

AperTO - Archivio Istituzionale Open Access dell'Università di Torino

A De Novo-Designed Type 3 Copper Protein Tunes Catechol Substrate Recognition and Reactivity

This is the author's manuscript

Original Citation:

Availability:

This version is available <http://hdl.handle.net/2318/1882080> since 2023-01-17T10:19:36Z

Published version:

DOI:10.1002/anie.202211552

Terms of use:

Open Access

Anyone can freely access the full text of works made available as "Open Access". Works made available under a Creative Commons license can be used according to the terms and conditions of said license. Use of all other works requires consent of the right holder (author or publisher) if not exempted from copyright protection by the applicable law.

(Article begins on next page)

A De Novo-Designed Type 3 Copper Protein Tunes Catechol Substrate Recognition and Reactivity

Fabio Pirro,^{*,[a]} Salvatore La Gatta,^{*,[a]} Federica Arrigoni,^[b] Antonino Famulari,^[c,d] Ornella Maglio,^[a,e] Pompea Del Vecchio,^[a] Mario Chiesa,^[c] Luca De Gioia,^[b] Luca Bertini,^[b] Marco Chino,^{*,[a]} Flavia Nastri,^{*,[a]} Angela Lombardi^{*,[a]}

Dedicated to Professor Vincenzo Pavone on the occasion of his retirement and in recognition of his remarkable contributions to peptide chemistry and metalloprotein models.

Abstract: De novo metalloprotein design is a remarkable approach to shape protein scaffolds toward specific functions. Here, we report the design and characterization of Due Rame 1 (DR1), a de novo designed protein housing a di-copper site and mimicking the Type 3 (T3) copper-containing polyphenol oxidases (PPOs).

To achieve this goal, we hierarchically designed the first and the second di-metal coordination spheres to engineer the di-copper site into a simple four-helix bundle scaffold. Spectroscopic, thermodynamic, and functional characterization revealed that DR1 recapitulates the T3 copper site, supporting different copper redox states, and being active in the O₂-dependent oxidation of catechols to o-quinones. Careful design of the residues lining the substrate access site endows DR1 with substrate recognition, as revealed by Hammett analysis and computational studies on substituted catechols. This study represents a premier example in the construction of a functional T3 copper site into a designed four helix bundle protein.

Introduction

Impressive progress has been made in engineering metal binding sites in designed and native protein scaffolds, to afford artificial metalloenzymes.^[1–8] Protein scaffolds can be handled as macromolecular ligands to build metal ion binding sites within micro-environments of increasing complexity to modulate their properties and afford a variety of activities.^[7–9] Recent results in this field reveal that endless possibilities of artificial metalloenzymes can be obtained, not only exceeding the performances of their natural counterparts but also showing innovative chemistry, unprecedented in nature.^[10–14]

The engineering of artificial metalloenzymes housing di-nuclear metal binding sites deserves special attention. Di-nuclear sites are widely used throughout biology to catalyze hydrolytic and redox-active processes.^[9,15–18] Differently from metalloenzymes containing well-defined prosthetic group such as hemes or iron-sulfur clusters, di-nuclear metal binding sites result from the interactions between metal ions and residue side chains in a precise spatial distribution of donor atoms as first shell ligands.^[9] The architecture of these sites holds further level of complexity, as it frequently encompasses bridging first-shell ligands connecting the two metal ions.

Type 3 (T3) copper-containing proteins are an illustrative example of di-nuclear sites. They comprise two copper ions, at a distance between 3.5 to 4.6 Å,^[18–20] each ligated by three His residues. Each His triad is provided by an antiparallel α -helix pair, with consequent formation of an uncommon crossed four-helix bundle. The di-cupreous *deoxy* form consists of two Cu⁺ ions with no bridging ligands in a trigonal planar geometry. The di-cupric *met*-form has two Cu²⁺ bridged by one or two hydroxides or water molecules (Figure 1a), whereas the *oxy*-form is bridged by the characteristic side-on (μ - η^2 : η^2) peroxy anion in the Cu₂O₂ core, in a quasi-tetragonal or trigonal bipyramidal geometries (Figure 1b).

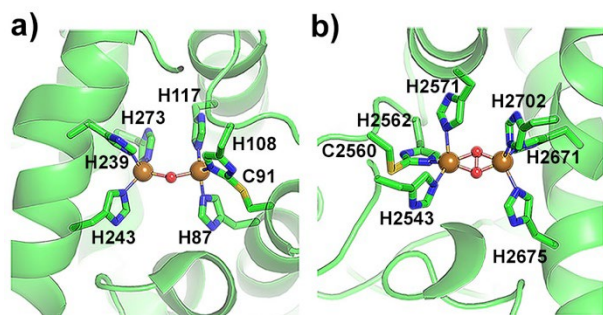


Figure 1 Active site structures of a) *met*-TYR from walnut leaves (*Juglans regia*, PDB ID: 5CE9,^[19] with the two copper ions bridged by a hydroxide anion; b) oxyhemocyanin from *Octopus* (PDB ID 1JS8) with the peroxide bridging the two copper ions in a $\mu\eta^2:\eta^2$ coordination mode.^[20] The 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.

In hemocyanins, the T3 site functions as dioxygen transporter, whereas the side-on Cu₂O₂ performs very demanding hydroxylation reactions as in tyrosinases (TYR), or oxidation of aromatic substrates as in catechol oxidases (CO). TYR and CO, together with aurone synthase (AUS), are generally classified as polyphenol oxidases (PPOs), and the identification of the molecular basis driving the T3 copper center toward mono- and/or diphenolase activity is still an unsolved puzzle.^[21]

As very recently highlighted by Lu et al.,^[4] the design of artificial oxidative metalloenzymes housing copper centers represents a great challenge, as it requires a protein scaffold able to: i) finely colocalize the two copper ions using two triads of histidine residues without

endogenous bridging ligands; ii) be flexible to change the coordination environment, depending on copper ions oxidation state; iii) enable the activation of molecular oxygen without ROS formation; iv) enable the substrate access to the metal site for catalysis.

We have approached this challenge through de novo design strategy and herein we report the results on Due Rame 1 (DR1, two copper in Italian), an artificial metalloenzyme housing a T3 copper site endowed with catechol oxidase activity. DR1 design started from DFs, de novo designed metalloproteins we previously developed, inspired by the di-iron-oxo proteins.^[22–24] The first member of the DF family, DF1, consists of an antiparallel four-helix bundle and binds two iron ions in a 4-Glu, 2-His coordination environment embedded into the protein core.^[22] Scaffold stability allowed increasing levels of complexity, thus obtaining analogues with a variety of activities and substrate/product specificity.^[24–27]

By using a combination of rational and computational design, we reshaped the first coordination sphere of DF1 active site to obtain a 6-His di-copper binding site. Second coordination sphere interactions were engineered through an extensive H-bond network involving the His ligands. Finally, the active site cavity was sculpted to create a suitable channel for substrate access.

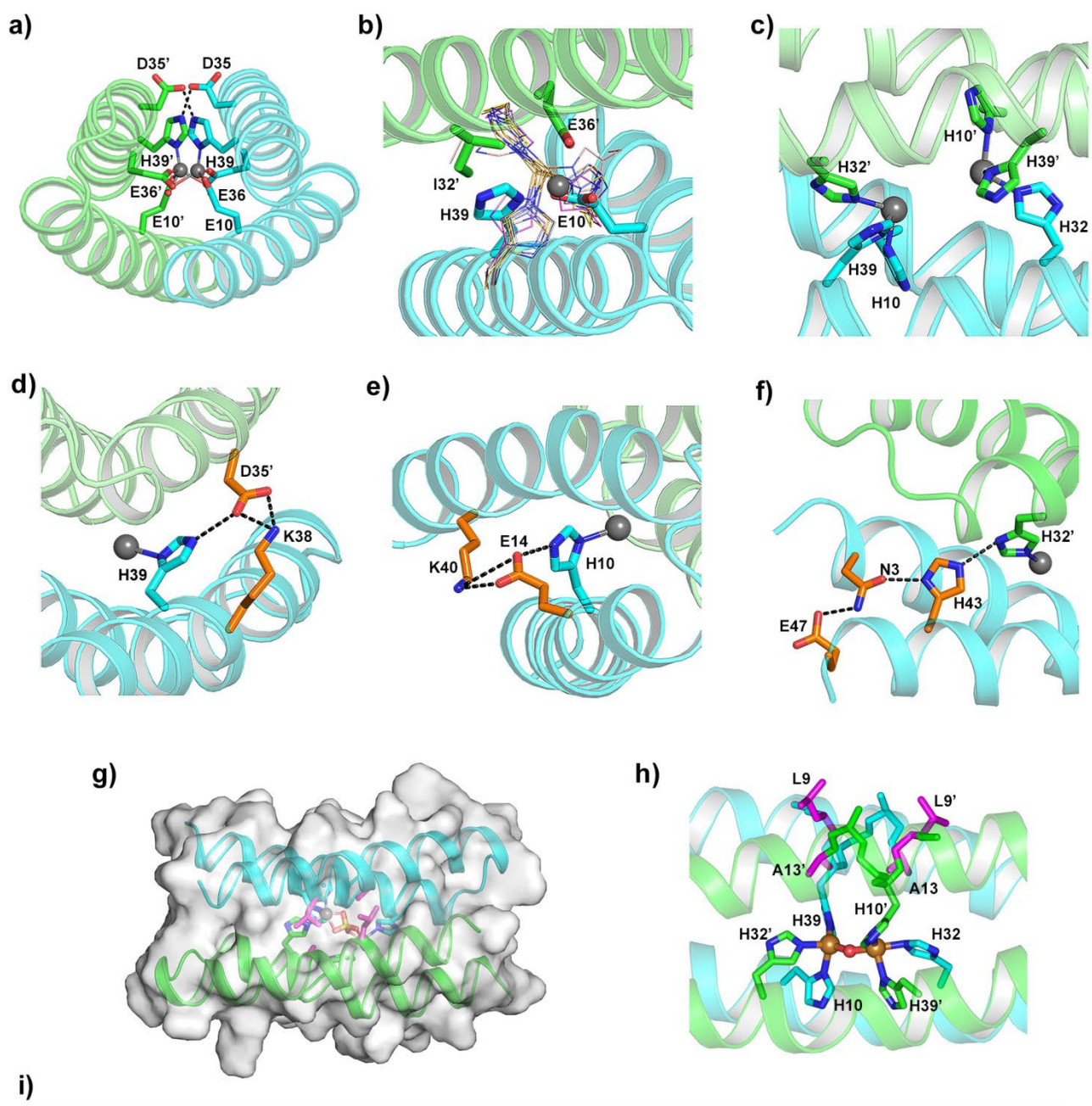
All the results demonstrated DR1 to behave as a well folded and stable four-helix bundle protein, able of replicating the T3 copper binding site. Indeed, di-Cu²⁺-DR1 displays diphenolase activity, catalyzing the oxidation of catechols to the corresponding quinones. Analysis of the kinetic parameters for differently substituted catechols revealed di-Cu²⁺-DR1 active site to preferentially interact with more hydrophobic substrates. The combination of Hammett structure–activity analysis and density functional theory (DFT) computations, to model the substrate interactions within the T3 copper site, provided evidence for a key role of the residues lining the active site pocket in tuning di-Cu²⁺-DR1 substrate specificity.

Results and Discussion

Engineering a T3 copper center into the four-helix bundle scaffold: first and second coordination spheres

To build the T3 site into the four-helix bundle scaffold, we started from the symmetric crystal structure of di-Zn²⁺-DF1 (PDB ID 1EC5).^[22] DF1 scaffold is comprised of two non-covalently associated helix-loop-helix motifs (α_2), with the primary ligands of the di-nuclear Glu₄His₂ site buried in the protein interior (Figure 2a). All the design process was performed preserving the C₂ 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, pointing inwards the bundle, and having the C α atoms located within 7 Å from the Zn²⁺ ions.^[28] Besides positions 10, 36 and 39, which coincide with the coordinating residues of the DF1 di-nuclear site, position 32' (numbers with prime refer to symmetry-related residues in the symmetric dimer) appeared as the only suitable candidate (Figure 2b). Next, we superimposed the x-ray structure of di-Zn²⁺-DF1 with the structures of several T3 copper proteins (Figure 2b and Table S1). To reduce the number of possible permutations, we kept His39 of the original di-metal site as the first member of the His triad. Then, we determined the best position for His mutation among Glu10, Glu36, Ile32' by iteratively superimposing the His39 to each His residues of the natural proteins. After this analysis, the best triad, in terms of ligand distance and coordination geometry, was His10, His32', and His39. Therefore, Glu10/10' and Ile32/32' were mutated to His. Notably, substitution of Ile to His in a corresponding position was previously adopted in 3His-DF_{SC}, a single-chain version of DFs, in order to reprogram the reactivity from oxidation of activated quinols to the N-oxygenation of anilines.^[27] To avoid competition for Cu binding, Glu36/36' were mutated to Ala, thus affording a 6His di-copper binding site in the homodimeric structure (Figure 2c).

Subsequently, we introduced hydrogen-bonded networks, crucial for stabilizing metal sites in de novo designed proteins.^[29,30] As successfully adopted in DF1 (Figure 2a), His39 was stabilized through the interaction of its N ϵ with Asp35'. This Asp is further involved in a salt bridge interaction with Lys38 (Figure 2d). His10 second shell interaction required Ile14Glu and Ile40Lys substitutions. Lys and Glu sidechains create a salt-bridge, which, by interacting with His N ϵ , may prevent the formation of undesired side chain/main chain H-bonds (Figure 2e). Finally, similarly to 3His-DF_{SC}, His32' was stabilized through an H-bond network, spanning almost the entire length of the bundle, by Leu43His, Leu3Asn and Leu47Glu substitution. (Figure 2f).



	Helix 1														Helix 2																																	
	a	b	c	d	e	f	g	a	b	c	d	e	f	g	a	b	c	d	e	f	g	a	b	c	d	e	f																					
DF1	D	Y	L	R	E	L	L	<u>K</u>	<u>L</u>	<u>E</u>	L	Q	<u>L</u>	I	K	Q	Y	R	E	A	L	E	<u>V</u>	<u>K</u>	L	P	V	L	A	K	I	L	E	<u>D</u>	<u>E</u>	<u>E</u>	<u>K</u>	<u>H</u>	I	E	W	L	E	T	I	L	G	
DF3	D	Y	L	R	E	L	L	<u>K</u>	<u>G</u>	<u>E</u>	L	Q	<u>G</u>	I	K	Q	Y	R	E	A	L	E	<u>Y</u>	<u>T</u>	<u>H</u>	N	P	V	L	A	K	I	L	E	<u>D</u>	<u>E</u>	<u>E</u>	<u>K</u>	<u>H</u>	I	E	W	L	E	T	I	L	G
DR1	D	<u>Y</u>	<u>N</u>	R	E	L	<u>I</u>	<u>K</u>	<u>L</u>	<u>H</u>	<u>E</u>	<u>Q</u>	<u>A</u>	<u>E</u>	<u>K</u>	<u>Q</u>	<u>L</u>	R	E	A	L	E	<u>Y</u>	<u>T</u>	<u>H</u>	N	P	V	L	A	K	<u>H</u>	<u>I</u>	<u>E</u>	<u>D</u>	<u>A</u>	<u>E</u>	<u>K</u>	<u>H</u>	<u>K</u>	<u>E</u>	<u>W</u>	<u>H</u>	<u>E</u>	<u>T</u>	<u>I</u>	<u>E</u>	<u>G</u>

Figure 2. DR1 design. a) active site structure of di-Zn²⁺-DF1 (PDB ID: 1ec5). b) superposition of DF1 metal binding site, with focus of the a and d residues within 7 Å from one Zn²⁺ ion, with one coordinating His triad from several T3 copper proteins (PDB IDs: see Table S1). Side chains and the metal ion of DF1 are represented as sticks and sphere, 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; g) Centroid structure of the most abundant cluster determined on DR1 trajectory by cluster analysis (See Supporting Information and Table S3); h) details of the di-Cu²⁺-DR1 active site as obtained by DFT calculations. His coordinating residues and Ala and Leu, shaping the access cavity, are illustrated; i) peptide sequence alignment of DR1 α₂ 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. In all panels, the two symmetry-related α₂ motifs are represented as cartoons in different colors (cyan and green) for clarity of display. The Zn²⁺ ions are reported as grey spheres, the Cu²⁺ ions as brown spheres.

Shaping active site accessibility and design validation

Having designed the putative 6His di-copper metal-binding site, we carefully evaluated the hydrophobic core repacking, the loop

definition, and the active site accessibility to face the tradeoff between protein stability and function.^[31] The primary and secondary sphere mutations, necessary for the T3 copper site design, extract a large thermodynamic cost,^[31] and required stabilizing substitutions at positions distant from the active site, both in the hydrophobic core and in the loop region.^[25,26]

To rebuild the hydrophobic core of the homodimer, a RosettaScripts repacking cycle^[32,33] led to the following mutations: Leu7Ile, Leu11Glu, Tyr17Leu, Leu33Ile (see Supporting Information). Afterwards, the idealized $\alpha_R-\alpha_L-\beta$ interhelical loop (Thr24-His25-Asn26) was preferred over the DF1 loop sequence (Val24-Lys25-Leu26). The selected loop sequence features a network of H-bonded sidechain/mainchain interactions and improves the stability/flexibility of the bundle, as previously observed in DF3 and more recently in DFP3.^[34,35]

The shape and accessibility of the active site cavity in DF proteins^[23,24] are mainly controlled by residues at positions 9/9' and 13/13', corresponding to *g* and *d* positions of the α_2 motif, respectively. In DF1, decreasing the bulk of the channel-lining residues, by Leu to Ala to Gly mutations, considerably improved di-metal center hydration and reactivity.^[23,24,36,37] Thus, we performed a comprehensive screening of the best pair of residues in positions 9 and 13 by Molecular Dynamics (MD) simulations (see Supporting Information). Leu, Ala and Gly residues were iteratively permuted starting from the most hydrophobic, but still accessible,^[38,39] Leu9Ala13 to the most flexible Gly9Gly13^[25,40]. To perform the MD simulation, copper(II) was substituted with zinc(II), considering that MD parameterization of zinc is particularly reliable in the CHARMM forcefield.^[41] Further, in both DF scaffolds and natural PPOs, zinc substitution does not lead to a significant variation of the orientation in the first coordination sphere.^[24-26] <https://www.nature.com/articles/ncomms5505>

The zinc-protein interactions were described with the 'non-bonded method' to allow 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 constraint the relative positioning of the zinc ions.^[42] The most favorable combination was Leu9Ala13 both in terms of the solvent-accessible surface area (SASA) at the access channel and of the overall stability of the folding, as calculated from the root mean square deviation (RMSD) of the C^o atom with respect to the starting model and the radius of gyration (R_g) (See Supporting Information, Table S2, Figure S1 and Figure S2). The calculated R_g (13.1 ± 0.1 Å) and the interhelical SASA (1060 ± 40 Å²) are almost identical to DF3 (13.1 ± 0.1 Å, 1071 ± 40 Å², respectively). In particular, the R_g value is consistent with those observed for globular all- α proteins with less than 100 residues.^[43] This supports that the overall four-helix bundle, represented in Figure 2g by the centroid of most abundant cluster in the simulation (see Supporting Information, Table S3), is conserved and stable along the dynamics.

To further validate the design protocol, a DFT geometry optimization was performed, replacing zinc and sulfate with copper in the DR1 MD relaxed model (see Supporting Information for further details). A μ -hydroxo di-copper cluster (met-form), together with the two His triads sidechains and other residues shaping the catalytic pocket, were modeled in the DFT calculation (Figure 2h and Figure S3). The substitution of zinc(II) with copper(II) and the addition of one hydroxide ion in 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 catalytic activity of T3 sites. The overall architecture of the optimized active site closely resemble the one observed for PPOs met-forms (Figure 1a, Figure S4, Table S4): both copper(II) ions retain the coordination of the two His triads and of the bridging OH⁻ ligand, with a Cu-N (N^o for His 10/10' and His 39/39', N^e for His 32/32') distance of 2.138 ± 0.113 Å, a Cu-OH distance of 2.013 ± 0.045 Å, and a Cu-Cu distance of 3.914 Å.

The final DR1 α_2 sequence with all the mutations with respect to DF1 and DF3 is reported in Figure 2i.

DR1 folding, copper binding and thermodynamic stability

Apo-DR1 was chemically synthesized by using Fmoc protocols, purified by reversed-phase HPLC and identified by ESI MS (observed mass: 5858.97 ± 0.01 Da, theoretical mass: 5858.96 Da, Figure S5). Size exclusion chromatography analysis revealed that DR1 assumes the designed homodimeric structure in both the apo and holo forms (Figure 3a). The elution peak corresponds to an apparent molecular weight of about 14.7 kDa, in good agreement with that expected for the α_2 homodimer (≈ 12 kDa), and with that of DF1 used as standard (Figure S6, Table S5). Circular dichroism (CD) analysis gave insights into DR1 folding upon copper-binding. CD spectrum of apo-DR1 at pH 7.5 (Figure 3b, red curve) exhibits double minima at 209 and 222 nm typical of a α -helix structure. Addition of Cu^{2+} to apo-DR1 led to an increase in the mean residue ellipticity (MRE) at 209 and 222 nm, indicating a further increase in the helical content (Figure 3b, blue curve). This finding proved that the binding of the metal ion to DR1 improves its global folding, as already observed for several members of the DF family.^[34,44] Cu^{2+} titration by CD (Figure S7), allowed to determine the stoichiometry of the interaction. Data were fitted to a general $P + nM \leftrightarrow PM_n$ equilibrium,^[26] considering DR1 as a dimer. A protein:metal ratio of 1:2 was obtained ($n = 1.91 \pm 0.05$), as expected.

Isothermal Titration Calorimetry (ITC) was used to determine the thermodynamic underpinnings of the copper binding to the protein. Figure 3c reports the binding isotherm obtained from integration of the raw ITC data (Figure S8), upon titrating a solution of 25 μ M DR1 with $CuSO_4$ 400 μ M. 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 μ M (Figure 3c).

The thermodynamic signature of the process at 298.15 K revealed a modest binding enthalpy contribution, ΔH_b is -2.5 ± 0.3 kJmol⁻¹, comparable with the energy of a single hydrogen bond (Figure 3c, inset). As a consequence, the binding process is entropically driven ($-\Delta S_b = -27 \pm 1$ kJmol⁻¹), as observed in other de novo designed copper proteins.^[45] The enthalpic gain related to the formation of histidine-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.

Spectroscopic characterization of the di-copper center

The UV-Vis spectrum of di- Cu^{2+} -DR1 (Figure 4a) is characterized by an absorption band centered at 621 nm. The observed molar extinction coefficient for di-copper site is ~ 100 M⁻¹ cm⁻¹ and can be attributed to a weak $d \rightarrow d$ transition, characteristic of tetragonal Cu^{2+} , as found in the met-form of hemocyanins,^[46] CO^[47] and TYR.^[48]

Di- Cu^{2+} -DR1 shows an axial electron paramagnetic resonance (EPR) spectrum (Figure 4b), typical of copper(II) in square planar field.^[18] The amount of EPR-active copper(II) species was evaluated to $40 \pm 10\%$ of total copper(II) (see Supporting Information, Figure S9). This demonstrates that a significant fraction of antiferromagnetically coupled EPR-silent T3 sites are present in DR1, as confirmed by DFT calculations (see below). Interestingly, upon O₂ exposure of the di- Cu^{2+} -DR1 (see Supporting Information, Figure S10), the fraction of EPR-active copper(II) species was reduced by 30%. These data indicate that O₂ binding to the di- Cu^{2+} site may affect the Cu...Cu distance, thus favoring antiferromagnetic coupling between the cupric ions.

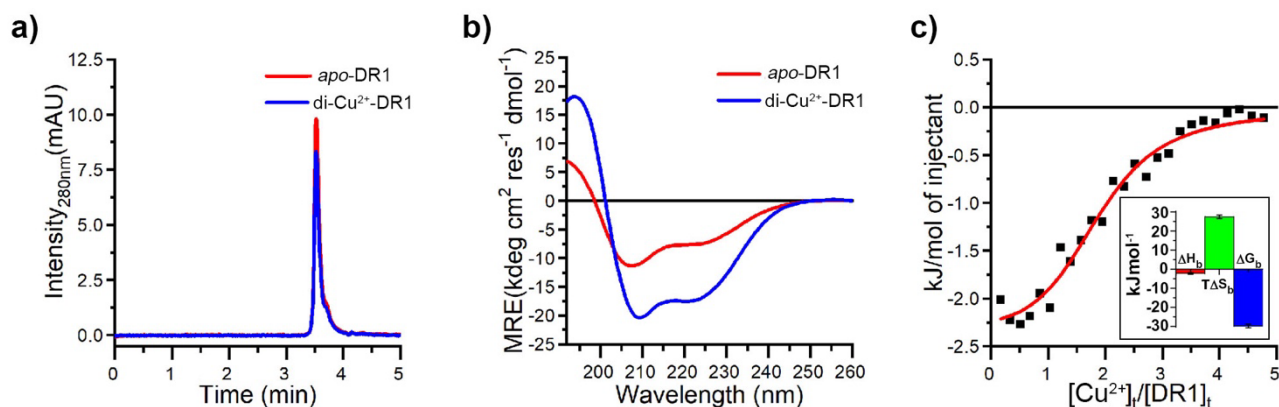


Figure 3 DR1 folding, copper binding and thermodynamic stability. a) Analytical gel filtration chromatograms, recorded at 280nm, of 100 μM DR1 (in 50 mM sodium phosphate, 300 mM NaCl pH 6.8) in the absence (red line) and in the presence (blue line) of 150 μM CuSO_4 . b) Ultraviolet CD spectra of DR1 (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 . c) integrated ITC data (black squares) as a function of the copper ions and DR1 dimer concentration ratio. The solid 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. Error bars represent the standard deviation.

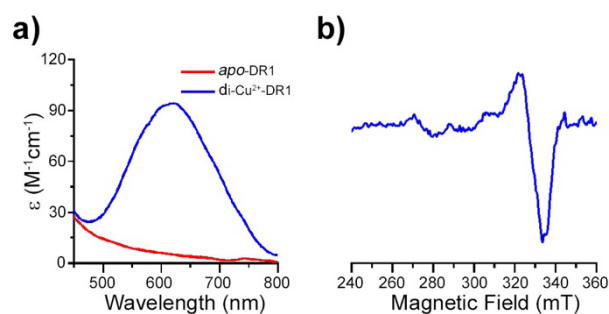


Figure 4 a) Visible absorption spectra of apo-DR1 (red line) and di- Cu^{2+} -DR1 (blue line), at 50 μM concentration, (ϵ is shown for di-copper site). b) CW-EPR spectrum of di- Cu^{2+} -DR1, with 30% of glycerol as glassing agent, at 300 μM copper concentration. The spectrum was recorded at 77 K. All the samples are in 10 mM HEPES 300 mM Na_2SO_4 pH 7.5.

Di-copper-DR1 diphenolase activity

Once demonstrated the ability of DR1 to bind copper and recapitulate the T3 site, its diphenolase activity was assayed. Addition of the model substrate 3,5-di-tert-butylcatechol (DTBC) to a solution of di- Cu^{2+} -DR1, caused the appearance in the UV spectrum of two bands around 410 nm and 660 nm (Figure 5a). The band at 410 nm was unambiguously attributed to 3,5-di-tert-butyl-*o*-benzoquinone (DTBQ) (see Figure S11), while the band at 660 nm is unrelated to the product and can be tentatively ascribed to the Cu(II) d→d transitions, red-shifted upon quinone binding (see *infra*). Analysis of the reaction mixture by NMR excluded the presence of any side-reaction product as only DTBQ was observed (see Supporting Information and Figure S12). A

complete conversion of DTBC to DTBQ was observed over 60 min (Figure 5b). No increase in DTBQ formation over the background was observed when the reaction was carried out with apo-DR1, with di- Zn^{2+} -DR1 or in sole buffer, thus confirming the catalysis to occur at the DR1 di-copper site (Figure S13a).

To get insights into the involvement of O_2 in the catalytic mechanism, the reaction was also carried out under anaerobic conditions (Figure S13b). After addition of DTBC to di- Cu^{2+} -DR1, ~ 1 equivalent (respect to the catalyst) of DTBQ was rapidly produced (Figure 5c dashed lines). Then, no increase in product formation was observed along time until restoring aerobic conditions (Figure 5c solid line). 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 addition of DTBC to di- Cu^{2+} -DR1 under anaerobic conditions (Figure S14). Addition of molecular oxygen re-oxidizes the di-copper center to complete the cycle, thus allowing DTBC to be promptly converted again to the *o*-quinone. Two possible mechanisms have been invoked for O_2 reduction catalyzed by di-copper sites: (i) one in which two molecules of water are produced, as observed in natural proteins, (ii) the other in which the reduction of O_2 leads to H_2O_2 formation, as frequently observed in small molecule mimics (Figure S15).^[49–51] More detailed mechanistic studies are needed to pinpoint the DR1-catalyzed activation of dioxygen, and are currently under course.

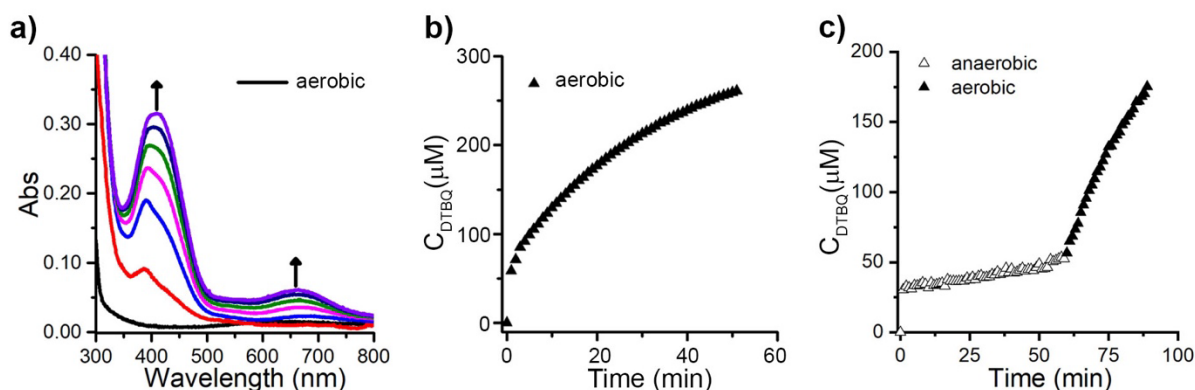


Figure 5 a) UV-Vis absorption spectra recorded every 10 minutes upon addition of 250 μM DTBC to 50 μM di- Cu^{2+} -DR1 in 10 mM HEPES 300 mM Na_2SO_4 , pH 7.5. b) DTBQ formation kinetic under aerobic conditions. c) DTBQ formation kinetic under anaerobic conditions (white symbol) and after the restoration of aerobic condition (black symbol). All spectra are subtracted for the protein background.

As already observed for several model compounds,^[51,52] upon DTBC oxidation, binding of DTBQ to the di-copper site may occur. Binding of DTBQ to di- Cu^{2+} -DR1 was ascertained by UV-Vis spectroscopy (Figure 6a) and saturation transfer difference (STD) NMR experiments (Figure 6b and 6c). Addition of DTBQ to the protein caused the appearance over time of the band at 660 nm (Figure 6a), already observed upon di- Cu^{2+} -DR1 catalyzed DTBC oxidation (Figure 5).

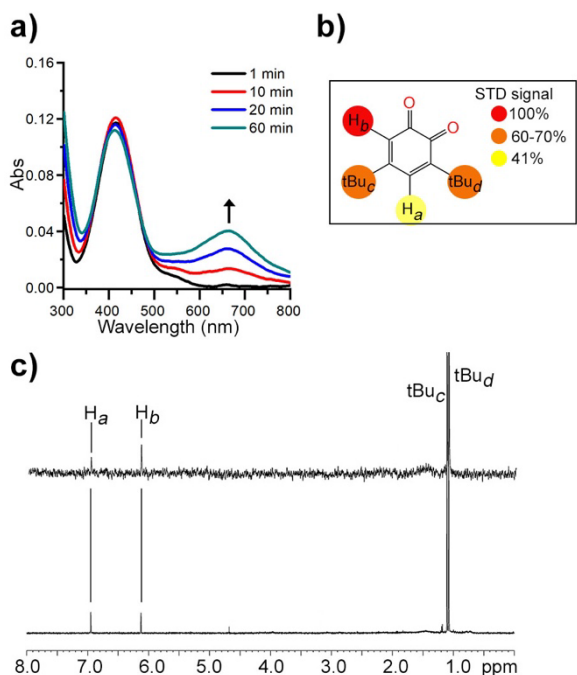


Figure 6 a) UV-Visible spectra recorded at different time upon addition of DTBQ to di- Cu^{2+} -DR1 (final concentration 150 μM and 100 μM , respectively) - in 10 mM HEPES 300 mM Na_2SO_4 , pH 7.5. All spectra are subtracted for the protein background. b) Structural formula of DTBQ with the relative degrees of saturation of the individual protons normalized to H_b proton. c) upper STD-NMR spectrum of 10 μM *met*-DR1 with 100 μM DTBQ (saturation time 5s); lower 1D ^1H NMR reference spectrum of DTBQ.

The STD spectrum, using 10-fold DTBQ excess with respect to the protein, exhibited narrow and well resolved lines (Figure 6c). Significant STD effects were observed for all the protons of

DTBQ, unequivocally confirming its binding to di- Cu^{2+} -DR1. The strongest STD effect was observed for proton H_b , whose value was set to 100%. Intensities for all the other protons were expressed relative to the H_b signal (Figure 6b, Figure S16 and Table S6). The *t*-butyl protons tBu_c and tBu_d exhibited similar saturation transfers ($\sim 70\%$). The least interacting proton with di- Cu^{2+} -DR1 was H_a with a STD of $\sim 40\%$. These data suggested 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 configuration, the DTBQ oxygen atoms would point towards the protein core.

The catalytic parameters for DTBC oxidation were next determined. In the presence of ambient oxygen, di- Cu^{2+} -DR1 followed Michaelis-Menten kinetics (Figure S17, Table S7) with values of 0.70 ± 0.06 mM and 65 ± 3 min^{-1} for K_m and k_{cat} , respectively ($k_{\text{cat}}/K_m = (9.2 \pm 0.4) 10^4$ $\text{M}^{-1} \text{min}^{-1}$)

Kinetic investigations using differently substituted di-phenols, as catechol and 4-*tert*-butylcatechol (4-TBC), revealed interesting di- Cu^{2+} -DR1 behaviors (Figure S17, Table S7). Though the removal of one substrate *tert*-butyl (*t*Bu) group resulted in a poor 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 *t*Bu 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 2h).

To further validate the substrate recognition capability of di- Cu^{2+} -DR1, Hammett analysis was performed using DTBC, 4-TBC, catechol, and 4-chlorocatechol. Compared to catechol as 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 7a). The correlation of the reaction kinetic to linear free energy relationship-based electronic substituent parameters was investigated using the Hammett parameter σ_p^- . The Hammett plot ($\log(\text{TOF})$ vs σ_p^- parameters^[53] was found to deviate from linearity ($R^2 = 0.87$) (Figure 7a). This suggests that other factors, such as steric effects of the substituents, also play an important role for catalytic activity, thanks to the constraints exerted by the protein matrix surrounding the active site. To evaluate the correlation of steric and electronic effects with the reaction kinetics of the different catechol substrates, the σ_p^- parameters were adjusted with the σ_v Charton's steric parameters.^[53] The adjusted parameter was found to be $\sigma' = \sigma_p^- -$

σ_v . A plot of $\log(\text{TOF})$ against σ' shows linear correlation ($R^2 = 0.99$) with a negative slope ($\rho = -1.6$) (Figure 7b). A similar ρ value was observed in natural T3 copper proteins^[54,55] and in several synthetic complexes, for which a positively charged transition state has been invoked in the rate-determining step.^[56] The good correlation found by correcting the σ_p^- parameter by steric volume clearly evidence the direct interaction of the protein matrix with substrates. This result further supports that the presence of tBu groups on the substrates determine favorable interactions within the active site, thus increasing the catalytic rate and lowering K_m values.

Di-copper-DR1 substrate recognition by DFT calculations

DFT calculations were performed to model the interaction of substrates within the DR1 pocket (see Figure S3 for details on the di-Cu²⁺-DR1 cluster model used). The di-copper site is in its ground state anti-ferromagnetically coupled met-form (favored over triplet solution by 1.8 kcal/mol at the B3LYP-D4 level), in agreement with EPR data. Differently substituted catechols were tested to better rationalize, on a stereo-electronic basis, the effect of tBu groups on binding affinities. Figure 8 reports the optimized structure of the three different adducts resulting from DTBC, 4-TBC and catechol binding to the di-copper site. Due to its shape and steric constraints, the DR1 cavity drives catechols towards one specific binding mode, more likely involving a single Cu center (Figure 8a and Figure S18).

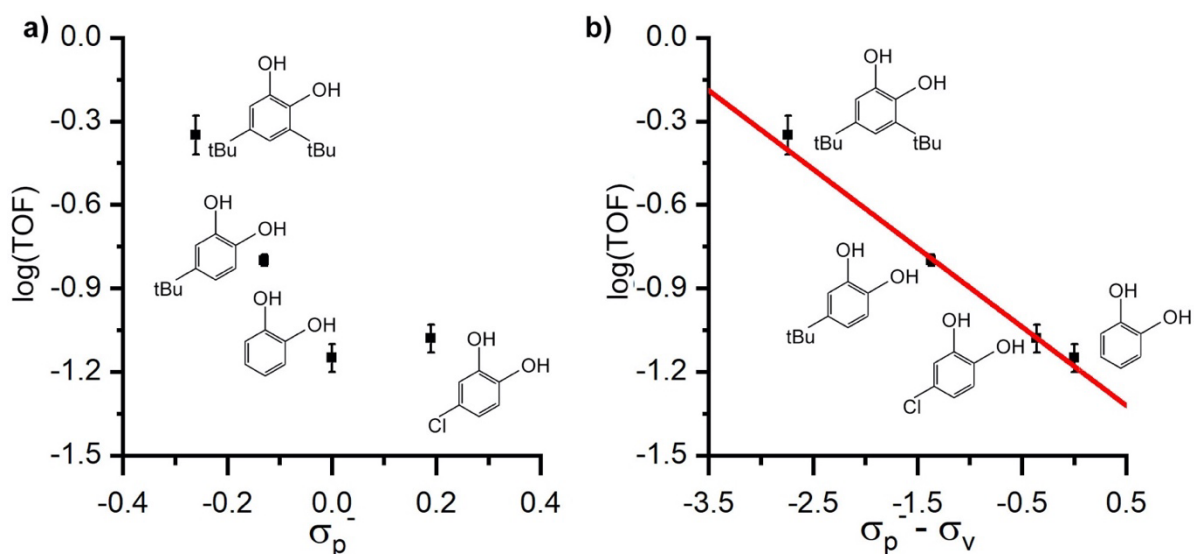


Figure 7 a) Hammett plot of values of $\log(\text{TOF})$ for oxidation of catechol substrates as a function of σ_p^- and b) as a function of $\sigma' = \sigma_p^- - \sigma_v$. The points are experimental points, measured for compounds of the structures shown adjacent to those points. Di-Cu²⁺-DR1 concentration was 30 μM in 10 mM HEPES, 300 mM Na₂SO₄, 20% DMF, pH 7.5 and substrate concentration was fixed at 550 μM . The data points were fitted with a linear function. Error bars were obtained from three independent measurements.

The calculated *relative* binding energies (see Figure S19 for details) clearly evidenced that the affinity decreases in the order DTBC > 4-TBC > catechol, nicely matching the experimental observation that the presence of bulky hydrophobic groups favors substrate binding. The observed trend is the same irrespective to the binding mechanism envisioned (see Figure S19). The removal of one tBu group, going from DTBC to 4-TBC, decreases affinity by ~4.5 kcal/mol, while a further tBu substitution with a hydrogen atom causes an additional and large affinity loss of ~8 kcal/mol.

Such differences can be explained in terms of both steric and electronic factors. One tBu (position 4 for 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 intercept L17 side chain (Figure 8b), 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 (Figure S20), reflecting a stronger binding for electron-rich (and thus more reducing) substrates.

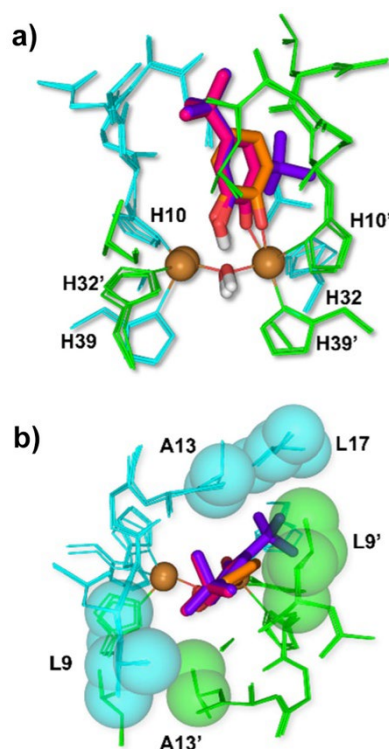


Figure 8 a) Front-view of the DFT-optimized adducts of DR1 pocket with DTBC (in purple), 4-TBC (in pink) and catechol (in orange) and b) top-view of catalytic pocket with the bound catechols, highlighting (as spheres) the hydrophobic side chains mostly involved in substrate recognition.

Conclusion

This work illustrates how de novo designed four-helix bundles can be manipulated to engineer diverse, catalytically active metal binding sites, thus obtaining excellent mimics of natural metalloproteins. The newly designed protein DR1 properly hosts the T3 copper site and displays diphenolase activity. The use of DF1 scaffold, patterned after natural di-iron proteins, allowed us to reproduce the 6His binding site, typical of PPO,^[18] in a well-defined and parametrized four-helix bundle.

This study represents a breakthrough in de novo metalloprotein design. Differently from T1 and T2 mononuclear copper sites, which have been widely constructed in a variety of de novo scaffolds,^[4,45,57–62] the T3 di-nuclear sites have been mainly studied in natural proteins^[63] and in small inorganic complexes.^[56,64] Very recently, Song and co-workers installed a di-copper site at the C₂-symmetric interface of the hexameric acetyltransferase, using the unnatural bipyridyl-alanine as a tight chelating residue.^[65]

Here, we set a new bar, not only by recapitulating the T3 site with the native histidines in the correct geometry, but also supporting its function by positioning a set of properly designed secondary sphere interactions, crucial for substrate binding and recognition.

All this was possible thanks to our deep understanding of the simple, versatile, and structurally stable DF framework, which allowed us to endow DR1 with substrate selectivity, a key feature of natural enzymes.

In conclusion, this study represents a step forward in the design of artificial metalloenzymes with demanding behaviors, demonstrating that simplified, model proteins can be effective for the development of new catalysts.

Acknowledgements

We are deeply grateful to Prof. Vincenzo Pavone for his interest in this work and for useful suggestions and fruitful discussion. We also wish to thank Dr. Rosario Oliva for helping with ITC data collection. This work was supported by Italian MUR, Project SEA-WAVE 2020BKK3W9, [CUP_E69J22001140005].

Keywords: artificial metalloenzymes • phenol oxidases • protein design • substrate selectivity • T3 di-copper site

- [1] H. J. Davis, T. R. Ward, *ACS Cent. Sci.* **2019**, *5*, 1120–1136.
- [2] F. Nastro, D. D'Alonzo, L. Leone, G. Zambrano, V. Pavone, A. Lombardi, *Trends Biochem. Sci.* **2019**, *44*, 1022–1040.
- [3] M. J. Chalkley, S. I. Mann, W. F. DeGrado, *Nat. Rev. Chem.* **2022**, *6*, 31–50.
- [4] 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.
- [5] T. Matsuo, T. Miyake, S. Hirota, *Tetrahedron Lett.* **2019**, *60*, 151226.
- [6] J. Wang, S. Lisanza, D. Juergens, D. Tischer, J. L. Watson, K. M. Castro, R. Ragotte, A. Saragovi, L. F. Milles, M. Baek, I. Anishchenko, W. Yang, D. R. Hicks, M. Expòsit, T. Schlichthaefer, J.-H. Chun, J. Dauparas, N. Bennett, B. I. M. Wicky, A. Muenks, F. DiMaio, B. Correia, S. Ovchinnikov, D. Baker, *Science* **2022**, *377*, 387–394.
- [7] J. Zhu, N. Avakyan, A. Kakkis, A. M. Hoffnagle, K. Han, Y. Li, Z. Zhang, T. S. Choi, Y. Na, C.-J. Yu, F. A. Tezcan, *Chem. Rev.* **2021**, *121*, 13701–13796.
- [8] K. J. Koebke, T. B. J. Pinter, W. C. Pitts, V. L. Pecoraro, *Chem. Rev.* **2022**, *122*, 12046–12109.
- [9] O. Maglio, F. Nastro, A. Lombardi, in *Ion. Interact. Nat. Synth. Macromol.*, John Wiley & Sons, Ltd, Hoboken, NJ, USA, **2012**, pp. 361–450.
- [10] S. Studer, D. A. Hansen, Z. L. Pianowski, P. R. E. Mittl, A. Debon, S. L. Guffy, B. S. Der, B. Kuhlman, D. Hilvert, *Science* **2018**, *362*, 1285–1288.
- [11] L. Leone, M. Chino, F. Nastro, O. Maglio, V. Pavone, A. Lombardi, *Biotechnol. Appl. Biochem.* **2020**, *67*, 495–515.
- [12] R. Stenner, J. W. Stevenson, A. Seddon, J. L. R. Anderson, *Proc. Natl. Acad. Sci.* **2020**, *117*, 1419–1428.
- [13] L. Leone, D. D'Alonzo, O. Maglio, V. Pavone, F. Nastro, A. Lombardi, *ACS Catal.* **2021**, *11*, 9407–9417.
- [14] M. Jeschek, R. Reuter, T. Heinisch, C. Trindler, J. Klehr, S. Panke, T. R. Ward, *Nature* **2016**, *537*, 661–665.
- [15] M. H. Sazinsky, S. J. Lippard, in *Sustain. Life Planet Earth Met. Mastering Dioxygen Chewy Gases* (Eds.: P.M.H. Kroneck, M.E. Sosa Torres), Springer International Publishing, Cham, Switzerland, **2015**, pp. 205–256.
- [16] T. J. Lawton, A. C. Rosenzweig, *J. Am. Chem. Soc.* **2016**, *138*, 9327–9340.
- [17] B. L. Greene, G. Kang, C. Cui, M. Bennati, D. G. Nocera, C. L. Drennan, J. Stubbe, *Annu. Rev. Biochem.* **2020**, *89*, 45–75.
- [18] 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.
- [19] A. Bijelic, M. Pretzler, C. Molitor, F. Zekiri, A. Rompel, *Angew. Chem. Int. Ed.* **2015**, *54*, 14677–14680.
- [20] M. E. Cuff, K. I. Miller, K. E. van Holde, W. A. Hendrickson, *J. Mol. Biol.* **1998**, *278*, 855–870.
- [21] M. Pretzler, A. Rompel, *Inorganica Chim. Acta* **2018**, *481*, 25–31.
- [22] A. Lombardi, C. M. Summa, S. Geremia, L. Randaccio, V. Pavone, W. F. DeGrado, *Proc. Natl. Acad. Sci.* **2000**, *97*, 6298–6305.
- [23] M. Chino, O. Maglio, F. Nastro, V. Pavone, W. F. DeGrado, A. Lombardi, *Eur. J. Inorg. Chem.* **2015**, *2015*, 3352–3352.
- [24] A. Lombardi, F. Pirro, O. Maglio, M. Chino, W. F. DeGrado, *Acc. Chem. Res.* **2019**, *52*, 1148–1159.
- [25] M. Faiella, C. Andreozzi, R. T. M. de Rosales, V. Pavone, O. Maglio, F. Nastro, W. F. DeGrado, A. Lombardi, *Nat. Chem. Biol.* **2009**, *5*, 882–884.
- [26] M. Chino, L. Leone, O. Maglio, D. D'Alonzo, F. Pirro, V. Pavone, F. Nastro, A. Lombardi, *Angew. Chem. Int. Ed.* **2017**, *56*, 15580–15583.
- [27] A. J. Reig, M. M. Pires, R. A. Snyder, Y. Wu, H. Jo, D. W. Kulp, S. E. Butch, J. R. Calhoun, T. Szyperki, E. I. Solomon, W. F. DeGrado, *Nat. Chem.* **2012**, *4*, 900–906.
- [28] J. S. Sodhi, K. Bryson, L. J. McGuffin, J. J. Ward, L. Wermisch, D. T. Jones, *J. Mol. Biol.* **2004**, *342*, 307–320.
- [29] S.-Q. Zhang, M. Chino, L. Liu, Y. Tang, X. Hu, W. F. DeGrado, A. Lombardi, *J. Am. Chem. Soc.* **2018**, *140*, 1294–1304.
- [30] S. I. Mann, A. Nayak, G. T. Gassner, M. J. Therien, W. F. DeGrado, *J. Am. Chem. Soc.* **2021**, *143*, 252–259.
- [31] O. Maglio, F. Nastro, V. Pavone, A. Lombardi, W. F. DeGrado, *Proc. Natl. Acad. Sci.* **2003**, *100*, 3772–3777.
- [32] F. DiMaio, A. Leaver-Fay, P. Bradley, D. Baker, I. André, *PLoS ONE* **2011**, *6*, e20450.
- [33] 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.
- [34] R. T. M. de Rosales, M. Faiella, E. Farquhar, L. Que, C. Andreozzi, V. Pavone, O. Maglio, F. Nastro, A. Lombardi, *JBIC J. Biol. Inorg. Chem.* **2010**, *15*, 717–728.
- [35] 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.
- [36] J. Kaplan, W. F. DeGrado, *Proc. Natl. Acad. Sci.* **2004**, *101*, 11566–11570.
- [37] R. A. Snyder, S. E. Butch, A. J. Reig, W. F. DeGrado, E. I. Solomon, *J. Am. Chem. Soc.* **2015**, *137*, 9302–9314.
- [38] L. Di Costanzo, H. Wade, S. Geremia, L. Randaccio, V. Pavone, W. F. DeGrado, A. Lombardi, *J. Am. Chem. Soc.* **2001**, *123*, 12749–12757.
- [39] A. Pasternak, J. Kaplan, J. D. Lear, W. F. DeGrado, *Protein Sci.* **2001**, *10*, 958–969.
- [40] W. F. DeGrado, L. Di Costanzo, S. Geremia, A. Lombardi, V. Pavone, L. Randaccio, *Angew. Chem. Int. Ed.* **2003**, *42*, 417–420.
- [41] 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.
- [42] Y. P. Pang, K. Xu, J. E. Yazal, F. G. Prendergas, *Protein Sci.* **2000**, *9*, 1857–1865.
- [43] M. Yu. Lobanov, N. S. Bogatyreva, O. V. Galzitskaya, *Mol. Biol.* **2008**, *42*, 623–628.
- [44] M. Chino, S.-Q. Zhang, F. Pirro, L. Leone, O. Maglio, A. Lombardi, W. F. DeGrado, *Biopolymers* **2018**, *109*, e23339.
- [45] 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.
- [46] E. I. Solomon, M. J. Baldwin, M. D. Lowery, *Chem. Rev.* **1992**, *92*, 521–542.
- [47] C. Eicken, F. Zippel, K. Büldt-Karentzopoulos, B. Krebs, *FEBS Lett.* **1998**, *436*, 293–299.
- [48] R. S. Himmelwright, N. Eickman, C. D. LuBien, E. Solomon, K. Lerch, *J. Am. Chem. Soc.* **1980**, *102*, 5378–5388.
- [49] A. Neves, L. M. Rossi, A. J. Bortoluzzi, B. Szpoganicz, C. Wiezbicki, E. Schwingel, W. Haase, S. Ostrovsky, *Inorg. Chem.* **2002**, *41*, 1788–1794.
- [50] K. Selmecki, M. Réglier, M. Giorgi, G. Speier, *Coord. Chem. Rev.* **2003**, *245*, 191–201.
- [51] I. A. Koval, K. Selmecki, 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.
- [52] A. Bencini, A. Dei, C. Sangregorio, F. Totti, M. G. F. Vaz, *Inorg. Chem.* **2003**, *42*, 8065–8071.
- [53] Corwin. Hansch, A. Leo, R. W. Taft, *Chem. Rev.* **1991**, *91*, 165–195.
- [54] H. W. Duckworth, J. E. Coleman, *J. Biol. Chem.* **1970**, *245*, 1613–1625.
- [55] J. L. Muñoz-Muñoz, M. del M. Garcia-Molina, F. Garcia-Molina, J. Berna, P. A. Garcia-Ruiz, M. Garcia-Moreno, J. N. Rodriguez-Lopez, F. Garcia-Canovas, *J. Mol. Catal. B Enzym.* **2013**, *91*, 17–24.
- [56] C. Citek, S. Herres-Pawlis, T. D. P. Stack, *Acc. Chem. Res.* **2015**, *48*, 2424–2433.
- [57] 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.
- [58] M. Tegoni, F. Yu, M. Bersellini, J. E. Penner-Hahn, V. L. Pecoraro, *Proc. Natl. Acad. Sci.* **2012**, *109*, 21234–21239.
- [59] K. J. Koebke, A. G. Tebo, E. C. Manickas, A. Deb, J. E. Penner-Hahn, V. L. Pecoraro, *JBIC J. Biol. Inorg. Chem.* **2021**, *26*, 855–862.
- [60] 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.
- [61] M. Flores, T. L. Olson, D. Wang, S. Edwardraja, S. Shinde, J. C. Williams, G. Ghirlanda, J. P. Allen, *J. Phys. Chem. B* **2015**, *119*, 13825–13833.
- [62] S. I. Mann, T. Heinisch, T. R. Ward, A. S. Borovik, *J. Am. Chem. Soc.* **2017**, *139*, 17289–17292.
- [63] I. Kipourous, A. Stańczak, M. Culka, E. Andris, T. R. Machonkin, L. Rulišek, E. I. Solomon, *Chem. Commun.* **2022**, *58*, 3913–3916.
- [64] A. Brinkmeier, R. A. Schulz, M. Buchhorn, C.-J. Spyrta, S. Dechert, S. Demeshko, V. Krewald, F. Meyer, *J. Am. Chem. Soc.* **2021**, *143*, 10361–10366.

[65] S.-M. Jung, M. Yang, W. J. Song, *Inorg. Chem.* **2022**, *61*, 12433-12441.