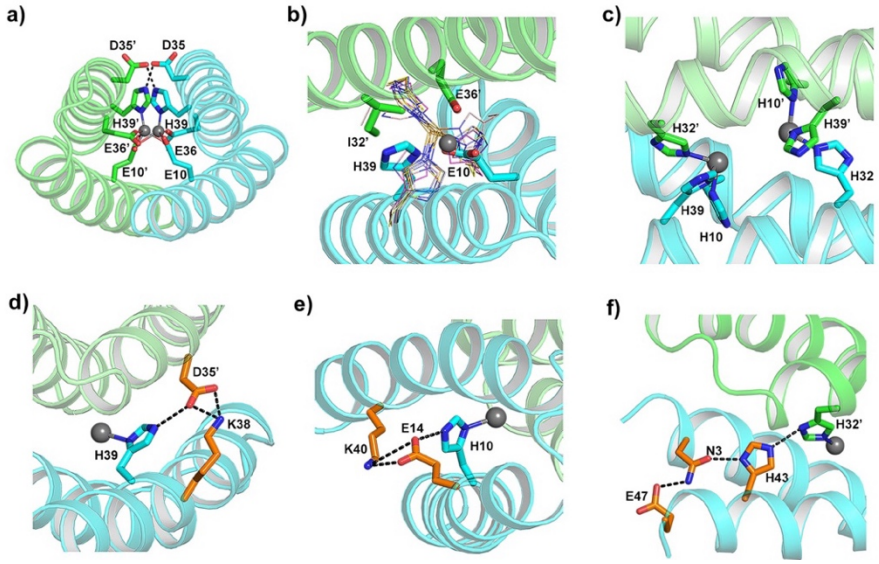


Figure 2.1. a) Active site structure of di-Zn^{2+}-DF1 (PDB ID: 1ec5). b) Superposition of DF1 metal-binding site, with focus on the *a* and *d* residues within 7 Å from one Zn^{2+} ion, with one coordinating His triad from several T3 copper proteins (PDB IDs: see Table 2.1). Side chains and the metal ion of DF1 are represented as sticks and spheres, respectively. The triad sidechains and the copper ion of the natural proteins are represented as lines. c) DR1 T3 di-copper site. d) H-bond network of His39. e) H-bond network of His10. f) H-bond network of His32. In all panels, the two symmetry-related α_2 motifs are represented as cartoons in different colors (cyan and green) for clarity of display. The Zn^{2+} ions are reported as grey spheres.

Table 2.1. T3 copper proteins selected for the superimposition with DF1.

PDB	Protein[a]	Organism	Form
4yd9	Hc	*T. pacificus*	*Oxy*
1js8	Hc	*E. dofleini*	*Oxy*
1ll1	Hc	*L. polyphemus*	*Met*
1bt2	CO	*I. batatas*	*Met*
2p3x	PPO	*V. vinifera*	*Met*
2y9w	PPO	*A. bisporus*	*Met*
5ce9	TYR	*J. regia*	*Met*

2.1.2 Shaping active site accessibility and design validation

After designing the putative 6His di-copper metal-binding site, the hydrophobic core repacking, loop definition, and active site accessibility were evaluated, in order to balance protein stability and function.[6] Indeed, the primary and secondary sphere mutations required for the T3 copper site design into the DF1 scaffold, impose a substantial thermodynamic cost[6] and necessitated the insertion of stabilizing substitutions at distant sites, namely in the hydrophobic core and loop region.[7,8] Reconstruction of the homodimer hydrophobic core was performed by RosettaScripts repackaging cycle[9,10] resulted in the following mutations: Leu7Ile, Leu11Glu, Tyr17Leu, and Leu33Ile. The idealized α_R–α_L–β interhelical loop (Thr24-His25-Asn26) was subsequently preferred over the DF1 loop sequence (Val24-Lys25-Leu26). As previously found in DF3 and more recently in DFP3,[11,12] the selected loop sequence possesses a network of H-bonded sidechain/mainchain interactions and enhances the bundle's stability/flexibility.

Residues at positions 9/9' and 13/13', corresponding to the *g* and *d* positions of the α_2 motif, influence the shape and accessibility of the active site cavity in DF proteins.[13,14] In DF1, all Leu residues at these sites afforded exceptional protein stability, but entirely restricted substrate access to the di-metallic site. By changing the channel-lining residues from Leu to Ala to Gly, the hydration and reactivity of the di-metal core were significantly enhanced.[13–16] Thus, a comprehensive screening of the best pair in positions 9 and 13 by Molecular Dynamics (MD) was performed.

Cu^{2+} was replaced with Zn^{2+} for the MD simulation, as the MD parameterization of zinc is particularly reliable in the CHARMM forcefield.[17] In addition, zinc substitution does not significantly alter the orientation of the first coordination sphere in DF scaffolds or natural PPOs.[7,8,14,18] The "non-bonded method" was used to characterize zinc-protein interactions to allow for molecular flexibility. To bypass the limitations of this method, caused by the lack of charged residues in the first coordination sphere, a sulfate anion was used as a bridging ligand to control the relative positioning of the zinc ions.[19]

DF3, which is catalytically active and presents the combination 9G13G, was used as an internal reference. To ensure that the structural and dynamic properties derived from the simulations are free from initial non-equilibrium effects, the following analyses do not include the equilibration time of each trajectory, unless otherwise stated. Specifically, 9L13A, 9L13G, 9A13A, 9A13G, 9G13A and 9G13G combinations were screened.

In detail, the calculation of Solvent Accessible Surface Area (SASA) around the access channel (from residue 8 to 15 of Chain A and B) was primarily analyzed to determine the best set of mutations. Table 2.2 reports the mean values and the deviation of the SASA along the production for all the analogs and DF3. Two distributions of the SASA were observed. 9L13A and 9L13G analogs showed a more hydrated interface, with a SASA higher than 1050 $Å^2$. All the other mutants had a SASA of ~960 $Å^2$, with only 9G13A showing a slight outlier behavior.

Table 2.2. Solvent Accessible Surface Area (SASA) calculated for DF3 and for the screened combinations at positions 9 and 13.

	DF3	9L13A	9A13A	9A13G	9G13A	9G13G	9L13G
SASA (Å²)	1071 ± 43	1064 ± 42	963 ± 47	942 ± 48	992 ± 62	953 ± 42	1097 ± 45

Thus, the subsequent analysis was carried out just on these two analogs, by evaluating their global folding and dynamic properties. The Root Mean Square Deviation (RMSD) computed on Cα atoms and Radius of gyration (R_g), calculated along the trajectories, are reported in Figure 2.2a and 2.2b, respectively. No differences were observed for the RMSD of the Cα atom (Figure 2.2a) and R_g values (Figure 2.2b) among 9L13G (12.9 ± 0.1 Å) and 9L13A and DF3 (13.1 ± 0.1 Å and 12.9 ± 0.1 Å, respectively).

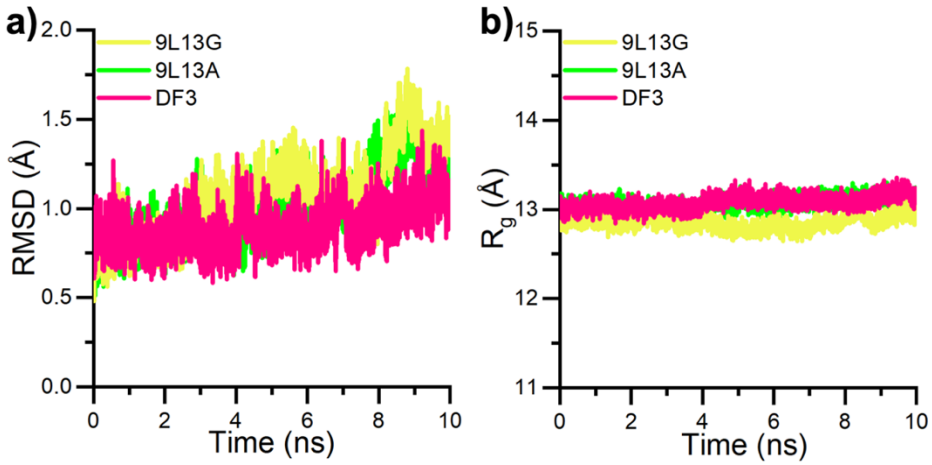


Figure 2.2. Time dependence of a) RMSD computed on Cα atoms and b) Radius of gyration (Rg). 9L13G, 9L13GA and DF3 are shown in yellow, green, and pink, respectively.

However, the root mean square fluctuation (RMSF) per residue at the Cα atom of 9L13G analog clearly differed from those of 9L13A and DF3 (Figure 2.3a, Figure 2.3c and Figure 2.3e, respectively). More specifically, the C-terminus of the helix2 showed a very high flexibility, including

residues involved in the H-bond network of the coordinating His32. Analyzing the trajectory, this behavior was caused by a kink in the helix2 with an α-helix to random transition in its terminus (Figure 2.3b) and justified the slightly lower Rg observed for the 9L13G analog (Figure 2.2b, in yellow). This conformational transition was not observed when position 13 was occupied by the helix-inducer Ala residue (Figure 2.3d). These results confirmed the thermodynamic toll of the Ala-to-Gly substitution on the four-helix bundle, reflecting the decreasing secondary structure propensity and hydrophobic driving force as observed from DF1-L13A to DF1-L13G.[6] With respect to DF3, the effect of Gly at position 13 in the helix 1 is more pronounced in the newly designed protein housing the T3 copper site, because of the numerous substitutions along the sequence. Indeed, DF3 did preserve its global folding in the same explored time, even though the presence of all Gly residues at positions 9 and 13 (Figure 2.3e).[7,11] Thus, all the above analyses suggested that adopting Leu at position 9 and Ala at position 13 would be the most favorable combinations of channel-lining residues to preserve global folding, without giving up a solvent accessible interface for the substrate binding.

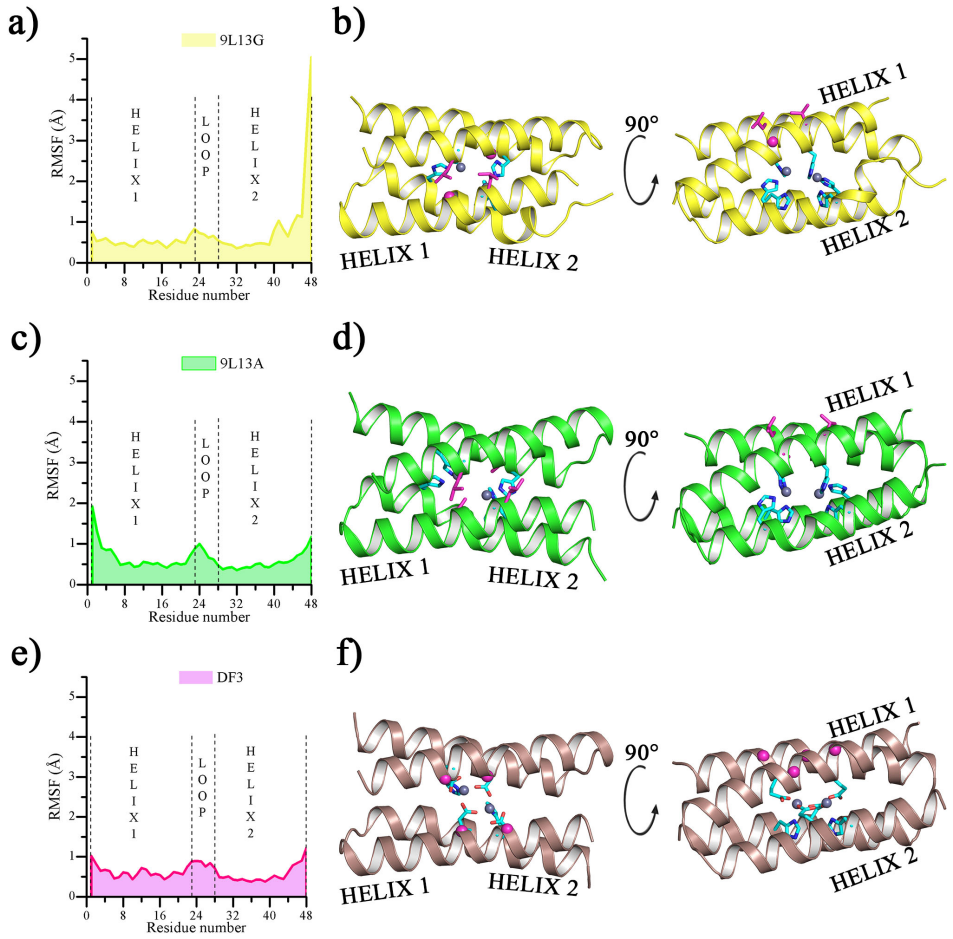


Figure 2.3. a) 9L13G analog RMSF computed on Cα atoms as a function of the residue position b) Final structure of 9L13G analog at the end of the trajectory. View of the inter-chain helix1-helix1 interface on the left and of the intra-chain helix1-helix2 interface on the right. c) 9L13A analog RMSF computed on Cα atoms as a function of the residue position. d) Final structure of 9L13A analog at the end of the trajectory. View of the inter-chain helix1-helix1 interface on the left and of the intra-chain helix1-helix2 interface on the right. e) DF3 RMSF computed on Cα atoms as a function of the residue position. f) Final structure of DF3 at the end of the trajectory. View of the inter-chain helix1-helix1 interface on the left and of the intra-chain helix1-helix2 interface on the right. Backbone, coordinating and channel-lining residues and the metal ions were represented as cartoons, sticks and spheres, respectively.

Finally, the 9L13A analog was further analyzed by cluster analysis. The trajectory was divided into 2-ps timeframes, generating a dataset of

5000 structures. The cluster analysis considered an RMSD value, calculated on the protein Cα atoms, of 0.075. This led to the formation of 51 clusters, with the first 10 clusters containing more than 92 % of the total number of structures in the trajectory (Table 2.3). The centroid of the first cluster (% abundance of 43.74) is reported in Figure 2.4.

Table 2.3. Structure abundance of the 10 clusters of 9L13A molecular simulation trajectory.

	Abundance	
Cluster	**(no. of structures)**	**% Abundance**
1	2187	43.74
2	804	16.08
3	448	8.96
4	336	6.72
5	230	4.6
6	196	3.92
7	136	2.72
8	110	2.2
9	82	1.64
10	72	1.44
Sum	4601	92.02

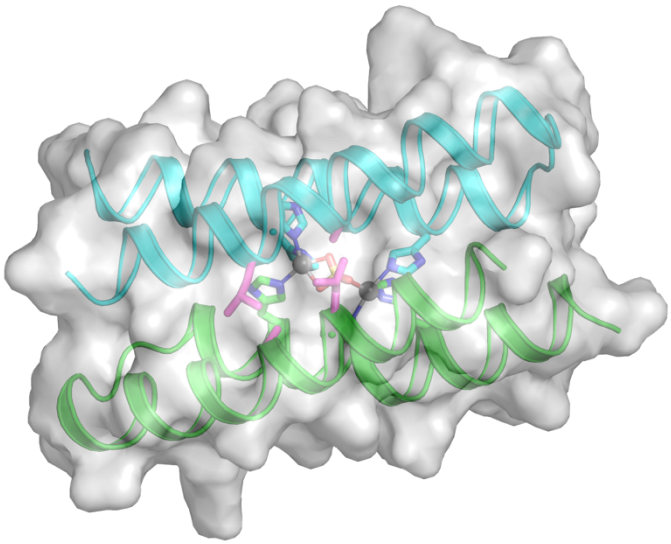


Figure 2.4. Centroid structure of the most abundant cluster determined on DR1 trajectory by cluster analysis (See Table 2.3).

To further validate the design protocol, in collaboration with the University of Milan (prof. Luca De Gioia) a DFT geometry optimization was performed, replacing zinc and sulfate with copper in the DR1 MD relaxed model (see Experimental Section for further details). A μ-hydroxo di-copper cluster (*met*-form), together with the two His triads side-chains and other residues shaping the catalytic pocket, were modeled in the DFT calculation (Figure 2.5a). The substitution of Zn^{2+} with Cu^{2+} and the addition of one hydroxide ion in the bridging position led to a shortening, upon geometry optimization, of the metal-metal distance by ≈ 1 Å. This indicates a significant flexibility of the coordination environment, which is a fundamental feature for the catalytic activity of T3 sites. The overall architecture of the optimized active site closely

resembles the one observed for PPOs *met*-forms (Figure 2.4b, 2.5c, 2.5d and Table 2.4): both Cu^{2+} ions retain the coordination of the two His triads and of the bridging OH. The final DR1 α_2 sequence with all the mutations with respect to DF1 and DF3 is reported in Figure 2.5e.

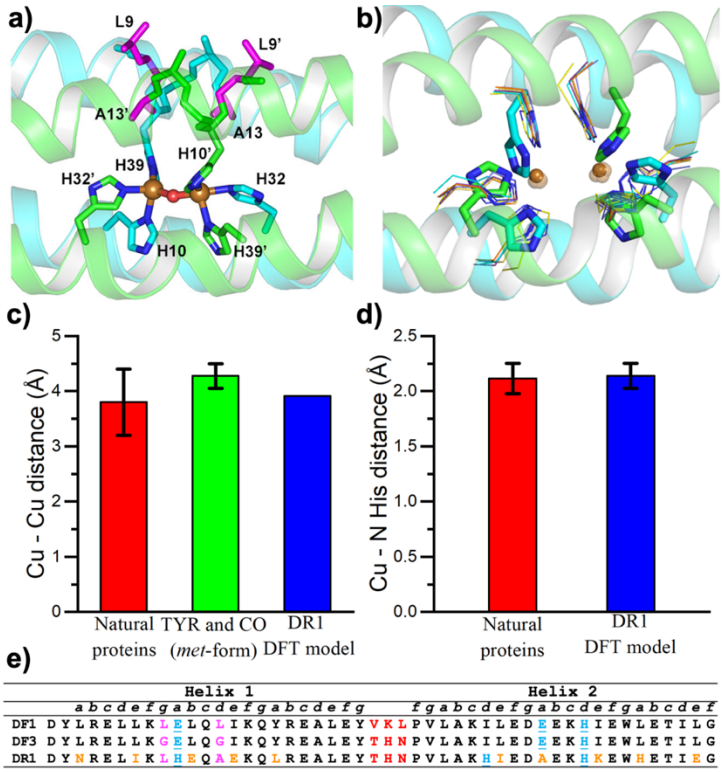


Figure 2.5. a) Details of the di-Cu^{2+}-DR1 active site as obtained by DFT calculations. His coordinating residues and Ala and Leu, shaping the access cavity, are illustrated. b) Superposition of the DFT optimized structure of di-Cu^{2+}-DR1 active site with the di-nuclear site of several T3 copper proteins (PDB IDs: see Table 2.4). The two symmetry-related α_2 motifs are represented as cartoons in different colors (cyan and green) for clarity of display. The Cu^{2+} ions are reported as brown spheres. c) Average Cu-Cu distance as observed in natural proteins (in red), average Cu-Cu distance as observed in the *met*-form of TYR and CO (in green) and Cu-Cu distance as derived from DFT geometry optimization of di-Cu^{2+}-DR1 active site (in blue). d) Average Cu-N distance as observed in natural proteins (in red), average Cu-N distance (Nδ for His 10/10' and His 39/39', Nε for His 32/32') as derived from DFT geometry optimization of di-Cu^{2+}-DR1 active site (in blue). e) Peptide sequence alignment of DR1 α_2 motif with DF1 and DF3. For all the sequences, coordinating residues are reported in cyan and underlined, residues lining the access site in magenta, and loop residues in red. Other DR1 substitutions with respect to DF1/DF3 are highlighted in orange.

Table 2.4. RMSD between T3 copper proteins and the final QM-minimized di-Cu^{2+}-DR1 model.

PDB	RMSD (Å)[a]
4yd9	0.94
1js8	0.93
1ll1	0.76
1bt2	1.08
2p3x	1.15
2y9w	0.99
5ce9	1.13

[a]Only the coordinating N atoms of His residues and copper ions were considered for the superimposition.

2.1.3 DR1 Folding, Copper Binding and Thermodynamic Stability

Using Fmoc methods, apo-DR1 was chemically synthesized, purified by Reversed-Phase High-Performance Liquid Chromatography (RP-HPLC), and identified by ESI MS (observed mass: 5858.97 ± 0.01 Da, theoretical mass: 5858.96 Da).

Size exclusion chromatography (SEC) analysis revealed that DR1 assumes the designed homodimeric structure in both the apo and holo forms (Figure 2.6a). The elution peak corresponds to a molecular weight of approximately 14.7 kDa (see Experimental Section), in good agreement with that expected for the α_2 homodimer (≈ 12 kDa), and with that of DF1 used as standard (Figure 2.6b).

2. Engineering a Type 3 Copper site into a De Novo Designed Scaffold

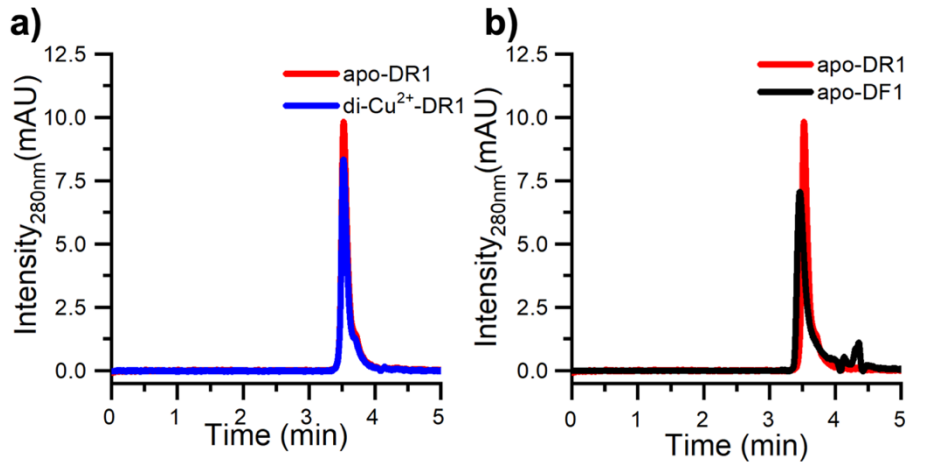


Figure 2.6. a) Analytical gel filtration chromatograms, recorded at 280 nm, of 100 μM DR1 (in 50 mM sodium phosphate, 300 mM NaCl pH 6.8) in the absence (red line) and in the presence (blue line) of 150 μM $CuSO_4$. b) Analytical gel filtration chromatograms, recorded at 280 nm, of 100 μM DR1 (red line) and 20 μM DF1 (black line) (in 50 mM sodium phosphate, 300 mM NaCl pH 6.8).

Circular dichroism (CD) analysis gave insights into DR1 folding upon copper-binding. The CD spectra of apo-DR1 at pH 7.5 (Figure 2.7a, red line) display double minima at 209 and 222 nm typical of a α-helix conformation. Cu^{2+} addition to apo-DR1 increased the mean residue ellipticity (MRE) at 209 and 222 nm, indicating an increase in helical content (Figure 2.7a, blue line). This finding demonstrated that the binding of the metal ion to DR1 increases its global folding, as was previously shown for a number of DF family members.[11,20] Cu^{2+} titration by CD (Figure 2.7b), following the change of the MRE at 222 nm, allowed to determine the stoichiometry of the interaction. The data were fit to a generic $P + nM \leftrightarrow PM_n$ equilibrium,[8] considering DR1 as a dimer. As expected, a protein-to-metal ratio of 1:2 was obtained ($n = 1.91 \pm 0.05$).

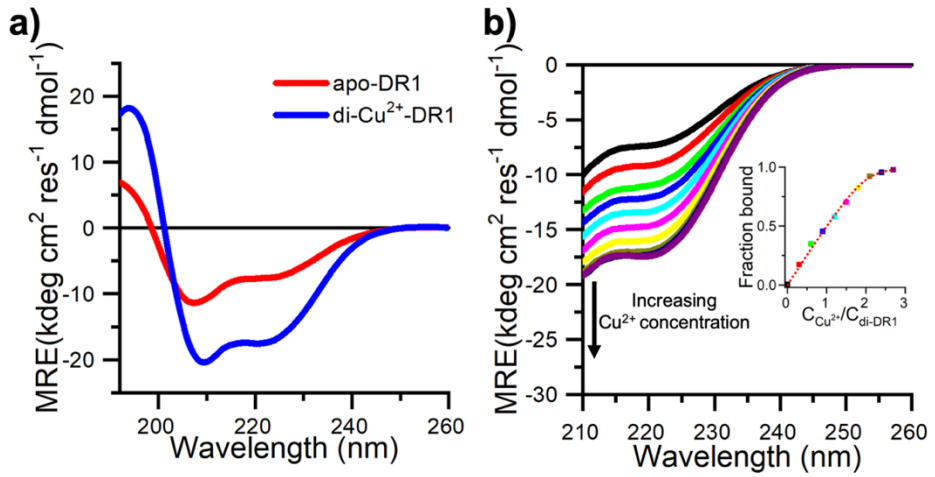


Figure 2.7. a) Ultraviolet CD spectra of DR1 (100 µM in 10 mM HEPES 300 mM Na_2SO_4 pH 7.5) in the absence (red line) and in the presence (blue line) of 100 µM $CuSO_4$. b) Ultraviolet CD spectra changes upon the addition of Cu^{2+} to DR1. Inset: the fraction bound, as calculated at 222 nm, is plotted against copper:protein ratio. The dashed red line represents the fit of the data to a binding isotherm with a 2:1 (metal:peptide) stoichiometry.

Isothermal Titration Calorimetry (ITC) was utilized to determine the thermodynamic underpinning for the copper binding to the protein. Figure 2.8 depicts the binding isotherm obtained from the integration of the raw ITC data (Figure 2.8a) after titrating a solution containing 25 µM DR1 with 400 µM $CuSO_4$. Data fitting by the independent binding model confirmed the expected binding stoichiometry of 2 Cu^{2+} ions per DR1 dimer, and gave a dissociation constant K_d value of (5 ± 2) µM (Figure 2.8b). The thermodynamic profile of the process at 298.15 K revealed a small contribution of binding enthalpy, ΔH_b is -2.5 ± 0.3 kJ mol^{-1}, equivalent to the energy of one hydrogen bond (Figure 2.8b, inset). As a result, the binding process is entropically driven ($-T\Delta S_b$= -27 ± 1 kJ mol^{-1}), as observed in other de novo designed copper proteins.[21] The enthalpic gain related to the formation of His-copper bonds slightly exceeds the thermodynamic cost of breaking the already established

favorable interactions both in the apo-protein and in the hydrated metal ion. However, the overall rearrangements probably give the entropic boost to the binding, likely releasing a relevant amount of water molecules.

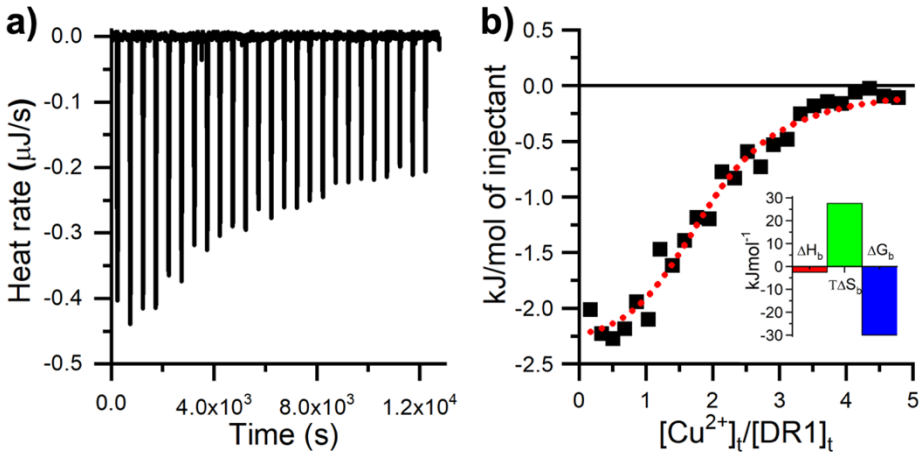


Figure 2.8. a) ITC raw data obtained from the titration of apo-DR1 with $CuSO_4$ in 10mM HEPES, 300 mM Na_2SO_4 pH 7.5 buffer solution. b) Integrated ITC data (black squares) as a function of the copper ions and di-DR1 concentration ratio. The dashed red line is the best fit of experimental data using the independent site binding model. Inset: Thermodynamic signature of the binding process: ΔH_b in red, $T\Delta S_b$ in green and ΔG_b in blue.

2.1.4 Spectroscopic Characterization of the Di-Copper Center

The UV-Vis spectrum of di-Cu^{2+}-DR1 (Figure 2.9a) is characterized by an absorption band centered at 621 nm. The observed molar extinction coefficient (ε) for the di-copper site is 100 M^{-1} cm^{-1} , which can be attributed to a weak d-d transition typical of tetragonal Cu^{2+}, as reported in the *met*-form of Hcs,[22] CO[23] and TYR.[24] Di-Cu^{2+}-DR1 shows an axial EPR spectrum (Figure 2.9b), typical of Cu^{2+} in tetragonal geometry.[25] 40 ± 10 % of total copper was determined to be EPR-active

(see Experimental Section for details). This reveals that DR1 contains a significant fraction of antiferromagnetically coupled EPR-silent T3 sites.

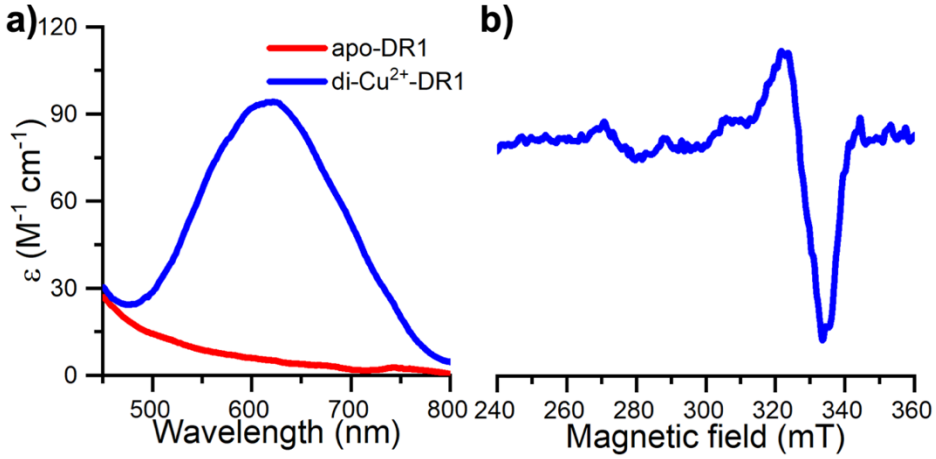


Figure 2.9. a) Visible absorption spectra of apo-DR1 (red line) and di-Cu^{2+}-DR1 (blue line), at 50 µM concentration, (ε is shown for di-copper site). b) X-band CW-EPR spectrum of di-Cu^{2+}-DR1, with 30% of glycerol as a glassing agent, at 300 µM copper concentration. The spectrum was recorded at 77 K. All the samples are in 10 mM HEPES 300 mM Na_2SO_4 pH 7.5.

Interestingly, upon O_2 exposure of the di-Cu^{+}-DR1, the fraction of EPR-active Cu^{2+} species was reduced by 40 % (figure 2.10). These results suggest that O_2 binding to the site may change the Cu-Cu distance, hence favoring antiferromagnetic coupling between cupric ions.

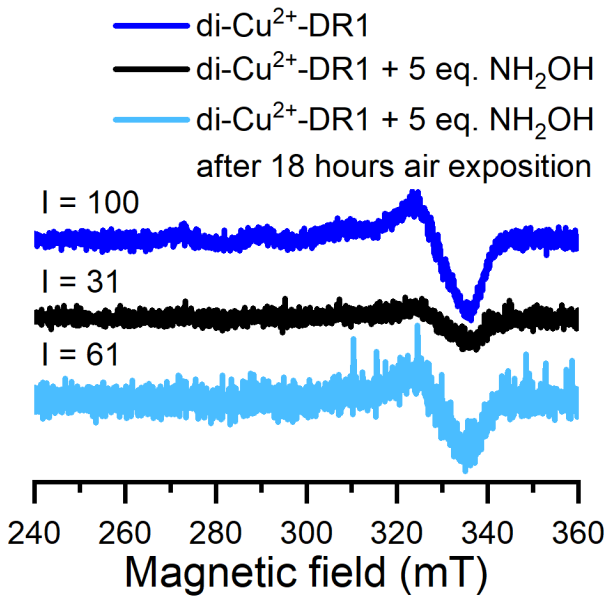


Figure 2.10. X-band CW-EPR spectra of di-Cu^{2+}-DR1 (blue line), upon reduction with hydroxylamine (black line), and subsequent exposure to air (cyan line). The integrated intensity of each spectrum (I), normalized to the intensity of the spectrum of di-Cu^{2+}-DR1 is reported. After re-oxidation, a loss of ≈ 40% of the initial intensity is observed.

2.1.5 Di-Copper-DR1 Diphenolase Activity: DTBC oxidation

Once DR1's capacity to bind copper and recapitulate the T3 site was established, its diphenolase activity was evaluated. The addition of the model substrate 3,5-di-*tert*-butylcatechol (DTBC) to a solution of di-Cu^{2+}-DR1 induced the formation of two bands near 410 nm and 660 nm in the UV spectrum (Figure 2.11a). The band at 410 nm was unambiguously assigned to 3,5-di-*tert*-butyl-o-benzoquinone (DTBQ) (see Experimental Section), but the band at 660 nm is unrelated to the product and can be provisionally attributed to the red-shifted Cu^{2+} d-d transitions upon quinone binding (see infra). NMR analysis of the reaction mixture ruled out the presence of any byproducts, as only DTBQ was detected (see Experimental Section). A complete conversion of DTBC to DTBQ was

observed over 60 min (Figure 2.11b). No increase in DTBQ production was detected over the background when the reaction was carried out with apo-DR1, di-Zn^{2+}-DR1, or in sole buffer, demonstrating that the catalysis occurs at the DR1 di-copper site (see Experimental Section).

In order to get insight into the role of oxygen in the catalytic mechanism, the reaction was also conducted in anaerobic conditions (Figure 2.11c). After DTBC was added to di-Cu^{2+}-DR1, 1 equivalent (relative to the catalyst) of DTBQ was rapidly produced (Figure 2.11d white symbols). Then, no rise in product formation was detected until aerobic conditions were restored (Figure 2.11d black symbols).

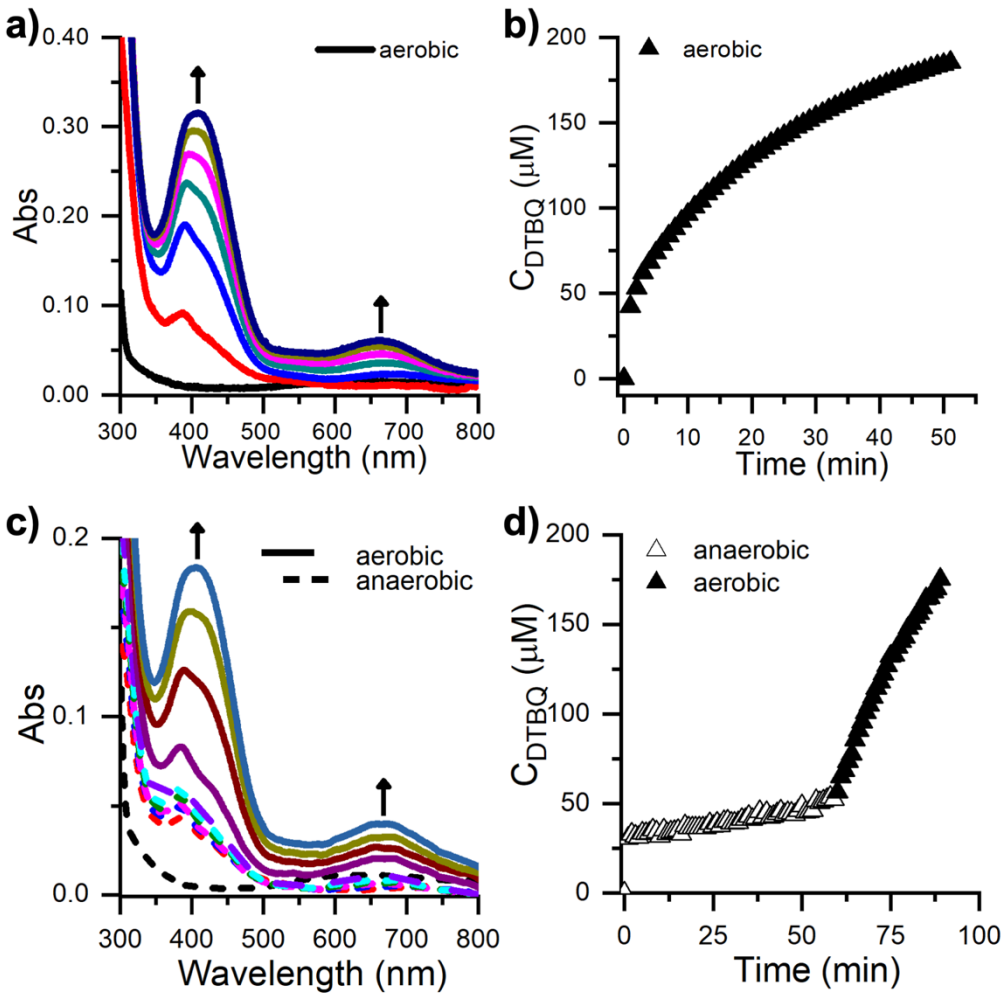


Figure 2.11. a) UV-Vis absorption spectra recorded every 10 minutes upon addition of 250 μM DTBC to 50 μM di-Cu^{2+}-DR1 in 10 mM HEPES 300 mM Na_2SO_4, pH 7.5. b) DTBQ formation kinetic under aerobic conditions. c) UV-Vis absorption spectra recorded every 10 minutes upon addition of 250 μM DTBQ to 50 μM met-DR1 in 10 mM HEPES 300 mM Na_2SO_4, pH 7.5 in anaerobic conditions (dashed lines) and after the restoration of aerobic condition (solid lines). d) DTBQ formation kinetic under anaerobic conditions (white symbol) and after the restoration of aerobic condition (black symbol). All spectra are subtracted for the protein background.

This finding suggests, similarly to natural proteins, a mechanism in which di-Cu^{2+}-DR1 reacts with DTBC thus producing one equivalent of the corresponding quinone with concomitant reduction of the *met-*

enzyme to the di-cupreous state. This hypothesis is consistent with the reduction of the EPR-active species upon the addition of DTBC to di-Cu^{2+}-DR1 under anaerobic conditions (Figure 2.12).

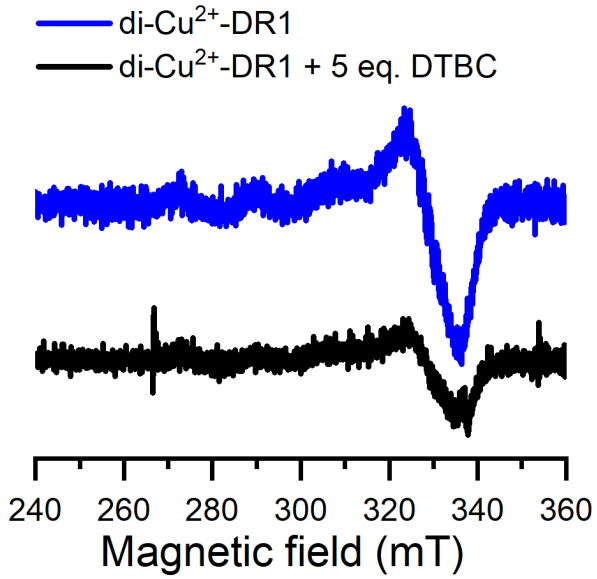


Figure 2.12. X-band CW-EPR spectra of di-Cu^{2+}-DR1 (blue line) and after the addition of 5 equivalents of DTBC (black line). The spectra were recorded at 77 K.

To complete the cycle, the addition of molecular oxygen re-oxidizes the di-copper core, allowing DTBC to be rapidly transformed again into o-quinone.

2.1.6 DTBQ interaction within the active site

As previously shown for several model compounds,[26,27] binding of DTBQ to the di-copper site may occur upon oxidation of DTBC. UV-Vis absorption spectroscopy and saturation transfer difference (STD) NMR measurements demonstrated the binding of DTBQ to di-Cu^{2+}-DR1. The addition of DTBQ to the protein resulted in the gradual development of

the band at 660 nm (Figure 2.13a), which had previously been detected with di-Cu^{2+}-DR1-catalyzed DTBC oxidation (Figure 2.11).

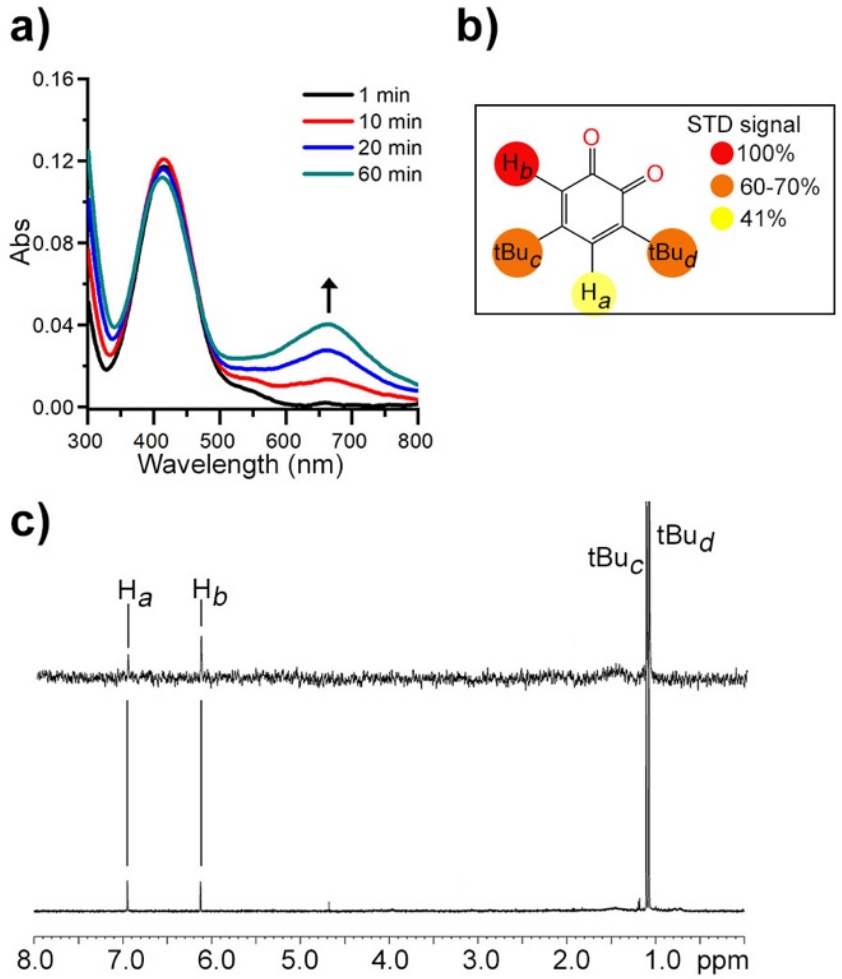


Figure 2.13. a) UV-Visible spectra recorded at a different time upon addition of DTBQ to di-Cu^{2+}-DR1 (final concentration 150 µM and 100 µM, respectively) in 10 mM HEPES 300 mM Na_2SO_4, pH 7.5. All spectra are subtracted for the protein background. b) Structural formula of DTBQ with the relative degrees of saturation of the individual protons normalized to Hb proton. c) upper STD-NMR spectrum of 10 µM di-Cu^{2+}-DR1 with 100 µM DTBQ (saturation time 5s); lower 1D ^1H NMR reference spectrum of DTBQ.

Using a 10-fold DTBQ excess relative to the protein, the STD spectrum displayed narrow and well-defined lines (Figure 2.13c).

Significant STD effects were detected for all the protons of DTBQ, conclusively proving its binding to di-Cu^{2+}-DR1. The strongest STD effect was seen for proton H_b, whose value was set to 100%. All other proton intensities were expressed relative to the H_b signal (Figure 2.13b and see Experimental Section for details). Similar saturation transfers (≈ 70%) were observed between the t-butyl protons tBu_c and tBu_d. H_a had the lowest interaction with di-Cu^{2+}-DR1, with STD of ≈ 40%. These data revealed proton H_b to point towards the di-copper site within the protein core, t-butyl groups to interact with the hydrophobic sidechain at the substrate-binding site, and proton H_a to be more solvent-exposed. From this structure, the DTBQ oxygen atoms would point toward the protein core.

2.1.7 Di-Cu^{2+}-DR1 kinetic parameters and substrate recognition by Hammett analysis

Next, the catalytic parameters for the oxidation of DTBC were determined. Di-Cu^{2+}-DR1 followed Michaelis-Menten kinetics (Figure 2.14a and Table 2.5) in the presence of ambient oxygen, with K_m and k_{cat} values of 0.70 ± 0.06 mM and 65 ± 3 min^{-1}, respectively (k_{cat}/K_m = (9.2 ± 0.4) 10^4 $M^{-1}min^{-1}$). Kinetic investigations using differently substituted di-phenols, such as 4-*tert*-butylcatechol (4-TBC) and catechol, revealed interesting di-Cu^{2+}-DR1 behaviors (Figure 2.14b, 2.14c and Table 2.5). Though the removal of one substrate *tert*-butyl (tBu) group has a negligible effect on K_m value (K_m for 4-*tert*-butylcatechol: 0.82 ± 0.08 mM), further lowering substrate steric hindrance, by removal of both tBu groups, determines a huge effect on K_m value (K_m for catechol: 39 ± 5 mM). The observed decrease in catechol affinity with respect to DTBC

suggests that the active site preferentially interacts with more hydrophobic substrates, thanks to the aliphatic residues lining the active site access channel (Figure 2.5a).

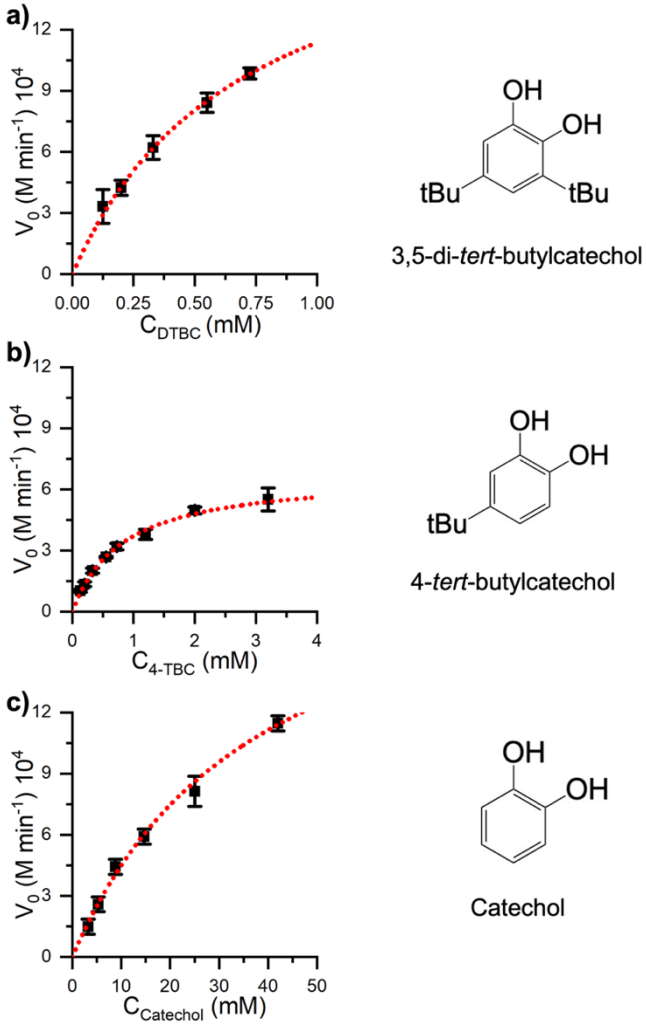


Figure 2.14. Diphenolase activity of di-Cu^{2+}-DR1. Initial rate dependence towards a) DTBC concentration; b) 4-TBC concentration; c) Catechol concentration. Reaction conditions were: 30 μM met-DR1 in 10 mM HEPES, 300 mM Na_2SO_4 pH 7.5, 20% DMF. The substrates' concentration varies from 0.125 mM to 0.725 mM for DTBC, from 0.125 mM to 3.2 mM for 4-TBC, and from 3.2 mM to 42 mM for catechol. Data were fitted according to Michaelis Menten equation (dashed red line).

Table 2.5. Kinetic parameters for the differently substituted catechols in the oxidation catalyzed by di-Cu^{2+}-DR1.

	K_m (mM)	k_{cat} (min^{-1})	k_{cat}/K_m (M^{-1} min^{-1})
DTBC	0.70 ± 0.06	65 ± 3	9.2 ± 0.4 10^4
4-TBC	0.82 ± 0.08	22.6 ± 1.2	2.7 ± 0.2 10^4
Catechol	39 ± 5	73 ± 5	1.89 ± 0.14 10^3

To further validate the substrate recognition capabilities of di-Cu^{2+}-DR1, DTBC, 4-TBC, catechol, and 4-chlorocatechol were utilized in a Hammett analysis. Compared to catechol as a reference substrate, the apparent catalytic reaction rate (TOF) was found to increase with the electron-releasing effect of the substituents, whereas electron-withdrawing groups gave smaller TOF values (Figure 2.15a). Using the Hammett parameter σ_p^-, the correlation between reaction kinetic and linear free energy relationship-based electronic substituent characteristics was investigated. It was found that the Hammett plot (log(TOF) vs σ_p^- parameters)[28] deviates from linearity (R^2 = 0.87). (Figure 2.15a). This suggested that other factors, such as the steric effects of the substituents, also play an important role in catalytic activity, thanks to the constraints exerted by the protein matrix surrounding the active site. To assess the relationship between steric and electronic effects and the reaction kinetics of the various catechol substrates, the σ_p^- parameters were modified using the σ_v Charton steric parameters.[28] The modified parameter was determined to be $\sigma' = \sigma_p^- - \sigma_v$. The correlation between log(TOF) and σ' is linear (R^2 = 0.99) with a negative slope (ρ = -1.6) (Figure 2.15b). A similar ρ value was observed in natural T3 copper proteins[29,30] and in several synthetic complexes,[31] for

which a positively charged transition state has been invoked in the rate-determining step. The good correlation found by correcting the σ_p^- parameter by steric volume is clear evidence of the direct interaction of the protein matrix with substrates. This result further supports that the presence of tBu groups on the substrates determines favorable interactions within the active site, thus increasing the catalytic rate and lowering K_m values.

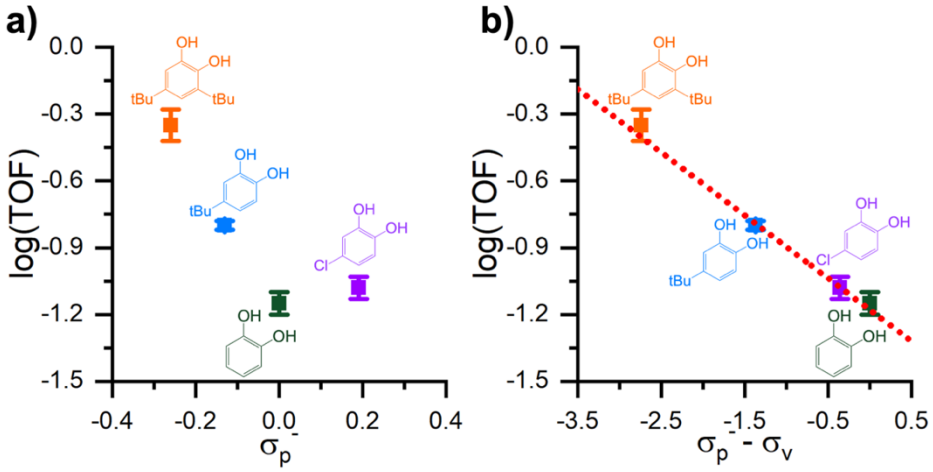


Figure 2.15. a) Hammett plot of values of log(TOF) for oxidation of catechol substrates as a function of σ_p^- and b) as a function of $\sigma' = \sigma_p^- - \sigma_v$. The points are experimental points, measured for compounds of the structures shown adjacent to those points. Di-Cu^{2+}-DR1 concentration was 30 μM in 10 mM HEPES, 300 mM Na_2SO_4, 20% DMF, pH 7.5 and substrate concentration was fixed at 550 μM. The data points were fitted with a linear function (dashed red line). Error bars were obtained from three independent measurements.

2.1.8 di-Cu^{2+}-DR1 substrate recognition by DFT calculations

To describe the interaction of substrates within the DR1 pocket, DFT calculations were undertaken (see Experimental Section for details on the utilized di-Cu^{2+}-DR1 cluster model). In agreement with EPR results, the di-copper site is in its ground state anti-ferromagnetically coupled *met*-form (preferred over triplet solution by 1.8 kcal/mol at the B3LYP-

D4 level). Differently substituted catechol substrates were tested to better rationalize, on a stereo-electronic basis, the effect of tBu groups on binding affinities. Figure 2.16 depicts the optimized structures of three distinct adducts resulting from the binding of DTBC, 4-TBC, and catechol to the di-copper site.

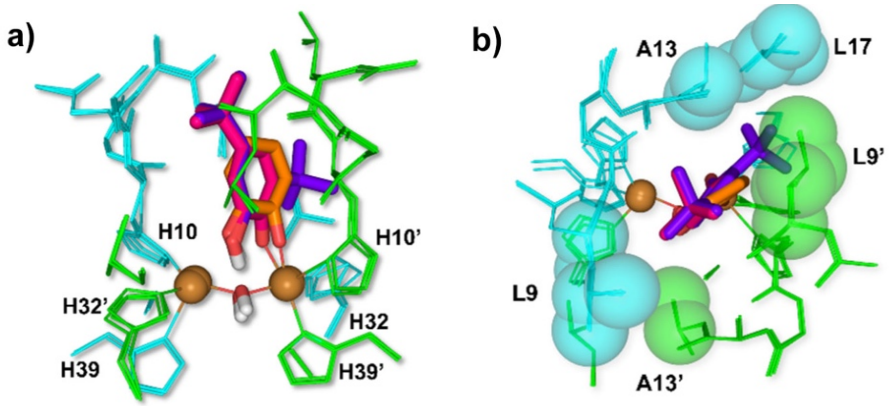


Figure 2.16. a) Front-view of the DFT-optimized adducts of DR1 pocket with DTBC (in purple), 4-TBC (in pink) and catechol (in orange) and b) top-view of catalytic pocket with the bound catechols, highlighting (as spheres) the hydrophobic side chains mostly involved in substrate recognition.

Due to its structure and steric limitations, the DR1 cavity directs catechol substrates towards a particular binding mode, which most likely involves a single Cu core (Figure 2.16a and Figure 2.17).

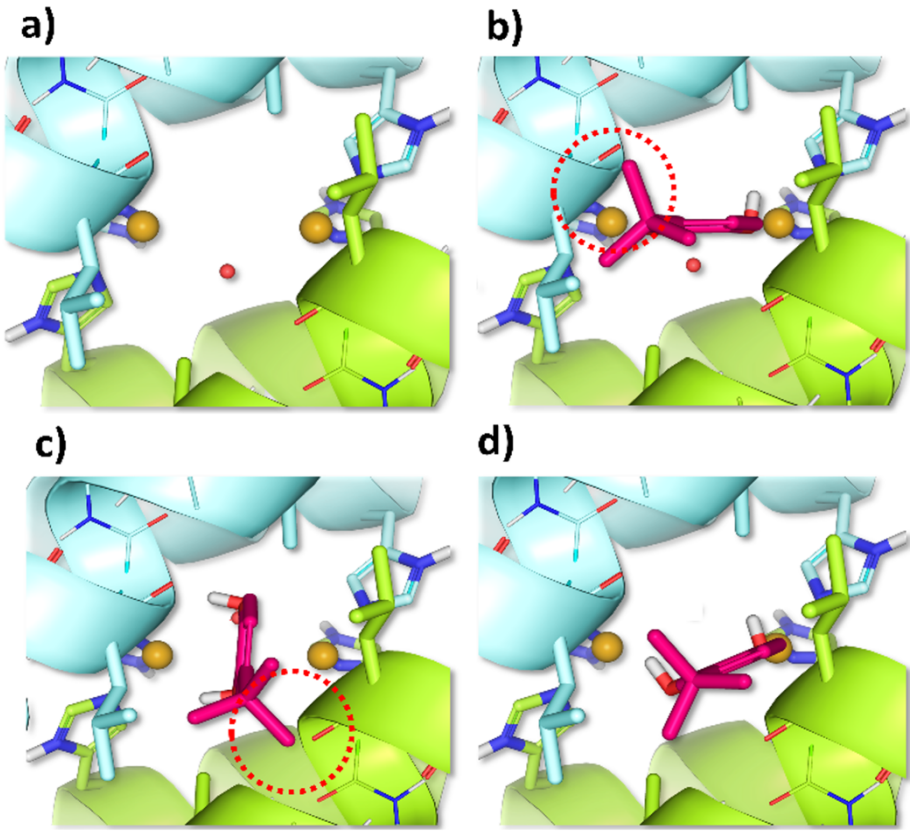


Figure 2.17. Interaction models of 4-TBC with di-Cu^{2+}-DR1 pocket. a) Top-view of the binding cavity; b) and c) substrates orientations impeded by steric clashes with the protein matrix (indicated by a red dotted circle); d) substrate orientation considered in QM calculations, that ensures the lower number of unfavorable contacts with the pocket. According to the accessible binding poses, it is highly unlikely for the substrate to interact with both metal centers (b) or perpendicularly with respect to the Cu-Cu bond (c), excluding some of the mechanisms previously proposed for COs.[32,33]

Since the exact mechanism for substrate binding for natural CO has not been unambiguously determined, two alternative routes were envisioned and simulated (in the absence of molecular oxygen and starting from the mono-hydroxide *met*-form, as proposed by Krebs)[34]: a) entailing the binding of singly-deprotonated substrates, assuming first deprotonation to be assisted, for instance, by nearby water molecules and b) in which deprotonation is assisted by the bridging hydroxide anion. In

any case, the same energetic trend is observed, thus it is reasonable to see a similar result independently from the mechanism. Such trend reflects the different number of tBu groups in the substrate, which affect the stereoelectronic features of the binding process. Besides forming more hydrophobic contacts with the pocket, the presence of tBu substituents also increases the reducing power of the substrate, further triggering its binding. Indeed, the latter is predicted to be accompanied by a significant negative charge transfer from the substrate to the di-copper center, which has been quantified through a NBO population analysis at the B3LYP-D4 level of theory (considering binding mechanism a) the CT values, in electrons, are: 1.06 for DTBC, 0.97 for 4-TBC and 0.90 for catechol).

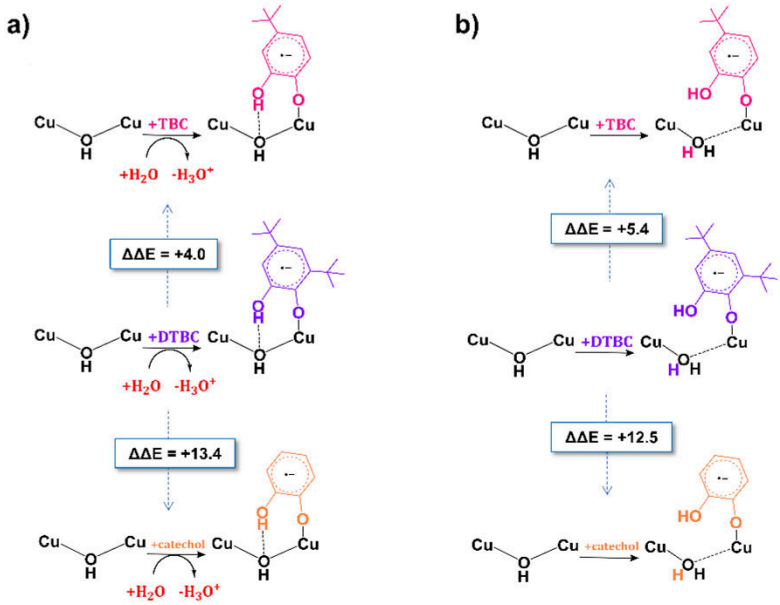


Figure 2.18. Relative binding energies (kcal/mol, BP86-D4 level) for the tested substrates to the met-form. a) route involving the binding of single-deprotonated substrates and assuming the first deprotonation to be assisted by nearby water molecules and b) in which deprotonation is assisted by the bridging hydroxide anion.

The calculated relative binding energies (Figure 2.18) clearly demonstrated that the affinity decreases in the order DTBC > 4-TBC > catechol, confirming the experimental result that bulky hydrophobic groups promote substrate binding. The observed trend is the same regardless of the envisaged binding mechanism. The removal of one tBu group, going from DTBC to 4-TBC, decreases affinity by ≈ 4.5 kcal mol^{-1}, while a further tBu substitution with a hydrogen atom causes an additional affinity loss of ≈ 8 kcal mol^{-1}.

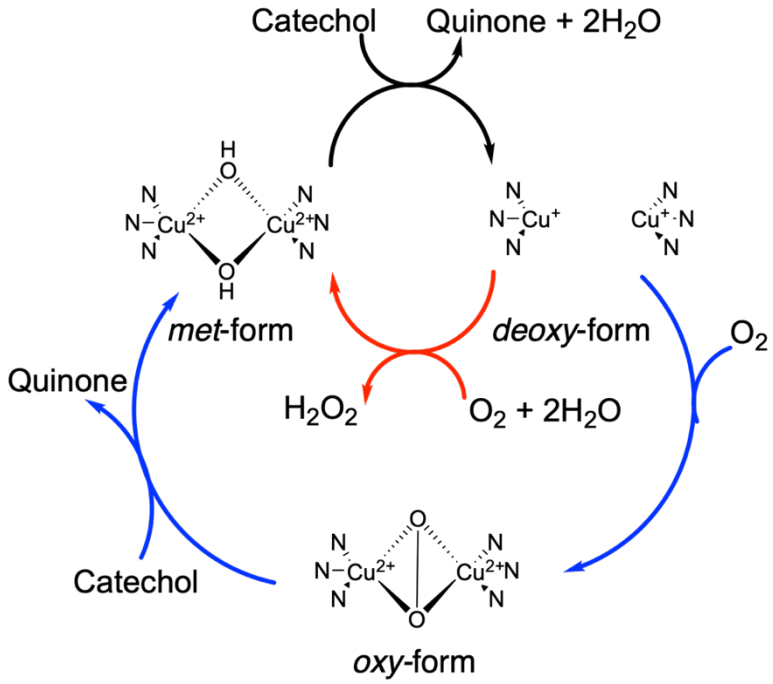
Such differences can be explained in terms of both steric and electronic factors. One tBu (position 4 for 4-TBC and 5 for DTBC) forms favorable dispersive contacts with residues L9/L9' and A13/A13', while the second tBu in DTBC (position 3) can also interact with the L17 side chain (Figure 2.16b), further stabilizing the binding adduct. The lack of these favorable hydrophobic interactions in catechol recognition explains its lower affinity for DR1, confirming the crucial role of positions 9 and 13 in substrate binding. Furthermore, the extent of charge transfer from the bound substrate to the di-copper site is larger for DTBC and 4-TBC rather than for catechol, reflecting a stronger binding for electron-rich (and thus more reducing) substrates.

2.1.9 Mechanistic studies of dioxygen activation

As described in paragraph 2.1.5, di-Cu^{2+}-DR1-displays diphenolase activity towards catechol substrates using molecular oxygen as cosubstrate. As previously reported for other small synthetic dicopper complexes,[26,35] two possible pathways have been proposed for O_2 reduction and subsequent catechol oxidation catalyzed by di-copper sites. The first common initiation step is the reduction of the copper ions with

the concomitant oxidation of one equivalent of catechol to quinone (Scheme 1, black line). The *deoxy*-form may then react with dioxygen to generate the Cu_2O_2 core (*oxy*-form). This intermediate may, in turn, oxidize another equivalent of catechol with the formation of the *met*-form resting state (Scheme 1, blue line). In the alternative pathway, the oxidation of *deoxy*-form is mediated by dioxygen with the concomitant formation of H_2O_2, and without the oxidation of another equivalent of catechol (Scheme 1, red line).

Scheme 1. Two possible pathways for dicopper center catalyzed oxidation of DTBC: the mechanism indicated by the blue line involves the formation of the Cu_2O_2 core (*oxy*-form); the mechanism indicated by the red line involves the *met*-form regeneration upon hydrogen peroxide production.



In order to gain mechanistic information, stopped-flow absorption spectroscopy was used to analyze the kinetic of 3,5-di-*tert*-butylcatechol (DTBC) oxidation. The product formation was followed at 415 nm.

All the progress curves (Figure 2.19a) showed an initial burst phase, upon the addition of the DTBC to *met*-DR1 (formerly called di-Cu^{2+}-DR1), followed by a severe slowdown in the conversion. For this reason, a kinetic model based on two steps to describe this behavior was proposed (Scheme 2): i) the equilibrium among *met*-DR1 and DTBC with *deoxy*-DR1 (formerly called di-Cu^{+}-DR1) and DTBQ, described by k_a and k_b as kinetics constants for the direct and inverse reactions, respectively; ii) the recycling oxidation of the enzyme from the *deoxy* to the *met*-form by molecular oxygen, with a catalase-like equilibrium, described by k_c and k_d as kinetics constants for the direct and inverse reactions, respectively.

Scheme 2. Kinetic model of DTBC oxidation by *met*-DR1.

$$1)\ met-\mathrm{DR1}+\mathrm{DTBC} \underset{k_b}{\overset{k_a}{\rightleftarrows}} deoxy-\mathrm{DR1}+\mathrm{DTBQ}$$

$$2)\ deoxy-\mathrm{DR1}+\mathrm{O_2} \underset{k_d}{\overset{k_c}{\rightleftarrows}} met-\mathrm{DR1}+\mathrm{H_2O_2}$$

In the Experimental Section, the mathematical description for Scheme 2 is reported. This model successfully reproduced the experimental data (Figure 2.19b) in the range of concentration explored and the following kinetic constants were obtained: $k_a = (4.9 \pm 0.3) \times 10^2$ $M^{-1}s^{-1}$, $k_b = (2.5 \pm 0.9) \times 10^4$ $M^{-1}s^{-1}$, $k_c = (0.84 \pm 0.09) \times 10^2$ $M^{-1}s^{-1}$, $k_d = (0.5 \pm 0.3) \times 10^3$ $M^{-1}s^{-1}$. Moreover, the equilibrium constant K (k_a/k_b) calculated from the model was in complete agreement with that determined in an anaerobic experiment, confirming the model robustness (see Experimental Section).

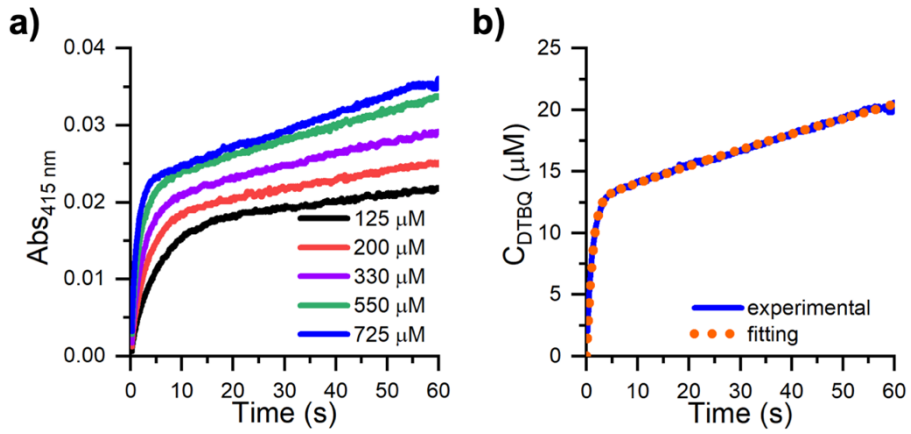


Figure 2.19. a) Stopped-Flow progress curves of DTBQ formation as measured at 415 nm and at different concentrations of DTBC. b) Graphical representation of the simulated curve (dashed orange line) superimposed on the progress curve acquired (solid blue line) by stopped-flow techniques at 725 µM of DTBC. The *met*-DR1 concentration was 30 µM in 10 mM HEPES, 300 mM Na_2SO_4, 20% DMF, pH 7.5.

To get insights into the involvement of H_2O_2 in the catalytic mechanism, an Amplex Red biochemical assay[36] to monitor the H_2O_2 production was performed. H_2O_2 concentration was determined at different reaction times (1, 10, 20, 30, 40 and 50 min) (Figure 2.20a). This analysis pointed out that, after one minute of reaction, a discrete amount of H_2O_2 can be observed (23 µM, approximately half the concentration of *met*-DR1 and 1/10 of DTBC concentration), and this amount is consumed to background levels within 30 minutes. Subsequently, stopped-flow analysis in the presence of catalase was carried out for a longer timeframe (3 minutes) (Figure 2.20b). Interestingly, the presence of catalase had an influence on the produced amount of quinone only on a longer timescale, specifically after the first minute of reaction progress. According to these data, it can be hypothesized that H_2O_2 is formed in a sub-stoichiometric amount during the first turnover and that it does not

accumulate over time nor reaches a significant steady-state concentration.

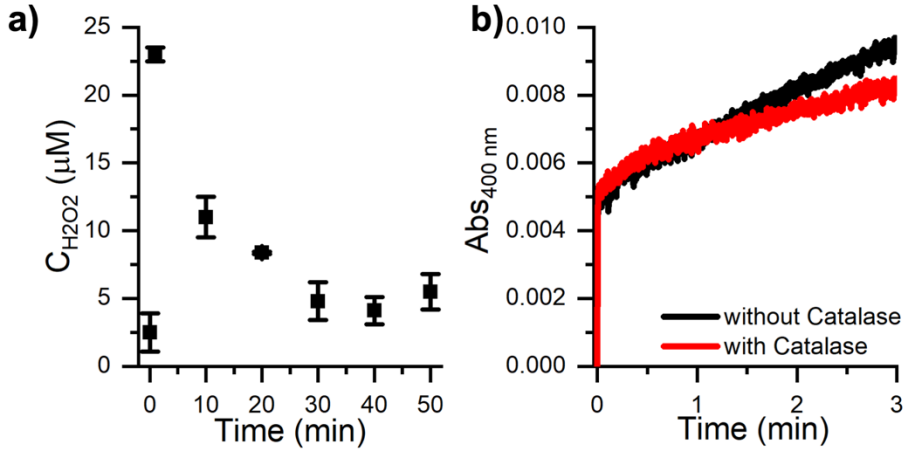


Figure 2.20. a) Hydrogen peroxide formation detected by Amplex Red ultrassay over time in presence of 50 μM *met*-DR1, 250 μM DTBC in 20 mM HEPES, 300 mM Na_2SO_4 (black squares). b) Stopped-Flow kinetic curve of quinone formation as measured at 400 nm in the absence (black line) and in the presence (red line) of 100 U/mL catalase. The *met*-DR1 concentration was 30 μM in 10 mM HEPES 300 mM Na_2SO_4 20% DMF pH 7.5.

It should be noted that the H_2O_2 could be either produced by off-track two-electrons reduction side reaction, or by the partial release of peroxo anion from the *oxy*-form. To try to detect the formation of the putative di-Cu^{2+}-DR1-oxygen species, *met*-DR1 (Figure 2.21, blue line) was fully reduced by hydroxylamine addition under nitrogen atmosphere. The reduction was confirmed by the disappearance of the d-d band at 621 nm typical of the *deoxy*-form of DR1 (Figure 2.21, black line).

Upon air exposure at room temperature, the *deoxy*-DR1 solution slowly developed a persistent dark green color. After 72 hours, the UV-Vis absorption spectrum showed three new spectral features at 300 nm, 398 nm and 600 nm (Figure 2.21, cyan line). In particular, the strong 300 nm band may be assigned to the π^*_o $(O_2^{2-}) \rightarrow d(x^2 - y^2)$ (Cu) charge transfer (CT) band. Further, the presence of a band at ≈ 400 nm has been

already observed in the studies reported by Stack et al. on *oxy*-TYR model compounds. In this analysis, UV spectroscopy supported by DFT calculation suggested that this band is expected for methyl-imidazole complexes and corresponds to a π^* (imidazole) → d(xy) (Cu) CT band.[31]

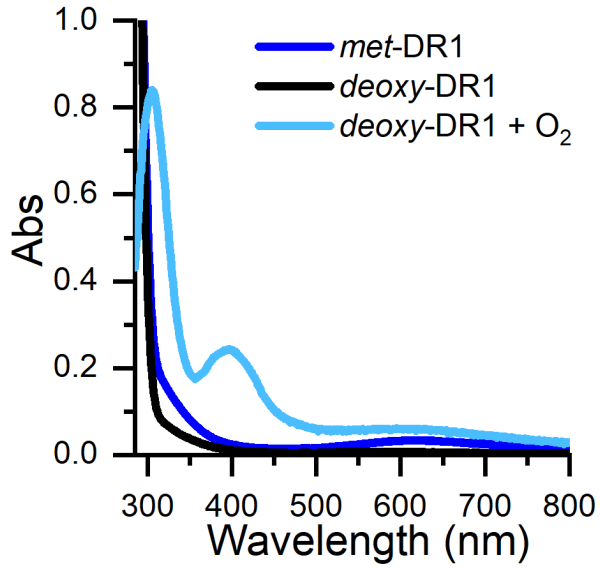


Figure 2.21. UV-Visible absorption spectra of 300 µM *met*-DR1 (blue line), 300 µM *deoxy*-DR1 in anaerobic conditions (black line) and after the restoration of aerobic conditions (cyan line). All spectra are acquired in 20 mM HEPES 300 mM Na_2SO_4 pH 7.5.

To better identify the nature of the intermediate DR1-oxygen complex, Raman spectroscopy was employed. Unfortunately, the region from about 600 to 800 cm^{-1} where the vibrations involving the oxygen-oxygen bond are observed, is featureless. This finding suggests that, if present, the intermediate di-Cu^{2+}-DR1-oxygen complex would be at a concentration below the detection limits of resonance Raman spectroscopy.

To shed light on the intermediates formed during catalysis, azide-binding experiments were performed on *met*-DR1. Indeed, azide is a

spectroscopic analog of peroxide, the natural substrate of T3 proteins, due to the striking similarity between their highest occupied orbitals. According to Solomon's classification[37] about endogenous ligands of dicopper site of type 3 copper proteins, azide belongs to group 2 ligands. While group 1 ligands (F^-, Cl^-, Br^-, $CH_3CO_2^-$) provide a ≈ 3 Å copper-copper distance that would sustain an endogenous protein bridge (and hence no EPR signal via antiferromagnetic coupling), group 2 ligands (N_3^-, SCN^-) can coordinate dicopper site to produce a > 5 Å copper-copper distance and destroy the endogenous bridge (producing an EPR detectable binuclear cupric site). Moreover, sodium azide, upon coordination to Cu(II), produces a complex with a moderate absorption in the near ultraviolet-visible region, due to N_3^- to Cu(II) charge transfer transition, whose relation with the azide coordination mode has been deeply analyzed.[38–40]

Addition of sodium azide to a solution of *met*-DR1 at pH 7.5 produces changes in both EPR and UV-Vis absorption spectra (Figure 2.22a and 2.22b). The UV-Vis absorption spectrum of *met*-DR1 + N_3^- shows a large, asymmetric band with a maximum at 380 nm and a shoulder around 450 nm (Figure 2.22a). The data were fitted with a multiple peak gaussian fitting, giving rise to two peaks at 370 nm (2000 $M^{-1}cm^{-1}$) and 420 nm (1000 $M^{-1}cm^{-1}$) (Fig 2.22a, dashed lines in green and orange, respectively). These bands should correspond to the components of the π^*_σ (N_3^-) → dx^2-y^2 (Cu) CT band, split in two by the C_2 symmetry of the system, with a μ-1,3 hapticity. The energies and the absorption coefficients agree with model compounds of copper dimers with a single μ-1,3 azide bridge and the natural dicopper proteins.[38] Moreover, N_3^- binding had an influence also on the d-d transition of *met*-DR1, causing a shift of ≈ 70 nm (from

621 nm to 690 nm), as observed in the natural T3 copper proteins Limulus Hc.[24]

The EPR spectrum (figure 2.22b), upon azide addition, shows a signal with increased intensity, related to the formation of an EPR-detectable complex.

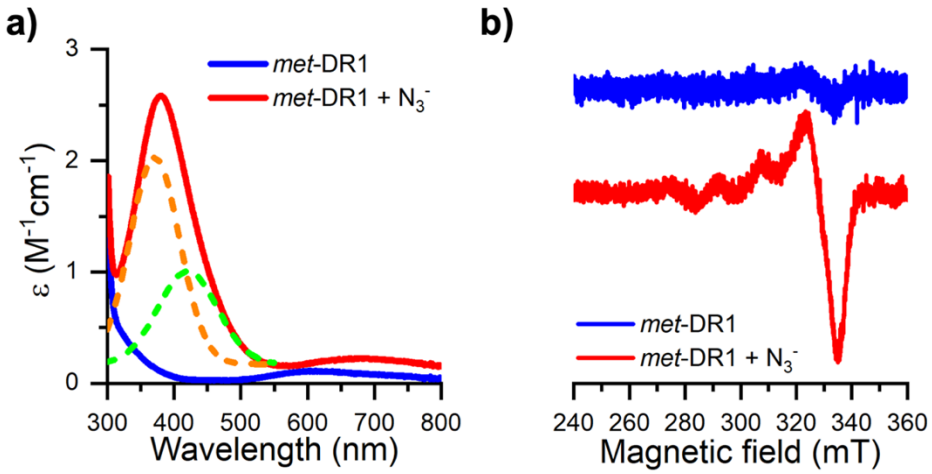


Figure 2.22. a) UV-Visible absorption spectra of 50 µM *met*-DR1 and 50 µM *met*-DR1 plus 0.8 M NaN_3, in blue and red, respectively. The gaussian component of the red line spectrum is shown with dashed orange and dashed green lines. c) X-band CW-EPR spectra of 50 µM *met*-DR1 (blue line) and upon addition of 0.8 M NaN_3. After NaN_3 addition and enhancement of ≈ 90% of the initial intensity is observed. All spectra are acquired in 20 mM HEPES 300 mM Na_2SO_4 pH 7.5.

These findings demonstrate that *met*-DR1 is able to bind and interact with exogenous ligands and suggest that the copper ions within the site may be involved in the formation of high-valent copper-oxygen intermediate during the catalytic cycle.

A detailed spectroscopic study, together with a deep analysis of the optimized experimental conditions to ascertain the presence of the DR1-oxygen complex, thus unraveling the catalytic mechanism of catechol oxidation, are under course.

2. Engineering a Type 3 Copper site into a De Novo Designed Scaffold

2.2 Experimental Section

2.2.1 Design

Sequence design of DR1 was accomplished with Accelrys Discovery Studio software.[41] The total symmetric di-Zn^{2+}-DF1 (PDB ID: 1ec5) crystal structure coordinates were relaxed with an adopted basis Newton-Raphson minimization cycle till convergence (gradient ≤ 0.1 kcal mol^{-1} $Å^{-1}$). Two distance restraints between zinc and His Nδ coordinating atom were adopted during all the minimization cycles to keep the crystallographic geometry of the metal-binding site. Di-Zn^{2+}-DF1 molecule was typed with all-atom Discovery Studio implementation of CHARMM force field,[42] a simplified Generalized Born method was adopted for solvation energies and forces,[43] long-range contributions were accounted with a 1.4 nm spherical cut-off with an exponential switching starting from 1.2 nm. From this minimized structured, all the subsequent mutations were introduced. The superposition of the designed model with the crystallographic structures of T3 copper proteins was performed with Pymol software (DeLano Scientific ltd). For the repacking of the hydrophobic core, RosettaScripts (v. 2016.32.58837) protocol was implemented.[9,10] The backbone, loop and active site sidechains were kept restrained in this phase of design not allowing the replacement of key residues and selecting amino acids depending on the environment (solvent accessible surface area, SASA) with a protocol of layer design.

2.2.2 Molecular Dynamics Simulation

All MD simulations were performed using NAMD2engine, with the CHARMM36 force field. Each designed structure was hydrated within a pre-equilibrated octahedral box with an entry edge distance of 10 Å with the CHARMM-GUI interface.[44] TIP3P water parameterization was used to describe the water molecules and the charge neutralized by potassium ions placed with the Monte Carlo method. The periodic electrostatic interactions were computed using particle mesh Ewald summation with a grid spacing smaller than 1 Å. The system has been first minimized by 10000 conjugate gradient steps and then equilibrated by using a linear temperature gradient, which heated up the system from 0 to 310.15 K in 5 ns. A constant temperature of 310.15 K was imposed by using Langevin dynamics with a damping coefficient of 1.0 ps. Constant pressure of 1 atm was maintained with Langevin piston dynamics, a 200-fs decay period, and 50-fs time constant. Free MD were performed up to 10 ns with a 2-fs integration time step using the RATTLE algorithm applied to all bonds.[42] All the mutations at the positions 9 and 13, to generate the different substrate-access analogues, were performed on the final model of the previous step of design. The first model of the NMR bundle was used as starting coordinates for DF3 MD simulation (PDB ID: 2kik). The zinc-protein interactions were described with the 'non-bonded method' to allow DR1 molecular flexibility. However, to overcome the limits of this method due to the absence of charged residues in the first coordination sphere, a sulfate anion was used as bridging ligand to implicitly constraint the position of the zinc ions. The trajectories were analyzed by the VMD molecular visualization suite.

A cluster analysis was performed to the molecular simulation trajectory of the final DR1 analogue. The GROMOS clustering algorithm,[45] implemented in g_cluster, was used to determine cluster memberships with a 0.75 Å RMSD as calculated on the Ca atoms. The centroid of the cluster was considered the structure with the greatest number of neighbors.

2.2.3 DFT calculations

Calculations have been carried out within the Density Functional Theory (DFT) framework, using the TURBOMOLE 4.1 suite of programs.[46] The pure exchange-correlation functional BP86[47,48] has been adopted, in conjunction with an SVP[49] basis set for geometry optimizations and with a TZVP[50] for subsequent single point calculations (at BP86/SVP optimized geometries) for a more accurate estimation of binding energies. The Resolution-of-Identity (RI) technique allowed to speed up geometry optimizations.[51] The Broken-Symmetry (BS) approximation has been used to model antiferromagnetic couplings, when needed.[52] In the case of *met*-form, the antiferromagnetic nature of the di-copper center was probed by performing single point calculations at the B3LYP/TZVP level[48,53] on BP86/SVP-optimized structures. The same approach was used to perform Natural Bond Orbital (NBO) analysis. D4 corrections[54] were added to geometry optimizations for an accurate treatment of dispersive forces. The hydrophobic environment within the protein active site has been modelled by setting a continuum dielectric with $\varepsilon = 4$, according to the COSMO approach.[55,56] Binding energies were calculated by neglecting the entropic contribution, since it can be reasonably approximated to a constant when considering the

binding of congeneric molecules. For this reason, relative binding energies are discussed rather than absolute ones.

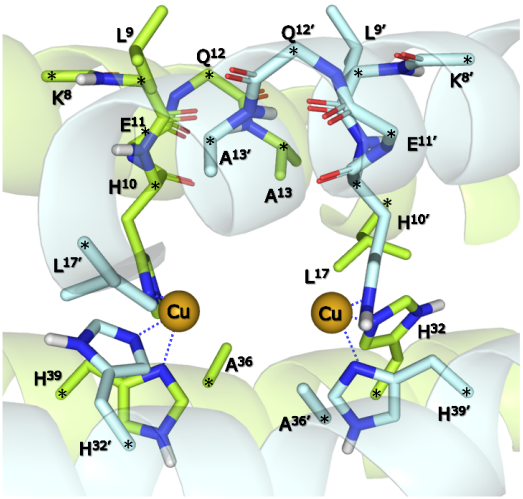


Figure 2.23. QM-cluster model used for DFT calculations. The model was built from the MD-relaxed DR1 structures, by including 18 symmetry-related residues, for a total of 252 atoms (excluding copper ions and substrates). Besides the Cu^{2+} coordination spheres composed by the two His triads, other residues have been added, that sterically define the pocket which accommodates the substrate. This would allow to rationalize the effect of both cavity shape and physico-chemical property on both binding modes and affinities. Only sidechains that are exposed at the cavity surface (i.e., that are supposed to interact with catechols) have been considered, while the others (K8, K8', E11, E11', Q18 and Q18') have been properly truncated. Alpha carbons (indicated by the * symbol) have been kept frozen to their initial position, to avoid artifacts arising from the absence of the explicit surrounding protein matrix.

2.2.4 Synthesis

The 48-mer peptide was synthesized by automatic solid-phase synthesis using an ABI 433A peptide synthesizer (Applied Biosystem, Foster City, CA, USA) with 0.25 mmol Fmoc-protocols on a 0.1 mmol scale. The N- and C-termini were acetylated and amidated, respectively. The resin used was the acid labile H-PAL ChemMatrix with a substitution of 0.31 mmol/g.

The following protected amino acids were used: Fmoc-Gly-OH, Fmoc-Glu(OtBu)-OH, Fmoc-Ile-OH, Fmoc-Thr(tBu)-OH, Fmoc-His(Trt)-OH, Fmoc-Trp(Boc)-OH, Fmoc-Lys(Boc)-OH, Fmoc-Ala-OH, Fmoc-Asp(OtBu)-OH, Fmoc-Leu-OH, Fmoc-Val- OH, Fmoc-Pro-OH, Fmoc-Asn(Trt)-OH, Fmoc-Tyr(tBu)-OH, Fmoc-Arg(Pbf)-OH and Fmoc- Gln(Trt)-OH.

Peptide chain assembly. Standard synthetic protocols were used for deprotection, activation, coupling and capping cycles.[8] Deprotection, coupling and capping steps were repeated with each subsequent amino acid, until the chain assembly was completed. The N-terminal amino group was acetylated with Ac2O/HOBt/DIEA solution in NMP. After synthesis completion, the resin was washed four times with DCM, NMP, isopropanol and methanol, and finally dried.

Cleavage and deprotection. The peptide cleavage from the resin, with concomitant side chain deprotection, was carried out using a solution of 95% TFA, 2.5% TIS, 2.5% H2O (v/v/v 40 mL). The reaction was carried out for one hour at 0 ° C and one hour at room temperature, under a slight magnetic stirring. The resin was filtered, and the solution was collected under vacuum in a glass flask, then the resin was washed again with undiluted TFA. The solution was concentrated to small volume and the crude peptide was precipitated immediately, by adding 5 volumes of cold diethyl ether respect to the peptide solution. After centrifugation, the supernatant was removed, and the precipitate was washed twice with three volumes of cold diethyl ether. The peptide was dried to remove diethyl ether, re-dissolved in water and lyophilized.

Peptide analysis and purification. The identity of the crude peptide was analyzed by analytical RP-HPLC, performed with a Shimadzu LC-10ADvp equipped with a SPDM10Avp diode-array detector. ESI-IT/TOF

spectra were recorded on a Shimadzu LCMS-IT-TOF system with ESI interface and Shimadzu LC-MS solution Workstation software for the data.

All analyses were performed on a Vydac C4 column (2.1 mm x 100 mm; 5µm), eluted with an H2O 0.05 % trifluoroacetic acid, TFA, (eluent A) and CH3CN 0.05 % TFA (eluent B) linear gradient, from 10 to 95% (solvent B), over 18.5 minutes, at 0.2 mL min^{-1} flow rate. The optimized MS parameters were selected as followed: CDL (curved desolvation line) temperature 250 °C; block temperature 250 °C; probe temperature 250 °C; detector gain 1.6kV; probe voltage +4.5kV; CDL voltage -15V. Nitrogen served as nebulizer gas (flow rate: 1.5 L min^{-1}).

Purifications were accomplished by preparative RP-HPLC with Shimadzu LC-8A connected to an SPD-20A Shimadzu UV-Vis spectrophotometric detector. Chromatographic column Vydac C4 (22 mm x 250 cm; 10 µm) was used at a flow rate of 23 mL·min^{-1}, with a linear gradient of eluent B from 10 to 80 % over 30 min, appropriately scaled up. The peak identity was verified by LC-MS. The pooled fractions, containing the desired products, were lyophilized. The peptide homogeneity was assessed by analytical HPLC and by ESI-IT-TOF mass spectrometry.

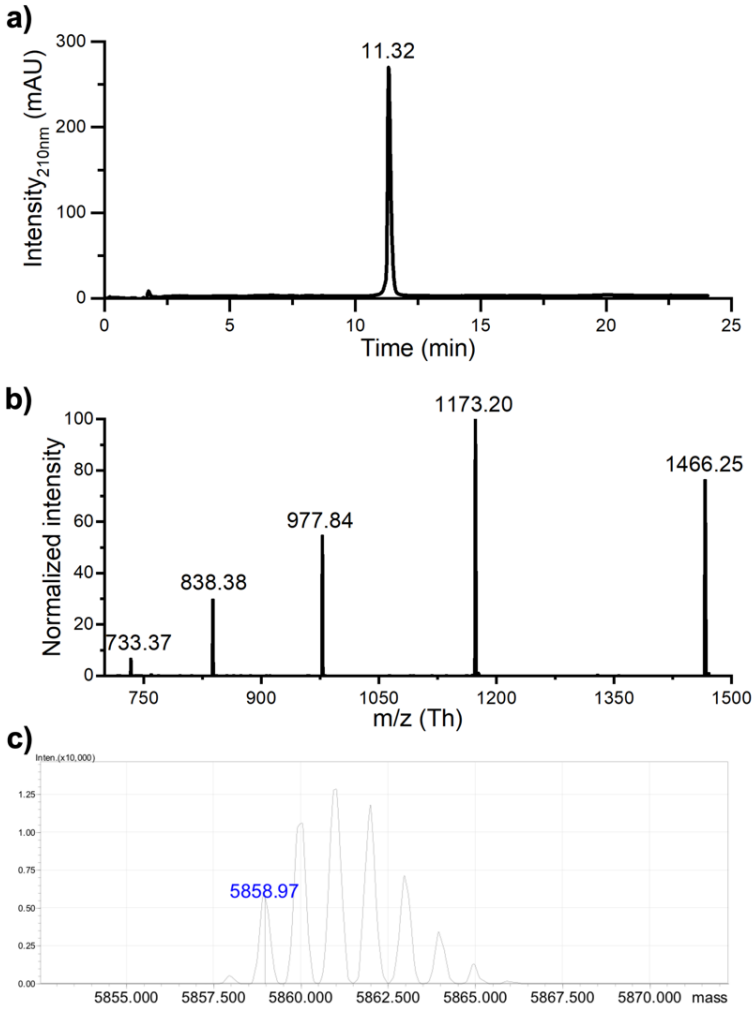


Figure 2.24. a) RP-HPLC chromatogram (210 nm trace) of pure DR1; b) ESI-TOF spectrum of pure DR1. The signal at m/z = 1466.25 Th corresponds to the $[M+4H^+]^{4+}$ ion (theoretical isotopic m/z: 1466.25 Th); the signal at m/z = 1173.20 Th corresponds to the $[M+5H^+]^{5+}$ ion (theoretical isotopic m/z: 1173.20 Th); the signal at m/z = 977.84 Th corresponds to the $[M+6H^+]^{6+}$ ion (theoretical isotopic m/z: 977.83 Th); the signal at m/z = 838.38 Th corresponds to the $[M+7H^+]^{7+}$ ion (theoretical isotopic m/z: 838.29 Th); the signal at m/z = 733.37 Th corresponds to the $[M+8H^+]^{8+}$ ion (theoretical isotopic m/z: 733.63 Th). c) Spectrum deconvolution gave (5858.97 ± 0.01) Da (theoretical isotopic mass: 5858.96 Da).

2.3.5 Gel Filtration Chromatography

Analytical gel filtration chromatography of DR1 (100 μM), (apo-form and in the presence of 1.5 equivalents $CuSO_4$, was performed on a Shimadzu LC-10ADvp, equipped with a SPDM10Avp diode-array detector. An analytical Yarra SEC-2000 column was equilibrated and eluted with a degassed solution, containing 50 mM sodium phosphate (pH 6.8) and 300 mM NaCl, at a flow rate of 0.35 ml min^{-1}. The absorbance at 280 nm was monitored. DF1 (20 μM) was also analyzed in the same experimental conditions, as standard. The injection volumes were 2 μL and 20 μL for DR1 and DF1 solutions, respectively.

SEC experiments were performed to obtain the apparent molecular weight (MW) of DR1 and DF1. Firstly, the calibration curve using the GE Low Molecular Weight (LMW) calibration kit was determined, calculating the partition coefficient K_{av} in function of log (MW). In particular, the partition coefficient K_{av} is:

$$K_{av} = \frac{V_e - V_0}{V_c - V_0}$$

where V_0 is the column void volume (2.02 mL, retention volume of Blue Dextran 2000 in Figure 2.25a), V_e is the elution volume of the protein or the standards and V_c is the geometric column volume (4.08 mL). The aprotinin (4.42 mL, elution volume) was retained by the column, since eluted after V_c, instead ovalbumin peak (3.00 mL, elution volume) was not well resolved. For these reasons, these proteins were excluded for the calibration curve. Finally, the apparent MW for DF1 and DR1 was determined, compatible with a dimeric oligomerization state (Figure 2.25b and Table 2.6).

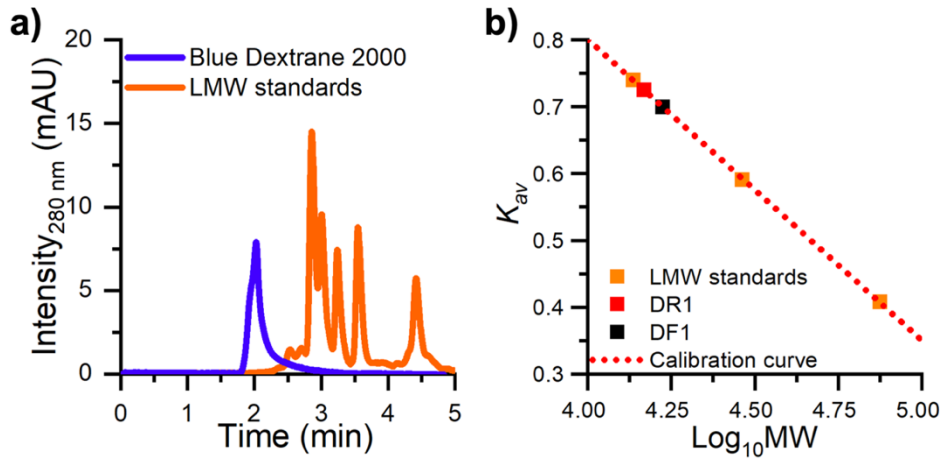


Figure 2.25. a) Analytical Gel Filtration chromatogram of the Blue Dextran 2000 (blue line) and of the LMW standards (orange line). b) Calibration curve (dashed red line) with standards proteins: Ribonuclease A, Carbonic Anhydrase and Conalbumin (in orange). DR1 and DF1 are displayed at the calculate log MW (in red and black, respectively).

Table 2.6. DR1 and DF1 gel filtration parameters, as obtained from standard samples and calibration curve (Figure 2.25b).

Protein	V_e (mL)	K_{av}	MW (kDa)
Ribonuclease A	3.55	0.74	13.7
Carbonic Anhydrase	3.24	0.59	29
Conalbumin	2.86	0.41	75
DR1	3.52	0.70	14.7
DF1	3.46	0.72	16.8

2.2.6 CD Spectroscopy

CD measurements were performed using a J-815 spectropolarimeter equipped with a thermostated cell holder (JASCO, Easton, MD, USA). CD spectra were collected at 25°C, from 260 to 190 nm at 0.2 nm

intervals with a 20 nm min-1 scan speed, at 1 nm bandwidth and at 16 s response. Cells of 0.1 mm path length were used in all measurements.

Metal ion titrations were carried out by adding small aliquots of freshly prepared aqueous stock solutions of copper sulfate to the protein solution (50 µM di-DR1, in 10 mM HEPES, 300 mM Na_2SO_4 buffer at pH 7.5), in a 1 mm quartz cell. Samples were allowed to incubate for 5 min before each measurement. DR1 binds Cu^{2+} in the expected 2:1 metal:protein ratio, with binding constant in the micromolar range. The equation for the CD data fitting was obtained by considering the following equilibrium for metal-binding:

$$P + nM \rightleftharpoons PM_n$$

Data were fitted as previously described.[8]

2.2.7 Isothermal Titration Calorimetry

ITC experiments were performed on a Nano ITC-III (TA Instrument) with a cell volume of 0.961 mL. Calorimetric titrations were carried out at 25 °C, by injecting 10 µL aliquots of a 400 µM $CuSO_4$ into a 25 µM of di-DR1 dimer solution for a total of 25 injections. Both the metal and peptide were dissolved in 10 mM HEPES, 300 mM Na_2SO_4 pH 7.5 buffer. Control experiments were performed by injecting copper into the buffer, to obtain the heat effects associated to dilution of each sample. The calorimetric enthalpy for injection was calculated, after correction for the dilution heat, by integration of the area under the peaks of ITC raw data and normalized per mol of injectant. The binding isotherm was fitted by using the model equation for independent binding site in the NanoAnalyze software (TA Instruments). The molar Gibbs energy and the molar entropy change were calculated from the classical relations:

$\Delta G_b = -RT \cdot \ln(K_b)$; $-T\Delta S_b = \Delta G_b - \Delta H_b$. The thermodynamic parameters were calculated from three independent titration experiments:
n = (1.82 ± 0.14); K_d = (5 ± 2) μM.

2.2.8 UV-Vis absorption spectroscopy

UV–Vis absorption spectra were recorded with a Cary Varian 60 Spectrophotometer, equipped with a thermostatic cell compartment (Varian, Palo Alto, CA, USA), using a quartz cuvette with 1 cm path length. Wavelength scans were performed at 25°C from 200 to 800 nm, with a 600 nm min^{-1} scan speed.

Both apo-protein and copper stock solutions were freshly prepared in H_2O milliQ. The stock solutions of $CuSO_4$ were prepared at 10 mM concentration. The initial DR1 concentration was determined spectrophotometrically by using ε_{280nm} = 7550 $M^{-1}cm^{-1}$ (per monomer). The addition of equimolar amount of $CuSO_4$ respect the apo-protein (monomer) into the buffer with apo-protein was performed slowly to avoid the precipitation of Cu^{2+} hydroxides. The formation of a d-d transition band at ≈ 621 nm (≈ 100 $M^{-1}cm^{-1}$ for dicopper site) has been used to test the correct insertion of copper into protein.

Stock solution of 50 mM NH_2OH was prepared by dissolving the solid HCl salt in 10 mM HEPES 300 mM Na_2SO_4 pH 7.5 as buffer. After the degassing of 300 μM di-DR1 solution (in 10 HEPES 300 mM Na_2SO_4 pH 7.5 as buffer) under a gentle flux of nitrogen for 60 minutes, 1 equivalent of $CuSO_4$ was added and after three minutes 5 equivalents of NH_2OH. Afterwards aerobic conditions were restored. After 72 hours, the solution was acidified to pH = 2 with concentrated H_2SO_4. The *met-* and *deoxy-* form spectrum was obtained by subtracting apo-DR1 spectrum to the one

recorded before and after the addition of NH_2OH, respectively. The spectrum of the *oxy*-form was obtained by subtracting the spectrum obtained after acidification to the spectrum recorded 72 hours after air exposure.

Stock solution of 5 M NaN_3 was prepared by dissolving the solid in 10 mM HEPES 300 mM Na_2SO_4 pH 7.5 as buffer and the pH of this solution was adjusted with H_2SO_4. Sodium azide was added to *met*-DR1 solution in 10 mM HEPES 300 mM Na_2SO_4 pH 7.5 as buffer to a final concentration of 50 µM *met*-DR1 and 0.8 M NaN_3.

2.2.9 EPR spectroscopy

CW-EPR experiments were performed on a Bruker EMX spectrometer (microwave frequency 9.44 GHz) equipped with a cylindrical cavity. All the spectra were recorded at 77 K and a microwave power of 0.68 mW, a modulation amplitude of 0.5 mT and a modulation frequency of 100 KHz. A 50 µM di-Cu^{2+}-DR1 solution, in 20 mM HEPES (pH 7.5), 300 mM Na_2SO_4 buffer was lyophilized and subsequently redissolved with 30% of glycerol, used as glassing agent, to a final copper concentration of 300 µM.

To determine the absolute amount of EPR active Cu^{2+} of DR1, a calibration curve was built. $CuSO_4$ standard solutions (50 µM, 100 µM, 200 µM and 500 µM) were prepared by dissolving copper sulfate in water-glycerol (30%) mixture. Upon performing baseline correction, the spectra were double-integrated from 250 to 350 mT and a standard curve correlating the EPR intensity with copper concentration was then generated (Figure 2.26). The EPR spectra of different DR1 concentrations were recorded using the same experimental conditions. Each experiment

was repeated twice. The copper spin concentration measured against copper sulphate standards was then inferred on the basis of the EPR signal intensity and the standard curve.

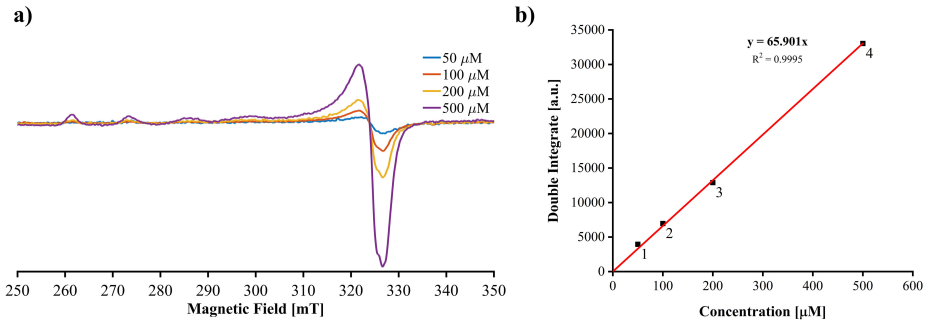


Figure 2.26. a) X-band CW-EPR spectra of $CuSO_4$ calibration curve solutions. b) Calibration curve of known title solutions of $CuSO_4$.

Reduction with hydroxylamine or DTBC was performed in anaerobic conditions, by adding 5 equivalents of reductant to the di-Cu^{2+}-DR1 solution. Catalyst re-oxidation was obtained by 18 h air exposure of the reduced samples.

Azide binding experiment was performed adding 38 µL of 5 M NaN_3 solution (in 10 mM HEPES 300 mM Na_2SO_4 pH 7.5) to 200 µL solution of 50 µM *met*-DR1 solution, obtaining a solution of 42 µM *met*-DR1 and 0.8 M N_3^-.

2.2.10 Products identification

A UV-vis spectrum of DTBQ commercially available reference sample (250 µM DTBQ in 10 mM HEPES 300 mM Na_2SO_4 pH 7.5) was recorded for comparison. The spectrum is characterized by a maximum at 410 nm (Figure 2.27).

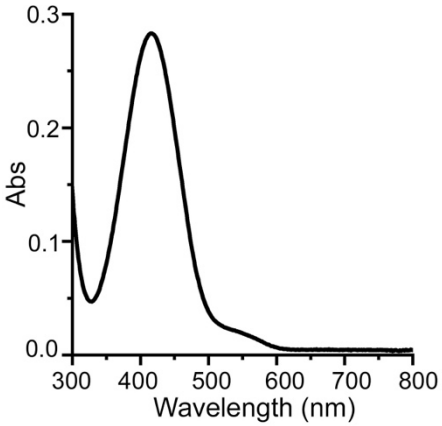


Figure 2.27. UV-Vis absorption spectrum of 250 μM DTBQ in 10 mM HEPES 300 mM Na_2SO_4 pH 7.5.

Product identity of the di-Cu^{2+}-DR1 catalyzed DTBC oxidation was confirmed by NMR analysis. To this end, the reaction mixture containing 50 μM di-Cu^{2+}-DR1 solution (0.600 mL, 10 mM HEPES 300 mM Na_2SO_4) and 5 equivalents of DTBC (stock solution 50 mM in methanol) was allowed to react for 1 hour. Afterward, the reaction was stopped by acidification and the mixture was extracted with 0.750 mL CDCl3. NMR spectrum of an authentic DTBQ sample in CDCl3 was also collected for comparison. Figure 2.28 reports the HSQC spectra of the extracted sample (a) and of the reference DTBQ (b), respectively. The general shape of the HSQC spectrum of the reaction sample extracted in CDCl3 is similar to that observed in the authentic DTBQ. The two spectra show the same resonances, thus allowing to unequivocally identify DTBQ as the reaction product and excluding the presence of any side-reaction product.

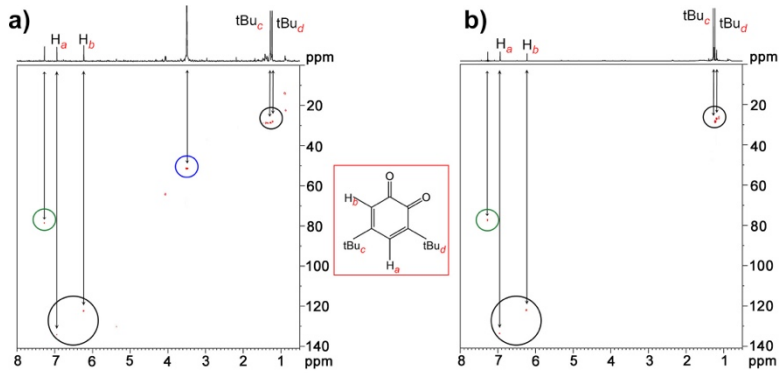


Figure 2.28. ^{1}H-^{13}C HSQC spectra of (a) DTBQ extracted in $CDCl_3$ and of (b) pure DTBQ. Black circles delimit DTQB signals (aliphatic and aromatic region), green and blue circles demarcate solvent residual peak ($CDCl_3$) and traces of CH_3OH, respectively. The DTBQ molecular structure, with protons labelled with red letters, is also shown (red contour).

^{1}H NMR ($CDCl_3$, 600 MHz): δ=1.23 (tBu_c, s); δ=1.28 (tBu_d, s); δ=6.22 (H_b, d, J = 2.3 Hz); δ=6.94 (H_a, d, J = 2.3 Hz).

2.2.11 Catalytic assays

Initial screening of the catalytic activity was performed in 10 mM HEPES 300 mM Na_2SO_4 pH 7.5 buffered solutions, at 25 °C and under magnetic stirring. Both protein and substrate stock solutions were freshly prepared. Stock solutions of DTBC were prepared by dissolving the solid in methanol to a final concentration of 50 mM. The catalytic assay was initialized by adding 5 equivalents of DTBC to 600 µL (in quartz cuvette with 1 cm path length) of 50 µM di-Cu^{2+}-DR1. Control catalytic assays were performed using 100 µM apo-DR1 or 50 µM di-Zn^{2+}-DR1 or the buffer solution (Figure 2.29).

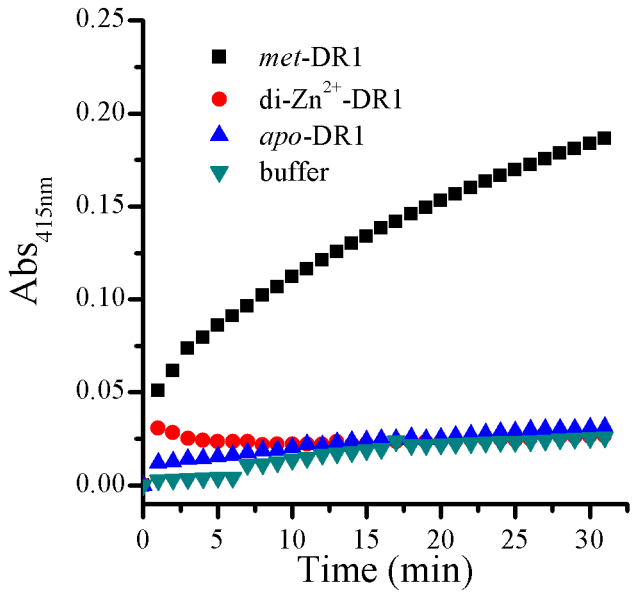


Figure 2.29. a) DTBQ formation kinetics in the presence of di-Cu^{2+}-DR1 (black symbols), di-Zn^{2+}-DR1 (red symbols), apo-DR1 (blue symbols) and only buffer (green symbols).

Stopped-flow kinetics were performed by using an SX20 stopped-flow spectrophotometer (Applied Photophysics) in single mixing mode, with a Xe Arc lamp, PEEK tubing, and a Fisher Scientific Isotemp 3016 water/ethanol bath for controlling temperature. Kinetic data were acquired for 60 s (10000 points per acquisition with external trigger). All kinetic measurements were performed at 25 °C in a 20 µL chamber with 10 mm path length. The instrument was blanked before each acquisition. The kinetic experiments were carried out in buffered solution containing 10 mM HEPES, 300 mM Na_2SO_4, 20% DMF, pH 7.5. Addition of DMF as cosolvent was required to ensure substrate solubility at saturating concentrations. The di-Cu^{2+}-DR1 concentration in all catalytic assays was fixed at 30 µM. Stock solution of substrates were prepared by dissolving the solid in DMF at 50 mM concentration and then diluting with buffer in the concentration range 0.250 mM - 1.45 mM for DTBC; 0.250 mM -

6.4 mM for 4-TBC; from 6.4 mM - 84 mM for catechol. Each point was collected at least 3 times. For the Hammett analysis the concentration of all the substrates was fixed at 0.55 mM. The ε at the maximum wavelength, reported in the following, were taken from the literature[7,57] or experimentally calculated: 1040 $M^{-1}cm^{-1}$ at λ=415 nm for 3,5-di-*tert*-butyl-o-benzoquinone, 1200 $M^{-1}cm^{-1}$ at λ =400 nm for 4-*tert*-butyl-o-benzoquinone, 1623 $M^{-1}cm^{-1}$ at λ =410 nm for o-benzoquinone, 1160 $M^{-1}cm^{-1}$ at λ =415 nm for 4-chloro-o-benzoquinone. The values for the Hammett (σ_p) and Charton (σ_v) parameters were taken from literature.[28] The data processing was performed with OriginPro 9.6.5 software. Initial rates of substrate oxidation were plotted as a function of substrate concentration and data points fitted according to the Michaelis-Menten equation, using OriginPro 9.6.5 software. From the data fitting with the Michaelis-Menten equation k_{cat} and K_M were calculated:

$$v = \frac{v_{max}[S]}{K_M + [S]}$$

Catalytic assays in presence of catalase (100 U/mL) and di-Cu^{2+}-DR1 (30 μM) and 4-TBC (3.2 mM) in 10 mM HEPES, 300 mM Na_2SO_4, 20% DMF, pH 7.5, were followed by stopped-flow spectroscopy.

2.2.12 Nuclear Magnetic Resonance spectroscopy

NMR experiments were carried out on a Bruker Avance 600-MHz equipped with a cryo probe at 298 K. Data acquisition and processing were analyzed using TOPSPIN 2.1 software (Bruker). For reaction product identification, NMR samples were dissolved in CDCl3 1H NMR spectra were registered by using 16 k data points. Heteronuclear single-quantum coherence (HSQC) experiments were performed in the ^{1}H-

detection mode by single-quantum coherence with proton decoupling in the 13C domain, using sensitivity improvement, in phase-sensitive mode with echo/antiecho gradient selection, and multiplicity editing during the selection step.[58] The experiments were performed using data points of 2048x300 points and setting one- bond heteronuclear coupling value to 145 Hz.

The NMR sample for Saturation Transfer Difference experiments (STD-NMR) contained 10 μM di-Cu^{2+}-DR1 in deuterated phosphate buffer (25 mM phosphate buffer, pD = 6.4) and DTBQ at a concentration of 100 μM (10-fold excess of ligand over protein). Time dependence of the saturation transfer was investigated using an optimized saturation times ranging from 0.2 to 5.0 s. The monoexponential function used to fit the data is defined as:

$$STD(t_{sat}) = STD_{max} \cdot (1 - e^{-k_{sat} \cdot t_{sat}})$$

Where $STD(t_{sat})$ is the STD intensity at a given saturation time, STD_{max} is the equilibrium STD intensity, and k_{sat} is the rate constant of saturation transfer. The values of STD_{max} and k_{sat} are calculated by least-square fitting to the data.

The epitope map of DTBQ binding to di-Cu^{2+}-DR1 was obtained by STD-NMR experiments. The term “binding epitope”[59] is used in the STD-NMR to refer to the hydrogens of the ligand that are closer to the protein upon binding. Figure 2.30 shows the STD amplification factor as a function of saturation time, using a 10-fold excess of DTBQ over protein.

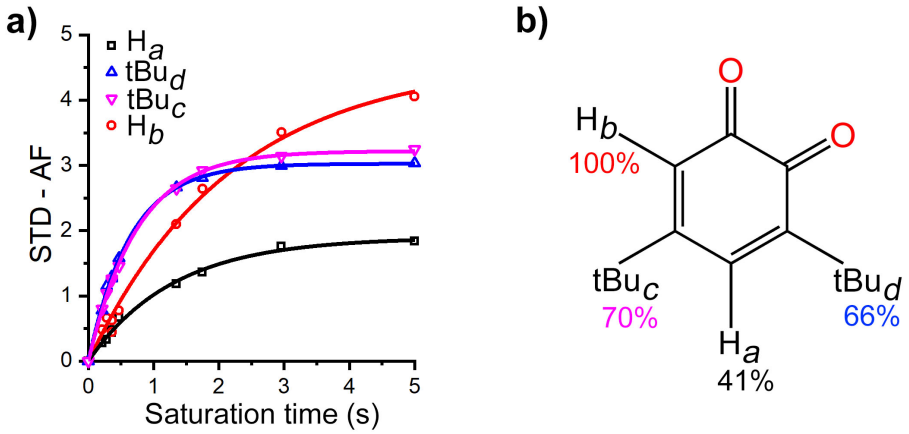


Figure 2.30. a) STD Amplification Factor as a function of saturation time. b) Binding epitope of DTBQ for its interaction with di-Cu2+-DR1, determined from STDmax normalized values reported in Table 2.7.

Table 2.7. Values of STDmax and ksat obtained by least-square fitting of the build-up curves in Figure 2.30.

	δ (ppm)	STD_{max}	k_{sat} (s^{-1})	STD_0	STD_{max} normalized
H_a	6.95	1.90 ± 0.05	0.77 ± 0.05	1.45 ± 0.10	0.41 ± 0.02
H_b	6.15	4.6 ± 0.2	0.46 ± 0.04	2.1 ± 0.2	1.00 ± 0.06
tBu_d	1.10	3.03 ± 0.02	1.55 ± 0.04	4.71 ± 0.13	0.66 ± 0.03
tBu_c	1.20	3.22 ± 0.02	1.33 ± 0.02	4.29 ± 0.07	0.70 ± 0.03

2.2.13 Amplex Red assay

Reaction conditions were the 50 μM *met*-DR1, 250 μM DTBC, 20 mM HEPES, 300 mM Na_2SO_4, pH 7.5.

Amplex Red reagent has been purchased from Thermo Fisher scientific (New York, US), and aliquoted at -20 °C in DMSO at 10 mM concentration according to the supplier's specifications. H_2O_2

determination has been performed by adapting the supplier's protocol to fully minimize the DTBC interference by means of a TECAN Spark plate reader (Tecan Trading AG, Switzerland) using a 96-well plate (Greiner non-binding PS Microplate, 96 Well, solid F-bottom (flat), chimney well with black sides). To 98 µL of working solution (50 µM Amplex Red, 0.1 U/mL HRP in 50 mM sodium phosphate buffer), 2 µL of either the standard or the reaction mixture were added and rapidly mixed for 1 minute. Fluorescence was monitored after 20 minutes from the addition. Standard solutions were prepared by consecutive dilutions of 100 µM H_2O_2 solutions in 20 mM HEPES pH 7.5, 300 mM Na_2SO_4, and a calibration curve was built (Figure 2.31). To this aim, a 3% (w/w) H_2O_2 stock solution was used, whose concentration was spectrophotometrically determined (ε_{240nm} = 43.6 $M^{-1}cm^{-1}$).

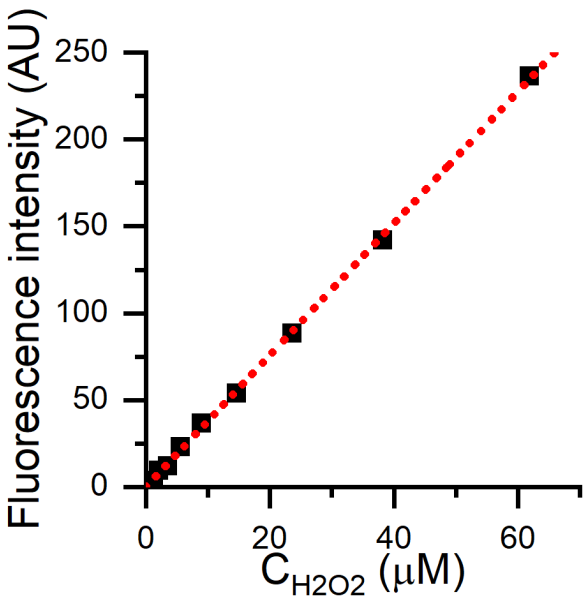


Figure 2.31. Calibration curve for H_2O_2 for Amplex Red assay.

2.2.14 Kinetic simulation of progress curves

Berkeley Madonna 9.1.19 software was employed for the curve fitting and simulations of the kinetic data. A fourth order Runge-Kutta method, with tolerance set at $1x10^{-10}$, was used to perform curve fitting. The kinetic constants obtained are the average of the kinetic constants calculated from the fitting of the data to each substrate concentration. The following kinetic model was tested.

$$met - \mathrm{DR1} + \mathrm{C} \underset{k_b}{\overset{k_a}{\rightleftarrows}} deoxy - \mathrm{DR1} + \mathrm{Q}$$

$$deoxy - \mathrm{DR1} + \mathrm{O_2} \underset{k_d}{\overset{k_c}{\rightleftarrows}} met - \mathrm{DR1} + \mathrm{H_2O_2}$$

Where C is the catechol (DTBC) and Q the corresponding quinone (DTBQ). The system of ordinary differential equations (ODEs) is as follows:

$$\frac{\mathrm{d}[deoxy - \mathrm{DR1}]}{\mathrm{dt}} = k_a[met - \mathrm{DR1}][\mathrm{C}] - k_b[deoxy - \mathrm{DR1}][\mathrm{Q}] - k_c[deoxy - \mathrm{DR1}][\mathrm{O_2}] + k_d[met - \mathrm{DR1}][\mathrm{H_2O_2}]$$

$$\frac{\mathrm{d}[\mathrm{C}]}{\mathrm{dt}} = -k_a[met - \mathrm{DR1}][\mathrm{C}] + k_b[deoxy - \mathrm{DR1}][\mathrm{Q}]$$

$$\frac{\mathrm{d}[\mathrm{Q}]}{\mathrm{dt}} = k_a[met - \mathrm{DR1}][\mathrm{C}] - k_b[deoxy - \mathrm{DR1}][\mathrm{Q}]$$

$$\frac{\mathrm{d}[\mathrm{O_2}]}{\mathrm{dt}} = -k_c[deoxy - \mathrm{DR1}][\mathrm{O_2}] + k_d[met - \mathrm{DR1}][\mathrm{H_2O_2}]$$

$$\frac{\mathrm{d}[\mathrm{H_2O_2}]}{\mathrm{dt}} = k_c[deoxy - \mathrm{DR1}][\mathrm{O_2}] - k_d[met - \mathrm{DR1}][\mathrm{H_2O_2}]$$

With boundary conditions:

[*met*-DR1] = $3.0 \cdot 10^{-5}$ (M) – [*deoxy*-DR1]

[*deoxy*-DR1] = 0 at t = 0 (s)

[C] = x at t = 0 (s)

[Q] = 0 at t = 0 (s)

$[O_2] = 2.5 \cdot 10^{-4}$ (M) at t = 0 (s)

$[H_2O_2]$ = 0 at t = 0 (s)

Where x is the concentration of DTBC substrate utilized: 125 µM, 200 µM, 330 µM, 550 µM, 725 µM.

The calculated equilibrium constant K_{cal} of the first equation of the proposed kinetic model is:

$$met\text{DR1} + \text{DTBC} \underset{kb}{\overset{ka}{\rightleftarrows}} deoxy\text{DR1} + \text{DTBQ} \qquad K_{cal} = \frac{k_a}{k_b} = 1.96 \cdot 10^{-2}$$

The equilibrium constant K_{exp} was also calculated experimentally with the following anaerobic experiment.

The DTBC oxidation catalyzed by *met*-DR1 was followed in anaerobic conditions by UV-Vis absorption spectroscopy in 10 mM HEPES, 300 mM Na_2SO_4, 20% DMF, pH 7.5. After one hour aerobic conditions were restored (Figure 2.32).

2. Engineering a Type 3 Copper site into a De Novo Designed Scaffold

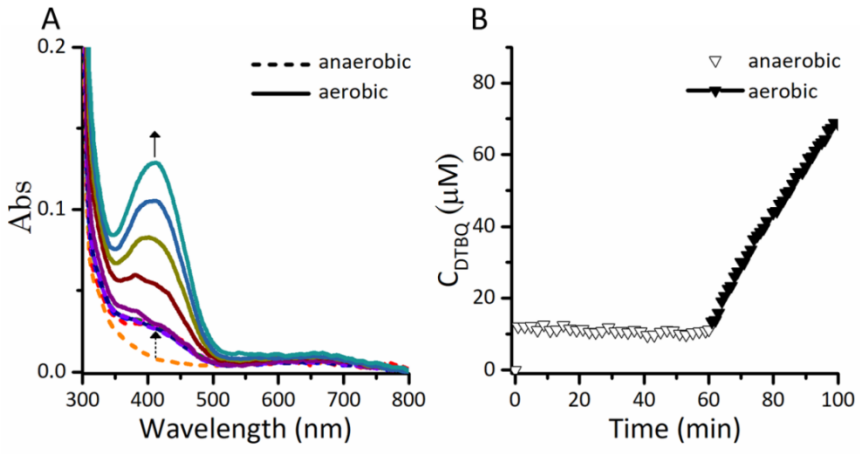


Figure 2.32. a) UV-Visible absorption spectra recorded every 10 minutes upon addition of 250 µM DTBQ to 50 µM *met*-DR1 in 10 mM HEPES, 300 mM Na_2SO_4, 20% DMF, pH 7.5 in anaerobic conditions (dashed lines) and after the restoration of aerobic conditions (solid lines). b) DTBQ formation kinetic in anaerobic conditions (white symbol) and after the restoration of aerobic conditions (black symbol).

Table 2.8. Experimental equilibrium values of concentration for the first reaction in the proposed kinetic model.

	***met*-DR1(µM)**	**DTBC (µM)**	=	***deoxy*-DR1(µM)**	**DTBQ (µM)**
Initial conditions	50	250		0	0
Variation	-13	-13		+13	+13
Equilibrium conditions	37	237		13	13

From the knowledge of initial concentrations (Table 2.8) and the ε of DTBQ, under experimental conditions, it was possible to estimate the experimental equilibrium constant K_{exp} of the reaction. The K_{exp} is $1.9 \cdot 10^{-2}$, in perfect agreement with K_{cal} ($(1.96 \pm 0.14) \cdot 10^{-2}$).

2.3 Conclusions

The first part of this PhD thesis has been devoted to the design, characterization and catalytic studies of DR1 (Due Rame 1, in Italian), an artificial metalloenzyme housing a Type 3 copper site endowed with catechol oxidase activity. To tackle this challenge, we reshaped the DF1 scaffold, an artificial diiron-protein previously developed. Starting from this protein, the first and second coordination spheres were hierarchically modified, in order to accommodate the different dimetal cofactor, as well as the residues lining the active site entrance, in order to modulate substrate binding. Spectroscopic, thermodynamic, and functional analysis demonstrated that DR1 resembles the T3 copper site, supporting different copper oxidation states and being active in the O_2-dependent oxidation of catechols to o-quinones. Hammett analysis and computational studies on substituted catechols show that DR1's substrate recognition is a result of fine design of the residues lining the substrate access site. Finally, a catalytic model was proposed and the significance of the reoxidation process in catalysis was emphasized. The results of this work demonstrate the feasibility of inserting a functional T3Cu site and tunes its activity into a de novo designed four-helix bundle protein. All this was possible thanks to the deep understanding of the simple, versatile, and structurally stable DF framework.

In conclusion, DR1 is a promising candidate for further redesign aimed at improving the catalytic properties and closing the step with the natural proteins.

References

[1] A. Lombardi, C. M. Summa, S. Geremia, L. Randaccio, V. Pavone, W. F. DeGrado, *Proc. Natl. Acad. Sci.* **2000**, *97*, 6298–6305.

[2] J. S. Sodhi, K. Bryson, L. J. McGuffin, J. J. Ward, L. Wernisch, D. T. Jones, *J. Mol. Biol.* **2004**, *342*, 307–320.

[3] A. J. Reig, M. M. Pires, R. A. Snyder, Y. Wu, H. Jo, D. W. Kulp, S. E. Butch, J. R. Calhoun, T. Szyperski, E. I. Solomon, W. F. DeGrado, *Nat. Chem.* **2012**, *4*, 900–906.

[4] S.-Q. Zhang, M. Chino, L. Liu, Y. Tang, X. Hu, W. F. DeGrado, A. Lombardi, *J. Am. Chem. Soc.* **2018**, *140*, 1294–1304.

[5] S. I. Mann, A. Nayak, G. T. Gassner, M. J. Therien, W. F. DeGrado, *J. Am. Chem. Soc.* **2021**, *143*, 252–259.

[6] O. Maglio, F. Nastri, V. Pavone, A. Lombardi, W. F. DeGrado, *Proc. Natl. Acad. Sci.* **2003**, *100*, 3772–3777.

[7] M. Faiella, C. Andreozzi, R. T. M. de Rosales, V. Pavone, O. Maglio, F. Nastri, W. F. DeGrado, A. Lombardi, *Nat. Chem. Biol.* **2009**, *5*, 882–884.

[8] M. Chino, L. Leone, O. Maglio, D. D'Alonzo, F. Pirro, V. Pavone, F. Nastri, A. Lombardi, *Angew. Chem. Int. Ed.* **2017**, *56*, 15580–15583.

[9] F. DiMaio, A. Leaver-Fay, P. Bradley, D. Baker, I. André, *PLOS ONE* **2011**, *6*, e20450.

[10] S. J. Fleishman, A. Leaver-Fay, J. E. Corn, E.-M. Strauch, S. D. Khare, N. Koga, J. Ashworth, P. Murphy, F. Richter, G. Lemmon, J. Meiler, D. Baker, *PLOS ONE* **2011**, *6*, e20161.

[11] R. T. M. de Rosales, M. Faiella, E. Farquhar, L. Que, C. Andreozzi, V. Pavone, O. Maglio, F. Nastri, A. Lombardi, *JBIC J. Biol. Inorg. Chem.* **2010**, *15*, 717–728.

[12] F. Pirro, N. Schmidt, J. Lincoff, Z. X. Widel, N. F. Polizzi, L. Liu, M. J. Therien, M. Grabe, M. Chino, A. Lombardi, W. F. DeGrado, *Proc. Natl. Acad. Sci.* **2020**, *117*, 33246–33253.

[13] M. Chino, O. Maglio, F. Nastri, V. Pavone, W. F. DeGrado, A. Lombardi, *Eur. J. Inorg. Chem.* **2015**, *2015*, 3352–3352.

[14] A. Lombardi, F. Pirro, O. Maglio, M. Chino, W. F. DeGrado, *Acc. Chem. Res.* **2019**, *52*, 1148–1159.

[15] J. Kaplan, W. F. DeGrado, *Proc. Natl. Acad. Sci.* **2004**, *101*, 11566–11570.

[16] R. A. Snyder, S. E. Butch, A. J. Reig, W. F. DeGrado, E. I. Solomon, *J. Am. Chem. Soc.* **2015**, *137*, 9302–9314.

[17] B. T. Miller, R. P. Singh, J. B. Klauda, M. Hodošček, B. R. Brooks, H. L. Woodcock, *J. Chem. Inf. Model.* **2008**, *48*, 1920–1929.

[18] M. Goldfeder, M. Kanteev, S. Isaschar-Ovdat, N. Adir, A. Fishman, *Nat. Commun.* **2014**, *5*, 4505.

[19] Y. P. Pang, K. Xu, J. E. Yazal, F. G. Prendergas, *Protein Sci.* **2000**, *9*, 1857–1865.
[20] M. Chino, S.-Q. Zhang, F. Pirro, L. Leone, O. Maglio, A. Lombardi, W. F. DeGrado, *Biopolymers* **2018**, *109*, e23339.
[21] D. Shiga, D. Nakane, T. Inomata, Y. Funahashi, H. Masuda, A. Kikuchi, M. Oda, M. Noda, S. Uchiyama, K. Fukui, K. Kanaori, K. Tajima, Y. Takano, H. Nakamura, T. Tanaka, *J. Am. Chem. Soc.* **2010**, *132*, 18191–18198.
[22] E. I. Solomon, M. J. Baldwin, M. D. Lowery, *Chem. Rev.* **1992**, *92*, 521–542.
[23] C. Eicken, F. Zippel, K. Büldt-Karentzopoulos, B. Krebs, *FEBS Lett.* **1998**, *436*, 293–299.
[24] R. S. Himmelwright, N. Eickman, C. D. LuBien, E. Solomon, K. Lerch, *J. Am. Chem. Soc.* **1980**, *102*, 5378–5388.
[25] E. I. Solomon, D. E. Heppner, E. M. Johnston, J. W. Ginsbach, J. Cirera, M. Qayyum, M. T. Kieber-Emmons, C. H. Kjaergaard, R. G. Hadt, L. Tian, *Chem. Rev.* **2014**, *114*, 3659–3853.
[26] I. A. Koval, K. Selmeczi, C. Belle, C. Philouze, E. Saint-Aman, I. Gautier-Luneau, A. M. Schuitema, M. van Vliet, P. Gamez, O. Roubeau, M. Lüken, B. Krebs, M. Lutz, A. L. Spek, J.-L. Pierre, J. Reedijk, *Chem. – Eur. J.* **2006**, *12*, 6138–6150.
[27] A. Bencini, A. Dei, C. Sangregorio, F. Totti, M. G. F. Vaz, *Inorg. Chem.* **2003**, *42*, 8065–8071.
[28] Corwin. Hansch, A. Leo, R. W. Taft, *Chem. Rev.* **1991**, *91*, 165–195.
[29] H. W. Duckworth, J. E. Coleman, *J. Biol. Chem.* **1970**, *245*, 1613–1625.
[30] J. L. Muñoz-Muñoz, M. del M. García-Molina, F. Garcia-Molina, J. Berna, P. A. Garcia-Ruiz, M. García-Moreno, J. N. Rodriguez-Lopez, F. Garcia-Canovas, *J. Mol. Catal. B Enzym.* **2013**, *91*, 17–24.
[31] C. Citek, S. Herres-Pawlis, T. D. P. Stack, *Acc. Chem. Res.* **2015**, *48*, 2424–2433.
[32] E. I. Solomon, U. M. Sundaram, T. E. Machonkin, *Chem. Rev.* **1996**, *96*, 2563–2606.
[33] M. Güell, P. E. M. Siegbahn, *JBIC J. Biol. Inorg. Chem.* **2007**, *12*, 1251–1264.
[34] C. Eicken, B. Krebs, J. C. Sacchettini, *Curr. Opin. Struct. Biol.* **1999**, *9*, 677–683.
[35] J.-P. Chyn, F. L. Urbach, *Inorganica Chim. Acta* **1991**, *189*, 157–163.
[36] O. Karakuzu, M. R. Cruz, Y. Liu, D. A. Garsin, *Bio-Protoc.* **2019**, *9*, e3409.
[37] E. I. Solomon, *Pure Appl. Chem.* **1983**, *55*, 1069–1088.
[38] J. E. Pate, P. K. Ross, T. J. Thamann, C. A. Reed, K. D. Karlin, T. N. Sorrell, E. I. Solomon, *J. Am. Chem. Soc.* **1989**, *111*, 5198–5209.
[39] L. Casella, O. Carugo, M. Gullotti, S. Garofani, P. Zanello, *Inorg. Chem.* **1993**, *32*, 2056–2067.
[40] F. Tuczek, E. I. Solomon, *Inorg. Chem.* **1993**, *32*, 2850–2862.
[41] A. D. Studio, Accelrys Discov. Studio Model. Environ. Release **2012**, 3.

[42] B. R. Brooks, C. L. Brooks, A. D. Mackerell, L. Nilsson, R. J. Petrella, B. Roux, Y. Won, G. Archontis, C. Bartels, S. Boresch, A. Caflisch, L. Caves, Q. Cui, A. R. Dinner, M. Feig, S. Fischer, J. Gao, M. Hodoscek, W. Im, K. Kuczera, T. Lazaridis, J. Ma, V. Ovchinnikov, E. Paci, R. W. Pastor, C. B. Post, J. Z. Pu, M. Schaefer, B. Tidor, R. M. Venable, H. L. Woodcock, X. Wu, W. Yang, D. M. York, M. Karplus, *J. Comput. Chem.* **2009**, *30*, 1545–1614.
[43] W. Im, M. S. Lee, C. L. Brooks, *J. Comput. Chem.* **2003**, *24*, 1691–1702.
[44] J. Lee, X. Cheng, J. M. Swails, M. S. Yeom, P. K. Eastman, J. A. Lemkul, S. Wei, J. Buckner, J. C. Jeong, Y. Qi, S. Jo, V. S. Pande, D. A. Case, C. L. Brooks, A. D. MacKerell, J. B. Klauda, W. Im, *J. Chem. Theory Comput.* **2016**, *12*, 405–413.
[45] W. R. P. Scott, P. H. Hunenberger, I. G. Tironi, A. E. Mark, S. R. Billeter, J. Fennen, A. E. Torda, T. Huber, P. Kruger, W. F. van Gunsteren, *J. Phys. Chem. A* **1999**, *103*, 3596–3607.
[46] R. Ahlrichs, M. Bär, M. Häser, H. Horn, C. Kölmel, *Chem. Phys. Lett.* **1989**, *162*, 165–169.
[47] J. P. Perdew, *Phys. Rev. B* **1986**, *33*, 8822–8824.
[48] A. D. Becke, *Phys. Rev. A* **1988**, *38*, 3098–3100.
[49] A. Schäfer, H. Horn, R. Ahlrichs, *J. Chem. Phys.* **1992**, *97*, 2571–2577.
[50] A. Schäfer, C. Huber, R. Ahlrichs, *J. Chem. Phys.* **1994**, *100*, 5829–5835.
[51] K. Eichkorn, F. Weigend, O. Treutler, R. Ahlrichs, *Theor. Chem. Acc.* **1997**, *97*, 119–124.
[52] L. Noodleman, *J. Chem. Phys.* **1981**, *74*, 5737–5743.
[53] C. Lee, W. Yang, R. G. Parr, *Phys. Rev. B* **1988**, *37*, 785–789.
[54] E. Caldeweyher, J.-M. Mewes, S. Ehlert, S. Grimme, *Phys. Chem. Chem. Phys.* **2020**, *22*, 8499–8512.
[55] M. M. Karelson, M. C. Zerner, *J. Phys. Chem.* **1992**, *96*, 6949–6957.
[56] A. Klamt, *J. Phys. Chem.* **1995**, *99*, 2224–2235.
[57] J. L. Muñoz, F. García-Molina, R. Varón, J. N. Rodriguez-Lopez, F. García-Cánovas, J. Tudela, *Anal. Biochem.* **2006**, *351*, 128–138.
[58] D. J. States, R. A. Haberkorn, D. J. Ruben, *J. Magn. Reson. 1969* **1982**, *48*, 286–292.
[59] B. Meyer, T. Peters, *Angew. Chem. Int. Ed Engl.* **2003**, *42*, 864–890.

3. Mimicking the Histidine Brace motif in De Novo Designed Proteins

---

Keywords: Artificial Metalloenzymes • Lytic Polysaccharide Monooxygenases • Protein Design • Electron Paramagnetic Resonance Spectroscopy • Histidine Brace

In this chapter, divided into two subsections, the de novo design, spectroscopic and functional characterization of two HB-containing metalloenzymes is reported. In particular, section A focus on miniLPMO, and section B focus on dHisB.

.

3A. Mimicking the Histidine Brace motif in De Novo Designed Proteins

3A.1 Results and Discussion

3A.1.1 Design strategy of miniLPMO

MiniLPMO is an engineered homodimeric small protein composed of two peptide monomers that self-associate to host 2 HB metal-binding sites per homodimer. The design started from the NMR solution structure of α2D (PDB ID: 1qp6).[1,2] α2D is a four-helix bundle formed by two 35-residue peptides, which dimerizes in an unusual "bisecting" topology (Figure 3A.1a).[3–5] The HB site of the natural AA10 LPMO from *Teredinibacter turnerae* (PDB ID: 6rw7)[6] was chosen as reference.

To install the metal-binding site into the best position of the α2D template, all possible backbone-dependent rotamers for the internal His (His107 of the natural LPMO) were superimposed on the α2D residues within 10 Å of the N-terminus. Ile 22 in α2D was determined to be the best position by placing the terminal His residue of HB very close to the original second residue on the N-terminal helix (Figure 3A.1b), consequently requiring only the removal of the first residue from α2D and minimal backbone rearrangement for the HB construction. Next, a backbone geometry optimization for terminal His position using the MASTER (*Method of Accelerated Search for Tertiary Ensemble Representatives*) software was performed. The best fragments obtained from the MASTER search (figure 3A.1c), ranked according to their RMSD from the query structure, were then individually examined and one of the best search results was implanted into α2D to reconstruct the N-terminal helix while ensuring optimal backbone geometry for the HB site.

Further, the following reasons led us to reconstruct the loop:

- four-helix bundles with bisecting topology are very sensitive to loop residues composition, as the, demonstrated by Ala 31 to Pro mutation in the ROP protein;[7]
- the removal of the two initial residues from α2D causes the loop to be more exposed to the solvent respect to starting structure;
- the loop region is close enough to the HB metal-binding site to inevitably exert an influence, both in terms of steric hindrance and primary/secondary coordination spheres.

The loop region was reconstructed by using MASTER to fill the gap in a query structure formed by two helix segments taken from the monomer of α2D. The best matches (Figure 3A.1d) were mostly similar in backbone geometry, and one match seemed particularly appropriate thanks to the presence of a Gln residue in position 16. Such residue is close enough to the metal site, thus resembling Gln167 side chain often found in the second coordination sphere of natural LPMO (*Mt*PMO3*).[8] This residue, sometimes replaced by a Glu in different strains, appears to be implicated in hydrogen bonding either to dioxygen or hydrogen peroxide molecules during the catalytic cycle, and is probably also important for specific interactions with the substrate polysaccharide.[9]

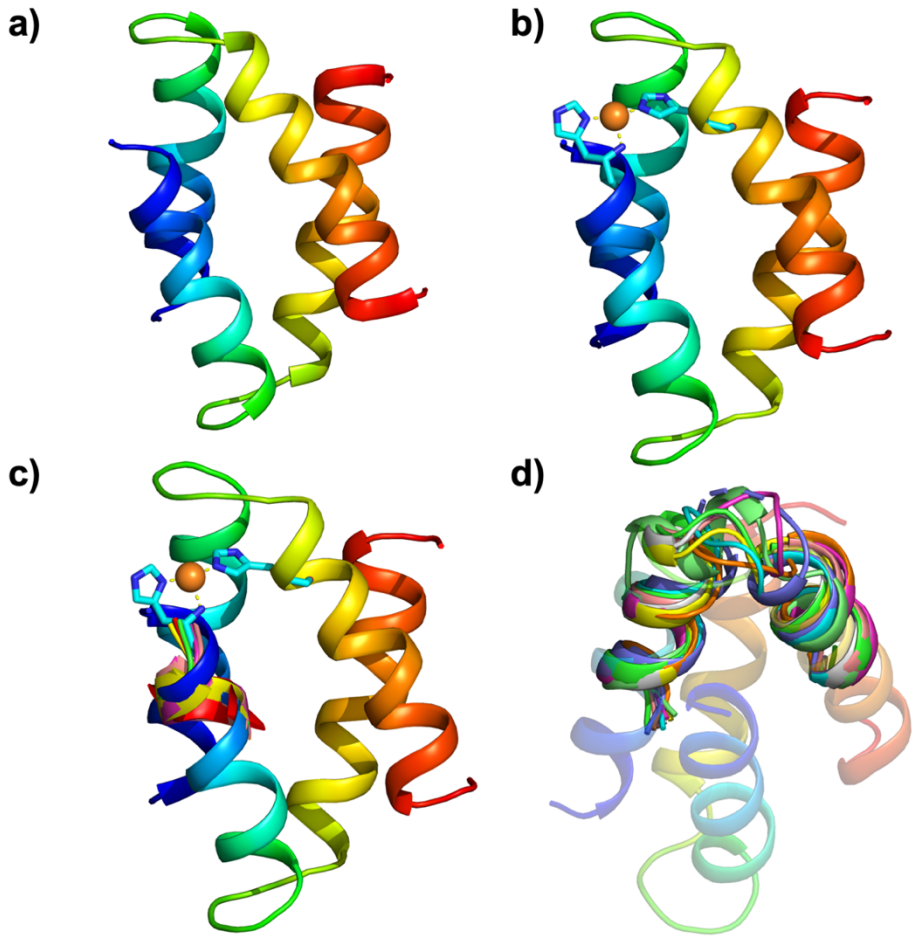


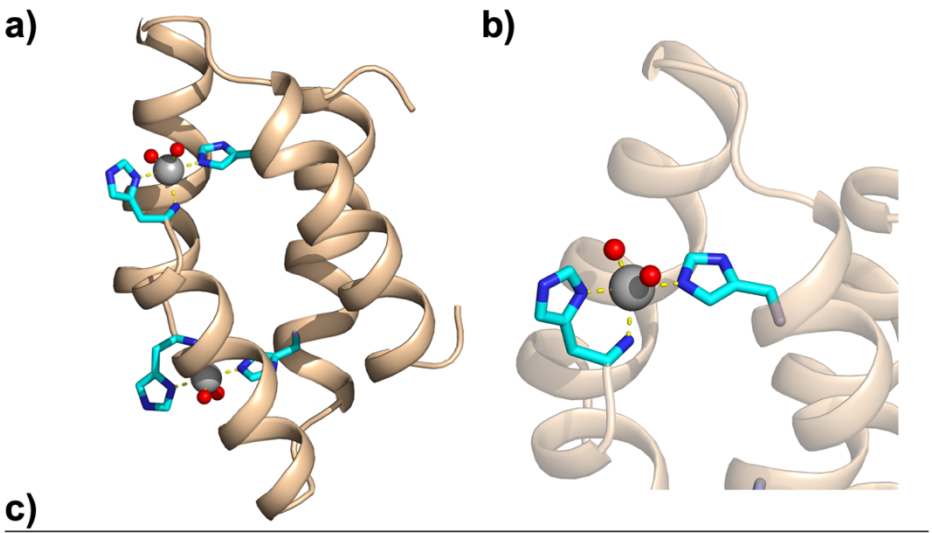
Figure 3A.1. a) α2D NMR structure (PDB ID: 1qp6, structure 3 of NMR bundle). b) Superposition of the best HB backbone-dependent rotamer (depicted as cyan sticks and copper as orange sphere) on the α2D structure. c) α2D N-terminal segment reconstruction by structural search. Only the five highest ranked matches are shown as cartoon. d) Loop redesign in α2D by structural search. Only the nine highest ranking matches are shown as cartoon. In all panels the cartoons are colored from blue (N terminal) to red (C terminal).

Finally, the backbone coordinates obtained by implanting the HB site, shortening and rebuilding the N-terminal helix and reconstructing the loop, were employed as starting point for computational design through the ROSETTA software suite. As a further validation method for the designed structure, MD simulations were performed on the lowest energy

sequences, as quantified by the Ref2015 score function in ROSETTA. Cu^{2+} was replaced with Zn^{2+} for the MD simulation, as the MD parameterization of zinc is relatively more reliable in the CHARMM36m forcefield. In addition, zinc substitution in natural LPMOs does not significantly alter the orientation of the first coordination sphere.[10] The "non-bonded approach" was used to characterize zinc-protein interactions to allow for molecular flexibility.

Coordination sphere analysis of homodimers during MD simulations suggested that long sidechain coordinating residues (as Glu and Gln) should be avoided in position 16 and 19. Indeed, shorter sidechain residues were preferred, able to form hydrogen bond with water in the first coordination sphere as found in natural LPMO, but simultaneously avoiding unwanted coordination to the metal. Asn in position 16 and Asp in position 19 were chosen and the final model of miniLPMO was obtained (Figure 3A.2a and 3A.2b). This sequence (Figure 3A.2c) was subsequently synthesized and characterized.

3A. Mimicking the Histidine Brace motif in De Novo Designed Proteins



		Helix 1																				Helix 2													
		a	*b*	*c*	*d*	*e*	*f*	*g*	*a*	*b*	*c*	*d*	*e*	*f*								*a*	*b*	*c*	*d*	*e*	*f*	*g*	*a*	*b*	*c*	*d*	*e*	*f*	*g*
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	
α2D	G	E	V	E	E	L	E	K	K	F	K	E	L	W	K	G	P	R	R	G	E	I	E	E	L	H	K	K	F	H	E	L	I	K	G
miniLPMO			H	E	R	A	W	R	D	F	K	S	G	I	A	A	G	N	E	A	D	H	K	R	F	V	E	E	F	I	N	A	V	N	G

Figure 3A.2. a) Zn^{2+}-miniLPMO minimized model with two HB per homodimer. b) zoom of the HB active site in Zn^{2+}-miniLPMO model (depicted as cyan sticks, zinc as grey sphere and water molecules as red spheres). c) Peptide sequence alignment of miniLPMO with α2D. For miniLPMO sequence, coordinating residues are reported in cyan.

3A.1.2 miniLPMO folding and oligomerization state

MiniLPMO monomeric sequence was synthesized by solid-phase peptide synthesis and purified to homogeneity by RP-HPLC (see Experimental Section for details).

CD spectroscopy was adopted to verify the helical global folding of miniLPMO at pH 7.5. At 10 μM miniLPMO concentration (Figure 3A.3a, blue line), the CD spectrum reveals a low helical content, being characterized by a deep minimum at λ = 201 nm and a MRE of 3.8 kdeg cm^2 res^{-1} $dmol^{-1}$. By increasing the miniLPMO concentration at 300 μM, the CD spectrum (Figure 3A.3a, orange line) changed, resembling a more regular conformation, as indicated by: (i) shift of the λ_{min} from 201 nm to

205 nm, (ii) shift of λ_0 towards higher wavelengths, (iii) increase in the MRE at 222 nm. These data assume an oligomerization process, as observed in the formation of other de novo designed coiled coils.[11] To this end, the oligomerization process stoichiometry and the thermodynamic driving force were evaluated. Fitting of the concentration dependence of the MRE at 222 nm (Figure 3A.3b), with a simple n-oligomerization model, gave a stoichiometry of (2.03 ± 0.11) and a K_{diss} of (7 ± 2) μM, similar to α2D, the template scaffold of miniLPMO.[1]

Even though miniLPMO is poorly folded, these findings demonstrate the correct oligomerization state, as predicted by design.

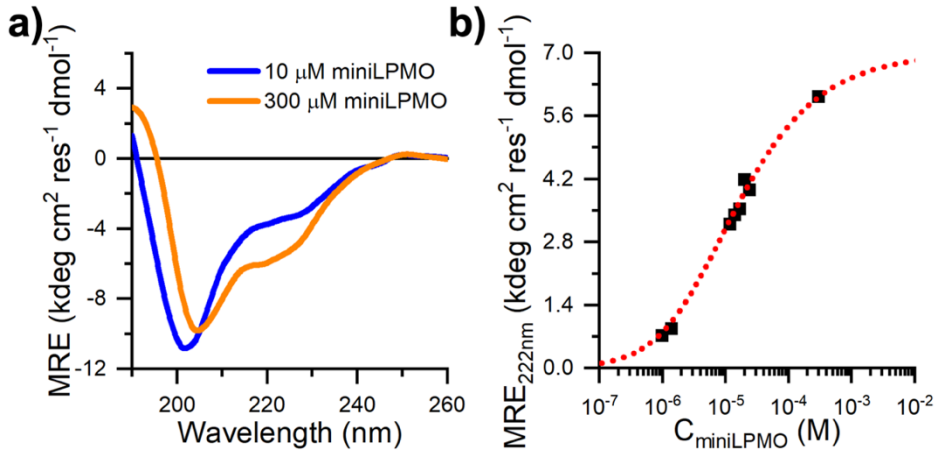


Figure 3A.3. a) Comparison of ultraviolet CD spectra of miniLPMO at 10 μM (blue line) and 300 μM (orange line) in 2 mM HEPES pH 7.5. b) Dependence of the MRE at 222 nm as a function of miniLPMO concentration in 2 mM HEPES at pH 7.5. The dashed red line represents the best fit for the oligomerization equation from monomer to dimer.

3A.1.3 miniLPMO copper binding

The binding of copper to miniLPMO has been characterized using EPR and UV-Vis absorption spectroscopy. The designed first coordination sphere of copper in the miniLPMO model contains

numerous ligating atoms with ionisable protons (e.g., His, NH_2, OH_2). To evaluate their pKa values, a pH titration in the range 2.0–11.0 (in 0.5 pH unit increments) was performed and followed by EPR and UV-Visible absorption spectroscopy. Different samples of Cu^{2+}-miniLPMO in a mixed buffer solution containing 15 mM MES, 15 mM HEPES, 15 mM CHES, 30% glycerol (for EPR measurements) were prepared. These buffers were chosen because of their weak interaction with metal ions and when used in combination, they can buffer a broad pH range. To adjust the pH of the solutions, diluted NaOH or H_2SO_4 solutions were added. The X-band CW-EPR spectra of the sample were obtained at each pH by rapidly cooling the sample to 77 K.

Copper-miniLPMO complex formation was found to be highly pH-dependent, as expected. Figure 3A.4a and 3A.4b show the EPR and visible absorption spectra recorded at different pHs, in the range pH 2.0-11.0.

Raising the pH from 2.0 to 11.0, it is observed a qualitatively change in the EPR spectra over the pH range (Figure 3A.4a), with all species displaying a spectral envelope consistent with a prominent $d(x^2 - y^2)$ Single Occupied Molecular Orbital (SOMO), indicating that the copper ion at all pHs has a ligand field dominated by σ-donation in its equatorial plane. Across the pH four different species are observed as they can be simulated with single set of spin Hamiltonian parameters (Table 3A.1). The EPR spectra at pH 2.0, 4.5, 7.0, and 11.0 were simulated with a single set of spin Hamiltonian parameters. Spectra taken at other pH values can be satisfactory simulated by a weighted combination of these spectra. The contribution of each spectrum was determined through least square fitting of the experimental and calculated spectra (see

Experimental Section for details), resulting in the determination of the speciation distribution curves and the pKa_s (Figure 3A.4c). Similar procedure was performed for UV-Vis absorption data (Figure 3A.4d and Figure 3A.4e)

The EPR and Visible absorption spectra recorded at pH 2.0 have the fingerprint of tetragonal aqua ion copper[12] (Figure 3A.4a and 3A.4b, dark green line in bold, and Table 3A.1) and herein referred as species 1. The first deprotonation event with subsequent coordination of copper to miniLPMO occurs after pH 3.0 and a dominant species is observed at pH 4.5 (Figure 3A.4a and 3A.4b, green line in bold). The pKa_s for this species obtained by EPR (Figure 3A.4c) and visible absorption spectroscopy (Figure 3A.4d) are very similar (4.0 and 4.3, respectively) and both involve two protons. This deprotonation event could be assigned to the deprotonation of two His residues, as the pKa is in good agreement with His pKa values in LPMOs.[13] For this species, herein referred as species 2, the EPR spectrum is simulated with a single set of g_3 and A_3 values (g_3 = 2.268; $|A_3|$ = 535 MHz and Table 3A.1) characteristic of a 2N2O coordination sphere. Moreover, the visible absorption spectra show a d-d transition band, whose λ_{max} and ε (Table 3A.1) are in good agreement with a $Cu(His)_2$ center with bound waters.[14]

As the pH rises, a further deprotonation event happens and a new species is formed (Figure 3A.4a and Figure 3A.4b, yellow line in bold). The pKa_2 value (5.5 as estimated in figure 3A.4c) could suggest that the amine nitrogen is included in the coordination sphere. Precipitation and scattering phenomena made precise assignment for this species and calculation of pKa and stoichiometry by visible absorption spectroscopy complicated, although a good simulation of the experimental EPR

spectrum of this species was obtained with g_3 and A_3 values (g_3 = 2.251; $|A_3|$ = 550 MHz and Table 3A.1) typical for a 3N1O coordination sphere and more in particular in good agreement with natural LPMOs' g values.[9,15] Starting from pH 8.5, the Cu^{2+}-miniLPMO complex becomes again soluble and a new species is fully formed at pH 11.0 (Figure 3A.4a and Figure 3A.4b, red line in bold). The formation of this species is accompanied by dramatic changes in the low field region of the EPR spectrum and a blue-shift in the d-d transition energy. A good simulation of the experimental EPR spectrum was obtained with the combination of a very large A_3 value (625 MHz) and a low g_3 value (2.169), typical for a 3N1O or 4N coordination sphere (Table 3A.1). Considering the g_3 value as a measure for the relative strengths of the ligand field in the equatorial plane of the copper, it is evident that the ligand field dramatically rises when going from species 2 to 4 (g_3 from 2.268 to 2.169). This finding is also confirmed by the high-energy shift of the d-d transition band. The pKa of this process is 8.6, calculated by EPR and visible absorption spectroscopy (Figure 3A.4c and 3A.4e, respectively), and by visible absorption interpretation data it involves two-protons. The pKa can be tentatively assigned to the deprotonation of the amine terminal to azanide (-NH) species,[15] of amide (backbone) or of a water molecule bounded on the copper ion. Based on these data, the formation of a species 4 with an azanide/amide and a hydroxide molecule in the first coordination sphere can be inferred.

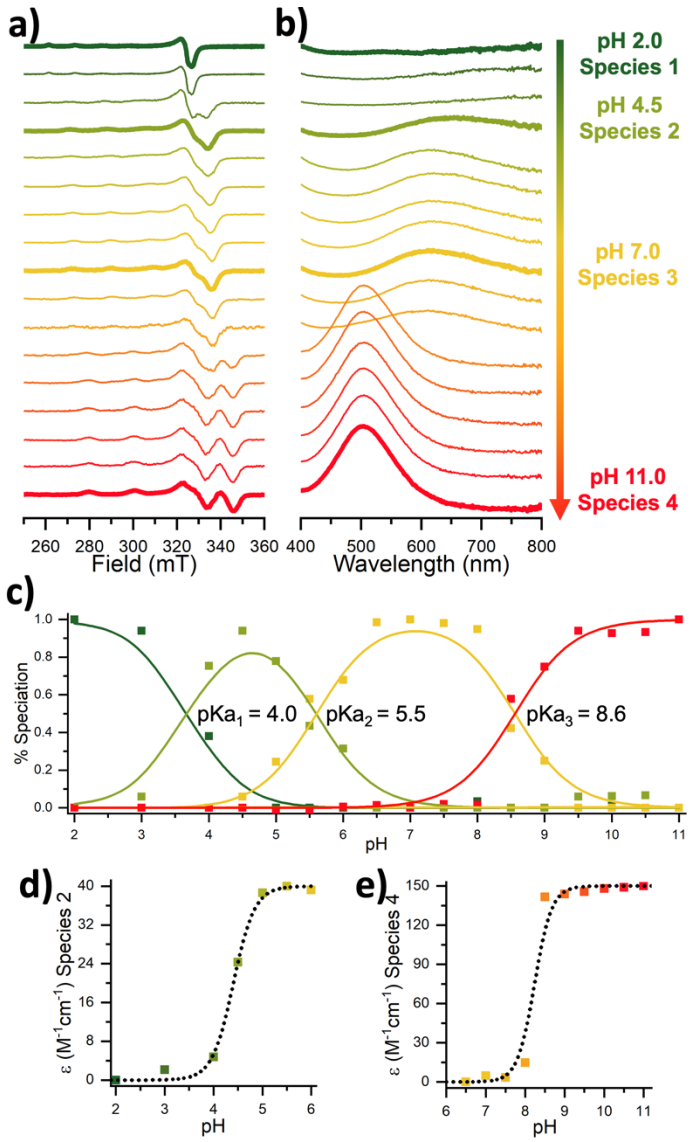


Figure 3A.4. a) Frozen solution X-band CW-EPR spectra (77 K) of Cu^{2+}-miniLPMO (0.3 mM) and b) visible absorption spectra of Cu^{2+}-miniLPMO (0.2 mM) between pH 2.0 and 11.0. Both EPR and Visible absorption spectra were recorded every 0.5 pH unit. The buffer chosen was the mixed solution of MES, HEPES and CHES (15 mM each. c) Speciation distribution curves of the different 4 species calculated by EPR spectroscopy. d) Speciation distribution curves for species 2 and e) for Species 4 calculated by visible absorption spectroscopy. The dashed lines represent the best fit for a classic acid-base equilibrium, considering for both a two-proton event.

Table 3A.1. Spin Hamiltonian parameters of the pH dependent species. Hyperfine tensor principal values are given in units of MHz.

Species	*g* values			Hyperfine coupling constants (MHz)			Coordination mode from P-B plot[16]	λ_{max} (nm) d-d electronic transition (ε in $M^{-1}cm^{-1}$)
	g_1	g_2	g_3	$\lvert A_1 \rvert$	$\lvert A_2 \rvert$	$\lvert A_3 \rvert$		
1 (pH 2.0)	2.080	2.080	2.415	5	5	410	4O	> 850
2 (pH 4.5)	2.055	2.055	2.268	60	60	535	2N2O	650 (50)
3 (pH 7.0)	2.052	2.058	2.251	40	40	550	3NO1	~620 (~60)
4 (pH 11.0)	2.040	2.045	2.169	35	50	625	3N1O/4N	505 (150)

3A.1.4 ^{14}N ENDOR and HYSCORE spectroscopy

The nitrogen superhyperfine coupling is too small to be detected by standard CW-EPR experiment. This coupling can however be measured using hyperfine spectroscopy. In particular, the strongly coupled ^{14}N of the imidazole ring (from His residues) are best obtained by means of ENDOR spectroscopy, whereas remote ^{14}N of the imidazole ring (from His residues) can be detected by HYSCORE (HYperfine Sublevel CORrElation) spectroscopy. When combined, these methods provide a detailed description of the copper coordination environment. ^{14}N ENDOR and HYSCORE spectra for species 2 (pH 4.5) and 4 (pH 11.0) were

acquired. Unfortunately, due to low solubility and precipitation phenomena, it was impossible to record the ^{14}N ENDOR and HYSCORE spectra for species 3 (pH 7.0).

The Q-band Davies ^{14}N ENDOR spectrum of species 2 (Figure 3A.5a), is well simulated with one kind of nitrogen with $|A|$ values, nuclear quadropolar coupling (e^2qQ/h) and asymmetry parameter (η) typical of a sp^2 nitrogen in equatorial position (Table 3A.2).[17–19]

Different results are obtained for species 4. The Q-band Davies ^{14}N ENDOR spectrum of the species 4 recorded at a magnetic field corresponding to the g_3 component (1084 mT, Figure 3A.5c) shows the presence of two overlapping doublets separated by $2\nu_N$ and centered at $A/2$. Orientationally selected experiments (data not shown) could be simulated considering two different nitrogen nuclei, with different hyperfine couplings (Table 3A.2). The presence of two different coordinated nitrogen nuclei together with the change in the g values of the EPR spectrum and the onset of g rhombicity, call for a change in the coordination geometry at this pH.

The X-band HYSCORE spectrum (Figure 3A.5b) is characterized by cross peaks in the (+,+) and (–,+) quadrant, arising from double quantum transitions associated to remote ^{14}N nuclei of the imidazole ring. No difference is observed at the two pHs, confirming the imidazole coordination at both pHs (Figure 3A.5d).

These findings confirm the completely different coordination environment between species 2 and species 4 and make possible the hypothesis of specie 3 with a HB-like coordination environment, even if further investigations are needed to confirm this hypothesis.

3A. Mimicking the Histidine Brace motif in De Novo Designed Proteins

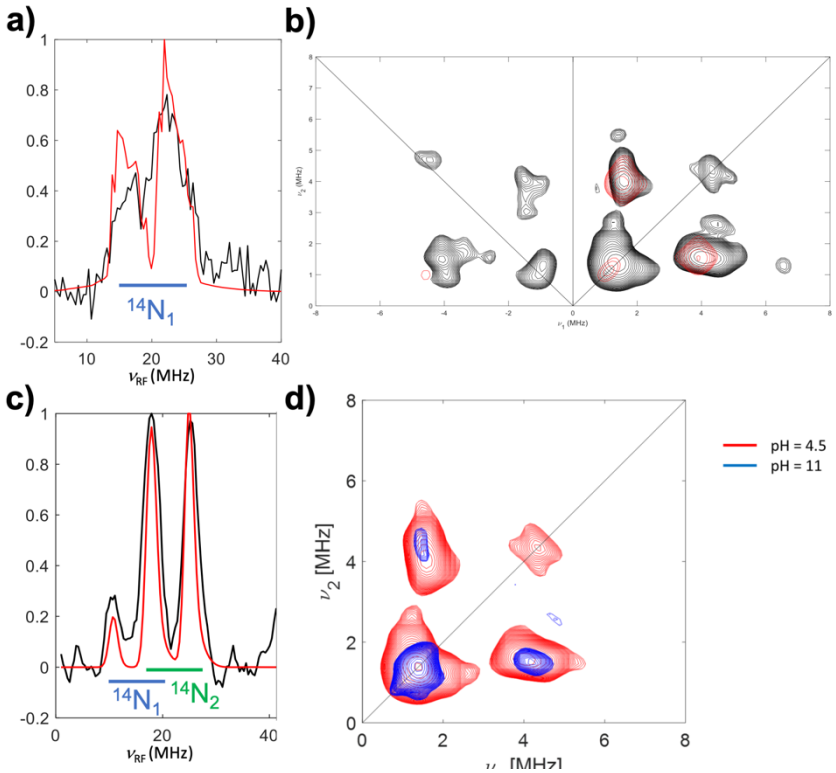


Figure 3A.5. a) Q-band Davies ^{14}N ENDOR spectra (20 K) recorded at 1167.9 mT and b) X-band 2D HYSCORE spectra (30 K) recorded at 336.8 mT of species 2 (pH 4.5). c) Q-band Davies ^{14}N ENDOR spectra (20 K) recorded at 1084 mT and d) Superposition of X-band 2D HYSCORE spectra (30 K) of species 2 (pH=4.5 red) and species 4 (pH 11.0, blue), recorded at 336.8 mT. The simulations are displayed in red.

Table 3A.2. ^{14}N Hyperfine and Nuclear Quadrupole Values calculated by ^{14}N ENDOR and HYSCORE spectroscopy of species 2 (pH 4.5) and 4 (pH 11.0). Hyperfine tensor principal values and nuclear quadrupole coupling are given in units of MHz.

	^{14}N ENDOR					**HYSCORE**				
Species	$\lvert A_1^N \rvert$	$\lvert A_2^N \rvert$	$\lvert A_3^N \rvert$	e^2qQ/h	η	$\lvert A_1^N \rvert$	$\lvert A_2^N \rvert$	$\lvert A_3^N \rvert$	e^2qQ/h	η
2 (pH 4.5)	36	36	45	2.2	0.2	1.6	1.6	2.5	1.5	1
4 (pH 11.0)	28	28	37	2	1	1.6	1.6	2.5	1.5	1
	43	43	67	3.5	1					

3A.1.5 Hypothetical coordination geometries of the species across the pH

The obtained CW-EPR signals are characteristic for Cu^{2+} monomeric species with a single electron in $d(x^2-y^2)$ orbital ($g_3 >> g_1 > g_2 > g_e$) and tetragonal-derived geometries.[20] From the analysis of the CW-EPR parameters derived from simulations (A and g tensors), it appears that miniLPMO mainly exists in 3 different species across the pH.

The g_3 and A_3 values of species 2 (pH 4.5) well fit within the range of values predicted by a canonical Peisach-Blumberg plot for a 2N2O coordination environment. These finding are confirmed by ^{14}N ENDOR analysis. A very likely scenario for species 2 is a tetragonal coordination environment with a Jahn-Teller effect (Species 2, Figure 3A.6). Similar results are expected for species 3 (pH 7.0) where the g_3 and A_3 parameters suggest a 3N1O coordination environment and the inclusion of amine in the copper first coordination sphere (Species 3, Figure 3A.6).

The hamiltonian spin parameters for species 4 (pH 11.0) are very distinct from those obtained for the other species, which is especially noticeable in the small value of g_3. The low value suggests that the ligand field around the copper in species 4 is very σ-donating. Furthermore, with an increasing rhombicity and ε (≈ 150 $M^{-1}cm^{-1}$), it is reasonable to assume a more distorted planar geometry with a strong σ-donating nitrogen atom (confirmed by ^{14}N ENDOR analysis) which can come from the backbone amide or the azanide (Specie 4, Figure 3A.6).

3A. Mimicking the Histidine Brace motif in De Novo Designed Proteins

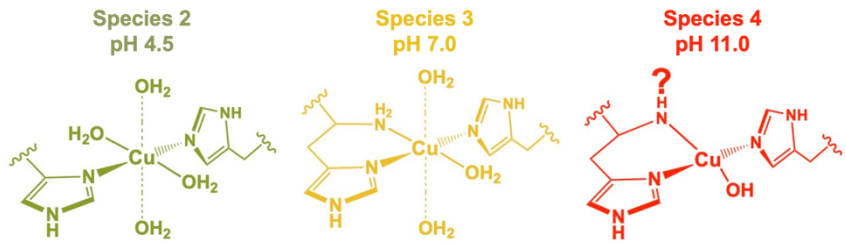
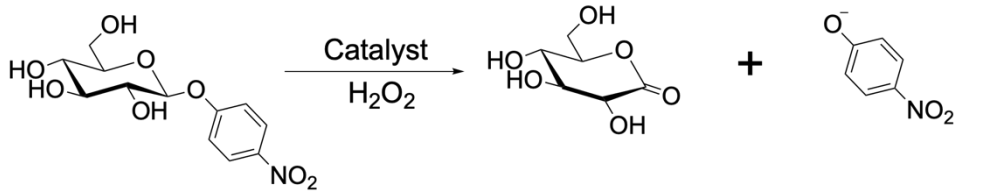


Figure 3A.6. Hypothetical protonation states and coordination geometries of species 2, 3 and 4.

3A.1.6 Hydrogen peroxide activation by Cu-miniLPMO

The LPMO-like reactivity of Cu^{2+}-miniLPMO was preliminary evaluated in the C-H bond activation, using 4-nitropheyl-β-D-glucopyranoside (PNPG) as model substrate and hydrogen peroxide as co-substrate (Scheme 3A.1). The formation of 4-nitrophenolate anion (PNP) was monitored under basic conditions by spectrophotometry (λ_{max} = 400 nm, ε_{400nm} = 18500 $M^{-1}cm^{-1}$).[21]

Scheme 3A.1. Oxidative Cleavage of PNPG using H_2O_2 as oxidant.



The initial rates and substrate conversion were found dependent on species acting as catalyst (Figure 3A.7). While the coordination environment of the species 2 and species 4 make the catalyst less active, the coordination environment of the species 3 makes the catalyst more active in this kind of reaction.

This finding suggests that an oxidative cleavage of the substrate is occurring. However, as formation of PNP may also occur upon hydrolytic pathway, more detailed investigations are needed to pinpoint the reaction mechanism and are currently under course.

Nevertheless, these catalytic data, together with the details gained from spectroscopic analysis, strongly support the presence of a HB-like motif at neutral pH.

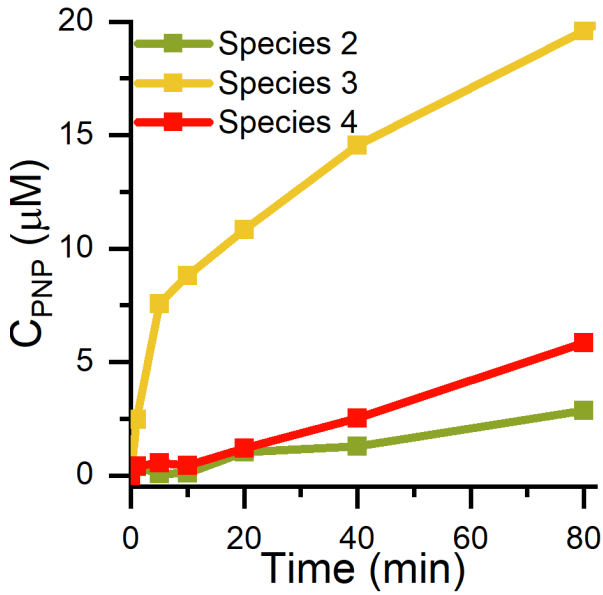


Figure 3A.7. PNP formation kinetics using different catalyst at different pH: Specie 2 (pH 4.5), Species 3 (pH 7.0), Species 4 (pH 11.0). The Cu^{2+}-miniLPMO concentration was fixed at 80 μM and PNPG concentration was fixed at 8 mM, in 15 mM MES, 15 mM HEPES, 15 mM CHES as buffer. The reaction was initiated by the addition of H_2O_2 to a final concentration of 80 mM. At fixed time intervals (square symbols) an aliquot of the solution was taken, diluted in 100 mM carbonate buffer pH 10.5 and PNP concentration monitored at 400 nm.

3A. Mimicking the Histidine Brace motif in De Novo Designed Proteins

3A.2 Experimental Section

3A.2.1 Design

To determine which position would be most appropriate to the inclusion of a HB site, all backbone-dependent rotamers of the internal His in a natural LPMO site (PDB ID: 6rw7) were superimposed to all residues whose Cα was within 10 Å of the N-terminal nitrogen atom in α2D. The superpositions were realized with the pair_fit function of the PyMOL software. Reconstruction of the N-terminal segment to optimize His1 position relative to the newly introduced HB metal-binding site was performed with the MASTER software. Query structure was composed of residues 4-8 from α2D in addition to the N-terminal His positioned by HB rotamer superposition (HEELEK). PDS format file for the search was generated with the createPDS application using the following command:

```
/home/computername/master-demo/master-v1.6/bin/createPDS
--type query --pdb query01.pdb --opdb query01_pds.pdb
```

The search was performed using the default database available from the same source as the MASTER program, composed of about 17000 non redundant single chain entries extracted from the PDB. RMSD was computed over all backbone atoms and cutoff value for matches was fixed to 0.5 Å. The search was performed with the command:

```
master --query query01.pds --targetList /MASTER_DB/list
--rmsdCut 0.5 --structOut Query01 --outType match --
bbRMSD --topN 1000 --matchOut query01.match --seqOut
query01.seq
```

The query structure for loop reconstruction was composed of two five-residue segments from α2D: residues 6-10 (LEKKF) and 22-26 (IEELH). A gap of eleven residues was left between the two segments to be filled by MASTER, reconstructing the loop region. The PDS format file was created with a command analogous to the one used for the N-terminal, while the search was performed with the following command:

```
master --query loop5_11_5.pds --targetList /MASTER_DB/list --rmsdCut 1.0 --structOut loop5_11_5 --outType wgap –gapLen '1-11' --bbRMSD --topN 1000 --matchOut loop5_11_5.match --seqOut loop5_11_5.seq
```

For both N-terminal and loop searches, the .seq files containing the sequences of all matches and the corresponding RMSD values from the query structure were analyzed to extract the frequency of different amino acids for each position. These statistics were used during the design step to inform the allowed residues for the loop region in all resfiles. In specific resfiles the statistics obtained from the MASTER search match were also used to inform the identity of allowed residues for N-terminal segment (positions 2 – 6 of modified structure). The resulting modified version of α2D, named α2D-C4, was used as the starting structure in the Rosetta design protocol.

3A.2.2 Molecular Dynamics simulations

MD simulations on the best structure from all variants generated subsequently were performed with the NAMD engine using the CHARMM36 force field. The input files for all simulations were created using CHARMM-GUI. The structures were solvated in an octahedral

water box with 10 Å edge distance. Solvent was explicitly modelled, and water molecules were described by TIP3P parametrization5. Chloride or potassium atoms were placed with Monte Carlo method in order to ensure electroneutrality. Long range electrostatic interactions were calculated using the Particle Mesh Ewald (PME) method with 1 Å grid spacing. The system was first minimized with 10000 conjugate gradient steps. Subsequently, the temperature was linearly raised from 0 to 320 K during 5 ns. The system was then equilibrated for 5 ns, and production was conducted for 10 ns in short simulations or 80 ns in extended simulations. Constant temperature was maintained through Langevin dynamics with damping coefficient 1.0 ps^{-1}. Constant pressure was imposed through Nose-Hoover Langevin piston with 50 fs piston period and 25 fs oscillation decay time. Integration step was 2 fs and all hydrogen bond length were fixed with the SHAKE method. VMD (Visual Molecular Dynamics) was used to perform data analysis and elaboration.

3A.2.3 Synthesis

The 33-mer peptide miniLPMO was synthesized by automatic solid-phase synthesis as previously reported in chapter 2. To preserve the N-terminal free, it was not acetylated, whereas the C-terminal was instead amidated. It was purified to homogeneity via RP-HPLC. Identity was ascertained by high resolution ESI-MS. (Figure 3A.8)

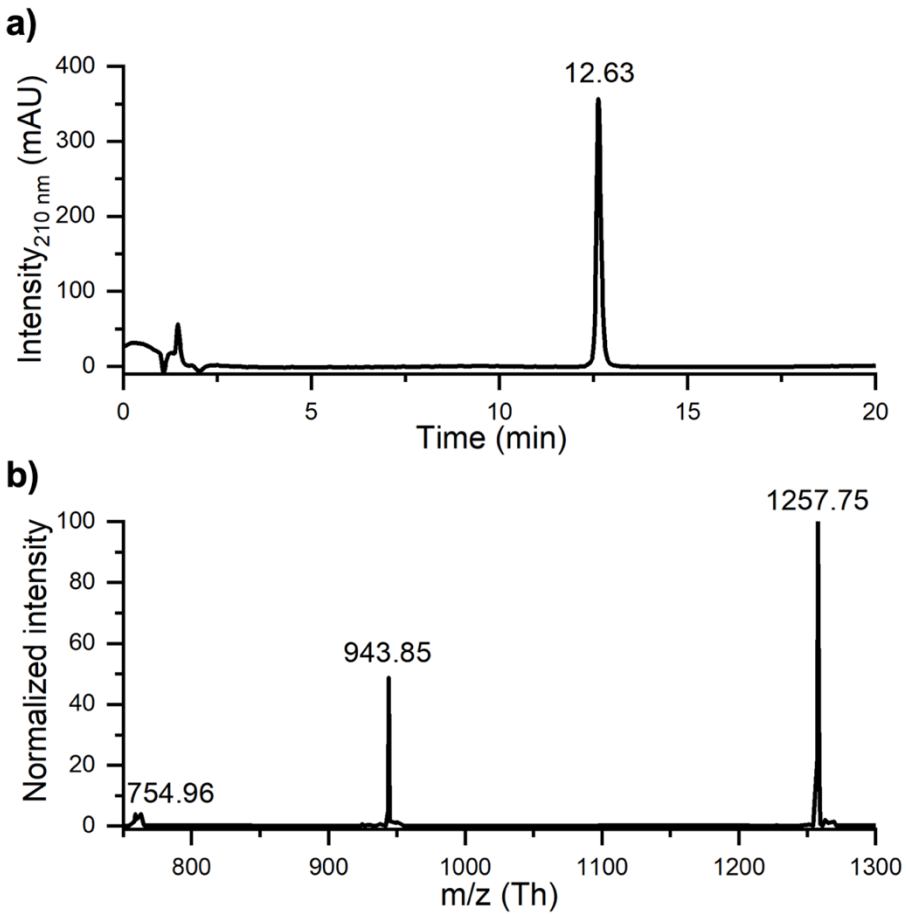


Figure 3A.8. RP-HPLC chromatogram (210 nm trace) of the pure miniLPMO; b) ESI-MS spectrum of miniLPMO. The signal at m/z = 1257.75 Th corresponds to the $[M+3H^+]^{3+}$ ion (theoretical average m/z: 1257.96 Th); the signal at m/z = 943.85 Th corresponds to the $[M+4H^+]^{4+}$ ion (theoretical average m/z: 943.72 Th); the signal at m/z = 754.96 Th corresponds to the $[M+5H^+]^{5+}$ ion (theoretical average m/z: 755.18 Th).

3A.2.4 CD Spectroscopy

CD measurements were performed using a J-815 spectropolarimeter equipped with a thermostated cell holder (JASCO, Easton, MD, USA). Spectra were collected at 25°C, from 260 to 190 nm at 0.2 nm intervals with a 20 nm min^{-1} scan speed, at 1 nm bandwidth and at 16 s response. Cells of 10, 1 and 0.1 mm path length were used depending on the

concentration of the peptide and the buffer was 2 mM HEPES pH 7.5. The experimental data of the concentration dependence, monitored by MRE at 222 nm, were globally fitted to a monomer-dimer equilibrium as previously described.[11]

3A.2.5 UV-Vis absorption spectroscopy

UV–Vis absorption spectra were recorded with a Cary Varian 60 Spectrophotometer, equipped with a thermostatic cell compartment (Varian, Palo Alto, CA, USA), using a quartz cuvette with 1 cm path length. Wavelength scans were performed at 25°C from 200 to 800 nm, with a 600 nm min^{-1} scan speed. Both apo-protein and copper stock solutions were freshly prepared in H_2O milliQ. The stock solutions of $CuSO_4$ were prepared at 10 mM concentration starting from the solid $CuSO_4 \cdot 5H_2O$. The initial miniLPMO concentration was determined spectrophotometrically by using ε_{280nm} = 5500 $M^{-1}cm^{-1}$.[22] The addition of equimolar amount of $CuSO_4$ respect the apo-protein into the buffer (MES, HEPES, CHES, 15 mM each) with apo-protein was performed slowly to avoid the precipitation of Cu^{2+} hydroxides.

Different samples of Cu^{2+}-miniLPMO (0.1 mM) in the mixed buffer solution were prepared. Samples from pH 2.0 to 11.0 (in 0.5 pH unit increments) were prepared by adding NaOH and H_2SO_4 solutions and a pH meter was employed to determine the final pH. These solutions were lyophilized and subsequently redissolved with H_2O milliQ to a final concentration of 0.2 mM.

3A.2.6 CW EPR spectroscopy

EPR analysis was performed at the University of Torino, in collaboration with prof. Mario Chiesa.

CW-EPR experiments were performed on a Bruker EMX spectrometer at X band (~ 9.44 GHz) equipped with a cylindrical cavity. All the spectra were recorded at 77 K and a microwave power of 0.68 mW, a modulation amplitude of 0.5 mT and a modulation frequency of 100 KHz. Different samples of Cu^{2+}-miniLPMO (0.1 mM) in a mixed buffer solution containing 15 mM MES, 15 mM HEPES, 15 mM CHES were prepared. The same preparation was employed as performed with the UV-Vis absorption spectroscopy. These solutions were lyophilized and subsequently redissolved with 30% of glycerol, used as glassing agent, to a final concentration of 0.3 mM. The samples at pH 4.5 and 11.0 were also used for Davies ^{14}N ENDOR and HYSCORE spectroscopy.

The EPR spectra at pH 2.0, 4.5, 7.0, and 11.0 were simulated with a single set of spin Hamiltonian parameters. Spectra taken at different pH values could be satisfactory simulated through a weighted combination of these spectra. All but the pH 11.0 spectrum contributions were systematically varied between 0 and 1 while minimizing the residual sum of squares between the experimental spectrum and the calculated spectrum. The contribution of each spectrum was determined by performing a stochastic search starting from a 1000 population size and using a nonlinear gradient-based optimization algorithm using the Microsoft Excel Solver tool. The pH 11.0 spectrum contribution was calculated by imposing 1 as the sum of the weighted contributions.

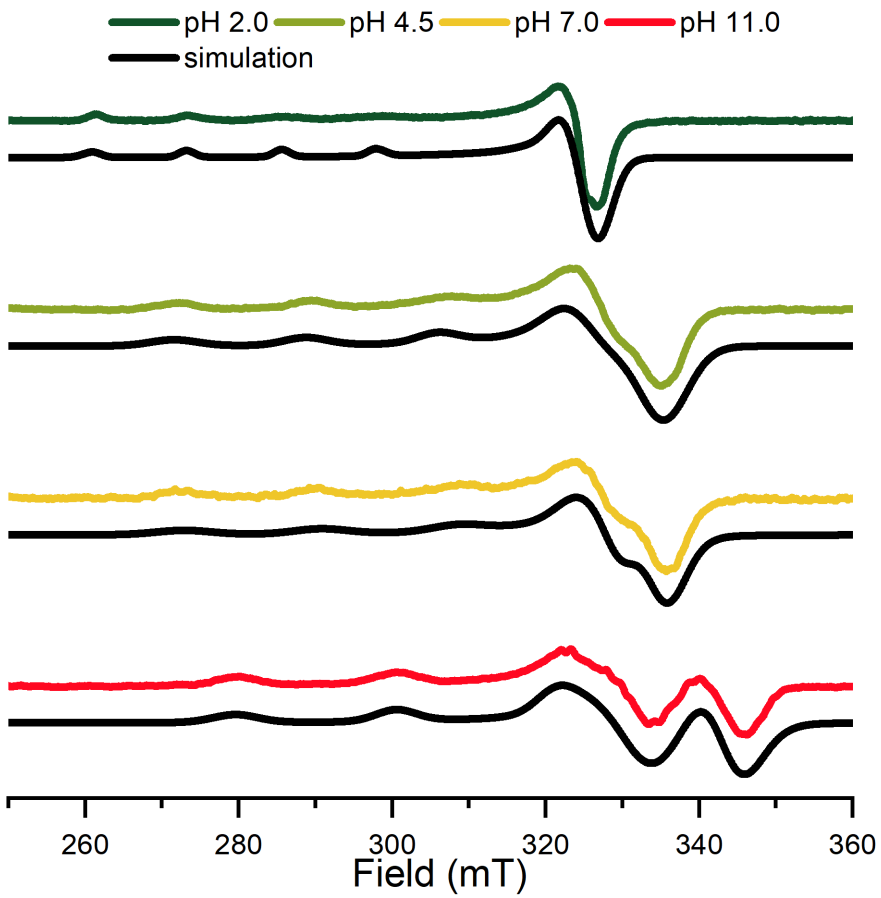


Figure 3A.9. Frozen solution X-band CW-EPR spectra (77 K) of Cu^{2+}-miniLPMO (0.3 mM) in a mixed buffer of MES, HEPES and CHES (15 mM each) at pH 2.0 (dark green line), pH 4.5 (green line), pH 7.0 (yellow line) and pH 11.0 (red line). The black traces are the simulated spectra.

Table 3A.3. EPR spin Hamiltonian parameters from simulations of X-band CW-Spectra for the species studied in this work.

		Species 1 (pH 2.0)	Species 2 (pH 4.5)	Species 3 (pH 7.0)	Species 4 (pH 11.0)
***g* values**	g_1	2.080	2.055	2.052	2.040
	g_2	2.080	2.055	2.058	2.045
	g_3	2.415	2.268	2.251	2.169
A_{Cu} **(MHz)**	$\lvert A_1 \rvert$	5	60	40	35
	$\lvert A_2 \rvert$	5	60	40	50
	$\lvert A_3 \rvert$	410	535	550	625
***g* strains**	g_1	0.03	0.05	0.03	0.05
	g_2	0.03	0.05	0.03	0.07
	g_3	0.02	0.05	0.06	0.04
A_{Cu} **strains (MHz)**	$\lvert A_1 \rvert$	1	1	1	1
	$\lvert A_2 \rvert$	1	1	1	1
	$\lvert A_3 \rvert$	1	10	1	10
Line widths		0.1 0.2	0.2 0.2	0.8 1.0	0.5 0.8

3A.2.7 Davies ^{14}N ENDOR spectroscopy

Q-band Davies ENDOR measurements were recorded on a Bruker Elexsys E580 spectrometer at 30 K and carried out using the following pulse sequence: π-T-π/2-τ-π-τ-echo. The experiments were done with mw pulse lengths of t_{π} = 32 ns, $t_{\pi/2}$ = 16 ns, and an interpulse time τ of 400 ns.

p The *RF* pulse length was set to 14 µs and a resolution of 440 points was adopted.

3A.2.8 HYSCORE spectroscopy

Six-pulse X-band HYSCORE experiments[23,24] were carried out with the pulse sequence $(\pi/2)_x-\tau_1-(\pi)_x-\tau_1-(\pi/2)_y-t_1-(\pi)_y-t_2-(\pi/2)_y-\tau_2-(\pi)_y-\tau_2-$echo, applying an eight-step phase cycle in order to eliminate unwanted echoes. The t_1 and t_2 time intervals were incremented in steps of 16 ns, starting from 100 ns to 4900 ns. Pulse lengths $t_{\pi/2}$ = 16 ns and t_π = 32 ns, and a 0.5 kHz shot repetition rate, and equal τ_1 and τ_2 values equal to 136 ns. The time traces of the HYSCORE spectra were baseline corrected with a third-order polynomial, apodized with a Hamming window, and zero-filled. After two-dimensional Fourier transformation, the absolute-value spectra were calculated. Spectra recorded with different τ_1 = τ_2 values (specified in the figure captions) were added to average out blind-spots. All of the EPR, ENDOR, and HYSCORE simulations were performed using the Easyspin software package,[25] running within the MathWorks MATLAB environment.

3A.2.9 Catalytic assays

Activity assays were performed in total volume of 500 µL placed in 1.5 mL vials. The concentration of the Cu^{2+}-miniLPMO was fixed at 80 µM in 15 mM MES, 15 mM HEPES, 15 mM CHES and 8 mM of the substrate PNPG. H_2O_2 stock solutions were prepared by diluting H_2O_2 (30%, v/v) in water, whose concentration was determined by UV- vis

absorption spectroscopy (λ_{max} = 240 nm; ε = 39.4 $M^{-1}cm^{-1}$). The reaction was initiated by the addition of H_2O_2, under magnetic stirrer, to a final concentration of 80 mM. The amount of PNP was determined at different times by taking an aliquot of the reaction solution and diluting it in 100 mM carbonate buffer pH 10.5. Quantification was performed measuring the absorbance at 400 nm (ε= 18500 $M^{-1}cm^{-1}$).

References

[1] D. P. Raleigh, S. F. Betz, W. F. DeGrado, *J. Am. Chem. Soc.* **1995**, *117*, 7558–7559.
[2] R. B. Hill, W. F. DeGrado, *J. Am. Chem. Soc.* **1998**, *120*, 1138–1145.
[3] R. B. Hill, J.-K. Hong, W. F. DeGrado, *J. Am. Chem. Soc.* **2000**, *122*, 746–747.
[4] R. B. Hill, W. F. DeGrado, *Structure* **2000**, *8*, 471–479.
[5] R. B. Hill, C. Bracken, W. F. DeGrado, A. G. Palmer, *J. Am. Chem. Soc.* **2000**, *122*, 11610–11619.
[6] Claire. A. Fowler, F. Sabbadin, L. Ciano, G. R. Hemsworth, L. Elias, N. Bruce, S. McQueen-Mason, G. J. Davies, P. H. Walton, *Biotechnol. Biofuels* **2019**, *12*, 232.
[7] N. M. Glykos, G. Cesareni, M. Kokkinidis, *Structure* **1999**, *7*, 597–603.
[8] E. A. Span, D. L. M. Suess, M. C. Deller, R. D. Britt, M. A. Marletta, *ACS Chem. Biol.* **2017**, *12*, 1095–1103.
[9] K. E. H. Frandsen, T. J. Simmons, P. Dupree, J.-C. N. Poulsen, G. R. Hemsworth, L. Ciano, E. M. Johnston, M. Tovborg, K. S. Johansen, P. von Freiesleben, L. Marmuse, S. Fort, S. Cottaz, H. Driguez, B. Henrissat, N. Lenfant, F. Tuna, A. Baldansuren, G. J. Davies, L. Lo Leggio, P. H. Walton, *Nat. Chem. Biol.* **2016**, *12*, 298–303.
[10] G. R. Hemsworth, B. Henrissat, G. J. Davies, P. H. Walton, *Nat. Chem. Biol.* **2014**, *10*, 122–126.
[11] M. Chino, S.-Q. Zhang, F. Pirro, L. Leone, O. Maglio, A. Lombardi, W. F. DeGrado, *Biopolymers* **2018**, *109*, e23339.
[12] E. I. Solomon, D. E. Heppner, E. M. Johnston, J. W. Ginsbach, J. Cirera, M. Qayyum, M. T. Kieber-Emmons, C. H. Kjaergaard, R. G. Hadt, L. Tian, *Chem. Rev.* **2014**, *114*, 3659–3853.
[13] F. L. Aachmann, M. Sørlie, G. Skjåk-Bræk, V. G. H. Eijsink, G. Vaaje-Kolstad, *Proc. Natl. Acad. Sci.* **2012**, *109*, 18779–18784.
[14] K. Krupa, M. Korabik, T. Kowalik-Jankowska, *J. Inorg. Biochem.* **2019**, *201*, 110819.
[15] P. J. Lindley, A. Parkin, G. J. Davies, P. H. Walton, *Faraday Discuss.* **2022**, *234*, 336–348.
[16] J. Peisach, W. E. Blumberg, *Arch. Biochem. Biophys.* **1974**, *165*, 691–708.
[17] R. J. Sundberg, R. B. Martin, *Chem. Rev.* **1974**, *74*, 471–517.
[18] H. L. Van Camp, R. H. Sands, J. A. Fee, *J. Chem. Phys.* **1981**, *75*, 2098–2107.
[19] H. Yokoi, *Biochem. Biophys. Res. Commun.* **1982**, *108*, 1278–1284.
[20] E. I. Solomon, M. J. Baldwin, M. D. Lowery, *Chem. Rev.* **1992**, *92*, 521–542.
[21] K. Selmeczi, M. Giorgi, G. Speier, E. Farkas, M. Réglier, *Eur. J. Inorg. Chem.* **2006**, *2006*, 1022–1031.

[22] C. N. Pace, F. Vajdos, L. Fee, G. Grimsley, T. Gray, *Protein Sci.* **1995**, *4*, 2411–2423.
[23] R. Song, Y. C. Zhong, C. J. Noble, J. R. Pilbrow, D. R. Hutton, *Chem. Phys. Lett.* **1995**, *237*, 86–90.
[24] B. Kasumaj, S. Stoll, *J. Magn. Reson.* **2008**, *190*, 233–247.
[25] S. Stoll, A. Schweiger, *J. Magn. Reson.* **2006**, *178*, 42–55.

3. Mimicking the Histidine Brace motif in De Novo Designed Proteins

3B.1 Results and Discussion

3B.1.1 Design strategy of dHisB

DHisB (designed Histidine Brace) is a heterodimeric small protein composed of two distinct peptides engineered to self-associate into a protein containing a HB metal-binding site. As the objective was a heterodimer with a specific orientation of the two monomers, negative design was employed to inhibit the formation of alternative topologies and other oligomeric states.[1–3]

The design started from the high-resolution structure of coil-Ser (PDB ID:1cos, Figure 3B.1a),[4] an antiparallel de novo designed three-helix bundle composed of three 29-residue monomers, widely used in the literature for the construction of artificial metalloproteins,[5,6] given its high stability and tolerance to mutations. This peptide assembles in aqueous solution in a noncooperative mechanism in which two random coil monomers join into a helical dimer. After binding a third helix, this helical dimer forms an antiparallel three-helix bundle. Based on these results, it is reasonable to assume that an α2 helical harpin could host a third helical peptide, in a similar fashion to the calcineurin binding receptor, previously designed by Lombardi, DeGrado and co-workers.[7]

Since coil-Ser is an antiparallel three-helix bundle, one end has two N-termini (blue sphere in Figure 3B.1a) and one C-terminus (red sphere in Figure 3B.1a). Thus, it was possible to join the N- and C-terminal groups from chain A and chain C, respectively (in raspberry and in green in Figure 3B.1a) through a loop and exploit the remaining N-terminal of chain B (in cyan in Figure 3B.1a) to house the terminal His of HB site.

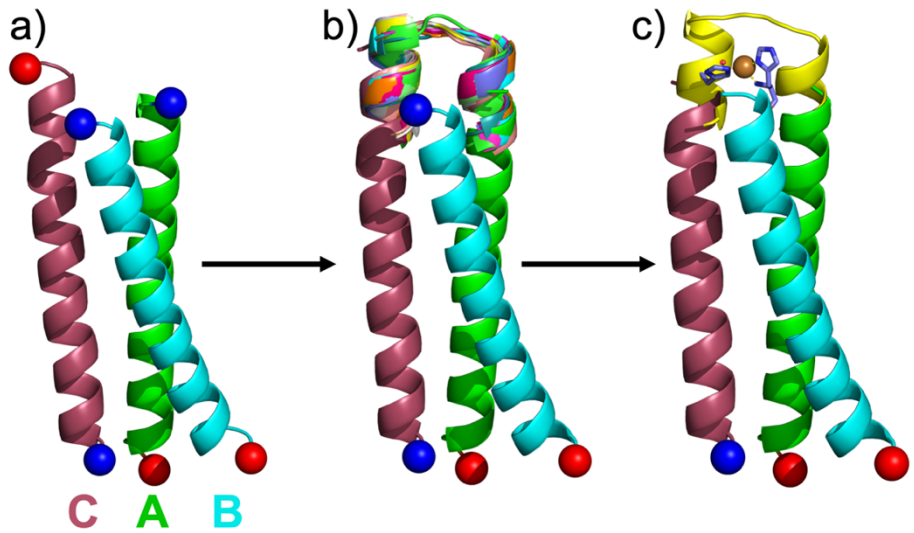


Figure 3B.1. a) Coil-Ser X-ray crystal structure (PDB ID: 1cos). For clarity, the homotrimer is shown with chain A in green, chain B in cyan and chain C in strawberry. N-terminal are depicted as blue spheres and C-terminal as red spheres. b) Superposition of loops, between chain A and chain C, generated by the structural search. c) Superposition of the best HB backbone-dependent rotamer (depicted as violet sticks and copper as brown sphere) on the coil-Ser structure modified with the newly loop (depicted as yellow cartoon). The HB was originally taken from LPMO crystal structure (PDB ID: 6rw7).

The loop was constructed between the antiparallel pair of coil-Ser helices A and C, performing a systematic search (MASTER)[8] to find the best fragment linking N- and C-termini (chain A and C, respectively). Query structures were built using residues 24-26 on chain C and residues 2-5 on chain A as anchoring points for the search, while allowing only fragments formed by less than 20 residues. From the analysis of the matches, a single cluster was identified that consists of four residue-long connections between the helices, which are conformationally similar (Figure 3B.1b). Amino acid preferences for the superimposed residues were recorded to inform the repacking routine about allowed amino acid types during computational sequence design.

The member of this cluster with lowest RMSD was selected and grafted onto the coil-Ser structure, completing the chimeric heterodimer containing the HB site.

To find the best orientation for the metal-binding site, the HB site of the natural AA10 LPMO from *Teredinibacter turnerae* (PDB ID: 6rw7) was chosen as reference. All possible backbone-dependent rotamers for the internal His (His107 of the natural LPMO) of the HB site were superimposed on residues within 10 Å of the N-terminal of the free-standing helix (chain B). During the iterative superposition process of the internal His rotamers, the distances between the Cα-Cα atoms of the just modeled terminal His from the HB site and the first four residues of chain B were calculated. All the superpositions resulting in a distance within 2 Å were considered as possible candidates for HB engineering, with the terminal His of the HB site positioned in place of the residue with the lowest Cα-Cα distance among the first four. This procedure yielded a handful of viable positions for the HB site within the coil-Ser scaffold, from which visual inspection was used to sort out the most reasonable. This was found to be either Leu26 or Trp2 on chain C and A, respectively, for the internal His of the HB site and Trp2 on chain B for the terminal His, thus truncating only one residue of chain B. Insertion of the internal His in position Trp2 (chain A), corresponding to an N3 position in *Schellman* $\alpha_L\beta$ *turn* or *Rose* $\alpha_L\beta$ *turn* was excluded because of the low intrinsic propensity of this position to house a His residue, whereas the C5 position (Leu26 on chain C) was preferred for the significantly higher His propensity.[9] The preferred conformation of the internal His in position 2 of chain A was confirmed later by preliminary computational structure prediction using AlphaFold2-multimer-v1,[10]

which was followed by minimization using Amber. The results showed that the imidazole ring of the internal His was positioned between helices A and C, making it unable to coordinate the copper ion, whereas internal His at position 26 of chain C preferred a conformation pointing directly toward the N-terminal. It should be noted that the loop inserted between helix A and C is neither a classical *Schellman* or *Rose* $\alpha_L\beta$ *turn* since the Pro31 is closer to a polyproline conformation than an α-helix.

Starting backbone coordinates for sequence design protocols that accommodate the N-terminal His on the B-chain helix were generated by grafting on the scaffold the top matching structural fragments from a second MASTER search. This step was crucial to ensure optimal backbone geometry for the HB site. Among the resulting fragments, only short α-helical matches were considered and the lowest Cα RMSD match was chosen to reconstruct the N-terminal region of chain B to host the terminal His. This procedure resulted in the formation of two peptide chains: a helix-loop helix motif (herein referred to as HLH), hosting the internal His, and single chain (herein referred to as chain H), which contains the N-terminal His.

Sequence design was firstly focused on residues at *g* and *e* positions, rationally choosing those that stabilize only the antiparallel heterodimeric topology by a "negative" design procedure, as previously reported.[11,12] Interacting interfacial positions were allowed to be both positively or negatively charged. A Monte Carlo search was implemented to find the identity of the residues in these positions, while ensuring destabilization of the competing topology. This resulted in a salt bridge network at the interface between chain A/C and chain B (Figure 3B.2a),

which was kept unaltered during sequence design protocol. Additionally, the overall formal charge of the two monomers (+5 for chain A/C and -1 for chain B) was chosen to disfavor homooligomers and encourage the assembly of the desired heterodimer through electrostatic complementarity. Subsequently, this structure underwent a flexible backbone sequence design process in order to build a complementary apolar surface between chain A/C and B. Moreover, all other positions, excluding hydrophobic core and charged interfacial residues were allowed to be any of the 20 natural amino acids, except for Cys, Met and His (in order to avoid undesired oxidization and/or alternative metal-binding sites). The overall design protocol generated 250 sequences. The top 10% lowest scoring sequences, as quantified by the Ref2015 score function in ROSETTA, were subjected to a structure prediction round by alphafold-multimer-v1. As a further validation method for the designed structure, MD simulations were performed on the sequences that showed the best agreement with the designed model (full-backbone RMSD < 1.5 Å). Cu^{2+} was replaced with Zn^{2+} for the MD simulation, as the MD parameterization of zinc is relatively accurate in the CHARMM36m forcefield. In addition, zinc substitution in natural LPMOs does not significantly alter the orientation of the first coordination sphere.[13]

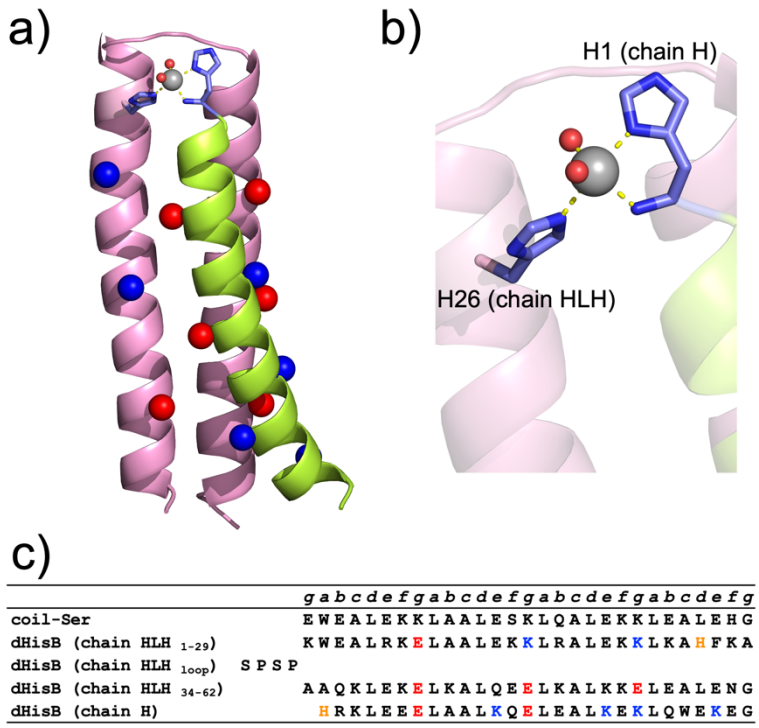


Figure 3B.2. a) dHisB final model. Chain HLH is depicted as pink cartoon, chain H is depicted as limon green cartoon, C_β of positive- and negative charged residues forming the salt bridge network between the two chains are reported in red and blue spheres. b) zoom of the HB active site in dHisB c) Peptide sequence alignment of chain HLH and chain H of dHisB with coil-Ser. For all the sequences, coordinating residues are reported in orange, positive- and negative charged residues forming the salt bridge network between the two chains are reported in red and blue, respectively.

The "non-bonded approach" was used to characterize zinc-protein interactions as a more stringent test over molecular flexibility.

The effect of multiple point mutations was ascertained by inspection of the MD trajectorie evolution. Trp residues, useful for spectroscopic analysis, were introduced in position 2 of chain HLH (which is equivalent to Trp2 position in coil-Ser), and in position 24 of chain B, which was found to interfere little on the overall fold. Residues close to the HB site

(position 34 and 35 in chain HLH and 4 in chain H) were also mutated to explore the effect on metal site stability. Larger hydrophobic residues at position 34 and 35 of chain HLH (Val, Leu, Ile) were found to have a detrimental effect on metal coordination geometry. Mutation of Leu to Phe at position 4 of chain H (a *d* heptad position)[14,15] was also found to be destabilizing for the overall fold, as expected from known amino acid occurrences in three-helix bundles.[16] Furthermore, reduced conformational flexibility of the N-terminal region in chain H was observed in the presence of an Arg residue at position 2, due to the interaction with Glu residues in positions 4 and 5.

Conformational analysis of designed heterodimers during MD simulations suggested that the Xxx-Pro-Ser-Pro motif (residues 30-33 in chain HLH) would confer a higher degree of conformational rigidity to the loop region. The identity of the first residue in the solvent-exposed loop was restricted to Ser, a small hydrophilic residue, to avoid the possibility of undesired metal coordination.

The structures of sequences containing these mutations were predicted using Alphafold2-multimer and relaxed using the Amber forcefield. Agreement to the desired structure, computed from backbone atom RMSD, and prediction confidence were used to identify the best sequences for chain HLH and chain H (Figure 3B.2c), which were subsequently synthesized and characterized.

3B.1.2 DHisB folding and oligomerization state

H and HLH peptide were chemically synthesized using Fmoc chemistry, then purified by RP-HPLC, and finally identified through ESI

MS. To preserve the N-terminal free, H was not acetylated, whereas the C-terminal was instead amidated.

The interaction between H and HLH and the stoichiometry was evaluated by SEC, RP-HPLC and CD spectroscopy. The SEC reveled one major peak when a 1:1 H:HLH solution was analyzed, both in absence and in presence of copper (Figure 3B.3a). As control, the injections of only H (Figure 3B.3b, green line) or only HLH (Figure 3B.3b, orange line) show peaks with retention times different from 1:1 H:HLH mix. The quantification of peptide content in 1:1 H:HLH mix solution by RP-HPLC resulted in a (0.9 ± 0.1) stoichiometry (see Experimental for details). These findings demonstrate that the dHisB assumes the designed heterodimeric structure both in apo and holo form, as predicted by the design procedure.

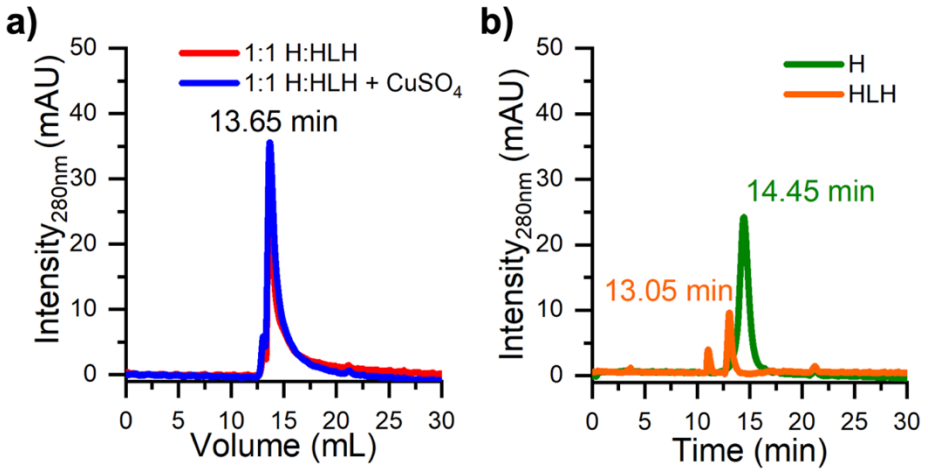


Figure 3B.3. a) SEC chromatograms, recorded at 280 nm, of (a) mix 1:1 H:HLH (both 300 μM) in the absence (red line) and in the presence (blue line) of 330 μM $CuSO_4$ and (b) 300 μM H (green line) or 300 μM HLH (orange line). The elution buffer was 50 mM HEPES, 150 mM NaCl, pH 7.0.

CD analysis gave further insights into the binding of H to HLH. Double minima at 208 and 222 nm in the CD spectra of HLH are characteristic of α-helical structure (Figure 3B.4a, dashed orange line). In comparison to coil-Ser, which was proven by X-ray crystallography to exist as a three-helix bundle, the MRE at 222 nm is slightly lower (31.7 kdeg cm^2 res^{-1} $dmol^{-1}$ for coil-Ser and 29.8 kdeg cm^2 res^{-1} $dmol^{-1}$ for HLH). The CD spectrum of HLH matches with the predicted helix-loop-helix pattern, in which about 94% of the residues adopt a helical conformation and 6% adopt a loop conformation.

Spectrum of H, in the absence of HLH, is consistent with a poor structured α-helix (Figure 3B.4a, dashed green line). Upon addition of H to 1 equivalent of HLH, H predominantly adopts a α-helix structure (with MRE_{222nm} from 14.3 to 28.1 kdeg cm^2 res^{-1} $dmol^{-1}$), with an ellipticity magnitude corresponding to about 3–4 helix turns (Figure 3B.4a, solid green line).

Interestingly, the sum of the spectra of H and HLH (Figure 3B.4b, dashed red line) differs from the spectra of the heterodimeric complex (Figure 3B.4b, solid red line). This characteristic demonstrates that there is a positive interaction between the two chains, resulting in a globally better folded structure. All these findings confirm the correct oligomerization state of the newly heterodimeric construct dHisB.

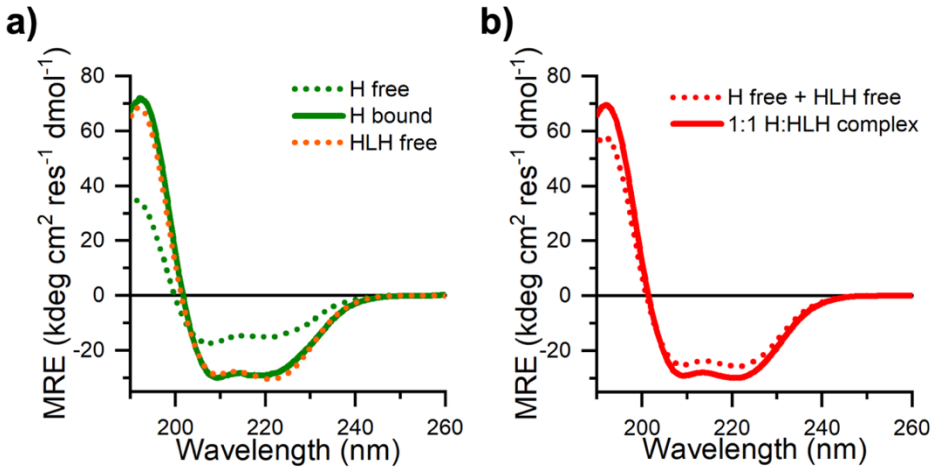


Figure 3B.4. a) Ultraviolet CD spectra of 20 μM H (dashed green line), 20 μM HLH (dashed orange line), difference spectrum resulting from subtraction of 20 μM 1:1 H:HLH complex, from that determined for 20 μM HLH alone (solid green line). b) Ultraviolet CD spectra of sum spectra of H and HLH (dashed red line) and 20 μM 1:1 complex of H with HLH (solid red line). The CD signals are expressed as MRE, using n (number of residues): 28 for H free, 62 for HLH free and 90 for 1:1 H:HLH complex). All the ultraviolet CD spectra were recorded in 2 mM HEPES pH 7.0.

3B.1.3 DHisB copper binding and spectroscopic analysis

Copper binding to dHisB was assessed by CW-EPR and UV-Visible absorption spectroscopy.

The copper reconstitution was performed adding 1.3 equivalent of copper to the 1:1 mix H:HLH, followed bySEC purification (see Experimental Section for details), from which any uncoordinated metal ions were separated from the purified Cu^{2+}-dHisB complex.

The first coordination sphere of copper in the Cu^{2+}-dHisB complex contains numerous ligating atoms with ionizable protons (e.g., His, NH_2, OH_2) whose pKa are unknown. To evaluate the pKa of ionizable protons, a pH titration in the range 2.0–13.0 was performed.

3. Mimicking the Histidine Brace motif in De Novo Designed Proteins

---

Different samples of Cu^{2+}-dHisB were prepared in a mixed buffer solution containing 10 mM MES, 10 mM HEPES, 10 mM CHES, and 10 mM CAPS. These buffers were selected because of their weak coordination propensity; moreover, when they are employed in tandem, they can buffer a wide pH range. To raise and lower the pH of the solutions, diluted solutions of NaOH and H_2SO_4 were added. Using a calibrated pH meter, the pH of the final solution at room temperature was measured by preparing 23 samples ranging from pH 2.0 to pH 13.0 in 0.5 pH unit increments. Fast cooling to 150 K was used to collect the X-band CW-EPR spectrum of the sample at each pH.

The spectra recorded at pH 2.0 and 2.5 show the characteristic tetragonal aqua ion copper EPR signal (Figure 3B.16a). The first deprotonation event with subsequent coordination of copper to dHisB occurs at pH 3.0.

Raising the pH from 3.0 to 13.0, a change in the EPR spectra is observed over the pH range (Figure 3B.5a), with all species displaying a spectral envelope consistent with a prominent $d(x^2 - y^2)$ SOMO, indicating that the copper ion at all pHs has a ligand field dominated by σ-donation in its equatorial plane.

At pH 3.0, as seen by a single set of copper hyperfine split peaks at low field (g_3 = 2.283, $|A_3|$ = 520 MHz and Table 3B.1), one species is shown to dominate the spectrum (Figure 3B.5a, violet line in bold). When the pH is raised, a new set of peaks appears in the low field region (g_3 = 2.249, $|A_3|$ = 550 MHz and Table 3.1), indicative of the appearance of a second unique species, the concentration of which is maximum at pH 5.5 (Figure 3B.5a, blue line in bold). A further increase in pH causes the development of a new set of peaks in the low field region (g_3 = 2.246,

$|A_3|$ = 550 MHz and Table 3B.1), indicating a third unique species. At a pH of 8.5, this third species is the dominating form (Figure 3B.5a, green line in bold). Dramatic changes in the peaks appear in the low field region (g_3 = 2.177, $|A_3|$ = 614 MHz and Table 3B.1) when the pH rises to 13.0 and together with a notable "overshoot" feature at high field, a fourth different species is formed (Figure 3B.5a, red line in bold).

The EPR spectra were convoluted by combining those at pH 3.0, 5.5, 8.5, and 13.0. The contribution of each spectrum was determined through least square fitting of the experimental and calculated spectra, resulting in the determination of the speciation distribution curves and the pKa_s (Figure 3B.5b). After the first deprotonation event and the copper binding to dHisB, the EPR spectrum at pH 3.0 is simulated with a single set of g_3 and A_3 values characteristic of a 1N3O coordination sphere (Species 1 in Figure 3B.5c and Table 3B.1). When the pH value increases, another loss of a proton occurs and a complex with g_3 and A_3 values characteristic of 2N2O coordination sphere is formed (Species 2 in Figure 3B.5c and Table 3B.1). Incidentally, two deprotonation processes must be involved under acidic conditions, a first pKa below 3.0, which could not be precisely evaluated, and a second deprotonation step centered at pH 4.0. These values are in good agreement with His pKa values in LPMOs around 3.5.[17] As the pH rises, a further deprotonation event happens and a new species is formed. The pKa_2 values suggests that the amine nitrogen is included in the coordination sphere (Species 3 in Figure 3B.5c). Indeed, a good simulation of the experimental EPR spectrum of this species was obtained with g_3 and A_3 values in good agreement with the values predicted for a 3N1O coordination sphere. A further increase

in pH sees the emergence of a species with peculiar g_3 and A_3 values. This species has pKa_3 = 10.5 and is fully formed at pH 13.0. The pKa_3 and the spin hamiltonian values suggests the deprotonation of an amide from backbone or the double deprotonation of an amine to azanide species (Species 4 in Figure 3B.5c) as already observed in LPMOs.[18,19] A good simulation of the experimental EPR spectrum was obtained with the inclusion of three nitrogen atom in the spin Hamiltonin and with the combination of a very large A_3 value and a low g_3 value (Table 3B.1), typical for a 3N1O or 4N coordination sphere.

Moreover, within the pH titration range, the species are in a condition of reversible equilibrium, as they can revert to each of the observed states, confirming the stability of Cu^{2+}-dHisB over a wide range of pH, without any leaching of Cu^{2+} ion (Figure 3B.16b).

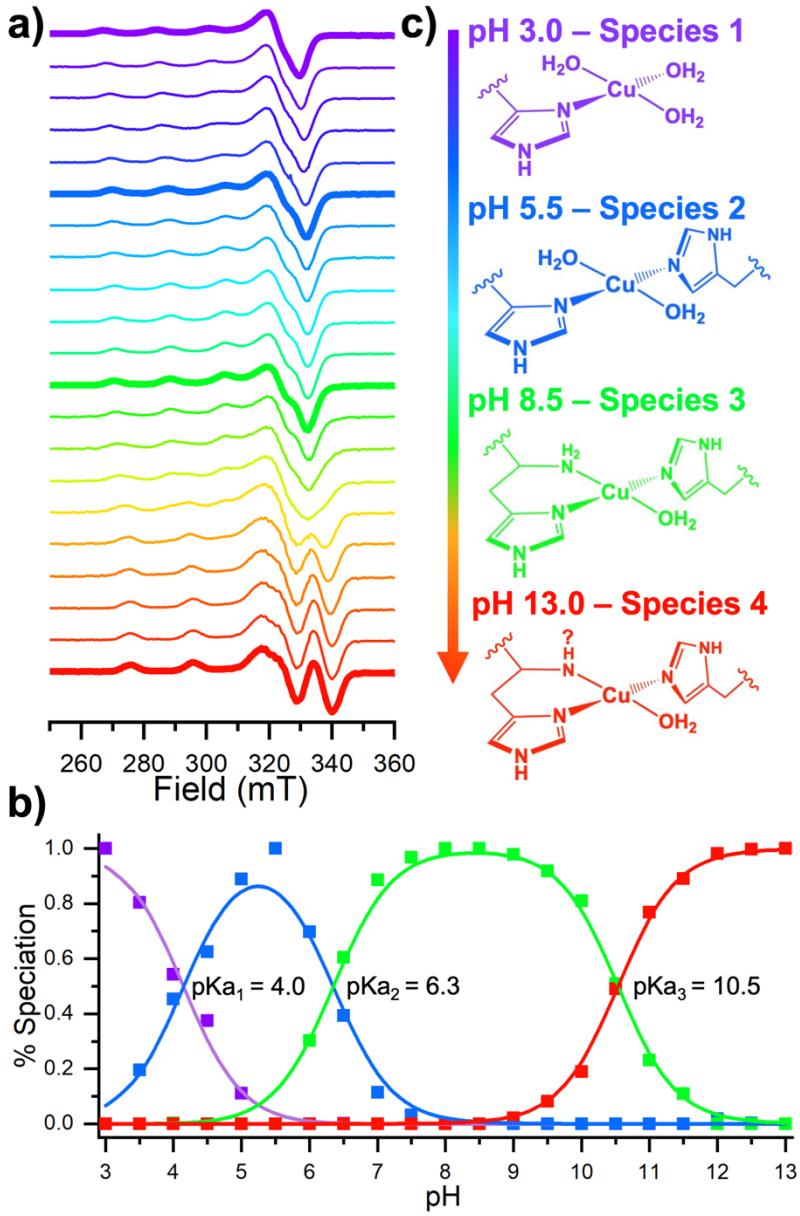


Figure 3B.5. a) Frozen solution X-band CW-EPR spectra (150 K) of Cu^{2+}-dHisB (0.4 mM) in a mixed buffer of MES, CHES, HEPES and CAPS (10 mM each) between pH 3.0 and 13.0, recorded every 0.5 pH unit. b) Speciation distribution curves of the different 4 species. c) Hypothetical protonation states of species 1 (pH 3.0), 2 (pH 5.5), 3 (pH 8.5) and 4 (pH 13.0).

Table 3B.1. Spin Hamiltonian parameters of the pH dependent species. Hyperfine tensor principal values are given in units of MHz.

Species	*g* values			Hyperfine coupling constants (MHz)			b^2 (%)	Coordination mode from P-B plot
	g_1	g_2	g_3	$\|A_1\|$	$\|A_2\|$	$\|A_3\|$		
1 (pH 3.0)	2.061	2.063	2.283	40	40	520	0.1	1N3O
2 (pH 5.5)	2.051	2.066	2.249	48	48	550	0.2	2N2O
3 (pH 8.5)	2.048	2.062	2.246	52	52	550	0.2	3N1O
4 (pH 13.0)	2.032	2.055	2.177	55	60	614	0.7	3N1O/4N

In all the simulations, the g values follow the pattern of $g_3 >> g_2 > g_1 > g_e$, which is evidence of a SOMO with a significant amount of $d(x^2 - y^2)$ character. The differences between g_1 and g_2 values, which apply to all species, point to the presence of some degree of $d(z^2)$ mixing. The conventional perturbative EPR formulas,[20] which yield the molecular orbital coefficients for the SOMO, a and b, is specified in Equation (3B.1) and may be used to calculate the degree of mixing, denoted by b^2 (Table 3B.1).

$$\psi\,(\mathrm{SOMO}) = a\psi\left(d_{x^2-y^2}\right) + b\psi(d_{z^2}) \qquad \text{Equation 3B.1}$$

Due to the minimal degrees of observed mixing, it is possible to assume that the ligand field around the copper in each of these species is close to axial four-fold symmetry. However, as various coordinating atoms may have identical ligand field characteristics, this is not always

indicative of a four-fold axial symmetry in the identity of the ligands. For species 4, a more mixing is evident. This species may exhibit a slightly more distorted geometry respect a square planar geometry.

With access to the d–d electronic transition energies, it is feasible to obtain more information about the nature of the ligand field (Figure 3B.6a). For comparison, visible absorption spectra at two different pH of *Bl*(AA10) (Figure 3B.6b), a natural LPMO, were recorded at University of York. *Bl*(AA10) does not contain a Tyr in the copper-first coordination sphere, and it is a good standard for the mimic Cu^{2+}-dHisB.

Table 3B.2 reports the energy and the molar absorption coefficient for the d-d electronic transition energies for Cu^{2+}-dHisB. The value of the natural *Bl*(AA10) LPMOare also for comparison. The visible absorption spectra of species 1 at pH 3.0 (Figure 3B.6a, violet line) and species 2 at pH 5.5 (figure 3B.6a, blue line) are similar and both show a d-d transition band characteristic of tetragonal Cu^{2+}, whose λ_{max} and ε are in good agreement with a Cu(His) and $Cu(His)_2$ with bound waters.[21] A blue shift is observed in the d-d transition band of the species 3 at pH 8.5 (Figure 3B.6a, green line), in perfect agreement with the values obtained by the spectrum of *Bl*(AA10) (Figure 3B.6b, green line) at the same pH and in general with other LPMOs.[22] Notably, the copper spin hamiltonian parameters of *Bl*(AA10) are close to those of the species 3.[19] Thus, taking together the copper spin hamiltonian values and visible absorption data, species 3 is a good spectroscopic mimics of the HB-containing natural proteins.

A more intriguing scenario supports species 4. Using g_3 value as a measure for the relative strengths of the ligand field in the equatorial

plane of the copper, it is clear that the ligand field rises dramatically when going from species 3 to 4 (g_3 from 2.246 to 2.177, Table 3B.1). Given the low value, it can be claimed that the ligand field around the copper in species 4 is extremely σ-donating, resulting in a substantial separation between the d($x^2 - y^2$) and d(xy) orbital energies. A copper-azanide interaction, which would result from the deprotonation of terminal amine to yield an $^-$NH ligating group, or a copper-deprotonated amide (from the backbone) interaction may be predicted for such a strong σ-donation. This hypothesis is in line with the pKa value (10.5) that could be assigned to the deprotonation of a coordinated amine moiety to -NH species[19,23] or amide (backbone) deprotonation[24,25] or both. A dramatic blue shift is observed in d-d energy transition of the species 4 at pH 13.0 (figure 3B.6a, red line), providing support to this hypothesis and in perfect agreement with the d-d energy transition of *BI*(AA10) at the same pH (Figure 3B.6b, red line). Furthermore, the 3-fold increase in the ε shows a less symmetric complex than species 1, 2 and 3, confirming the high degree of mixing of d(z^2) orbital to the SOMO respect the other species across the pH.

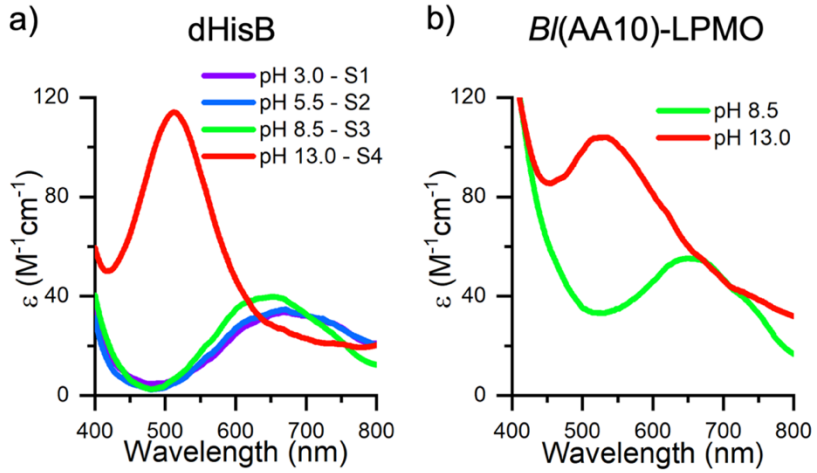


Figure 3B.6. a) Visible absorption spectra of species 1 (violet line), species 2 (blue line), species 3 (green line) and species 4 (red line). The concentration of Cu^{2+}-dHisB was fixed at 0.4 mM in 10 mM MES, 10 mM HEPES, 10 mM CHES, 10 mM CAPS. b) Visible absorption spectra of 0.2 mM *BI*(AA10) at pH 8.5 (green line) and pH 13.0 (red line) in 50 mM phosphate buffer.

Table 3B.2. d-d electronic transition energies (λ_{max}) and molar absorption coefficient (ε) for Cu^{2+}-dHisB and *BI*(AA10) at different pH.

Species	λ_{max} d-d electronic transition (nm)	ε molar absorption coefficient ($M^{-1}cm^{-1}$)
1 (pH 3.0)	670	30
2 (pH 5.5)	670	30
3 (pH 8.5)	640	35
4 (pH 13.0)	515	110
***BI*(AA10) (pH 8.5)**	650	45
***BI*(AA10) (pH 13.0)**	525	100

3B.1.4 Role of N-Terminal Amine in the copper binding

As further prove of the N-terminal amine involvement in the copper coordination, the H chain acetylated on the N-terminal amine was synthesized. Respect the species 3 of Cu^{2+}-dHisB, an increase in g_3 value is observed for Cu^{2+}-dHisB_Acetylated (from 2.246 to 2.259, Table 3B.3) consistent with less strengths of the ligand field in the equatorial plane of the copper, as the terminal amine is not anymore available for the coordination. While the species 3 of Cu^{2+}-dHis was not influenced by the presence of chloride ions, the acetylated species was very sensitive to the presence of chloride in solution. Addition of the chloride leads to the formation of a distinctive superhyperfine coupling in the high field region. To improve the spectral resolution of the narrow superhyperfine lines, second-derivative analysis was performed (Figure 3B.7, black line). The best simulation (Figure 3B.7, orange line) was obtained with the inclusion of 2 nitrogen ligands with an isotropic coupling of about 37 and 39 MHz, and with a chloride ion with an isotropic coupling of about 48 MHz. The obtained values for nitrogen and chloride ligands are compatible with the values from literature.[26,27] Other investigations are under course to further characterize the first coordination sphere . However, these spectroscopic results strongly support that dHisB coordinates Cu^{2+} ion using also the N-terminal amine, thus forming the HB motif and resembling natural LPMOs.

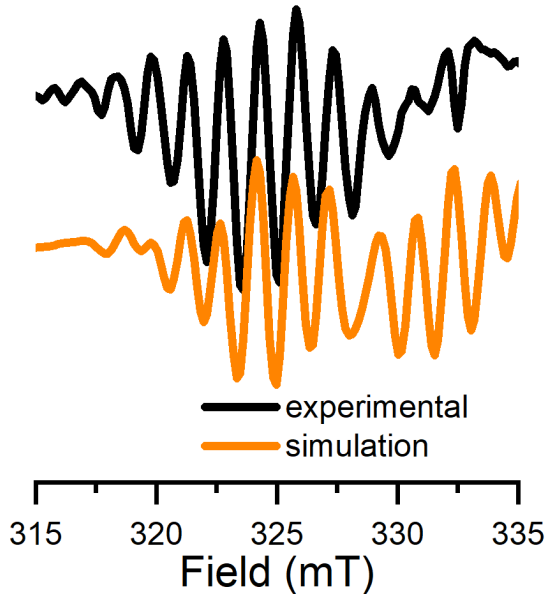


Figure 3B.7. Experimental (black line) and simulation (orange line) of the second-derivative X-band CW-EPR spectrum of a Cu^{2+}-dHisB_Acetylated (0.2 mM) in 50 mM HEPES, 250 mM NaCl, pH 7.0.

3B.1.5 dHisB reduction and Azide Binding

As described in the previous paragraph, Cu^{2+}-dHisB recapitules the spectroscopic features of natural LPMOs. Thus, it was tested its reactivity towards O_2, as natural LPMOs. It is evident from literature that LPMO function requires a reductant, and that various small molecule reducing compounds, such as the widely utilized Asc, as well as redox enzymes, such as CDH, can reduce LPMOs and trigger the O_2/H_2O_2 activation.[28,29] To mimic the LPMO reactivity, Cu^{2+}-dHisB should be able to cycle between Cu^{2+} and Cu^{+} under mild conditions. To test the possible reduction of Cu^{2+}-dHisB to Cu^{+}-dHisB by a reducing agent, EPR spectroscopy was employed as this technique is diagnostic of Cu^{2+}.

The Cu^{+}-dHisB was obtained from a degassed solution of Cu^{2+}-dHisB (whose EPR spectrum was previously recorded, Figure 3B.8, blue line) and subsequently transferred to a glovebox, where it was reduced by addition of 2 equivalents of Asc. The EPR spectra was mostly silent (Figure 3B.8, black line), confirming copper reduction. When exposed to air, Cu^{+} complex oxidizes again to the cupric form, as assessed by EPR spectrum (Figure 3B.8, cyan line).

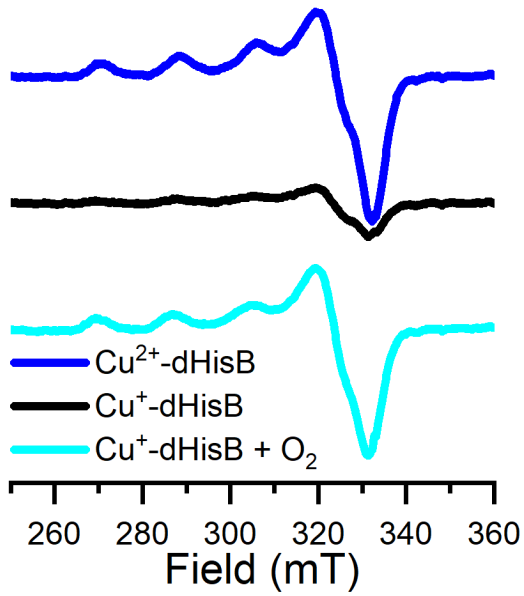


Figure 3B.8. X-band CW-EPR spectra of Cu^{2+}-dHisB (blue line), upon reduction with Asc (black line), and subsequent exposure to air (cyan line). The initial concentration of Cu^{2+}-dHisB is 0.4 mM in 10 mM MES, 10 mM HEPES, 10 mM CHES, 10 mM CAPS pH 7.0. After re-oxidation no loss of the initial intensity is observed.

Generation of copper-oxygen species, upon molecular oxygen exposition of Cu^{+}-complexes is well-documented in synthetic systems in organic solvents and at low temperature.[30] No clearly differences are observed in the EPR spectrum of Cu^{+}-dHisB upon air exposition respect the EPR spectrum before Asc addition. Therefore, the formation of a

putative copper-oxygen species, if present, exists below the detection limits of EPR spectroscopy.

Next, to demonstrate copper ion accessibility to exogenous ligands, sodium azide was employed as a spectroscopic probe. Azide is a spectroscopic analogue of superoxide (and more in general of dioxygen species), due to the striking similarity between their highest occupied orbitals. Anions, like azide, that bind to Cu^{2+} typically can cause changes in the EPR spectra, providing helpful indications for empirical spectroscopic/structural correlations. These changes are typically accompanied by LMCT transitions that appear in the visible or near-UV regions. As already studied in natural copper proteins[31] and model compounds,[32] the stereochemistry and the overall charge of the copper-complex and the coordination position occupied by the azide (equatorial vs axial) play significant roles in determining the spectroscopic features of the azide complexes and, more generally, the intrinsic affinity of azide for a given copper center.

Addition of sodium azide to a solution of Cu^{2+}-dHisB at pH 7 produces changes in both EPR and UV-Vis absorption spectra (Figure 3B.9). In the former, addition of a 2100-fold excess of azide gives a shift in g_3 from 2.246 to 2.242 and A_3 from 550 MHz to 510 MHz (Table 3B.3) with no change in the spectral envelope (Figure 3B.9a). In the UV-Vis absorption spectra (Figure 3B.9b), a single moderate-intense band ($\varepsilon \approx 1300\ M^{-1}cm^{-1}$) appears at 373 nm, ascribable to an azide to copper LMCT. The g_3 and A_3 values and the UV-Vis spectroscopic data demonstrate that Cu^{2+}-dHisB can bind small molecule, mimics of dioxygen species. By comparison with literature data, the azide should occupy the axial position in the first

coordination sphere of copper.[32] This finding is corroborated by a small increase in the molar absorption coefficient of the d–d transition absorption, with a small red-shift in the λ_{max}, that can be interpreted by assuming a change in Cu^{2+} geometry from square planar to square pyramidal, and by a low azide binding constant ((40 ± 4) M^{-1}, inset Figure 3B.9b).

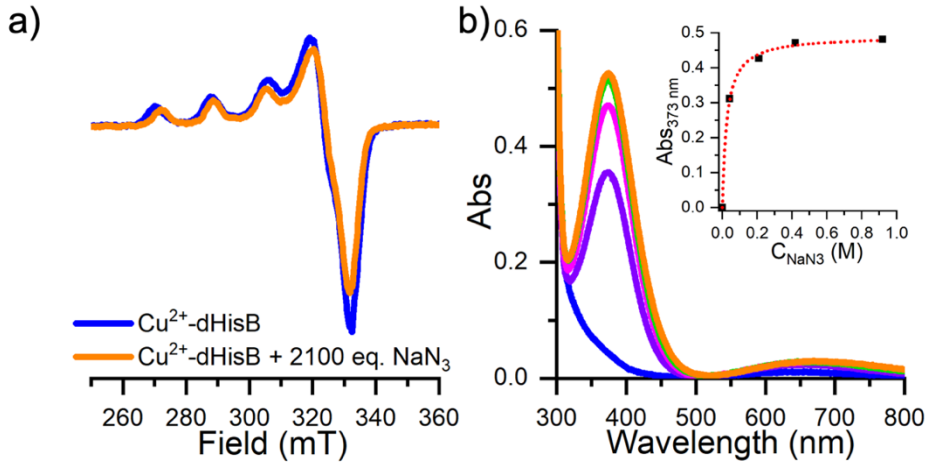


Figure 3B.9. Azide binding to Cu^{2+}-dHisB. a) X-band CW-EPR spectra of Cu^{2+}-dHisB (blue line), and after addition of 2100 equivalents of NaN_3 (orange line). b) Titration of Cu^{2+}-dHisB with NaN_3 followed by UV-Vis absorption spectroscopy (0 equivalents blue line, 100 equivalents violet line, 500 equivalents pink line, 1000 equivalents green line, 2100 equivalents orange line). Inset shows the plot of the absorbance at 373 nm as function of NaN_3 concentration. The absorbance is corrected for the protein background. The black squares are the experimental data, which have been fitted with a single binding model (dashed red line). Cu^{2+}-dHisB concentration in all experiments was 0.42 mM in 10 mM MES, 10 mM HEPES, 10 mM CHES, 10 mM CAPS pH 7.0.

3B.1.6 Dioxygen activation by Cu-dHisB

Decoding the LPMOs' activity is a difficult task. The polymeric structure of natural LPMO substrates and the wide range of released products make the assessment of the activity difficult. LPMO and a polymeric substrate are typically incubated under established assay conditions for a specific period (generally days), after which the results

are quantified using a calibrated HPLC technique. Low molecular weight cleavage products are the only reaction products detectable by HPLC because they are soluble in both the reaction buffer and the HPLC eluent. LPMOs use O_2 or H_2O_2 as cosubstrate for monooxygenases or peroxygenase reaction, respectively. A reductant provides the electrons necessary for a monooxygenase reaction to take place, or for the *in situ* production of H_2O_2 that is later utilized in a peroxygenase reaction, thus driving the LPMO reaction. It is widely recognized that in the absence of substrate, hydrogen peroxide can be generated as a result of the oxidase activity of LPMOs; this is commonly referred to as the "uncoupled reaction".[22] The monitoring of enzyme-dependent H_2O_2 production can be utilized to identify active LPMOs, in a fast and inexpensive LPMO assay.

The Horseradish Peroxidase (HRP)/Amplex Red test is an extensively used fast quantitative method for detecting H_2O_2 in living organisms,[33] cell cultures,[34] cell extracts,[35] and other enzymatic processes.[36] Kittl and coworkers used this assay for this first time to monitoring the H_2O_2 production in presence of O_2 and a reductant catalyzed by LPMO.[22] Using Amplex red conversion to resorufin in the presence of HRP, the assay provides an indirect measurement of peroxide concentration. The HRP/Amplex Red method is based on the single-electron oxidation of Amplex Red by HRP, which is H_2O_2-dependent. The resulting Amplex Red radicals are unstable, and two of them undergo a dismutation reaction that result in the regeneration of one Amplex Red molecule and the production of the highly absorbent (λ_{max} = 570 nm) and highly fluorescent (λ_{ex} = 570 nm, λ_{em} = 585 nm) reporter molecule, resorufin

(Figure 3B.10). Therefore, the generation of resorufin is 1:1 proportional to the amount of H_2O_2 present in the reaction. However, possible side-reactions, as the Amplex Red radical reduction by reductant (side-reaction I in figure 3B.10) or direct reductant peroxidation, catalyzed by HRP (side-reaction II in figure 3B.10), may interfere with this method.

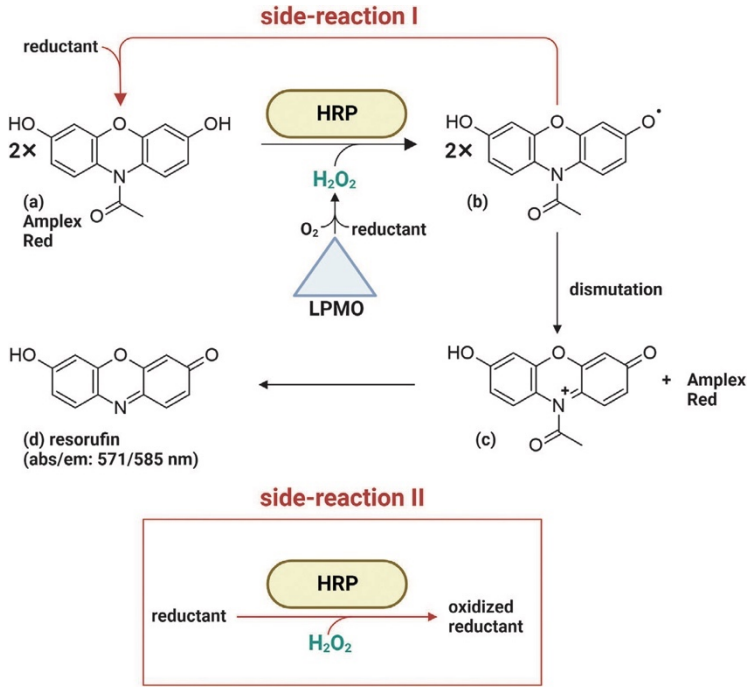


Figure 3B.10. Schematic illustration of the HRP/Amplex Red assay conducted in the presence of LPMO and a reducing agent, with the two possible side-reactions. Adapted with permission from reference.[37] Copyright 2023 Elsevier Inc.

To gain insight into the O_2-activation activity of Cu^{2+}-dHisB, H_2O_2 production was measured *in situ* in presence of different reductants. The progress curves show that the apparent H_2O_2 production rate is high with Asc, making it the most efficient reductant for Cu^{2+}-dHisB (Figure 3B.11a). Next, the H_2O_2 production rate by Cu^{2+}-dHisB, in the presence of Asc, was compared with $CuSO_4$ (Figure 3B.11b). In this case, the concentration of H_2O_2 generated was evaluated using a calibration curve

(Figure 3B.18), in which standard solutions of H_2O_2 were reacted with HRP and Amplex Red solution. Standard solutions for H_2O_2 quantification were supplemented with the reductant to consider potential side-reactions, as discussed earlier. The apparent H_2O_2 production rate, calculate from the linear parts of the progress curves in Figure 3.11b, resulted $(2.8 \pm 0.1) \cdot 10^{-2}$ s^{-1} and $(8.9 \pm 0.1) \cdot 10^{-3}$ s^{-1} for Cu^{2+}-dHisB and $CuSO_4$, respectively. Surprisingly, Cu^{2+}-dHisB shows a higher apparent H_2O_2 production respect to $CuSO_4$ and comparable to that of natural LPMOs.[37]

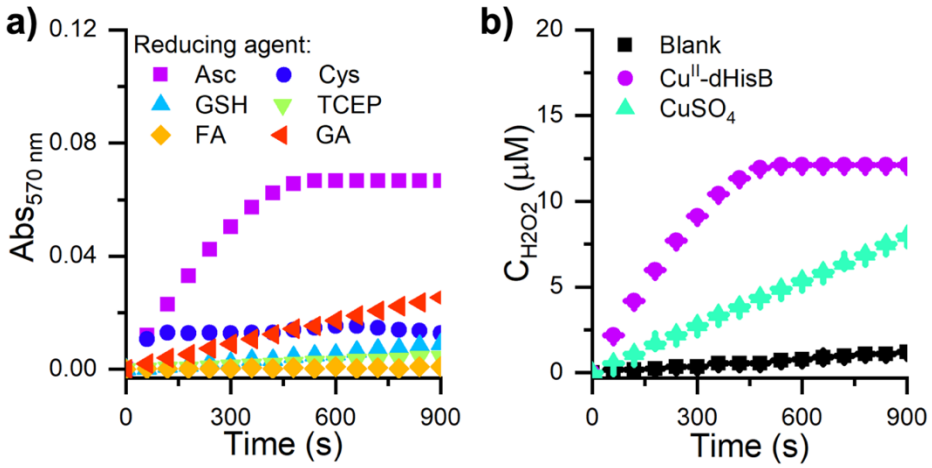


Figure 3B.11. a) Progress curved, monitored at 570 nm, showing the accumulation of resorufin molecule. Reactions were performed with 1 μM Cu^{2+}-dHisB and 25 μM of either Asc (violet), Cys (blue), GSH (cyan), TCEP (green), FA (yellow) and GA (red). b) H_2O_2 accumulation in a reaction containing 25 μM Asc and buffer (black) or 1 μM $CuSO_4$ (light green) or 1 μM Cu^{2+}-dHisB (violet). All the reaction were performed combining two solutions in a 1:1 ratio. The first solution contained 2 μM Cu^{2+}-dHisB or 2 μM $CuSO_4$ in 50 mM HEPES, 150 mM NaCl, pH 7.0 while the second solution contained 10 U/mL HRP and 100 μM Amplex Red in 20 mM phosphate, 150 mM NaCl pH 7.0. The reaction was initiated by the addition of reductant to a final concentration of 25 μM.

To explain these different apparent rates in H_2O_2-production between Cu^{2+}-dHis and free copper, different reactions should be taken into account. The first reaction is the reduction of Cu^{2+} to Cu^{+} (Equation

3B.2), followed by binding of O_2 to Cu^+ to form a superoxide complex, $[CuO_2]^+$ (Equation 3B.3). H_2O_2 can be directly formed by two-electron reduction of O_2 without release of ROS into solution (Equation 3B.4). Alternatively, $[CuO_2]^+$ could dissociate into Cu^{2+} and O_2^- anion. H_2O_2 could be formed in a reaction between O_2^- and the reductant (Equation 3B.5) or through O_2^- dismutation (Equation 3B.6), eventually catalyzed by Cu^{2+}.

$$Cu^{2+} + e^- \rightarrow Cu^+ \qquad \text{Equation 3B.2}$$

$$Cu^+ + O_2 \rightarrow [CuO_2]^+ \qquad \text{Equation 3B.3}$$

$$[CuO_2]^+ + e^- + 2H^+ \rightarrow Cu^{2+} + H_2O_2 \qquad \text{Equation 3B.4}$$

$$O_2^- + RedH_2 \rightarrow H_2O_2 + Red^+ \qquad \text{Equation 3B.5}$$

$$O_2^- + 2H^+ \xrightarrow{Cu^{2+}} O_2 + H_2O_2 \qquad \text{Equation 3B.6}$$

A very likely scenario could be the dismutation of O_2^- catalyzed by Cu^{2+}-dHisB, as already observed in other copper de novo designed proteins, containing a Cu(Nitrogen)$_3$ center.[38] As a further possibility, the first coordination sphere ligands may be redox active, facilitating electron transfer to the superoxide in the $[CuO_2]^+$ complex, similar to a mechanism recently reported by Walton and coworkers.[39]

3B.2 Experimental Section

3B.2.1 Design

A procedure analogous to the one use for miniLPMO design, reported in the experimental section of chapter 3A, was also implemented for dHisB.

3B.2.2 Molecular Dynamics simulations

All MD simulations were performed at 330 K, without any explicitly defined geometric constraint on the metal site, with N-term of chain A acetylated, C-term of chain A amidated and C-term of chain B amidated. Coordinates were written to .dcd file every 5000 integration steps (10 ps). Octahedral box was built with 9 Å edge distance from the center. First coordination sphere was saturated with three H_2O molecules in all cases. RMSD and R_g were calculated on all backbone atoms. All the other details for the MD simulations have been described in the experimental section of chapter 3A.

3B.2.3 Synthesis

The 28-mer peptide H, the acetylated H, and the 62-mer peptide HLH were synthesized by automatic solid-phase synthesis as previously reported in chapter 2. Only for the peptide H, to preserve the N-terminal free, H it was not acetylated, whereas the C-terminal was instead amidated. It was purified to homogeneity via RP-HPLC. Identity was ascertained by high resolution ESI-MS (Figure 3B.12, 3B.13, 3B.14).

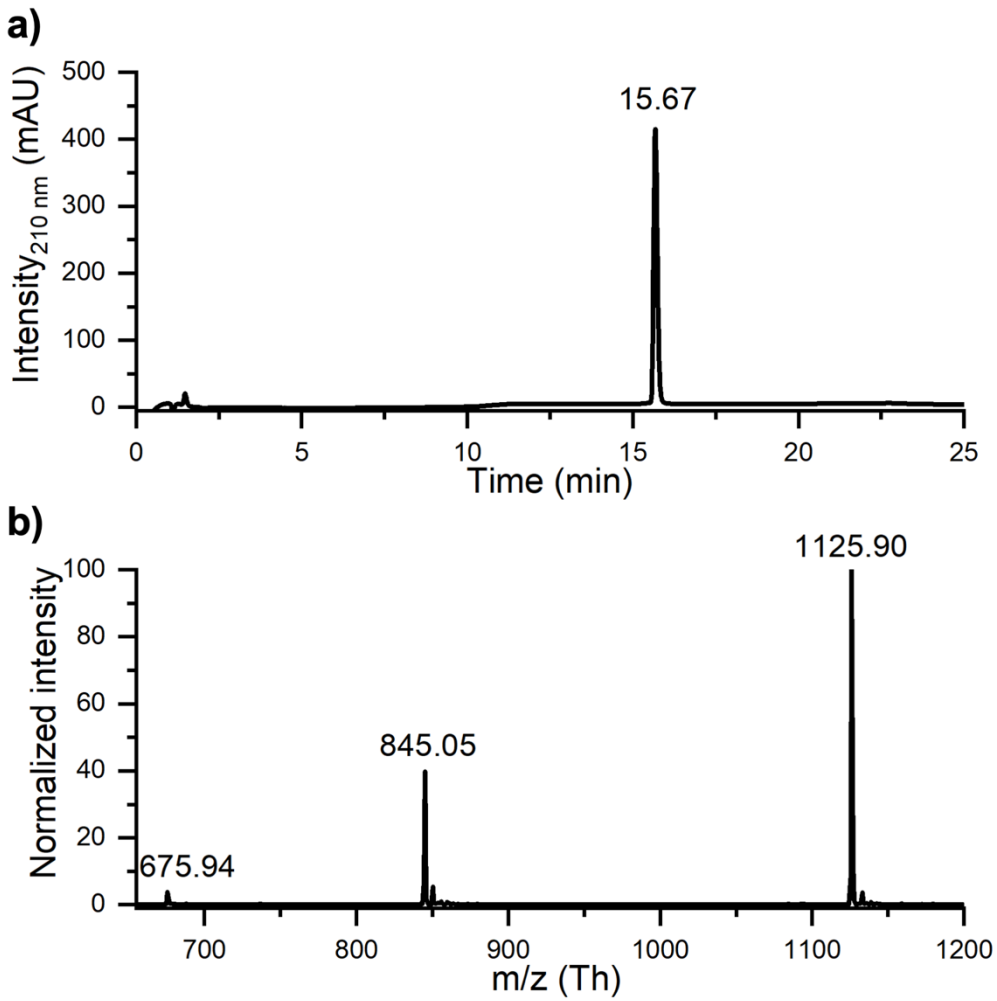


Figure 3B.12. RP-HPLC chromatogram (210 nm trace) of the pure H; b) ESI-MS spectrum of pure H. The signal at m/z = 1125.90 Th corresponds to the $[M+3H^+]^{3+}$ ion (theoretical average m/z: 1125.95 Th); the signal at m/z = 845.05 Th corresponds to the $[M+4H^+]^{4+}$ ion (theoretical average m/z: 844.72 Th); the signal at m/z = 675.94 Th corresponds to the $[M+5H^+]^{5+}$ ion (theoretical average m/z: 675.97 Th).

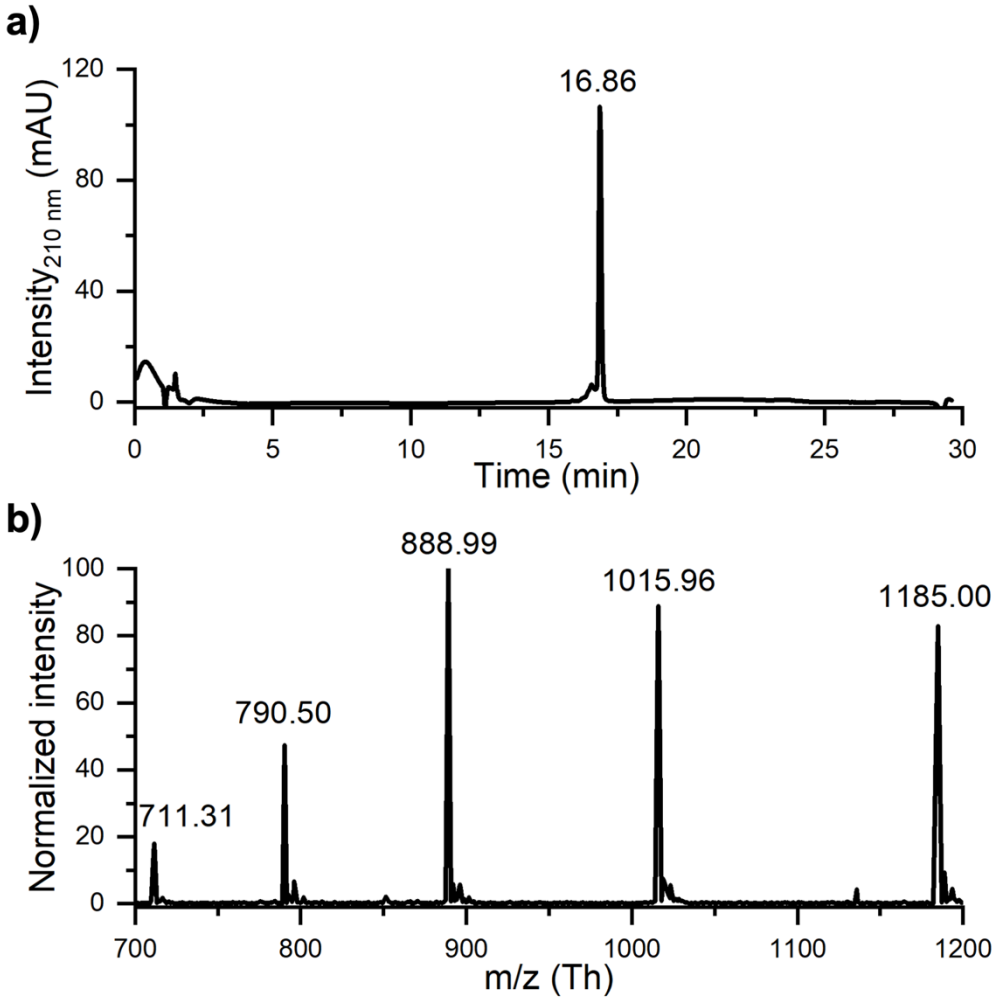


Figure 3B.13. RP-HPLC chromatogram (210 nm trace) of the pure HLH; b) ESI-MS spectrum of pure HLH. The signal at m/z = 1185.00 Th corresponds to the $[M+6H^+]^{6+}$ ion (theoretical average m/z: 1185.19 Th); the signal at m/z = 1015.96 Th corresponds to the $[M+7H^+]^{7+}$ ion (theoretical average m/z: 1016.02 Th); the signal at m/z = 888.99 Th corresponds to the $[M+8H^+]^{8+}$ ion (theoretical average m/z: 889.14 Th); the signal at m/z = 790.50 Th corresponds to the $[M+9H^+]^{9+}$ ion (theoretical average m/z: 790.46 Th); the signal at m/z = 711.31 Th corresponds to the $[M+10H^+]^{10+}$ ion (theoretical average m/z: 711.52 Th). Spectrum deconvolution gave (7105 ± 1.0) Da (theoretical average mass: 7105.10 Da).

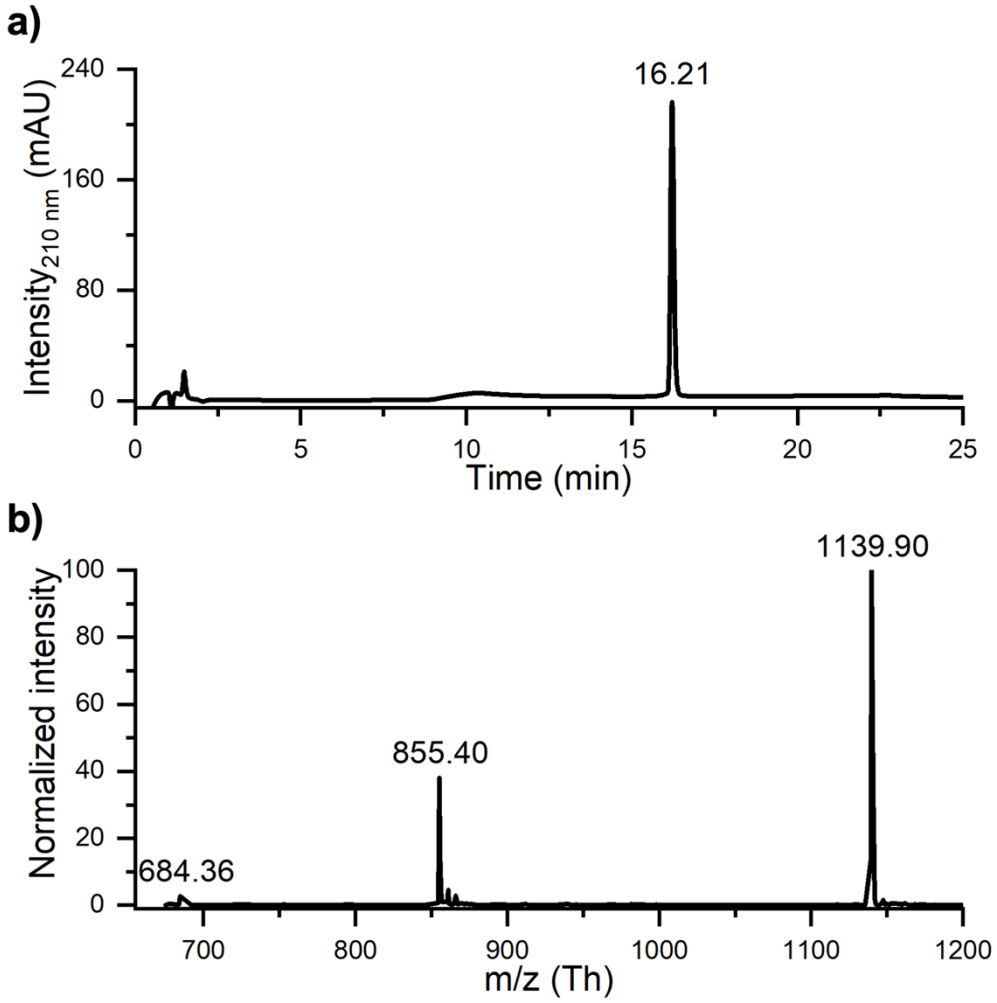


Figure 3B.14. RP-HPLC chromatogram (210 nm trace) of the pure H_Acetylated; b) ESI-MS spectrum of pure H_Acetylated. The signal at m/z = 1139.90 Th corresponds to the $[M+3H^+]^{3+}$ ion (theoretical average m/z: 1139.96 Th); the signal at m/z = 855.40 Th corresponds to the $[M+4H^+]^{4+}$ ion (theoretical average m/z: 855.22 Th); the signal at m/z = 684.36 Th corresponds to the $[M+5H^+]^{5+}$ ion (theoretical average m/z: 684.38 Th).

3B.2.4 Size-exclusion chromatography

Analytical SEC was carried out on an ÄKTA pure (GE) fitted with a Superdex 75 10/300 and eluted at 6°C with 50 mM HEPES, 150 mM

NaCl, pH 7.0 at a 0.2 mL/min flow rate. The dHisB complex was formed combining two solutions in a 1:1 ratio. The first solution contained 600 μM H while the second solution contained 600 μM HLH. To this mixed solution (300 μM H + 300 μM HLH), 330 μM of $CuSO_4$ was added to obtain the holo form.

Molecular weight of the sample was determined based on a calibration curve obtained with standards (Figure 3B.15a). The GE Low Molecular Weight (LMW) calibration kit, containing Conalbumin (75 kDa), Ovalbumin (44 kDa), Carbonic anhydrase (29 kDa), Ribonuclease A (13.7 kDa) and Aprotinin (6.5 kDa), all at a concentration of 1 mg/ml, was used for calibration.

A value of ≈ 14 kDa was found for the molecular weight, compatible with the formation of a dimer. To ascertain the correct oligomerization state (dimer), the following analysis was performed to calculate the molar ratio between H and HLH. 500 μL of mix solution containing 1.2 mM H, 1.2 mM HLH and 1.5 mM $ZnSO_4$ was injected into Superdex 75 16/600 SEC column. An aliquot of the main SEC peak was acidified and 10 μL of this solution were injected into a reversed-phase Vydac C4 column. The concentration of peptide H was then calculated by integrating the peak and comparing the area to a previously acquired external standard calibration curve (Figure 3B.15b). The HLH peptide concentration may be then easily extrapolated by subtracting the H peptide contribution to the 280 nm absorbance of the acidified SEC fraction (ε_{280nm} H = 5500 $M^{-1}cm^{-1}$; ε_{280nm} HLH = 5500 $M^{-1}cm^{-1}$). The registered absorbance was 0.61, and the HPLC-derived H peptide concentration of 52 uM led to a H/HLH molar ratio of 0.9 ± 0.1.

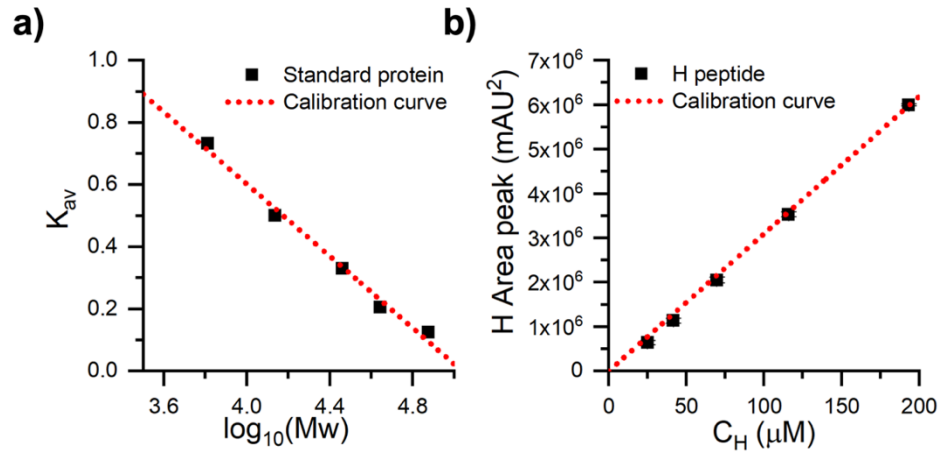


Figure 3B.15. a) Calibration curve, acquired with analytic SEC, with standard proteins: Conalbumin (75 kDa), Ovalbumin (44 kDa), Carbonic anhydrase (29 kDa), Ribonuclease A (13.7 kDa) and Aprotinin (6.5 kDa). b) Calibration curve, acquired with RP-HPLC, with different concentration of H peptide.

3B.2.5 CD spectroscopy

CD measurements were performed using a J-815 spectropolarimeter equipped with a thermostated cell holder (JASCO, Easton, MD, USA). CD spectra were collected at 25°C, from 260 to 190 nm at 0.2 nm intervals with a 20 nm min-1 scan speed, at 2.5 nm band width and at 16 s response. A quartz cell of 1 mm path length was used. MRE were calculated using the equation:

$$\mathrm{MRE} = \frac{\theta_{obs}}{10\,l\,C\,n}$$

in which θ_{obs} is the ellipticity measured in millidegrees, l is the path length of the cell in centimeters, C is the concentration in moles per liter, and n is the number of residues of the peptide (28 for H, 62 for HLH and 90 for H-HLH dimer). The difference CD spectrum of 20 μM H bound was obtained from the subtraction of 20 μM 1:1 complex of H with HLH spectrum, from that determined for 20 μM HLH alone.

3B.2.6 Copper reconstitution

H, acetylated H, HLH and copper stock solutions were freshly prepared in H_2O milliQ. The stock solutions of $CuSO_4$ were prepared at 10 mM concentration starting from the solid $CuSO_4 \cdot 5H_2O$. For H, acetylated H and HLH, the concentration was determined spectrophotometrically by using ε_{280nm} = 5500 $M^{-1}cm^{-1}$.[40] 500 µL of a solution containing 1.0 mM H (or acetylated H), 1.0 mM HLH, 1.3 mM $CuSO_4$ was prepared in 50 mM HEPES, 150 mM NaCl. The pH of this solution was adjusted to pH 7.0 adding small amount of NaOH. After 30 minutes of incubation at room temperature, excess copper was removed by HiLoad SEC. HiLoad SEC was carried out on an ÄKTA pure (GE) fitted with a Superdex 75 16/600 and eluted at 6 °C with 50 mM HEPES, 150 mM NaCl, pH 7.0 at 1 mL min^{-1} flow rate. The eluted fractions corresponding to the main SEC peak were then concentrated with a 3 kDa MWCO VivaSpin centrifuge concentrator and stored at -80 °C, prior further experiments.

3B.2.7 CW-EPR spectroscopy

EPR analysis was performed at University of York (UK) under the supervision of prof. Paul Walton.

X-band CW-EPR spectra of copper loaded dHisB samples in frozen solution were acquired on a Bruker micro EMX spectrometer operating at ≈ 9.30 GHz and 150 K, with a modulation amplitude of 4 G and microwave power of 10.02 mW. EPR spectra were collected at different pH values. The buffer of the Cu^{2+}-dHisB stock solution was exchanged

with a mixed buffer solution composed of 10 mM MES, 10 mM HEPES, 10 mM CHES, 10 mM CAPS with a 3 kDa MWCO VivaSpin centrifuge concentrator. The stock solution of Cu^{2+}-dHisB was divided into 23 independent aliquots (each at 0.4 mM Cu^{2+}-dHisB) and each was brought to the desired pH value by the addition of concentrated solutions of NaOH and H_2SO_4. The pH of the each resulting solution at room temperature was determined using a calibrated FisherbrandTM AccumetTM AE150 Benchtop pH Meter, from which 21 samples from pH 2.0 to 13.0 were prepared in 0.5 pH unit steps. At each pH, the sample was rapidly cooled to 150 K and its X-band CW-EPR spectrum collected. The spectra at pH 2.0 and 2.5 show the characteristic signal of tetragonal aqua ion copper (Figure 3B.16a). To verify that dHisB in complex with copper is stable over a wide range of pH, EPR spectra were collected on sample starting at pH 7.0 and then raising to pH 13.0, and then recollected on the same sample at a pH 7.0 again, indeed restoring the previous species. (Figure 3B.16b).

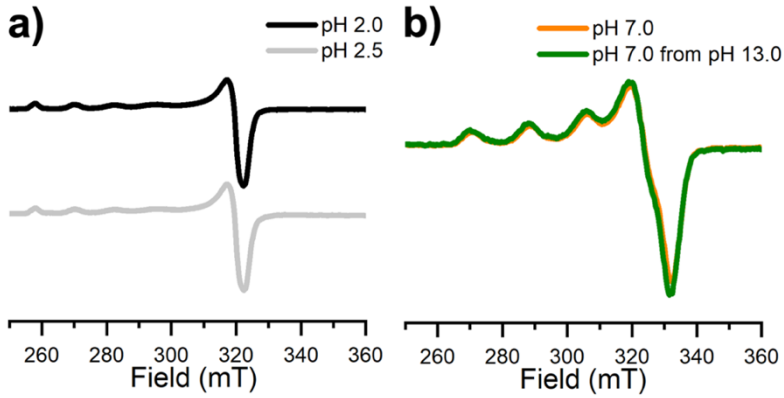


Figure 3B.16. a) Frozen solution X-band CW-EPR spectra (150 K) of Cu^{2+}-dHisB (0.4 mM) in a mixed buffer of MES, CHES, HEPES and CAPS (10 mM each) at pH 2.0 (black line) and pH 2.5 (grey line). The EPR spectra show the characteristic signal of tetragonal aqua ion copper. b) Frozen solution X-band CW-EPR spectra (150 K) of Cu^{2+}-dHisB (0.4 mM) in a mixed buffer of MES, CHES, HEPES and CAPS (10 mM each) at pH 7.0 (orange line) and pH 7.0 coming from pH 13.0 (green line).

For azide binding experiments, a stock solution of 5 M NaN_3 was prepared by dissolving the solid in 10 mM MES, 10 mM HEPES, 10 mM CHES, 10 mM CAPS and the pH of this solution was adjusted to 7.0 with H_2SO_4. From the 5 M NaN_3 solution stock, 5 different diluted stocks were prepared in order to add the same volume to 5 independent Cu^{2+}-dHisB solutions. In this was the Cu^{2+}-dHisB concentration is fixed (0.42 mM) and the concentration of sodium azide is changed (0 mM, 42 mM, 210 mM, 420 mM, 882 mM).

For reduction experiments, the reduced Cu^{+} state of dHisB was generated by reaction of the complex with 2 equivalents of Asc inside a N_2 atmosphere glovebox. The excess of Asc was then removed via buffer exchange with a 3 kDa MWCO VivaSpin centrifuge concentrator. All of the solutions used inside the N_2 atmosphere glovebox were degassed by freeze–pump–thawing on a Schlenk line (water and buffer) or by purging the solution with N_2 for 30 min (dHisB). dHisB re-oxidation was obtained by ≈ 24 h air exposure of the reduced sample.

Each CW-EPR spectra were obtained by subtracting the spectrum of Cu-dHisB with blank (only buffer). The simulations of CW-EPR spectra were performed using the Easyspin software package,[41] running within the MathWorks MATLAB environment.

The EPR spectra at pH 3.0, 5.5, 8.5, and 13.0 were simulated with a single set of spin Hamiltonian parameters. Spectra taken at different pH values could be satisfactory simulated through a weighted combination of these spectra. All but the pH 13.0 spectrum contributions were systematically varied between 0 and 1 while minimizing the residual sum of squares between the experimental spectrum and the calculated

spectrum. The contribution of each spectrum was determined by performing a stochastic search starting from a 1000 population size and using a nonlinear gradient-based optimization algorithm using the Microsoft Excel Solver tool. The pH 13.0 spectrum contribution was calculated by imposing 1 as the sum of the weighted contributions.

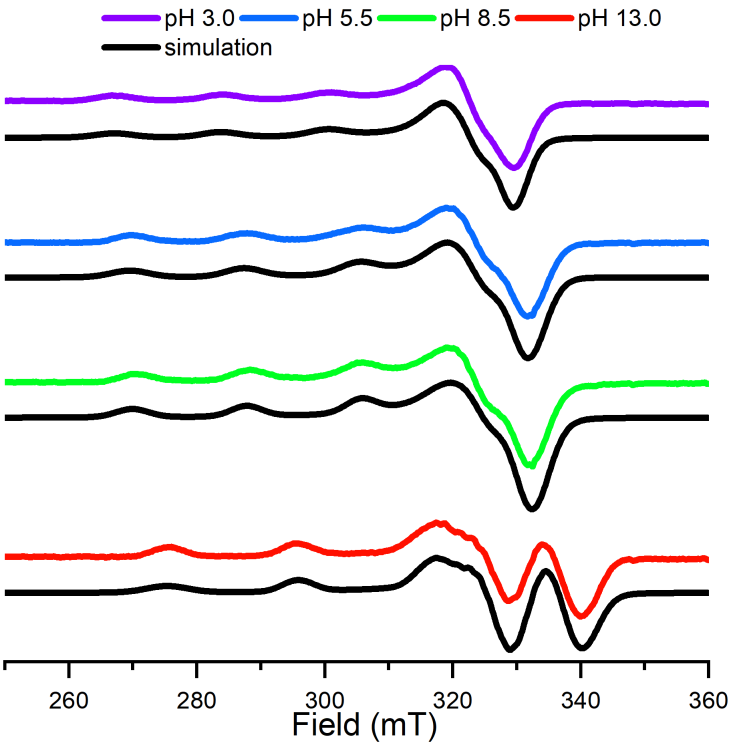


Figure 3B.17. Frozen solution X-band CW-EPR spectra (150 K) of Cu^{2+}-dHisB (0.4 mM) in a mixed buffer of MES, CHES, HEPES and CAPS (10 mM each) at pH 3.0 (violet line), pH 5.5 (blue line), pH 8.5 (green line) and pH 13.0 (red line). The black traces are the simulated spectra.

Table 3B.3. EPR spin Hamiltonian parameters from simulations of X-band CW-EPR spectra for the species studied in this work.

		Species 1 (pH 3.0)	Species 2 (pH 5.5)	Species 3 (pH 8.5)	Species 4 (pH 13.0)	dHisB + 2100 eq. NaN_3 (pH 7.0)	dHisB Acetylated (pH 7.0)
g **values**	g_1	2.061	2.051	2.048	2.032	2.048	2.048
	g_2	2.063	2.066	2.062	2.055	2.062	2.052
	g_3	2.283	2.249	2.246	2.177	2.242	2.259
A_{Cu} **(MHz)**	$\lvert A_1 \rvert$	40	48	52	55	52	48
	$\lvert A_2 \rvert$	40	48	52	60	52	58
	$\lvert A_3 \rvert$	520	550	550	614	510	550
g **strains**	g_1	0.03	0.04	0.04	0.001	0.04	/
	g_2	0.03	0.04	0.04	0.001	0.04	/
	g_3	0.05	0.05	0.04	0.001	0.04	/
A_{Cu} **strains (MHz)**	$\lvert A_1 \rvert$	1	1	1	10	1	7
	$\lvert A_2 \rvert$	1	1	1	10	1	7
	$\lvert A_3 \rvert$	1	1	1	100	1	7
SHF							
A_N **isotropic (MHz)**		/	/	/	30 41 50	/	37 39 48 (Cl^-)
Line widths		0.8 0.8	0.1 0.3	0.8 0.8	0.9 0.9	0.8 0.8	0.7 0.7

3B.2.8 UV-Vis absorption spectroscopy

UV–Vis absorption spectra were recorded with a Shimadzu UV-1800 UV/Visible Scanning Spectrophotometer, using a quartz cuvette with 1 cm path length. Wavelength scans were performed at 25°C from 200 to 800 nm, with a 600 nm min^{-1} scan speed. The same solutions used for EPR analysis were employed for UV-Vis absorption analysis. The UV-Vis spectra of *Bl*(AA10), expressed and purified at University of York, in 50 mM phosphate buffer were recorded at pH 8.5 and pH 13.0.

For azide binding experiments, the same solutions used for EPR analysis were employed for UV-Vis absorption analysis. This analysis allowed to evaluate the binding constant for equilibrium:

$(Cu^{2+}\text{-dHisB})\text{-}H_2O + N_3^- \leftrightarrows (Cu^{2+}\text{-dHisB})\text{-}N_3^- + H_2O$

Data were fitted according to the following equation:

$$A_{373nm} = \frac{K_a A_\infty [N_3^-] + A_0}{(1 + K_a[N_3^-])}$$

where K_a is the association constant, A_∞ is the absorption of the complex at infinite concentration of N_3^- and A_0 is the absorption of the starting aqua-complex.

3B.2.9 Amplex Red assay

The oxygen reactivity of Cu^{2+}-dHisB was measured by HRP/Amplex Red assay in 96-well plates (total volume of 100 µL) using a Perkin Elmer EnSpire Multimode plate reader and monitoring wavelength at 570 nm. All the reaction were performed combining two solutions in a 1:1 ratio. The first solution contained 2 µM Cu^{2+}-dHisB or 2 µM $CuSO_4$ in 50 mM HEPES, 150 mM NaCl, pH 7.0 while the second solution contained

10 U/mL HRP and 100 μM Amplex Red in 20 mM phosphate, 150 mM NaCl pH 7.0. The reaction was initiated by the addition of Asc to a final concentration of 25 μM. The reactions were performed in triplicate. To calculate the amount of H_2O_2 production, a calibration curve (Figure 3B.18) was obtained using standard samples of H_2O_2, prepared in the same buffer solution used for the experiment (with the addition of 25 μM Asc), at concentrations of 0 μM, 1 μM, 5 μM, 10 μM and 20 μM. H_2O_2 stock solutions were prepared by diluting H_2O_2 (30%, v/v) in water, whose concentration was determined by UV- vis absorption spectroscopy (λ_{max} = 240 nm; ε = 39.4 $M^{-1}cm^{-1}$).

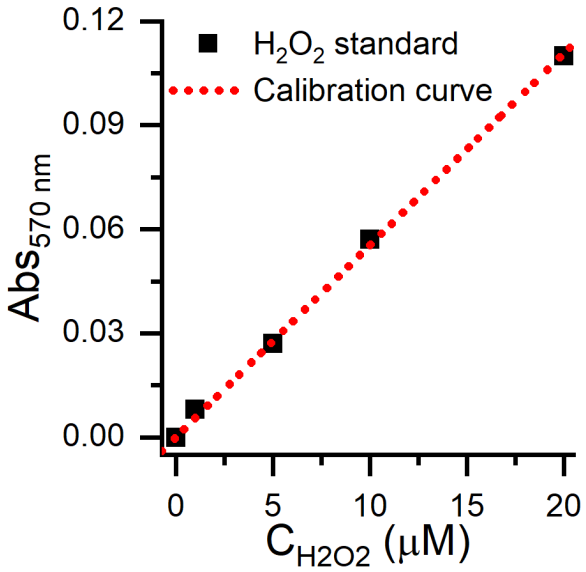


Figure 3B.18. Calibration curve for H_2O_2 for Amplex Red assay.

3.3. Conclusions

This part of the thesis has been devoted to the design, spectroscopic and catalytic characterization of miniLPMO and dHisB, as models of LPMOs. As starting scaffold, the use of proteins structurally close to the natural enzyme was avoided, considering that all natural LPMOs share a core formed by a β-sandwich that includes the metal-binding residues in its surface. Scaffold selection was therefore steered towards proteins with helical secondary structure. Scaffolds with high α-helical content are particularly convenient because they tend to have a lower propensity to form insoluble oligomers. Another restriction imposed on the hypothetical scaffold concerns the length of the polypeptide chain that was limited to about 50 amino acids or less, for avoiding synthetic problems and low yield.

Both de novo designed protein scaffolds, for miniLPMO a quite unnatural homodimeric four-helix bundle and for dHisB a heterodimeric three-helix bundle, were found able to host an HB copper-binding site. To the best of our knowledge, these models represent the first de novo designed proteins containing the HB copper-binding site

The correct oligomerization state and folding for both proteins were confirmed by CD spectroscopy and SEC analysis. Unlike naturally occurring proteins, these models allowed for a detailed spectroscopic characterization over a wide pH range. CW/pulsed-EPR and UV-Vis absorption spectroscopy demonstrated that the copper binding site of both proteins can exist in different protonation states. Based on the available data, the species at neutral pH are proposed to have the same first coordination sphere of natural LPMOs for both models.

More interestingly, the two models not only recapitulate the spectroscopic signature of the natural counterparts but were also found to activate both H_2O_2 and O_2. Next steps will be devoted to more complete catalytic studies for both models.

References

3B.3 References

[1] S. F. Betz, W. F. DeGrado, *Biochemistry* **1996**, *35*, 6955–6962.
[2] W. F. DeGrado, C. M. Summa, V. Pavone, F. Nastri, A. Lombardi, *Annu. Rev. Biochem.* **1999**, *68*, 779–819.
[3] R. B. Hill, D. P. Raleigh, A. Lombardi, W. F. DeGrado, *Acc. Chem. Res.* **2000**, *33*, 745–754.
[4] B. Lovejoy, S. Choe, D. Cascio, D. K. McRorie, W. F. DeGrado, D. Eisenberg, *Science* **1993**, *259*, 1288–1293.
[5] G. R. Dieckmann, D. K. McRorie, D. L. Tierney, L. M. Utschig, C. P. Singer, T. V. O'Halloran, J. E. Penner-Hahn, W. F. DeGrado, V. L. Pecoraro, *J. Am. Chem. Soc.* **1997**, *119*, 6195–6196.
[6] G. R. Dieckmann, D. K. McRorie, J. D. Lear, K. A. Sharp, W. F. DeGrado, V. L. Pecoraro, *J. Mol. Biol.* **1998**, *280*, 897–912.
[7] A. Lombardi, J. W. Bryson, G. Ghirlanda, W. F. DeGrado, *J. Am. Chem. Soc.* **1997**, *119*, 12378–12379.
[8] J. Zhou, G. Grigoryan, *Protein Sci.* **2015**, *24*, 508–524.
[9] S. J. Lahr, D. E. Engel, S. E. Stayrook, O. Maglio, B. North, S. Geremia, A. Lombardi, W. F. DeGrado, *J. Mol. Biol.* **2005**, *346*, 1441–1454.
[10] R. Evans, M. O'Neill, A. Pritzel, N. Antropova, A. Senior, T. Green, A. Žídek, R. Bates, S. Blackwell, J. Yim, O. Ronneberger, S. Bodenstein, M. Zielinski, A. Bridgland, A. Potapenko, A. Cowie, K. Tunyasuvunakool, R. Jain, E. Clancy, P. Kohli, J. Jumper, D. Hassabis, *bioRxiv* **2022**, 2021.10.04.463034.
[11] C. M. Summa, M. M. Rosenblatt, J.-K. Hong, J. D. Lear, W. F. DeGrado, *J. Mol. Biol.* **2002**, *321*, 923–938.
[12] N. F. Polizzi, W. F. DeGrado, *Science* **2020**, *369*, 1227–1233.
[13] G. R. Hemsworth, B. Henrissat, G. J. Davies, P. H. Walton, *Nat. Chem. Biol.* **2014**, *10*, 122–126.
[14] S. Y. Lau, A. K. Taneja, R. S. Hodges, *J. Biol. Chem.* **1984**, *259*, 13253–13261.
[15] K. T. O'Neil, W. F. DeGrado, *Science* **1990**, *250*, 646–651.
[16] J. P. Schneider, A. Lombardi, W. F. DeGrado, *Fold. Des.* **1998**, *3*, R29–R40.
[17] F. L. Aachmann, M. Sørlie, G. Skjåk-Bræk, V. G. H. Eijsink, G. Vaaje-Kolstad, *Proc. Natl. Acad. Sci.* **2012**, *109*, 18779–18784.
[18] J.-P. Bacik, S. Mekasha, Z. Forsberg, A. Y. Kovalevsky, G. Vaaje-Kolstad, V. G. H. Eijsink, J. C. Nix, L. Coates, M. J. Cuneo, C. J. Unkefer, J. C.-H. Chen, *Biochemistry* **2017**, *56*, 2529–2532.

[19] P. J. Lindley, A. Parkin, G. J. Davies, P. H. Walton, *Faraday Discuss.* **2022**, *234*, 336–348.
[20] A. A. Gewirth, S. L. Cohen, H. J. Schugar, E. I. Solomon, *Inorg. Chem.* **1987**, *26*, 1133–1146.
[21] K. Krupa, M. Korabik, T. Kowalik-Jankowska, *J. Inorg. Biochem.* **2019**, *201*, 110819.
[22] R. Kittl, D. Kracher, D. Burgstaller, D. Haltrich, R. Ludwig, *Biotechnol. Biofuels* **2012**, *5*, 79.
[23] M. R. McDonald, F. C. Fredericks, D. W. Margerum, *Inorg. Chem.* **1997**, *36*, 3119–3124.
[24] C. J. K. Wijekoon, T. R. Young, A. G. Wedd, Z. Xiao, *Inorg. Chem.* **2015**, *54*, 2950–2959.
[25] M. Seal, S. G. Dey, *Inorg. Chem.* **2018**, *57*, 129–138.
[26] M. Iwaizumi, T. Kudo, S. Kita, *Inorg. Chem.* **1986**, *25*, 1546–1550.
[27] K. E. H. Frandsen, T. J. Simmons, P. Dupree, J.-C. N. Poulsen, G. R. Hemsworth, L. Ciano, E. M. Johnston, M. Tovborg, K. S. Johansen, P. von Freiesleben, L. Marmuse, S. Fort, S. Cottaz, H. Driguez, B. Henrissat, N. Lenfant, F. Tuna, A. Baldansuren, G. J. Davies, L. Lo Leggio, P. H. Walton, *Nat. Chem. Biol.* **2016**, *12*, 298–303.
[28] B. Wang, Z. Wang, G. J. Davies, P. H. Walton, C. Rovira, *ACS Catal.* **2020**, *10*, 12760–12769.
[29] G. R. Hemsworth, *Essays Biochem.* **2023**, EBC20220164.
[30] C. E. Elwell, N. L. Gagnon, B. D. Neisen, D. Dhar, A. D. Spaeth, G. M. Yee, W. B. Tolman, *Chem. Rev.* **2017**, *117*, 2059–2107.
[31] D. M. Dooley, C. E. Cote, *Inorg. Chem.* **1985**, *24*, 3996–4000.
[32] L. Casella, M. Gullotti, G. Pallanza, M. Buga, *Inorg. Chem.* **1991**, *30*, 221–227.
[33] O. Karakuzu, M. R. Cruz, Y. Liu, D. A. Garsin, *Bio-Protoc.* **2019**, *9*, e3409.
[34] L. A. Maddalena, S. M. Selim, J. Fonseca, H. Messner, S. McGowan, J. A. Stuart, *Biochem. Biophys. Res. Commun.* **2017**, *493*, 246–251.
[35] S. Chakraborty, A. L. Hill, G. Shirsekar, A. J. Afzal, G.-L. Wang, D. Mackey, P. Bonello, *Methods* **2016**, *109*, 105–113.
[36] V. Mishin, J. P. Gray, D. E. Heck, D. L. Laskin, J. D. Laskin, *Free Radic. Biol. Med.* **2010**, *48*, 1485–1491.
[37] A. A. Stepnov, V. G. H. Eijsink, in *Methods Enzymol.*, Academic Press, **2022**.
[38] E. Mathieu, A. E. Tolbert, K. J. Koebke, C. Tard, O. Iranzo, J. E. Penner-Hahn, C. Policar, V. Pecoraro, *Chem. – Eur. J.* **2020**, *26*, 249–258.
[39] J. Zhao, Y. Zhuo, D. Diaz, M. Shanmugam, A. Telfer, P. Lindley, D. Kracher, T. Hayashi, L. Seibt, F. Hardy, O. Manners, T. Hendison, K. Hollywood, S. Diaz-Moreno, N. Scrutton, M. Tovborg, G. Davies, P. Walton, D. Heyes, A. Green, *The Histidine-Brace Active Site of a Copper Monooxygenase Is Redox Active and Forms Part of an in-Built Enzyme Repair Mechanism*, In Review, **2022**.

References – 3B. Mimicking the Histidine Brace motif in De Novo Designed Proteins

[40] C. N. Pace, F. Vajdos, L. Fee, G. Grimsley, T. Gray, *Protein Sci.* **1995**, *4*, 2411–2423.
[41] S. Stoll, A. Schweiger, *J. Magn. Reson.* **2006**, *178*, 42–55.

4. Conclusions and Perspectives

This PhD thesis demonstrates the feasibility of de novo designed helical bundles to engineer T2Cu and T3Cu catalytically active copper binding sites and obtaining exceptional replicas of natural metalloproteins.

Engineering artificial metalloenzymes housing di-nuclear metal binding sites is not a trivial task. Differently from metalloenzymes containing well-defined prosthetic group, the architecture of these sites is much more intricate as it results from a defined geometry of the first shell ligands, usually also arranged in a bridging fashion connecting the two metal ions. As in T3Cu PPO the di-copper site cycles between different functional oxidation states, the protein scaffold must be able to host the site and tolerate these variations.

Common structural motif of the T2Cu active site of LPMOs and pMMOs, which both catalyzes the oxidation of exceedingly strong C-H bond, is the HB copper-binding site. The presence of this peculiar metal-binding motif raised questions concerning its role. Therefore, the development of mimics, which reproduce the HB and are sufficient stable for structural and functional studies, is also a challenging task.

The results on the three model developed successfully demonstrate the reliability of the approach used in this thesis. The newly designed protein DR1 appropriately hosts the T3 copper site and exhibits diphenolase activity. Careful design of the residues lining the substrate access endows DR1 with substrate recognition, as demonstrated by Hammett analysis and computational studies. This finding is particularly meaningful because such modulated catalytic activity

represents a hallmark of natural enzymatic activity. Further spectroscopic characterization led us to propose a simple but highly informative kinetic model for dioxygen activation of this synthetic enzyme, highlighting hydrogen peroxide as an intermediate.

For LPMO model proteins the de novo design procedure started with the selection of appropriate protein scaffolds, suitable to the introduction of HB metal-binding site. Besides satisfying the structural requirements for the HB site, potential candidates were also selected for generally desirable features in an artificial protein, such as water solubility and conformational stability. Thus, secondary structures lacking β-sheet segments or highly unordered regions were preferred, in order to minimize aggregation and subsequent precipitation.

Along this line, both the scaffolds selected for miniLPMO and dHisB are based on α-helices. For miniLPMO, a HB binding site was designed into a quite unnatural de novo designed four-helix bundle. In the latter, dHisB, the HB binding site has been hosted into a “designable” de novo designed three-helix bundle. For both models an in-depth spectroscopic analysis was performed through the combination of different spectroscopic techniques. Spectroscopic and functional characterization revealed these simple models can recapitulate the HB features.

. In conclusion, this PhD thesis set a new bar, by implanting specific outer shell interactions and peculiar copper binding sites in non-native, versatile, and stable scaffolds. In the endeavor to develop synthetic metalloenzymes for the degradation and conversion of biomass into second-generation fuels, these simple models represent a significant step forward.

5. List of Abbreviations

AA: Auxiliary Activity

Amplex Red: 10-acetyl-3,7-dihydroxyphenoxazine

Asc: Ascorbic acid

AUS: Aurone Synthase

bmTYR: Tyrosinase from Bacillus megaterium

Bpy: 2,2′-Bipyridine

CAPS: 3-(Cyclohexylamino)propane-1-sulfonic acid

CAZy: Carbohydrate-Active Enzymes

CBM: Carbohydrate-Binding Module

CBP21: Carbohydrate Binding Protein 21

CD: Circular Dichroism

CDH: Cellobiose Dehydrogenase

CHES: 2-(Cyclohexylamino)ethane-1-sulfonic acid

CO: Catechol Oxidase

CT: Charge Transfer

CuA: Copper A

CuZ: Copper Z

CW-EPR: Continuous Wave-Electron Paramagnetic Resonance

DF: Due Ferri

DFT: Density Functional Theory

dHisB: designed Histidine Brace

DMP: 2,6-dimethoxyphenol

DR1: Due Rame

DTBC: 3,5-di-tert-butylcatechol

ENDOR: Electron Nuclear Double Resonance

5. List of Abbreviations

EPR: Electron Paramagnetic Resonance

ET: Electron Transfer

EXAFS: X-ray Absorption Fine Structure

FA: Ferulic acid

FT-IR: Fourier-Transform Infrared Spectroscopy

GA: Gallic acid

GSH: L-Glutathione

HAA: Hydrogen Atom Abstraction

HATU: 1-[Bis(dimethylamino)methylene]-1H-1,2,3-triazolo[4,5-b]pyridinium 3-oxide hexafluorophosphate

HB: Histidine Brace

Hc: Hemocyanin

HCO: Heme-Copper Oxidase

HCTU: O-(1H-6-Chlorobenzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate

HEPES: 2-[4-(2-Hydroxyethyl)piperazin-1-yl]ethane-1-sulfonic acid

HPLC: High-Performance Liquid Chromatography

HppE: (S)-2-hydroxypropylphosphonic acid epoxidase enzyme

HRP: Horseradish Peroxidase

HYSCORE: Hyperfine Sublevel Correlation

ibCO: Catechol Oxidase from Ipomoea batatas

IZ: Isoleucine Zipper

jrTYR: Tyrosinase from Juglans regia

LMCT: Ligand-to-metal Charge Transfer

LPMO: Lytic Polysaccharide Monooxygenase

MASTER: Method of Accelerated Search for Tertiary Ensemble Representatives

MCO: Multicopper Oxidase

MD: Molecular Dynamics

MES: 2-morpholin-4-ylethanesulfonic acid

MRE: Mean Residue Ellipticity

MtPMO3*: Lytic Polysaccharide Monooxygenase subtypecloned from the filamentous ascomycete from Myceliophthora thermophila

N2OR: Nitrous Oxide Reductase

NiR: Nitrite Reductase

PDB: Protein Data Bank

pMMO: particulate Methane Monooxygenase

PNP: 4-nitrophenolate anion

PNPG: 4-nitropheyl-β-D-glucopyranoside

PPO: Polyphenol Oxidase

PQQ: Pyrroloquinoline-quinone

QM/MM: Quantum Mechanics/Molecular Mechanics

RMSD: Root Mean Square Deviation

RMSF: Root Mean Square Fluctuation

ROS: Reactive Oxygen Species

RP-HPLC: Reversed-Phase High-Performance Liquid Chromatography

SASA: Solvent Accessible Surface Area

SEC: Size-Exclusion Chromatography

sMMO: soluble Methane Monooxygenase

SOD: Superoxide Dismutase

SOMO: Single Occupied Molecular Orbital

T1Cu: Type 1 Copper

T1D: Type 1 Depleted

5. List of Abbreviations

T2Cu: Type 2 Copper

T2D: Type 2 Depleted

T3Cu: Type 3 Copper

TBC: 4-tert-butylcatechol

TCEP: Tris(2-carboxyethyl)phosphine

TFA: Trifluoroacetic acid

TNC: Trinuclear center

TYR: Tyrosinase

vvCO: Catechol Oxidase from Vitis vinifera

WSC: Wall Stress Component

XANES: X-ray Absorption Near Edge Structure

XAS: X-ray Absorption Spectroscopy