Artificial Metalloenzymes as Catalysts for Oxidative Lignin Degradation

M. V. Doble[†], A. G. Jarvis[†‡]*, A. C. C. Ward[†], J. D. Colburn[†], J. P. Götze[†], M. Bühl[†], P. C. J. Kamer[†¶]

[†] School of Chemistry, University of St Andrews, North Haugh, St Andrews, Fife KY16 9ST, United Kingdom.
[‡] School of Chemistry, University of Edinburgh, David Brewster Road, King's Buildings, Edinburgh, EH9 3FJ, United Kingdom.
[¶] Bioinspired Homo- & Heterogeneous Catalysis, Leibniz Institute for Catalysis, Albert-Einstein-Strasse 29a, 18059, Rostock, Germany
* Corresponding author: amanda.jarvis@ed.ac.uk

Abstract

We report novel artificial metalloenzymes (ArMs), containing tris(pyridylmethyl)amine (TPA), for the atom economic oxidation of lignin β-O-4 model compounds, using hydrogen peroxide. The protein scaffold alters the selectivity of the reaction from a low yielding cleavage reaction when using the parent Fe-tpa complex to a high yielding benzylic alcohol oxidation when using the complex incorporated into a protein scaffold, SCP-2L A100C. Engineering the protein scaffold to incorporate glutamic acid was found to improve the ArM activity, showing that rational design of the protein environment using metal binding amino acids can be a first step towards improving the overall activity of an artificial metalloenzyme.

Graphical Abstract
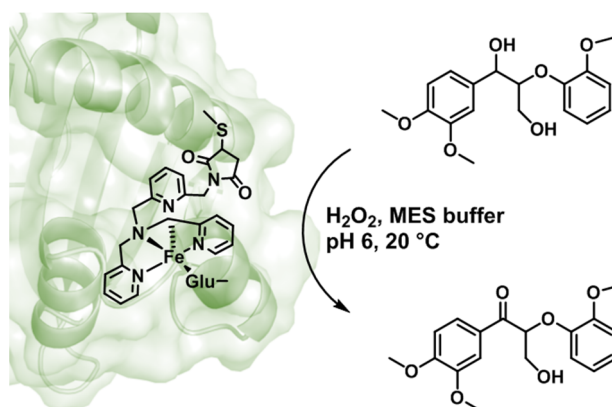
Keywords: Lignin, Artificial metalloenzymes, Catalytic oxidation

Synopsis: We report novel catalysts – artificial metalloenzymes – for use in the green oxidation of lignin model compounds under mild aqueous conditions.

Introduction

The rapid depletion of fossil fuels as a resource for energy and chemicals production necessitates a future change to an increase in the use of biomass, ideally through the development of biorefineries that provide a wide range of products. The most abundant

renewable resource in this context is lignocellulosic biomass, which is much more plentiful than alternative feedstocks such as sugars, starch, oils and fats.[1] Another advantage is that it does not compete with food production. Lignocellulose, consisting of 40-50% cellulose, 16-33% hemicellulose and 15-30% lignin,[2] is therefore viewed as an abundant future source of chemicals.[3,1] Whilst the technologies for the exploitation of cellulose and hemicellulose are well established, lignin is currently underutilised with only one commercial process in operation – the preparation of vanillin.[4] The major hurdle for the complete exploitation of lignocellulosic biomass, including lignin, is the efficient depolymerisation of lignin.[5,2] In the biorefinery, cellulose and hemicellulose can provide sugars, alcohols and alkane based products, whilst lignin has the potential to replace oil as the major natural source of aromatic compounds. Therefore methods to overcome the depolymerisation limitation are required.

Traditional chemical methods can breakdown lignin, with catalytic methods being sought to provide selective methods that are sustainable.[6] The lignin polymer contains a number of different linkages with the β-O-4 unit the highest in
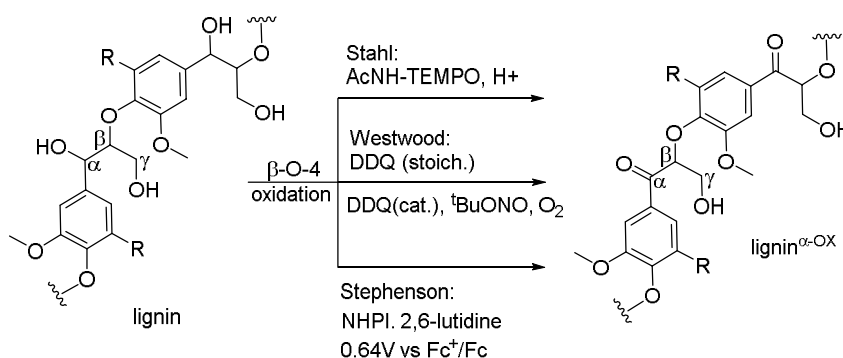
Scheme 1: Previously reported methods to Lignin^{OX}.[9,7,10]

abundance (approximately 60%) in hardwood lignins.[2] A number of successful approaches have targeted selective oxidation of the benzylic alcohol in the β-O-4 linkage to give Lignin$^{\alpha\text{-}OX}$ as a common intermediate (Scheme 1). These include catalytic methods such as electrocatalysis,[7] acetamido-TEMPO mediated catalytic benzylic oxidation,[8,9] and both catalytic and stoichiometric methods using DDQ.[10,11] There have also been examples of preparation of Lignin$^{\gamma\text{-}OX}$ by organocatalytic oxidation of the γ-hydroxyl group.[12] Both types of Lignin^{OX} can be broken down to numerous different products depending on the choice of reagents, making it a promising material to target for the biorefinery.[10] For example, using formic acid/sodium formate up to 52 wt% of aspen lignin could be recovered a low molecular weight aromatics. Additionally, it has been used as a promising starting point for biomaterials.[13]

In nature, enzymes such as the lignin peroxidases found in white rot fungi are known to breakdown the lignin polymer over time, to give a variety of products using simple oxidants such as hydrogen peroxide and oxygen.[14] Whilst isolated enzymes have been found to catalyse lignin depolymerisation,[15] many enzymes are only tested in standard oxidation assays and show inconsistent activity with lignin model compounds, often requiring phenolic groups for activity.[16] A number of biomimetic methodologies have been developed for lignin depolymerisation including the use of metalloporphyrins,[17] and Schiff-base complexes.[6,18] A general drawback of biomimetic oxidative methodologies is that often the catalysts are degraded over time by oxidation. Artificial Metalloenzymes (ArMs) provide a different

approach from biomimetic chemistry to bridge the gap between naturally occurring enzymes and chemical methods for tackling challenging reactions.[19] There have been many examples of ArMs being used in oxidation reactions such as alcohol oxidation, sulfoxidation and C=C and C-H bond oxygenation.[20] A recent example showed that the protein scaffold in an ArM can increase the stability of metal-hydroperoxido complexes, which have been proposed as key intermediates in biological and synthetic oxidations.[21]

We have been interested in developing artificial metalloenzymes to enhance the selectivity of various chemical transformations. We were interested in using this approach to identify selective methods of lignin depolymerisation with the focus on targeting clean and sustainable routes to either Lignin[OX] or low molecular weight aromatics using green oxidants such as hydrogen peroxide. We hypothesised that the protein scaffold of the ArM would allow us to control the selectivity of oxidation and subsequent cleavage reactions.

Experimental Section (Materials and Methods)

Preparation of modified proteins

The proteins were obtained via expression from *E. coli* Rosetta(DE3) cells and purification as described in the literature.[22, 23] SCP-2L A100C F94E/H were obtained by site directed mutagenesis and expressed and purified using the same procedure as ACP-2L A100C. Further details and details of the cofactor preparation are given in the Supporting Information (SI). The proteins were modified as follows: To a solution of the protein (50-100 µM) in Mes buffer (50 mM 2-(N-morpholino)ethanesulfonic acid (Mes), 20 mM NaCl, pH 6.0) was added the desired maleimide-containing cofactor (5-10 eq., added as a stock solution in MeCN) and shaken overnight at 18 °C. The addition was followed by mass spectrometry and more cofactor added if incomplete conversion was observed. Excess cofactor is removed by exchanging the buffer three times using centrifugal concentrators with a 10,000 MWCO (Millipore Amicon Ultra Regenerated Cellulose). Characterisation of the ArMs by mass spectrometry was performed on a Waters Alliance HT 2795 equipped with a Micromass LCT-TOF mass spectrometer, using positive electrospray ionisation and applying a Waters MASSPREP® On-line Desalting 2.1x10 mm cartridge using a gradient of 1% formic acid in H2O to 1% formic acid in acetonitrile. ESI-MS results were analysed by MassLynx V. 4.0 and its MaxEnt algorithm.

Catalytic reactions

The modified protein, **SCP-2L A100C-X** in MES Buffer (61 µM, 1 mL, 50 mM Mes, 20 mM NaCl, pH 6.0) was incubated with the Fe(OTf)$_2$.2MeCN (1 eq.) whilst shaking (60 rpm) for 20 mins to pre-form the metalloenzyme. The substrate, **5** (1.23 µmol) in 50µl MeCN was added to the metalloenzyme solution (catalyst loading 5 mol%). Naphthalene (1.23 µmol) was added as the internal standard for analysis. H$_2$O$_2$ (1 eq.) was then added to the reaction mixture in one addition and the reaction mixture shaken for 60 minutes. To stop the reaction 500 µl DCM was added, the reaction mixture was centrifuged for 2 mins at 13000 rpm to separate and extract the organics. Analysis of the organic layer was carried out using HPLC (see SI).

Computation

Modelling studies were started from the X-ray structure of the A100C mutant (hereto referred as WT) of the PDB structure 1IKT (SCP-2-like domain).[24] A preliminary MD simulation was run for the pristine WT (with free Cys100) to produce suitable starting geometries for adding the ligand. After determining the protonation states of titratable residues, supplying hydrogen atoms and immersing the system in an equilibrated droplet of TIP3P water molecules (including chloride ions to neutralize the system), an MD simulation was performed in the NVE ensemble at 300 K, for 500 ps with a 0.5 fs step size, using the CHARMM force field.[25,26,27,28] The target Cys100 residue is slightly buried below the protein surface initially, but thermal movement allows Cys100 to move towards the surface every few picoseconds, allowing the ligand to bind (see Figure S39 in the ESI). A structure with surface-exposed Cys100 was chosen to construct ligand-protein initial structures, manually adding the ligand C to the protein framework, orientating the ligand to face the solvent and adding an $M(H_2O)_3$ fragment in *fac* orientation to the aza-crown ether. Following equilibration in TIP3P water molecules (performed in two stages in NVT and NPT ensembles), production-runs were conducted for 2 ns in the NPT ensembles for both S and R stereoisomers at 300 K, with a 2 fs timestep. These simulations were performed under periodic boundary conditions in a cubic box (using the particle mesh Ewald treatment of long-range electrostatics) using the Gromacs suite of programs.[29] For further details (including ligand and metal parameterisation) see the SI.

Results and Discussion

Iron, manganese and copper complexes are amongst some of the most well-known oxidation catalysts, both chemically and biologically as metalloenzymes. Previous work by Stahl had shown that classic metal complexes such as Cu/bipyridine and Fe/phenanthroline are able to oxidise and in some cases break the β-O-4 linkage.[8] However, in most cases both oxidation of the alcohols and cleavage of the lignin was observed with relatively low yields. We hypothesised that an artificial metalloenzyme approach to lignin cleavage would prove beneficial in terms of both selectivity and activity of the reaction. The covalent anchoring of a cofactor within a protein scaffold is an effective method to obtain artificial metalloenzymes with the desired primary metal environment for the reaction of interest. Previously our groups have used the Steroid Carrier Protein 2L (SCP 2L) as a protein scaffold for ArM design. A single reactive cysteine was introduced to the protein using site-directed mutagenesis, and could then be modified using maleimide-containing cofactors. This allowed the introduction of the desired cofactor for the reaction i.e. phosphines for Rh-catalysed hydroformylation[23] or nitrogen cofactors such as **B** and **C** for Cu-catalysed Diels-Alder reactions.[30] Following the success of this approach, we decided to continue with this protein and approach for our initial work on lignin oxidation. Taking inspiration from the ligands used in homogenous oxidative catalysis, a library of nitrogen cofactors containing maleimide linkers was prepared (Figure 1). Cofactors **A**[31], **B**[30], and **C**[32] were prepared according to literature procedures.
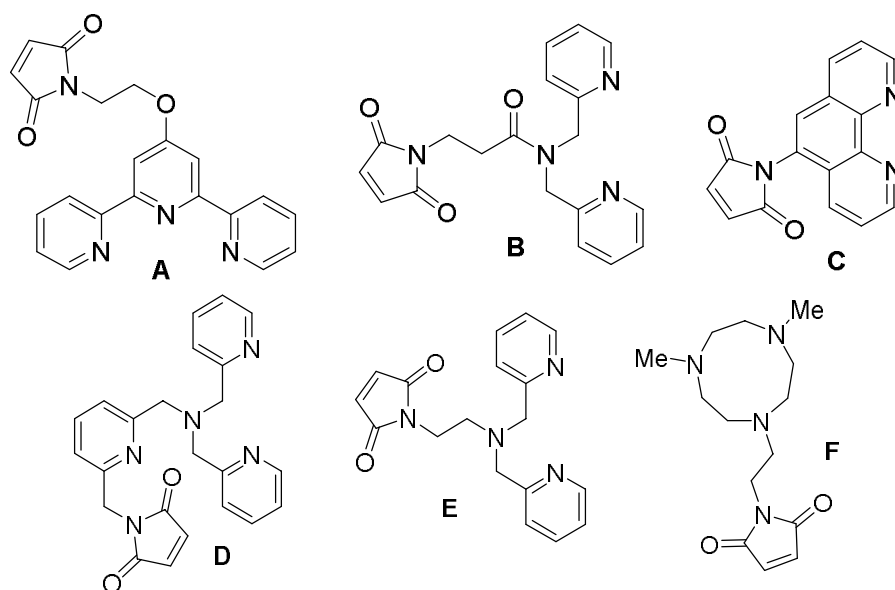
Figure 1: Organic cofactors used in this work.

A further three cofactors were prepared (**D**-**F**) in moderate to good yields using classical organic transformations. Cofactor **D** was prepared via a nucleophilic substitution reaction between dipicolylamine and compound **2** as shown in Scheme 2. Compound **2** was prepared from 2,6-bis(chloromethyl)pyridine and the protected maleimide **1**.[33] It was found that the maleimide needed to be protected by a Diels-Alder reaction with furan to avoid side reactions of the maleimide in the presence of free halide ions. A final retro-Diels-Alder reaction gave the product. Cofactors **F** and **E** were prepared in a similar manner via the nucleophilic substitution of N-(2-bromoethyl)-exo-7-oxabicyclo[2.2.1]hept-5-ene-2,3-dicarboximide(**4**),[34]
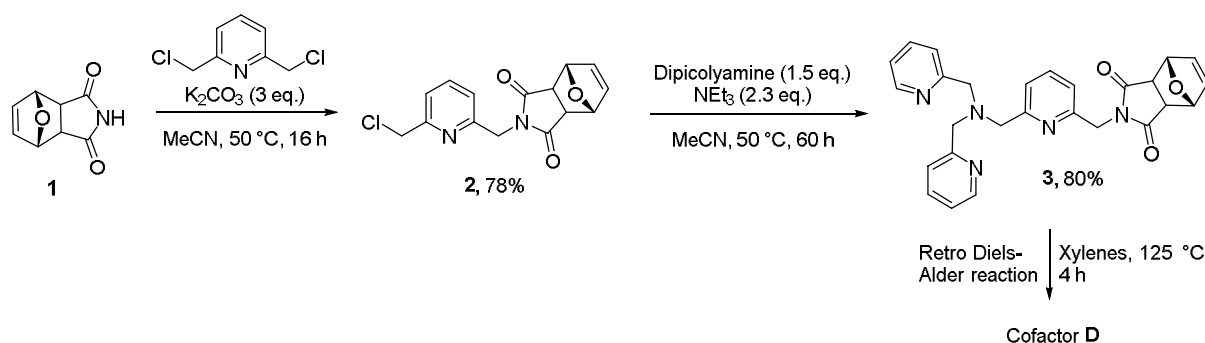

*Scheme 2: Preparation of Cofactor **D**.*

with Me$_2$TACN (1,4-dimethyl-1,4,7-triazacyclononane) and dipicolylamine respectively.

The modification of SCP 2L A100C with cofactors **A**-**F** was followed by MS, and in all cases the reaction went to completion under the standard conditions (50 mM MES, 20 mM NaCl, pH 6, rt, 5-10 eq. of cofactor) to give the singularly modified proteins.[30] Using cofactor **F**, which has no appreciable UV absorbance itself, no changes in the CD spectrum of the modified protein SCP-2L A100C-**F** were observed compared to SCP-2L A100C (see SI, Figure S27). This indicates that the modification of the cysteine does not disrupt the protein fold. It was unclear if the Bradford assay would accurately estimate the protein

concentrations due to the inclusion of synthetic hydrophobic cofactors. In order to undertake catalytic work with the ArMs, an accurate method for determining concentration was investigated. Comparing concentration results from the standard Bradford assay or using absorbance at A_{280} (SCP 2L A100C-**D** $\varepsilon = 15227$ M^{-1}cm^{-1}, calculated from theoretical ε of SCP-2L A100C and the experimentally observed extinction coefficient of **D**), with results from Quantitative Amino Acid analysis for SCP-2L A100C-**D**, showed that A_{280} was more accurate and was therefore used for all cofactors (see SI for full details).

Oxidation of β-O-4 lignin model compounds

Lignin itself is a very complex biopolymer and therefore initial work on lignin oxidation and depolymerisation generally uses model compounds that simplify the main components. Several model compounds have been used in the literature from basic aryl ethers to synthetic polymers.[6] For this work we chose to start with compound **5**, which contains the main components of the β-O-4 linkage including the γ-hydroxyl moiety. With the modified protein scaffolds in hand, an initial screening of the iron triflate complexes of the ArM library in the oxidation of lignin model **5** was carried out (Scheme 3 and Table 1). Our initial conditions (Buffer: 50 mM Mes, 20 mM NaCl, pH 6.0) were chosen based on our experience that the protein is more stable i.e. less chance of precipitation, under these conditions, than for example with Tris buffers at pH 7.4. The metal complexes were pre-formed by the addition of 1 eq. of Fe(OTf)$_2$ to the modified protein solution and incubating for 20 minutes before catalysis was carried out.
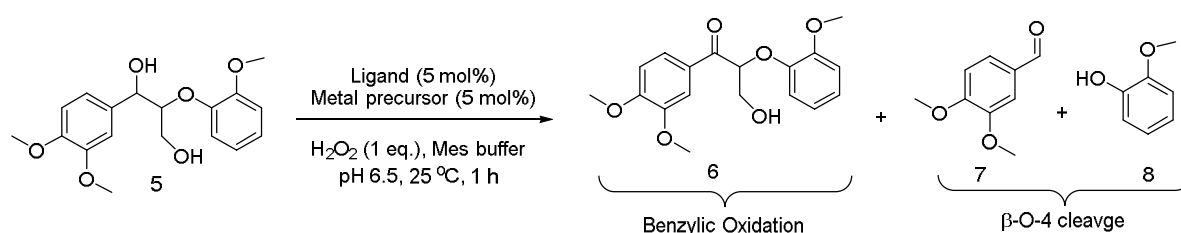
Scheme 3: Scheme showing the observed products from either oxidation of the model compound or from cleavage of the β-O-4 linkage.

*Table 1: Artificial metalloenzyme catalysed oxidation of lignin model **5**.[a]*

Entry	Scaffold/Ligand	Metal	Conversion[b] (%)	Yield (%)[b]		
				6	7	8
1	SCP-2L A100C-**D**	Fe(OTf)$_2$.2MeCN	100	85	-	-
2	SCP-2L A100C-**D**	FeCl$_3$	0	-	-	-
3	SCP-2L V83C-**D**	Fe(OTf)$_2$.2MeCN	85	70	-	-
4	SCP-2L A100C-**E**	Fe(OTf)$_2$.2MeCN	35	28	-	-
5	SCP-2L V83C-**E**	Fe(OTf)$_2$.2MeCN	0	-	-	-
6	SCP-2L A100C-**C**	Fe(OTf)$_2$.2MeCN	0	-	-	-

7	SCP-2L A100C-**A**	Fe(OTf)$_2$.2MeCN	15	12	-	-
8	SCP-2L A100C-**B**	Fe(OTf)$_2$.2MeCN	0	-	-	-
9	SCP-2L A100C-**F**	Fe(OTf)$_2$.2MeCN	0	-	-	-
10	Cofactor **D** [c, e]	Fe(OTf)$_2$.2MeCN	45	8	20	10
11	Cofactor **D** [d, e]	Fe(OTf)$_2$.2MeCN	0	-	-	-
12	WT SCP-2L	Fe(OTf)$_2$.2MeCN	0	-	-	-
13	WT SCP-2L + Cofactor **D**	Fe(OTf)$_2$.2MeCN	0	-	-	-
14	-	Fe(OTf)$_2$.2MeCN	0	-	-	-
15	-	-	0	-	-	-

[a] Conditions: Catalyst preformed *in situ* by addition of the metal precursor (1eq.) to the modified protein solution (61 µM, 50 mM Mes, 20 mM NaCl, pH 6.0). Reactions carried out at 25 °C with **5** (1.23 µmol), H$_2$O$_2$ (1 eq.) and metalloprotein in Mes buffer (5 mol%, 61 µM, 1 mL, Mes buffer: 50 mM Mes, 20 mM NaCl, pH 6.0). [b] Conversions and yields were determined by HPLC with internal standard (Naphthalene), in triplicate. [c] Reaction carried out with **5** (1.23 µmol), H$_2$O$_2$ (1 eq.) in MeCN. [d] In buffer solution 50 mM Mes buffer 20mM NaCl pH 6.0. [e] Stirring at 60 rpm, at rt.

No activity was observed with either the methyl-tacn based cofactor, **F**, or the phenanthroline based cofactor, **C**. The pyridyl cofactors, **D**, **E**, and **A** provided the oxidised product **6**, in variable yields. Changing the electronics of the central nitrogen atom in **E** by linking through an amide bond, cofactor **B**, was detrimental to catalysis resulting in no activity. Overall, the TPA based cofactor, **D**, showed the most promise giving full conversion under the initial conditions.

Control reactions with cofactor **D** showed that the protein scaffold is changing the reactivity of the parent Fe-tpa complex. In acetonitrile, cofactor **D** + Fe(OTf)$_2$ showed moderate conversion and a mixture of oxidation, **6** and cleavage products, **7** and **8** (Table 1, entry 10). Repeating the reaction in buffer rendered the catalyst inactive (Table 1, entry 11). However, in the ArM activity is retained despite the aqueous environment and increased selectivity is obtained, with **6** as the sole product being observed. This is attributed to the formation of a unique active site containing the cofactor, as no conversion is observed when Fe(OTf)$_2$ is added directly to the WT SCP-2L protein potentially leading to random iron-protein interactions. Additional evidence is provided by the change in the extent of conversion under the same conditions when changing the site of modification within the protein scaffold. SCP-2L V83C-**D** gave lower conversion and yield, suggesting the protein scaffold plays a role either in controlling substrate positioning/access or potentially though amino acid interactions that stabilise the catalytic metal center.

We also explored if other metal precursors could be utilised. Due to the presence of excess chloride ions in the buffer, we expect that the iron(II) chloride complex will be the most likely form of the metalloenzymes, and that different non-coordinating anions will not play a role in the actual catalysis. No activity was observed when FeCl$_3$ was used as the precursor, suggesting that the active species cannot be accessed directly from the LFeCl$_3$ complex.

Lowering the ArM loading led to decreased conversions (see SI, Table S4) beyond that expected by first order kinetics suggesting that we were no longer at the optimal ArM concentrations for catalysis To obtain further kinetic details about the system, the initial rates were determined at different substrate concentrations (0.6-2.8 mM). Across these concentrations saturation kinetics were not observed, instead the reaction was found to be first order in substrate concentration as expected for enzymes at low substrate concentration (see SI). The turnover numbers of the ArM are relatively low (<20 TONs), indicating a catalyst with a limited lifetime, and that catalyst decomposition is occurring under the reaction conditions. Current work is looking to improve the ArM stability using rational design based on the SCP super family to identify beneficial mutations for stability.

Substrate scope
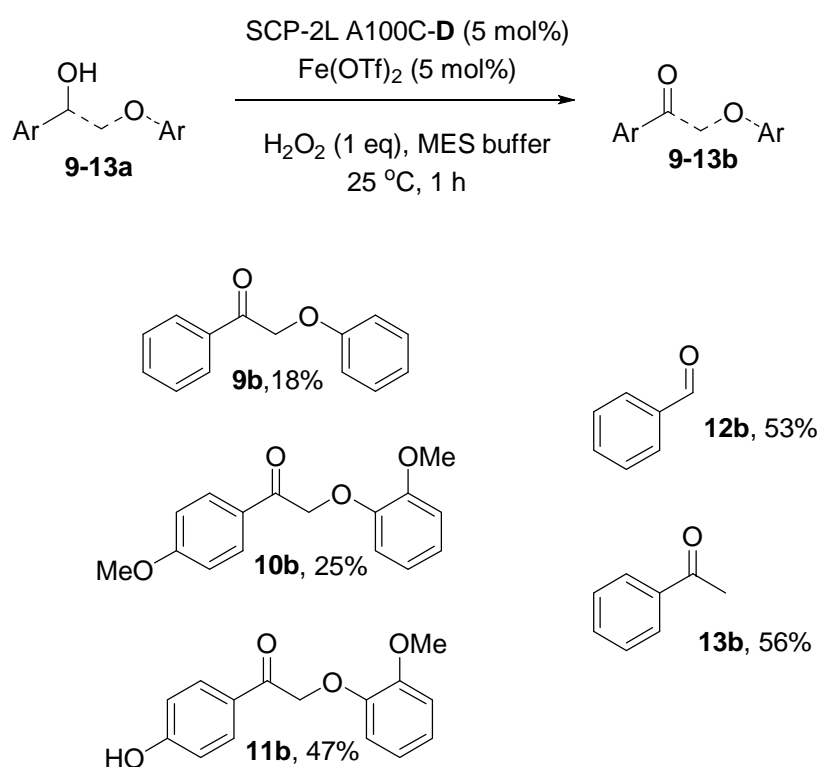
Scheme 4: Substrate scope for ArM catalysed benzylic oxidation. Yields determined by HPLC.

A range of model compounds have been designed to mimic the β-O-4 linkage of lignin and utilised in the literature to explore how the different substituents in lignin may impact on the reactivity of the β-O-4 linkage.[2] A variety of lignin model compounds (**9a**-**11a**) were subjected to our optimised conditions (Scheme 4). Removing the γ-hydroxyl moiety led to incomplete conversion under the standard conditions. Introducing a phenol moiety increased conversion over the unsubstituted or methoxy substituted model compounds, albeit still with lower yields than observed with model **5**. We also wished to see if the catalyst would work more generally for benzylic oxidation. Therefore, SCP-2L A100C-**D** was tested in the oxidation of benzyl alcohol (**12a**) and 1-phenylethanol (**13a**). Both primary and secondary benzylic alcohols are tolerated in the reaction giving moderate yields of the corresponding aldehyde (**12b**) or ketone (**13b**). Potentially the reduction in yield of the simple model

compounds could be due to reduced water solubility of the lignin model. Acetonitrile has been used as a co-solvent in oxidation reactions with ArMs[31] in the literature to aid substrate solubilisation. The addition of MeCN as a co-solvent was tolerated in the reaction with compound **5** to 30% MeCN albeit with reduced yields (Table 2, entry 2 vs 1). Increasing the co-solvent ratio to 50% led to very low levels of activity (Table 2, entry 3).

A key limitation in most methodology developed on lignin model compounds is the insolubility of lignin in most solvents. DMSO and dioxane are often solvents of choice for reactions with lignin. Both DMSO and dioxane were also shown to be tolerated as co-solvents under the standard reaction conditions, although substantial decreases in yield were observed (Table 2 entry 1 vs entries 4-6).

*Table 2: Co-solvent tolerance in the oxidation of lignin model compound **5**.*

Entry	Solvent (Ratio)	Conversion[b] (%)	Yield of 6[b] (%)
1	MES	100	85
2	MES : MeCN (7:3)	54	41
3	MES : MeCN (1:1)	12	5
4	MES : Dioxane (7:3)	38	35
5	MES : Dioxane (1:1)	0	0
6	MES : DMSO (7:3)	57	53

Conditions: Catalyst preformed *in situ* by addition of the metal precursor (1eq.) to the SCP-2L A100C-**D** solution (61 μM, 50 mM Mes, 20 mM NaCl, pH 6.0). Reactions carried out at 25 °C with **5** (1.23 μmol), H_2O_2 (1 eq.) and metalloprotein in a mixture Mes buffer: solvent system (5 mol%, 61 μM, Mes buffer: 50 mM Mes, 20 mM NaCl, pH 6.0). [b] Conversions and yields were determined by HPLC with internal standard (Naphthalene), in triplicate.

With a successful method in place for the oxidation of the benzylic alcohol in the lignin model systems, and having shown that dioxane and DMSO could be tolerated, we went on to see if SCP-2L A100C-**D** would oxidise lignin from Douglas Fir sawdust to give the Lignin^{OX} intermediate. Whilst changes consistent with oxidation were observed under the reaction conditions, almost identical changes were observed using hydrogen peroxide alone.

Computational studies of ArMs

UV-visible studies of SCP-2L A100C-**D** showed that 1 equivalent of Fe(OTf)$_2$ bound to the cofactor (see SI, Figures S30 and S31). Further attempts to characterise the obtained ArM with MS were unsuccessful and to date we have been unable to obtain crystal structures of any of the modified proteins or their metal complexes. Therefore, in parallel to this work, classical molecular dynamics (MD) simulations were carried out to probe the position of the cofactors in the protein and any interactions that could be observed.

Cofactor **F** was chosen for this work as it is fairly rigid and will have close to symmetrical coordination to one face of the octahedral M^{2+} ion. The main purpose of these simulations is to explore the structural integrity of the protein upon functionalisation and to map the regions that the metal complex tethered to the enzyme would be able to visit (and potentially interact with). Because ligand **F** has the same ethylene tether and the same coordination number as ligand **E**, which can show some activity (Entry **4**, Table 1), the results of these simulations should be informative and potentially transferable to the more active catalysts identified in this work. A complication for the computational work is the possibility of two stereoisomers on modification of the cysteine. Due to the complexity of proteins there is no method to predict if one stereoisomer will be favoured due to the protein environment or if a racemic mixture is formed. MD simulations were therefore carried out with each stereoisomer, and as expected slightly different distances and interactions were observed for each stereoisomer. In both cases, however, the structural fold of the peptide is essentially maintained, as is the hydrophobic channel, which remains mostly closed due to a loose contact between two lipophilic residues (Phe93 and Ile104). During parts of the simulations, the metal complex appeared to approach the Phe93/Phe94 cluster (for Phe94 see the black curve in Figure 2 between 0-500 ps). Because Phe93 was involved in maintaining the structure of the hydrophobic channel (which may be important for the overall structural integrity), Phe94 was chosen as a potential target for mutations in order to increase the stability of the metal cofactor. The mutations that were subsequently realised experimentally are detailed in the following section.

Mutants to increase cofactor stability of SCP-2L A100C

Amino acids that can coordinate to metals such as His or Asp are known to stabilise $M-O_2$ and $M-OOH$ intermediates either directly or through hydrogen bonding in a number of enzymes.[35,36] Adding acetic acid to oxidation reactions catalysed by iron complexes has also been shown to improve yields, presumably though similar stabilisation of iron intermediates.[37] The mutants SCP-2L A100C F94E and F94H were considered most promising and were expressed, purified and modified with cofactor **D** as above. These were then tested in the oxidation of compound **5**. The introduction of His at position 94 had no influence on the reaction (Table 3, entry 3 and 6 vs entry 1 and 4); however, the introduction of Glu had a positive influence on the reactivity increasing the observed yield (Table 3, entry 2 and 5 vs entry 1 and 4). The increase was most remarkable with cofactor **E**, where the yield doubled on the introduction of Glu (Table 3, entry 5 vs entry 4.). Interestingly, small increases were also observed when one equivalent of acetic acid was simply added as an additive (Table 3, entries 7 and 9 vs entry 1). This effect was increased on the addition of further equivalents of acetic acid (Table 3, entry 8). These results match observations seen in the literature with small molecule catalysts.[37]

Table 3: Lignin oxidation with mutant ArMs
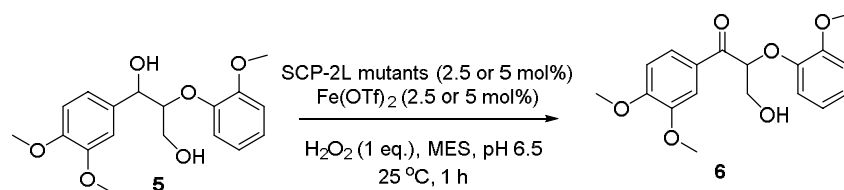
Entry	Scaffold/Ligand	AcOH eq	Conversion (%)	Yield of 6 (%)
1	SCP-2L A100C-**D**[a]	-	35	31
2	SCP-2L A100C F94E-**D**[a]	-	47	43
3	SCP-2L A100C F94H **D**[a]	-	35	29
4	SCP-2L A100C-**E**[b]	-	35	28
5	SCP-2L A100C F94E-**E**[b]	-	59	54
6	SCP-2L A100C F94H-**E**[b]	-	35	30
7	SCP-2L A100C-**D**[a]	1	40	36
8	SCP-2L A100C-**D**[a]	5	53	48
9	SCP-2L A100C-**E**[b]	1	45	41

Conditions: Catalyst preformed *in situ* by addition of the metal precursor (1eq.) to the modified protein solution (61 µM, 50 mM, 20 mM NaCl, pH 6.0). Reactions carried out at 25 °C with **5** (1.23 µmol), H_2O_2 (1 eq.) and metalloprotein in Mes buffer (1 mL, Mes buffer: 50 mM, 20 mM NaCl, pH 6.0). Conversions and yields were determined by HPLC with internal standard (Naphthalene), in triplicate. [a] 2.5 mol% SCP-2L A100C-F94X-**D**, and 2.5 mol% Fe(OTf)$_2$, [b] 5 mol% SCP-2L A100C-F94X-**E** and 5mol% Fe(OTf)$_2$

To rationalise these findings, we performed MD simulations for a suitable model for the Glu mutant, expecting to see an enhanced contact between the carboxylate side chain and the metal complex, or possibly even a spontaneous coordination of this moiety to the metal. For consistency with the WT simulations, the same ligand **F** was used, and for simplicity Phe94 was mutated into Asp rather than Glu, because the shorter Asp side chain has less conformational degrees of freedom to explore. Again, both R and S stereoisomers were modelled. While in no case, on the timescale of our simulations, the carboxylate displaced a water ligand and coordinated to the metal, on average the carboxylate tended to be significantly closer to the metal than the phenyl group in the WT (see red curve in Figure 2). Occasionally the carboxylate moiety would come within H-bonded distance to a water ligand coordinated to the metal (e.g. between 1.7 - 2 ns in Figure 2, see Figure 3 for a snapshot from this region), clearly poised for a stabilising interaction and, possibly, an eventual coordination to the metal (modelling of which would require longer time scales and/or constrained free-energy simulations that are beyond the scope of the present paper). The strategy to engineer mutants that can potentially bind to the metal, thus providing an additional stabilisation, is thus fully borne out by these simulations.
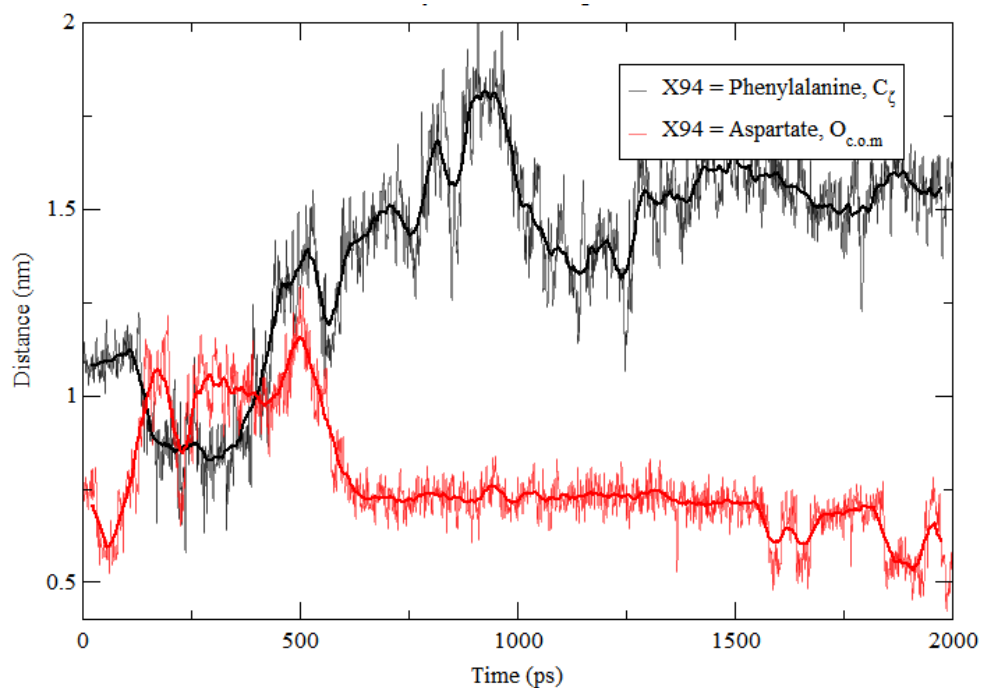
Figure 2: Fluctuations of the distances between the metal and selected points of the side chain of residue 94 during 2 ns of MD of the protein-ligand complex (results shown for the S-isomer, see Figure SE for the R isomer); in black: WT (distance to C_ζ of Phe94), in red: F94D mutant (distance to the centre of mass of the two carboxylate O atoms, which, when Phe and Asp residues are overlaid, is very close to the position of C_ζ).
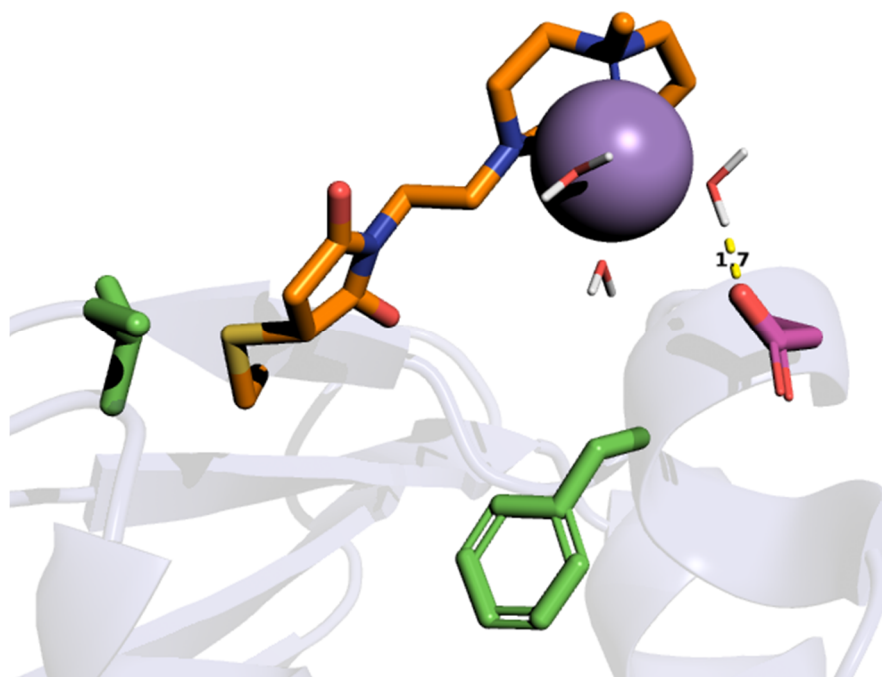

Figure 3: Snapshot from the simulation of the F94D mutant showing a close contact between the carboxylate side chain and the metal fragment (distance in Å).

Overall, we have shown that ArMs can be used to change the selectivity of chemical reactions compared to the parent small molecule catalysts, providing selective catalysts. In this case the incorporation of the TPA ligand within a protein scaffold gave rise to selective benzylic alcohol oxidation. Current work is focused on understanding more about the iron intermediates in the reaction and improving the protein scaffold to increase its stability enabling a greater range of reactions to be exploited using these and related metalloenzymes. The observation that the conversion can be increased by mutations in the surrounding protein environment, suggests that future genetic optimisation of the protein will lead to more active enzymes that can be used for larger scale processes and ultimately within the biorefinery.

Supporting Information

Synthesis of Ligands (Experimental and NMR), Protein Expression and Purification, Artificial metalloenzyme preparation (Mass spectrometry, Circular Dichroism, Determination of Concentration), Metal coordination studies, Model compound oxidation (General procedure, HPLC conditions, Metal loading, Kinetics), Further computational information.

References

[1] S. R. Collinson, W. Thielemans, The catalytic oxidation of biomass to new materials focusing on starch, cellulose and lignin, *Coord. Chem. Rev*. **2010**, 254, 1854-1870.

[2] J. Zakzeski, P. C. Bruijnincx, A. L. Jongerius, B. M. Weckhuysen, The Catalytic Valorization of Lignin for the Production of Renewable Chemicals, *Chem. Rev*. **2010**, 110, 3552-3599.

[3] F. S. Chakar, A. J. Ragauskas, Review of current and future softwood kraft lignin process chemistry, *Ind. Crops. Prod*. **2004**, 20, 131-141.

[4] M. Fache, B. Boutevin, S. Caillol, Vanillin Production from Lignin and Its Use as a Renewable Chemical, *ACS Sustainable Chem. Eng*. **2016**, 4, 35-46.

[5] P. J. Deuss, K. Barta, J. G. de Vries, Homogeneous catalysis for the conversion of biomass and biomass-derived platform chemicals, *Catal Sci. Technol*, **2014**, 4, 1174-1196.

[6] M. D. Karkas, B. S. Matsuura, R. M. Monos, G. Magallanes, C. R. J. Stephenson, Transition-metal catalyzed valorization of lignin: the key to a sustainable carbon-neutral future, *Org. Biomol. Chem*, **2016**, 14, 1853-1914.

[7] I. Bosque, G. Magallanes, M. Rigoulet, M. D. Kärkäs, C. R. J. Stephenson, Redox Catalysis Facilitates Lignin Depolymerisation, *ACS Cent. Sci.* **2017**, *3*, 621-628.

[8] A. Rahimi, A. Azarpira, H. Kim, J. Ralph, S. S. Stahl, Chemoselective Metal-Free Aerobic Alcohol Oxidation in Lignin, *J. Am. Chem. Soc*, **2013,** 135, 6415-6418.

[9] A. Rahimi, A. Azarpira, J. J. Coon, S. S. Stahl, Formic-acid-induced depolymerization of oxidized lignin to aromatics, *Nature*, **2014**, 515, 249-252.

[10] H. Guo, D. M. Miles-Barrett, A. R. Neal, T. Zhang, C. Li, N. J. Westwood, Unravelling the enigma of lignin[OX]: can the oxidation of lignin be controlled? *Chem. Sci.* **2018**, 9, 702-711.

[11] C. S. Lancefield, O. S. Ojo, F. Tran, N. J. Westwood, Isolation of Functionalized Phenolic Monomers through Selective Oxidation and C-O Bond Cleavage of the β-O-4 Linkages in Lignin, *Angew. Chem.* **2015**, *127*, 260-264.

[12] S. Dabral, J. G. Hernández, P. C. J. Kamer, C. Bolm. Organocatalytic Chemoselective Primary Alcohol Oxidation and Subsequent Cleavage of Lignin Model Compounds and lignin, *ChemSusChem* **2017**, *10*, 2707-2713.

[13] V. K. Thakur, M. K. Thakur, P. Raghavan, M. R. Kessler, Progress in Green Polymer Composites from Lignin for Multifunctional Applications: A Review, *ACS Sustainable Chem. Eng.* **2014**, *2*, 1072-1092.

[14] R. ten-Have, P. J. Teunissen, Oxidative mechanisms involved in lignin degradation by white-rot fungi. *Chem. Rev*. **2001**, 101, 3397-3413.

[15] G. M. M. Rashid, C. R. Taylor, Y. Liu, X. Zhang, D. Rea, V. Fülöp, T. D. H. Bugg. Identification of Manganese Superoxide Dismutase from *Sphingobacterium sp.*T2 as a Novel bacterial Enzyme for Lignin Oxidation, *ACS. Chem. Biol.* **2015**, *10*, 2286-2294.

[16] M. E. Brown, T. Barros, M. C. Y. Chang, Identification and Characterization of a Multifunctional Dye Peroxidase from a Lignin-Reactive Bacterium, *ACS Chem. Biol.* **2012**, *7*, 2074-2081.

[17] C. Zhu, W. Ding, T. Shen, C. Tang, C. Sun, S. Xu, Y. Chen, J. Wu, H. Ying, Metallo-Deuteroporphyrin as a Biomimetic Catalyst for the Catalytic Oxidation of Lignin to Aromatics, *ChemSusChem* **2015**, *8*, 1768-1778.

[18] S. Son, F. D. Toste, Non-Oxidative Vanadium-Catalyzed CO Bond Cleavage: Application to Degradation of Lignin Model Compounds *Angew. Chem. Int. Ed.* **2010**, *49*, 3791-3794.

[19] F. Schwizer, Y. Okamoto, T. Heinisch, Y. Gu, M. M. Pellizzoni, V. Lebrun, R. Reuter, V. Köhler, J. C. Lewis, T. R. Ward, Artificial metalloenzymes: Reaction scope and optimization strategies, *Chem. Rev.* **2018**, *118*, 142-231.

[20] M. V. Doble, A. C. C. Ward, P. J. Deuss, A. G. Jarvis, P. C. J. Kamer, Catalyst design in oxidation chemistry; from $KMnO_4$ to artificial metalloenzymes, *Bioorg. Med. Chem.* **2014**, *22*, 5657-5677.

[21] S. I. Mann, T. Heinisch, T. R. Ward, A. S. Borovik, Peroxide Activation Regulated by Hydrogen Bonds within Artificial Cu Proteins, *J. Am. Chem. Soc.* **2017,** 139, 17289-17292.

[22] P. J. deuss, G. Popa, C. H. Botting, W. Laan, P. C. J. Kamer, Highly Efficient and Site-Selective Phosphane Modification of Proteins through Hydrazone Linkage: Development of Artificial Metalloenzymes, *Angew. Chem. Int. Ed.* **2010**, *49*, 5315-5317.

[23] A. G. Jarvis, L. Obrecht, P. Deuss, W. Laan, E. K. Gibson, P. P. Wells, P. C. J. Kamer, Enzyme Activity by Design: An Artificial Rhodium Hydroformylase for Linear Aldehydes, *Angew. Chem. Int. Ed.* **2017**, *56*, 13596-13600.

[24] A. M. Haapalainen, D. M. van Aalten, G. Meriläinen, J. E. Jalonen, P. Pirilä, R. K. Wierenga, J. K. Hiltunen, T. Glumoff, Crystal structure of the liganded SCP-2-like domain of human peroxisomal multifunctional enzyme type 2 at 1.75 A resolution, *J. Mol. Biol.* **2001**, *313*, 1127-1138.

[25] A. D. MacKerell, D. Bashford, M. Bellott, R. L. Dunbrack, J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, *et al*. All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins. *J. Phys. Chem. B* **1998**, *102*, 3586–3616.

[26] A. D. Mackerell, M. Feig, C. L. Brooks, Extending the treatment of backbone energetics in protein force fields: Limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulation. *J. Comput. Chem.* **2004**, 25, 1400–1415.

[27] N. Foloppe, A. D. MacKerell, All-Atom Empirical Force Field for Nucleic Acids: I. Parameter Optimization Based on Small Molecule and Condensed Phase Macromolecular Target Data. *J. Comput. Chem.* **2000**, *21*, 86–104.

[28] R. B. Best, X. Zhu, J. Shim, P. E. M. Lopes, J. Mittal, M. Feig, A. D. MacKerell, Optimization of the additive CHARMM all-atom protein force field targeting improved sampling of the backbone ϕ, ψ and side-chain χ1and χ2Dihedral Angles. *J. Chem. Theory Comput.* **2012**, *8*, 3257–3273.

[29] B. Hess, C. Kutzner, D. van der Spoel, E. Lindahl, GROMACS 4: algorithms for highly efficient, load balanced, and scalable molecular simulations. *J. Chem. Theory Comput.* **2008**, *4*, 435–447.

[30] P. J. Deuss, G. Popa, A. M. Z. Slawin, W. Laan, P. C. J. Kamer, Artificial Copper Enzymes for Asymmetric Diels–Alder Reactions, *ChemCatChem* **2013**, *5*, 1184-1191.

[31] C. Zhang, P. Srivastava, K. Ellis-Guardiola, J. C. Lewis, Manganese terpyridine artificial metalloenzymes for benzylic oxygenation and olefin epoxidation, *Tetrahedron*, **2014**, *70*, 4245-4249.

[32] S. A. Trammell, H. M. Goldston, Jr., P. T. Tran, L. M. Tender, D. W. Conrad, D. E. Benson, H. W. Hellinga, Synthesis and Characterization of a Ruthenium(II)-Based Redox Conjugate for Reagentless Biosensing, *Bioconjugate Chem*. **2001**, *12*, 643–647.

[33] D. M. Gooden, 6-(4-(2-{2-[2-(2-Hydroxy-ethoxy)-ethoxy]-ethoxy}-ethyl)-10-oxa-4-aza-tricyclo[5.2.1.02,6]dec-8-ene-3,5-dione), *Molbank* **2009**, M638.

[34] D. Fujita, K. Suzuki, S. Sato, M. Yagi-Utsumi, E. Kurimoto, Y. Yamaguchi, K. Kato, M. Fujita, Synthesis of a Bridging Ligand with a Non-denatured Protein Pendant: Toward Protein Encapsulation in a Coordination Cage, *Chem. Lett.* **2012**, *41*, 313-315.

[35] J. S. Olsen, A. J. Mathews, R. J. Rohlfs, B. A. Springer, K. D. Egeberg, S. G. Sligar, J. tame, J-P. Renaud, K. Nagai, The role of the distal histidine in myoglobin and haemoglobin, *Nature* **1988**, *336*, 265-266.

[36] E. A. Span, D. L. M, Suess, M. C. Deller, R. D. Britt, M. A. Marletta, The Role of the Secondary Coordination Sphere in a Fungal Polysaccharide Monooxygenase, *ACS Chem Biol.* **2017**, *12,* 1095-1103.

[37] M. S. Chen, M. C. White, A Predictably Selective Aliphatic C–H Oxidation Reaction for Complex Molecule Synthesis, *Science* **2007**, *318*, 783-787.