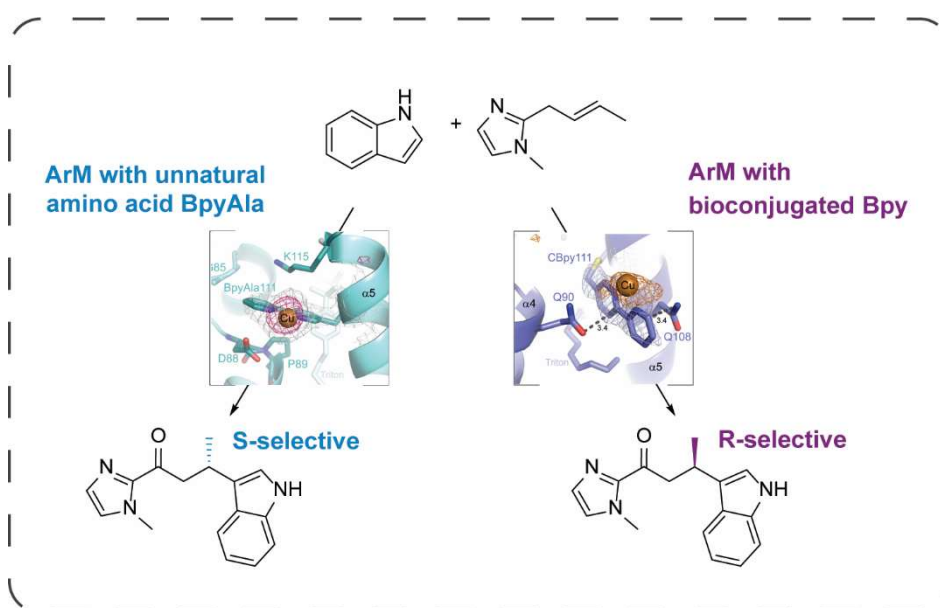


Using BpyAla to generate Cu Artificial metalloenzymes: a catalytic and structural study

E. Klemencic¹, R. C. Brewster¹, H. S. Ali^{1,2}, J. M. Richardson³, A. G. Jarvis^{1*}

¹School of Chemistry, Joseph Black Building, Kings Buildings, University of Edinburgh, David Brewster Road, Edinburgh, EH9 3FJ. ² Current address: Chemistry Research Laboratory, Department of Chemistry and the INEOS Oxford Institute for Antimicrobial Research, University of Oxford, 12 Mansfield Road, Oxford OX1 3TA, UK. ³ School of Biological Sciences, Swann Building, University of Edinburgh, Edinburgh, EH9 3BF. *Corresponding author: amanda.jarvis@ed.ac.uk

Abstract



Artificial metalloenzymes (ArMs) have emerged as a promising avenue in the field of biocatalysis, offering new reactivity. However, their design remains challenging due to the limited understanding of their protein dynamics and how the introduced cofactors alter the protein scaffold structure. Here we present the structures and catalytic activity of novel copper ArMs capable of (*R*)- or (*S*)-stereoselective control, utilizing a steroid carrier protein (SCP) scaffold. To incorporate 2,2'-Bipyridine (Bpy) into SCP, two distinct strategies were employed: either Bpy was introduced as an unnatural amino acid (2,2'-bipyridin-5-yl)alanine (BpyAla) using amber stop codon expression or *via* bioconjugation of bromomethyl-Bpy to cysteine residues. The resulting ArMs proved to be effective at catalysing an enantioselective Friedel-Crafts reaction with SCP_Q111BpyAla achieving the best selectivity with an enantioselectivity of 72% *ee* (*S*). Interestingly, despite using the same protein scaffold, different attachment strategies for Bpy at the same residue (Q111) led to a switch in the enantiopreference of the ArM.

X-ray crystal structures of SCP_Q111CBpy and SCP_Q111BpyAla ArMs with bound Cu(II) ions unveiled crucial differences in the orientation of the catalytic centre. Combining structural information, alanine scanning studies, and computational analysis shed light on the distinct active sites of the ArMs, clarifying that these active sites stabilise the nucleophilic substrate on different sides of the electrophile leading to the observed switch in enantioselectivity. This work underscores the importance of integrating structural studies with catalytic screening to unravel the intricacies of ArM behaviour and facilitate their development for targeted applications in biocatalysis.

Keywords: Artificial Metalloenzymes, unnatural amino acids, Friedel-Crafts, non-canonical amino acids, protein structure.

Introduction

Artificial metalloenzymes (ArMs) represent an avenue to new-to-nature reactions^{1–4}. These systems expand the biocatalytic toolbox by using transition metal complexes tethered to protein scaffolds; notable examples of chemical transformations using ArMs include transfer hydrogenation⁵, hydroformylation^{6,7}, *in vivo* metathesis⁸, lignin oxidation⁹, Friedel-Crafts alkylation¹⁰ and other cross-coupling reactions¹¹.

Site-specific coordination of synthetic metal complexes to proteins plays a crucial role in the development of stereoselective artificial metalloenzymes (ArMs), wherein the protein scaffold forms a secondary coordination sphere around the reactive centre providing stereocontrol^{12,13}. To synthesise ArMs, researchers have used many different approaches from simply leveraging native metal-binding activity found in certain proteins¹⁴ to combining synthetic chemistry and proteins, including using variants of natural cofactors, such as the iron-binding protoporphyrin IX, with alternative transition metals^{15,16}. More synthetic approaches to site-specific incorporation of metal complexes include the utilisation of supramolecular binding, with several different tethered transition metal complex systems explored to date, the most notable being the biotin-streptavidin (Sav) system¹⁷. Bioconjugation of metal complexes *via* covalent attachment with unique reactive amino acid residues such as cysteine or azidophenylalanine is another widely used approach to prepare ArMs¹⁸. *In vivo* ligand incorporation offers an attractive alternative where metal-binding unnatural amino acids such as (2,2'-bipyridin-5-yl)alanine (BpyAla) can be selectively introduced to the protein scaffold directly using genetic code expansion technologies, allowing for more streamlined enzyme engineering^{19,20}.

Within these different approaches to ArM design, several scaffolds have been repeatedly used with various attachment strategies for transformation into novel catalysts. These scaffolds have been described as 'privileged' scaffolds, in a manner analogous with privileged ligands in homogeneous catalysis. The most prominent examples are the LmrR scaffold, which allows supramolecular, bioconjugation, and incorporation of unnatural amino acids²¹, and Sav, which has been explored in dative coordination²², alongside the more common supramolecular approach¹⁷. Despite these advances, few comparative studies have been conducted which included direct comparisons of the modification strategy for site-selective metal coordination.

This makes it difficult to discern the role of protein modification upon ArM reactivity and predict the most effective route to ArM assembly for a desired application.

The Kamer and Jarvis groups have extensively studied the human steroid carrier protein (SCP-2L)²³, a single-domain 13.5 kDa protein containing a hydrophobic tunnel, as a scaffold for ArM design. It has been exploited in the design of ArMs for selective hydroformylation using Rh-phosphine organometallic complexes⁶, which showed high activity and selectivity in the production of long chain linear aldehydes, under aqueous conditions. Moreover, engineering the scaffold for improved thermostability incurred a five-fold increase in TON compared to the wild type⁷. This protein has also been used in ArMs for the selective oxidation of lignin model compounds⁹, for enantioselective Cu-catalysed Diels-Alder reactions²⁴, and artificial photoenzyme design²⁵. Other carrier proteins such as adipocyte lipid binding protein (ALBP)²⁶, ferric hydroxamate uptake protein component A (FhuA)²⁷, and maltose binding protein²⁸ have been used in the design of ArMs suggesting that proteins with a carrier function may also serve as privileged scaffolds. In this work, the potential of SCP-2L to serve as a privileged protein scaffold is explored, through studying the copper-catalysed 1,4-addition of indole to enones.

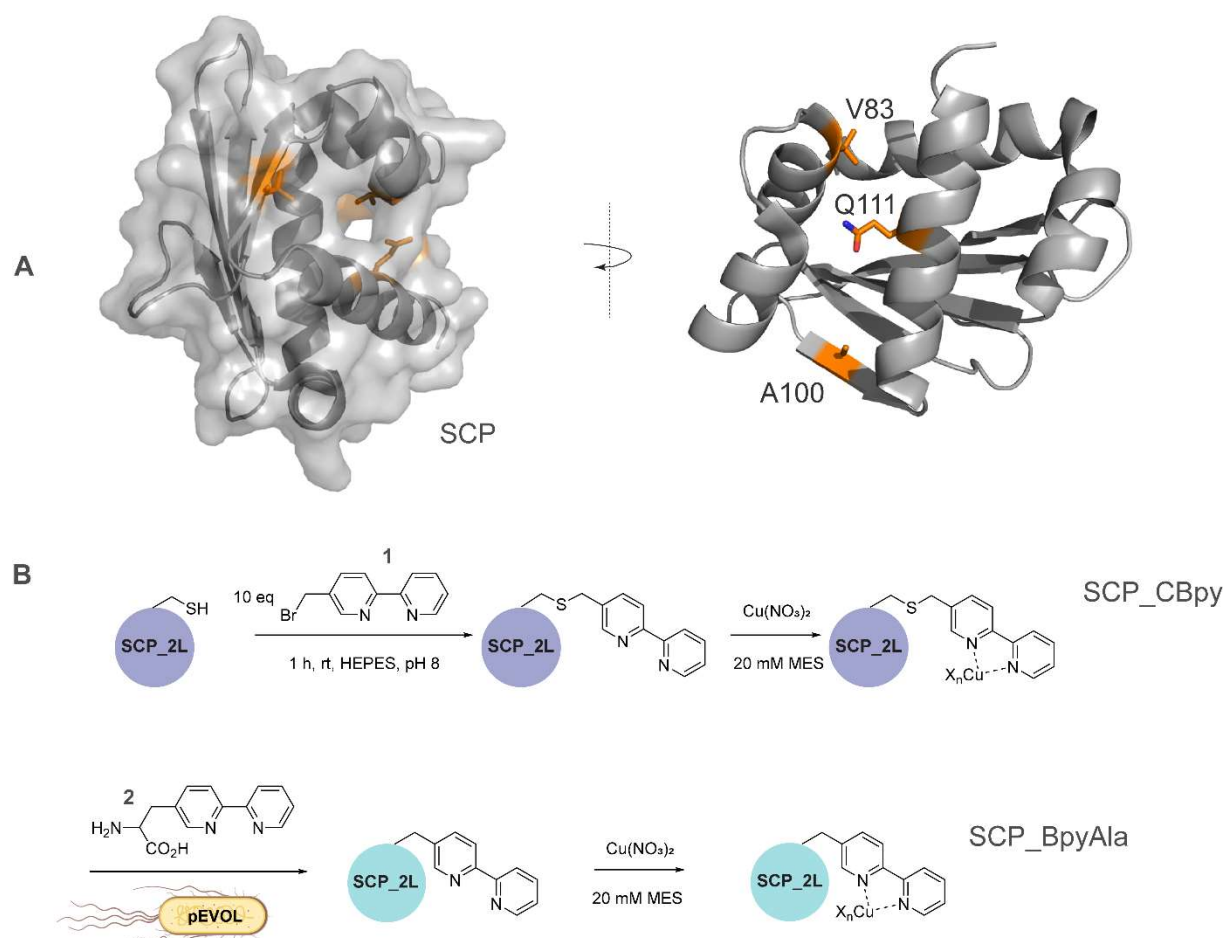


Figure 1: **A** SCP-2L and the residues (orange) used for site-selective ligand attachment (PDB 1IKT). **B** Comparison of modification approaches for SCP_{Cu} ArM design: (Upper panel) *In vitro*

bioconjugation of unique Cys residues with **1** before addition of Cu²⁺ ions in a two-step protocol; (Lower panel) one-step *in vivo* site-selective incorporation of the unnatural amino acid BpyAla **2**.

Bpy is a well-studied ligand for transition metals and is utilised in the coordination sphere of many catalytic complexes²⁹. It can be introduced into a protein via the bioconjugation of maleimide or bromomethylbipyridines at cysteine residues (Figure 1). Bpy is also one of very few metal ligands that has been incorporated into the genetic code by amber stop codon suppression. Here two methods for site-selective incorporation of bipyridine into the SCP-2L scaffold were compared, and the copper-catalysed 1,4-addition of indole to enones was used to analyse ArM activity. Differences in stereoselectivities were identified with results rationalised using ArM crystal structures and DFT simulations.

Results and discussion

Preparation of SCP-Cu artificial metalloenzymes

The steroid carrier protein type 2 like (SCP-2L) domain of the human multifunctional enzyme 2 (MFE-2) has previously been used as a protein scaffold for both regioselective and enantioselective reactions using covalent modification through the introduction of unique cysteine residues³⁰. A number of residues in the scaffold's apolar tunnel (V83, A100, Q111) have previously been identified as being amenable to mutation and modification⁶. Their location along the apolar tunnel provides an ideal starting point for the introduction of bipyridine into the protein scaffold to create an active site within a protein pocket. Three different sterol carrier proteins, each containing a unique cysteine at position Q111C, A100C or V83C, were prepared as described previously⁶. Briefly, pEHISTEV plasmids containing the SCP gene inserted after an *N*-terminal His₆ tag with TEV cleavage site were transformed into Rosetta DE3 *E. coli* cells. His₆-tagged proteins were expressed in production broth media and purified by Ni-affinity chromatography with the His₆ tag subsequently removed using TEV protease; a further purification step by Ni-affinity chromatography gave the final proteins in good yields (20-70 mgL⁻¹ of media). Bipyridine was introduced into the protein scaffold through bioconjugation of the cysteine residues with 10 equivalents of 5-bromomethyl-2,2'-bipyridine **1** in HEPES buffer at pH 8 (Figure 1B). The optimal conversion was approximately 70-85% modified protein, with minimal secondary modification (<5 % secondary modification; the remaining mass balance is unreacted protein, see Figure 2 and ES1a Fig. S4).

In comparison to bioconjugation, genetic code expansion using stop codon suppression provides a method of directly incorporating Bpy as the amino acid BpyAla **2** during protein expression (Figure 1B)³¹. The synthesis of BpyAla was optimised to allow BpyAla production on scales of 10 grams without the need for column chromatography (see ES1a for full details).

The SCP-2L gene was codon-optimised for *E. coli* and prepared by gene synthesis. Amber stop codons were introduced at A100, V83 or Q111 by site-directed mutagenesis. *E. coli* BL21 (DE3) cells were co-transformed with the pEVOL-BpyAla plasmid, which contains the orthogonal *MjTyrRS/MjtRNA^{Tyr}* genes³¹, and the pET28 plasmid carrying the SCP gene with a C-terminal TEV cleavable His₆ tag. The genes were expressed in the presence of 0.5 mM BpyAla²⁰ and the

resulting proteins purified by Ni-affinity chromatography with the His₆ tag removed using TEV-protease. Typical yields were 5-15 mg L⁻¹ in LB media, representing a decrease of 20-80% compared to the yields of SCP-2L without BpyAla. Proteins containing unnatural amino acids are known to have lower expression yields, due to incomplete stop codon suppression giving truncated protein.¹⁹ Mass spectrometry on the purified proteins confirmed the successful incorporation of BpyAla at residue positions 100, 111, or 83 (Figure 2 and ES1a Fig. S2).

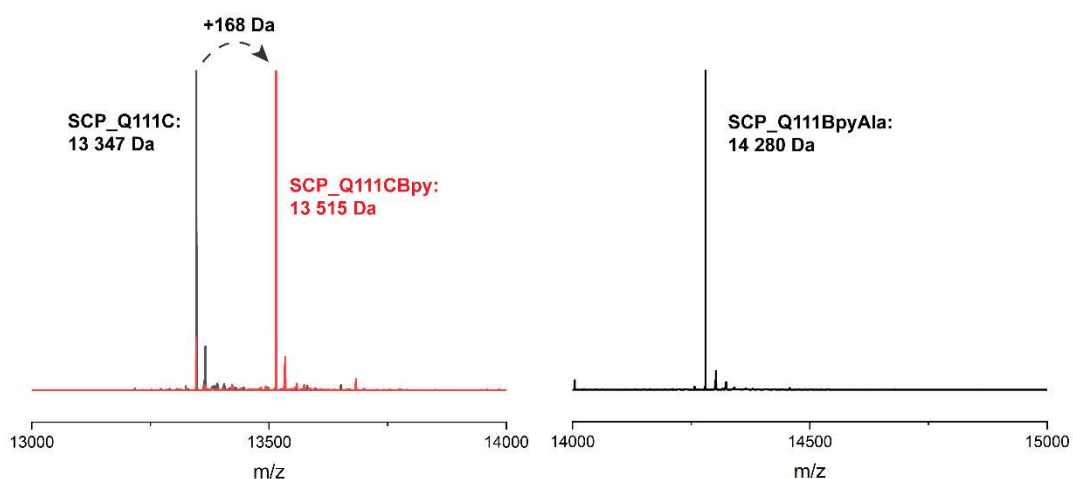


Figure 2: Mass spectra of SCP ArMs. Left panel: mass spectrum of SCP_Q111C (black line: expected mass: 13347 Da) and bioconjugation products from incorporation of **1** into SCP_Q111C (red line; expected mass after bioconjugation: 13515 Da, 10% remained unmodified). Right panel: incorporation of BpyAla during translation produced a single species – SCP_Q111BpyAla (expected mass after purification: 14280 Da. Note: The mass of the BpyAla proteins is not 46 Da lower than the SCP CBpy proteins as might be expected due to the absence of the -CH₂S- linker, but higher due to the presence of additional amino acids from the TEV site remaining at the C-terminus.).

To obtain the Cu-metalloproteins, one equivalent of Cu(NO₃)₂ was added directly to the Bpy-containing SCP proteins. Copper binding was confirmed by UV-Vis spectroscopy which clearly showed a red shift in the π - π^* transition of the Bpy ligand in the presence of Cu(II), consistent with previous reports on copper binding to Bpy-containing proteins²⁰. Titration experiments confirmed that, at the concentrations of interest (20-100 μ M), Cu(II) bound to the Bpy-containing SCP proteins with \sim 1:1 stoichiometry (see ES1a, Figures S5 and S6). Copper binding was also confirmed by ICP-MS analysis (see ES1a).

Structural analysis

Whilst, the UV-vis spectroscopy and ICP-MS studies showed copper binding to the ArMs (see ES1a), they give no information about the precise environment of the copper. We therefore looked to X-ray crystallography to obtain structural understanding of the newly created active sites within the metalloproteins. Crystals of SCP_Q111CBpy and SCP_Q111BpyAla were soaked in solutions of Cu(II) ions and their structures determined by X-ray crystallography

(Figure 3A and 3B). Triton-X-100 was needed for crystallisation, similar to wt SCP-2L (PDB ID: 1IKT)²², the apolar tunnel is occupied by Triton X-100 in the structures of Cu(II)-bound SCP_Q111CBpy and SCP_Q111BpyAla (Figure 3).

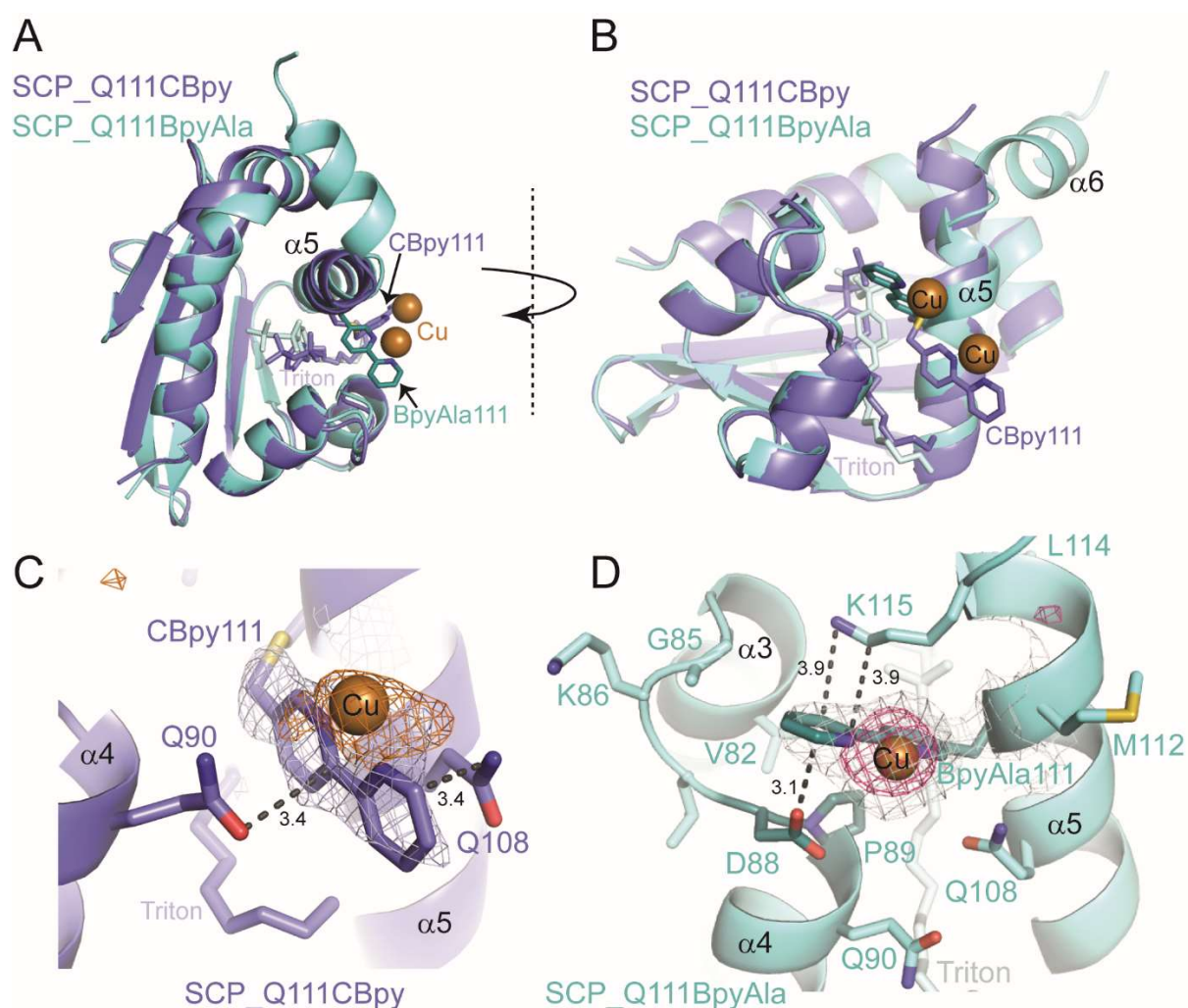


Figure 3: Crystal structures of Cu(II)-bound SCP_Q111CBpy and SCP_Q111BpyAla. **A)** Top view of superimposed structures of SCP_Q111CBpy (purple) and SCP_Q111BpyAla (cyan) looking down the tunnel, with Triton bound. The atoms of CBpy, BpyAla and Triton are shown as sticks. Orange spheres represent the Cu(II) ions. **B)** Side-view of the superimposed structures. **C)** Close-up view of CBpy111 in SCP_Q111CBpy structure. The anomalous difference map (contoured at 3σ) is shown as an orange mesh. **D)** Close-up view of BpyAla111 in the SCP_Q111BpyAla structure. A Cu omit map (contoured at 3.6σ) is shown as a pink mesh. In C and D the grey mesh represents the $2F_o - F_c$ electron density map (contoured at 1σ).

The structure of Cu(II)-bound SCP_Q111CBpy was determined to 1.52 Å resolution, and clearly confirmed the incorporation of 2,2'-bipyridine at C111, which is positioned at the centre of the C-terminal helix, $\alpha 5$. The bipyridine rings are sandwiched in a parallel fashion between the amide side-chain atoms of Q108 (one turn up on α -helix 5) and Q90 on α -helix 4 (Figure 3C). The outer and inner pyridine rings are 3.4 Å from Q108 and Q90 respectively, a distance that favours formation of stabilising aromatic π -amino electrostatic interactions. α -Helix 5 is slightly shifted relative to its position in the wt SCP_2L structure (PDB ID 1IKT) (Figure S9 in

ES1a) and is more flexible than the remainder of the structure, with B-factors ranging from 52-96 Å² from the N- to the C-terminus of the helix (N103-L120), compared to 75 Å² for the protein overall. The B-factor of the Bpy moiety is 92 Å², suggesting flexibility in its position. Additional electron density was observed in the 2F_o-F_c map around the solvent-exposed face of the Bpy moiety, and a single peak in an anomalous difference map confirmed this is a bound Cu(II) ion (Figure 3C). The Cu(II) atom lies in the same plane as the two pyridine rings and is 2 Å from the two pyridine ring nitrogens, and thus we assume that the Cu(II) ion adopts an octahedral coordination geometry. The high B-factor (108 Å²) of the Cu(II) is consistent with its position in the flexible α-helix 5. Only one water molecule could be modelled around the Cu(II) ion, but the positions of Q108 and Q90 suggest that their amide side-chain atoms may play a role in coordinating the other (unmodelled) waters in the coordination sphere.

The structure of Cu(II)-bound SCP_Q111BpyAla was determined to 2.51 Å resolution (Figure 3B and 3D) and revealed a different orientation of the bipyridine moiety compared to that seen in the Cu(II)-bound SCP_Q111CBpy structure. Despite the lower resolution, the electron density for the BpyAla bipyridine rings, α-helix 5 and the additional C-terminal amino acids (120-128) was very clear, indicative of a more rigid conformation for these parts of the protein. The co-planar BpyAla bipyridine rings are wedged into a shallow pocket formed by the turn between G85 and P89 at the edge of the apolar tunnel. This conformation is stabilised by van der Waals interactions between C1 of BpyAla and both G85 H_α (3.35 Å) and D88 H_α (3.36 Å), C2 of BpyAla and H_δ of P89 (2.87 Å: 3.36 Å to the C_δ) and C3 of BpyAla and H_γ of P89 (2.74 Å: 3.26 Å to the C_γ) (Figure 3D). In addition, the side-chains of D88 and K115 (one turn down on α-helix 5) lie on either face of the bipyridine and may stabilise the BpyAla conformation by aromatic-π electrostatic interactions.

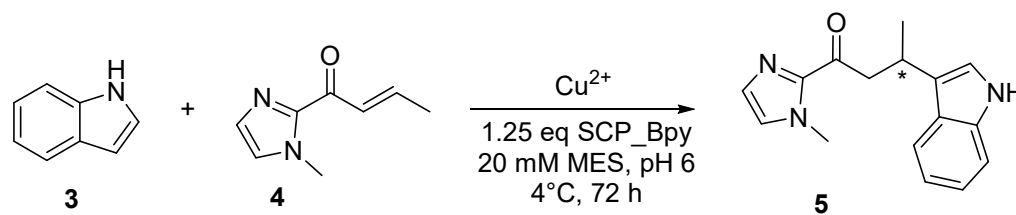
The 2F_o-F_c electron density map showed a single intense peak adjacent to the BpyAla bipyridine corresponding to a Cu(II) ion coordinated to N1 and N2 of BpyAla (Cu-N bond distance of 2 Å) and four water molecules in an octahedral arrangement (Figure S10 in ES1a). The B-factors for the Bpy moiety and the Cu(II) ion (43 Å² and 57.91 Å², respectively) were lower in this structure compared to the bipyridine rings in the Cu(II)-bound Q111CBpy structure, indicating that the catalytic environment is more rigid.

Taken together the structural analyses confirm the incorporation of Bpy at position 111 and the binding of Cu(II) to the Bpy moieties in both SCP ArMs. They also revealed marked differences in the environments of the Bpy and bound Cu(II) ions in the SCP_Q111BpyAla and SCP_Q111CBpy structures. Moreover, the Bpy moiety at position 111 is more rigid when incorporated as the unnatural amino acid BpyAla than via bioconjugation at a cysteine.

Catalytic screening

The catalytic activities of Bpy-containing SCP metalloenzymes were evaluated in the Cu-catalysed Friedel–Crafts alkylation of 5-methoxy-1H-indole **3** with 1-(1-methyl-1H-imidazol-2-yl)but-2-en-1-one **4**, resulting in the product **5**, the benchmark reaction for Cu(II) catalysis (Table 1)³². Enantioselective product formation was only observed in the presence of the Bpy-containing SCP-metalloenzymes.

Table 1: Copper catalysed enantioselective Friedel-Crafts alkylation using bipyridine containing SCP proteins.



Entry	Ligand/protein	yield of 5 (%) ^a	e.r. ^b (<i>R</i> : <i>S</i>)
1	none	76 (±6)	50:50
2	BpyAla(rac)	50 (±2)	50:50
3	wt SCP_2L	43 (±5)	50:50
4 ^c	SCP_A100CBpy	42 (±1)	66:34
5 ^c	SCP_V83CBpy	28 (±1)	52:48
6 ^d	SCP_Q111CBpy	25 (±2)	64:36
7	SCP_A100BpyAla	45 (±3)	51:49
8	SCP_V83BpyAla	45 (±8)	63:37
9	SCP_Q111BpyAla	42 (±7)	20:80
10	SCP_Q111CBpy + Triton	30 (±8)	57:43
11	SCP_Q111BpyAla + Triton	32 (±3)	23:77
12	SCP_A100BpyAla + Triton	21 (±3)	52:48
13	SCP_V83BpyAla + Triton	16 (±5)	57:43

Conditions: The reaction was carried out using 9 mol% of $\text{Cu}(\text{NO}_3)_2$ (90 μM), a small excess of 1.25 equivalents of SCP (112.5 μM to ensure that all $\text{Cu}(\text{II})$ ions are bound, 1 mM of **3** and 2.5 mM of **4** in MES buffer (20 mM, 150 mM NaCl, pH = 6.0) for 3 days at 4 °C giving **5**. All data are averages of experiments done in triplicate. Standard deviations are shown in brackets. a) Yield obtained by HPLC using 2-phenylquinoline as internal standard. b) e.r. determined using chiral HPLC. c) $\text{Cu}(\text{NO}_3)_2$ 72 μM . d) $\text{Cu}(\text{NO}_3)_2$ 63 μM . *S* and *R* were assigned using chiral HPLC³³.

SCP ArMs with BpyAla showed a higher yield (42-45%) compared to the equivalent Cys-coupled Bpy SCP ArMs. Among the novel ArMs, SCP_Q111BpyAla displayed the greatest enantioselectivity with an e.r. of 20:80 (±1), favouring the *S* enantiomer, and a yield of 42% (Table 1, entry 9). The BpyAla moiety in this SCP ArM is situated at the centre of one side of the tunnel, in the middle of α -helix 5 (Figure 1A and 3A). By contrast, the ArMs with BpyAla situated at either end of the tunnel showed either no enantioselectivity in the case of SCP_A100BpyAla (Table 1, entry 7), or lower enantioselectivity with a preference for the *R* enantiomer (Table 1, entry 8). We hypothesise that the lower selectivity at the entries to the tunnel could be due to increased flexibility and space leading to less defined binding pocket, compared to position Q111 in the centre of the tunnel. While enantioselectivity towards the *R* enantiomer was observed with all three Cys mutants, it was limited with the best result observed for SCP_A100CBpy which gave a 66:34 ratio (Table 1, entries 4-6).

Both enantiomers of **5** can be obtained in the Friedel-Crafts alkylation using SCP-ArMs with Bpy attached to residue 111, but by using different attachment strategies (Table 1, entries 6 and 9). The crystal structures of SCP_Q111BpyAla and SCP_Q111CBpy provide a structural indication for the observed differences in the enantioselectivities of these ArMs [almost 2x higher e.e., from 64:36 (*R*) for SCP_Q111CBpy to 20:80 (*S*) for SCP_Q111BpyAla]. The longer

linker to the protein backbone in SCP_Q111CBpy, which has an additional $-\text{CH}_2-\text{S}-$ compared to SCP_Q111BpyAla, allows it to extend out from α -helix 5 and adopt a more exposed position on the surface of the protein, compared to the BpyAla which is closer to the scaffold protein. In addition, the longer linker confers greater flexibility (as reflected in the B-factors), explaining the lower enantiomeric excess. Alignment of the Cu(II)-bound SCP_Q111BpyAla and SCP_Q111CBpy structures (Figures 3A and 3B) highlights the different positions and opposite orientations of the Bpy moieties, despite being incorporated at the same amino acid.

Computational modelling of metalloenzyme-catalysed Friedel-Crafts alkylation

In order to gain more information on how the ArMs control enantioselectivity, computational modelling was used to probe the stereochemical outcomes for the Cu-catalysed Friedel-Crafts reaction. The crystal structure coordinates were used for the computations. As these included Triton-X100, the ArM-catalysed Friedel-Crafts reactions were repeated in the presence of 2 eq. Triton-X100 to ascertain if this was likely to change the catalysis and thus if the crystal structures are a reasonable structural starting point. In the case of SCP_Q111BpyAla and SCP_Q111CBpy, the same enantioselectivities were observed as reactions without Triton-X100, with only a slight decrease in yield suggesting that the structures are a valid model to use to rationalise the differences in enantioselectivity. However, for SCP_A100BpyAla and SCP_V83BpyAla there was a significant decrease in activity, which could be caused by Triton-X100 blocking the availability of Cu-Bpy to the substrates as it binds within the protein tunnel (Table 1, entries 10-13) ⁷.

The first step of the reaction is the conjugate addition of indole **3** to enone **4** to form intermediate (**Int1**), via a transition state (**TS1**) (Scheme S1 ES1b). The second step is the product formation by the protonation reaction via a second transition state (**TS2**). To investigate the enantioselectivity of the Friedel-Craft alkylation, we looked to prior work using cluster models for metalloenzymes-catalyzed reactions ^{34,35}. The same QM-cluster model technique along with density functional theory (DFT) methodology was utilised to study our reaction. Three active site models were created for analysis (Figure S1 ES1b). Model A represents a reaction catalysed by Cu(II)-2,2'-bipyridine, in the absence of protein, whereas models B and C represent reactions catalysed by Cu(II)-bound SCP_Q111BpyAla and SCP_Q111CBpy, respectively. The computational analysis on Model A (Figure S11 ES1b) showed that the conjugate coupling step is both the rate-determining step as well as the enantioselective step, and the formation of the keto product is most likely. This agrees well with the experimental results (Table 2) as well as previous studies ³⁶, therefore only the first step was modelled going forwards. To create models B and C, the indole **3** and enone **4** substrates, along with Cu(II) ions, were incorporated into the crystal structures of Cu(II)-bound SCP_Q111BpyAla and SCP_Q111CBpy, respectively. The PDB files of the apo SCP_ArMs were prepared as described in the ES1b, and **3** and **4** were docked using AutoDock Vina ³⁷, with the lowest energy docked poses used for molecular dynamics (MD) simulations (see ES1b for detailed methods and results).

For SCP_Q111BpyAla catalysed reactions (Model B), the reactant complexes in two conformations: *proS* and *proR* were selected and their geometries optimized to give $\text{Re}_{\text{proS,B}}$ and $\text{Re}_{\text{proR,B}}$ as shown in Figure 4A and 4B. The $\text{Re}_{\text{proS,B}}$ structure is lower in energy by 5.4 kcal mol⁻¹ ($\Delta E + \text{ZPE}$) than $\text{Re}_{\text{proR,B}}$ and therefore $\text{Re}_{\text{proS,B}}$ represents the more favourable substrate binding orientation.

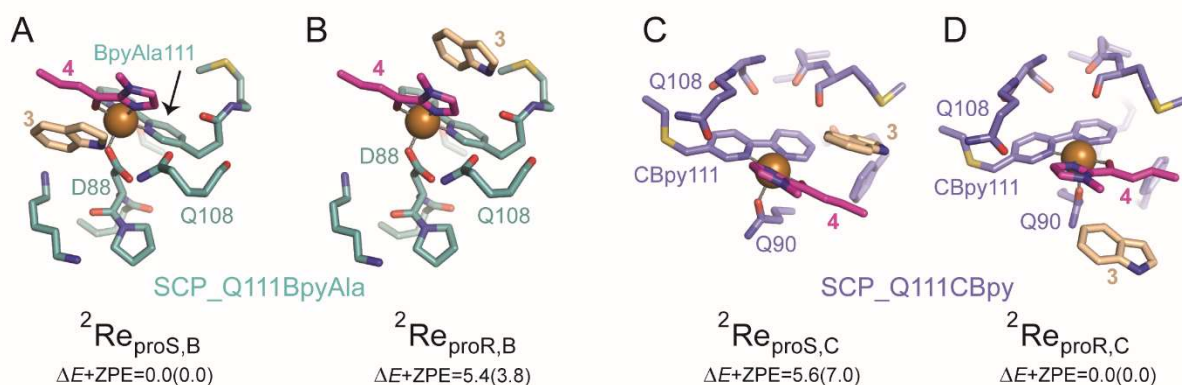


Figure 4: Optimized geometries of SCP_Q11BpyAla catalysed reactions (Model B): A) ${}^2\text{Re}_{\text{proS,B}}$ and B) ${}^2\text{Re}_{\text{proR,B}}$, and SCP_Q111CBpy catalysed reactions (Model C): C) ${}^2\text{Re}_{\text{proS,C}}$ and D) ${}^2\text{Re}_{\text{proR,C}}$. The *proS* and *proR* subscripts indicate the substrate-binding configuration and the superscript 2 represents the doublet spin state. Relative energies (in kcal mol⁻¹) are UB3LYP/BS2//UB3LYP/BS1 values with zero-point energy (ZPE) included. Orange spheres represent the Cu(II) ions.

Next, the Friedel-Craft (FC) alkylation reaction catalysed by SCP_Q111BpyAla, using $\text{Re}_{\text{proS,B}}$ and $\text{Re}_{\text{proR,B}}$ reactants as the starting structures, was calculated. Firstly, the conjugate addition of indole **3** to enone **4**, where substrate **3** configuration is either *proS* or *proR* was tested. The C-C coupling transition states ($\text{TS1}_{\text{proS,B}}$ and $\text{TS1}_{\text{proR,B}}$) of **3** with **4** leads to an intermediate for each configuration: $\text{Int1}_{\text{proS,B}}$ and $\text{Int1}_{\text{proR,B}}$ respectively. The calculated potential energy landscape for the conjugate coupling reaction for Model B is represented, along with the optimized geometric structures of these transition states, in Figure 5A. The energy barrier for $\text{TS1}_{\text{proS,B}}$ (10.1 kcal mol⁻¹) is lower than $\text{TS1}_{\text{proR,B}}$ (15.5 kcal mol⁻¹) (values relative to the stable configuration $\text{Re}_{\text{proS,B}}$). The transition states were characterized by an imaginary frequency $i308\text{ cm}^{-1}$ and $i319\text{ cm}^{-1}$ for $\text{TS1}_{\text{proS,B}}$ and $\text{TS1}_{\text{proR,B}}$ respectively. Furthermore, in transition state $\text{TS1}_{\text{proS,B}}$, indole **3** has strong water bridge interactions with D88 that make it more stable than the $\text{TS1}_{\text{proR,B}}$, which has a weak hydrogen bonding interaction with M112. Subsequently, the transition states are relaxed to their respective intermediates ($\text{Int1}_{\text{proS,B}}$ and $\text{Int1}_{\text{proR,B}}$), which are characterized by lower relative energy values: 1.65 and 7.92 kcal mol⁻¹, respectively. Thus, the *proS* configuration pathway is energetically favoured over the *proR* configuration pathway. Taken together these analyses explain the preferred enantioselectivity of SCP_Q111BpyAla (e.r. 14:86, *R*:*S*) that we observed experimentally (Table 2).

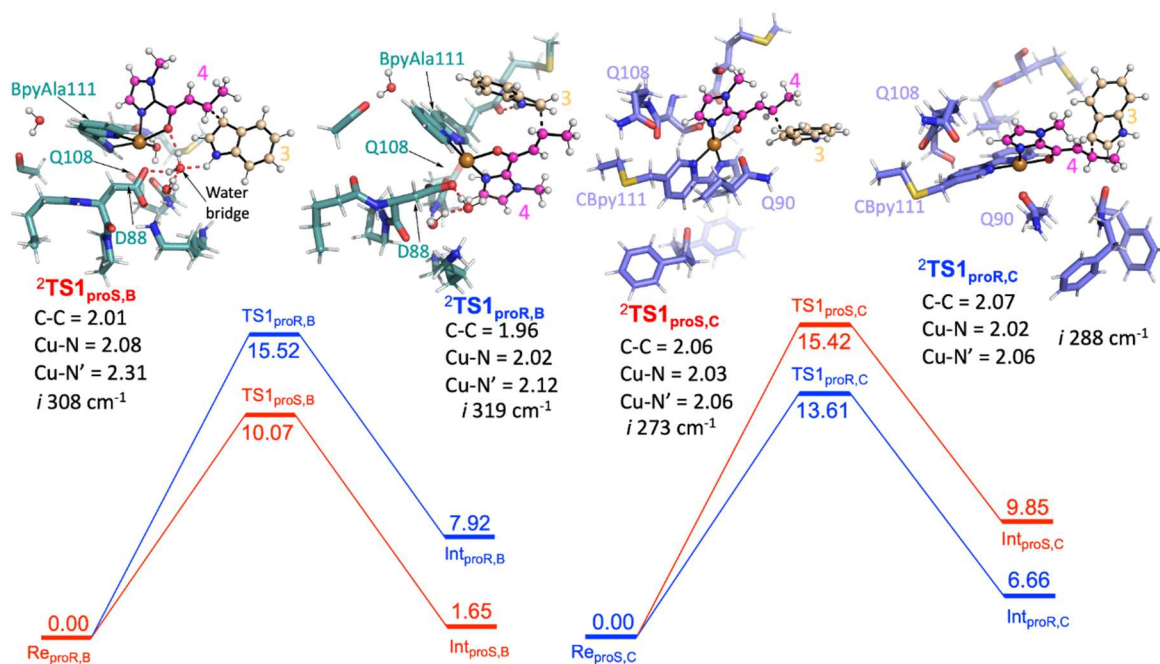


Figure 5: Potential energy landscape for the Friedel-Crafts alkylation by **A)** SCP_Q111BpyAla and **B)** SCP_Q111CBpy. Energy values are in kcal mol⁻¹. The optimized geometric structures of the transitions are shown on the left and right of free energy landscape with their important bond lengths in Å. The imaginary frequency represents the stretching vibration of C-C bonds between **3** and **4** for each transition state.

In a similar manner, the FC alkylation reaction mechanism catalysed by SCP_Q111CBpy (Model C) was explored. The reactant complexes in both *proS* and *proR* configurations were selected as the starting structures. The optimized geometries of these reactant complexes $\text{Re}_{\text{proS,C}}$ and $\text{Re}_{\text{proR,C}}$ are represented in Figures 4C and 4D. The free energy of the $\text{Re}_{\text{proS,C}}$ structure is higher than the $\text{Re}_{\text{proR,C}}$ by 5.6 kcal mol⁻¹ ($\Delta E + \text{ZPE}$), indicating that the reactant complex of model **C** in *proR* configuration shows the strongest substrate-bound pose and hence signifies the favoured reactant orientation.

Finally, the FC reaction pathway of Cu(II)-bound SCP_Q111CBpy was explored, using $\text{Re}_{\text{proS,C}}$ and $\text{Re}_{\text{proR,C}}$ reactants as the starting structures. The calculated potential energy landscape for the conjugate coupling reaction for model **C** is represented, along with the optimized geometric structures of these transition states, in Figure 5B. The lowest energy barrier (13.6 kcal mol⁻¹) was obtained for the *proR* C-C coupling transition state ($\text{TS1}_{\text{proR,C}}$) while the energy barrier for the *proS* transition state ($\text{TS1}_{\text{proS,C}}$) was higher (15.4 kcal mol⁻¹). The transition states are characterized by the presence of an imaginary frequencies $i288 \text{ cm}^{-1}$ and $i273 \text{ cm}^{-1}$ for $\text{TS1}_{\text{proR,C}}$ and $\text{TS1}_{\text{proS,C}}$ respectively. Furthermore, in the $\text{TS1}_{\text{proR,C}}$ transition state a hydrophobic interaction from M105 was observed with **3**, while no such interaction of **3** was observed in $\text{TS1}_{\text{proS,C}}$. The system then relaxed to the low energy state intermediates $\text{Int1}_{\text{proR,C}}$ and $\text{Int1}_{\text{proS,C}}$, which are characterized in the local minima state by the presence of all real frequency values indicating stable structures. The $\text{Int1}_{\text{proR,C}}$ has lower energy (6.7 kcal mol⁻¹) than the $\text{Int1}_{\text{proS,C}}$ configuration (9.8 kcal mol⁻¹), which favours the *proR* configuration pathway

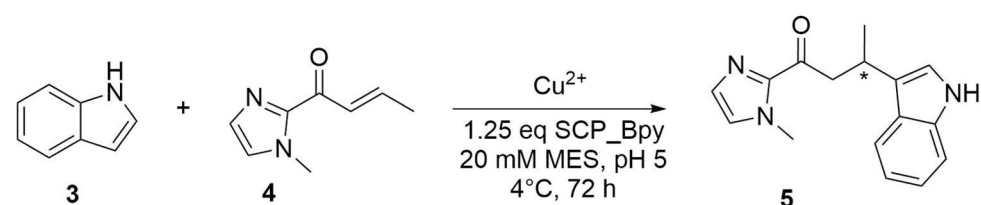
over the *proS* pathway. This is consistent with the experimentally observed preference to produce the *R* enantiomer by SCP_Q111CBpy (Table 1).

Overall, the substrates' activation in FC alkylation mechanism with SCP_Q111BpyAla favours the formation of *proS* product over *proR* while with SCP_Q111CBpy favours *proR* product formation over *proS*. However, the products distribution was achieved via a competitive pathway, so a mixture of both enantiomers of the product is predicted, which agrees well with our experimental results. The larger difference in the relative energies for the SCP_Q111BpyAla TS intermediates matches the experimental observation of improved enantioselectivity when using SCP_Q111BpyAla as the catalyst over the use of SCP_Q111CBpy

Engineering SCP_Q111BpyAla

Taking the most promising ArM (SCP_Q111BpyAla) forwards, we chose to use a structure-based alanine scanning approach to see if any of the residues identified from either the computational or structural work impacted activity or selectivity³⁸. Each amino acid was substituted individually with alanine via site-directed mutagenesis, and the yields and enantioselectivities in Friedel-Crafts reactions catalysed by these mutant SCP ArMs were determined (Table 2). These reactions were carried out at pH 5, as slightly improved enantiomeric ratios were observed compared to those performed at pH 6 (Table 2 entry 1 vs. Table 1 entry 9) matching previous observations by the Roelfes group³⁹.

Table 2: Enantioselectivity in Friedel-Crafts reactions catalysed by SCP_Q111BpyAla mutant ArMs.



Entry	Mutant of SCP_Q111BpyAla	Yield of 5 ^a (%)	e.r. ^b (R:S)
1	none	43 (±2)	14:86
2	Q108A	38 (±7)	28:72
3	V82A	33 (±2)	24:76
4	F34A	21 (±1)	31:69
5	M112A	33 (±4)	19:81
6	K115A	41 (±5)	27:73
7	D88A	54 (±10)	18:82

Conditions: Reactions were carried out using 9 mol% of Cu(NO₃)₂ (90 μM), a small excess of 1.25 equivalents of SCP (112.5 μM to ensure that all Cu(II) ions are bound, 1 mM of **3** and 2.5 mM of **4** in MES buffer (20 mM, 150 mM NaCl, pH = 5.0) for 3 days at 4 °C giving **5**. All data are averages of experiments done in triplicate. Standard deviations are shown in brackets. a) Yield obtained by HPLC using 2-phenylquinoline as internal standard; standard deviations are shown in brackets b) e.r. determined using chiral HPLC.

None of the mutations screened showed substantial variations (i.e. a complete drop in reactivity or selectivity). Indeed, the mutation with the biggest difference was F34A which was

chosen due to its location nearby and within the protein hydrophobic pocket (Table 2, entry 4). SCP_Q111BpyAla F34A exhibited substantial precipitation suggesting that disruption to the protein core reduced stability, leading to the observed low activity and selectivity. A similar rationale was also used for the choice of V82A which gave lower activity and selectivity but not to the same extent (Table 2, entry 3 and 4). The crystal structure suggested that K115 and D88 may play a role in stabilising the bipyridine position within the protein through aromatic- π electrostatic interactions. Mutating D88 to alanine gave no meaningful change in enantioselectivity (Table 2, entry 7), whilst K115A showed a reduction in enantioselectivity (Table 2, entry 6). The more flexible nature of the lysine side chain means it is less clear if this can be attributed to structural changes as opposed to some role with substrate binding. The computational work revealed no interaction of the substrates with K115. In contrast, a close contact between the side chain of D88 and the copper atom was observed in the modelled transition states and D88 was shown to participate in water bonding networks with the substrates. Both these interactions would be disrupted on mutation of D88 to alanine, suggesting that either these interactions are of minor importance in catalysis or alternative residues such as Q108 could replace the hydrogen bonding interactions with water. Whilst Q108A was shown to reduce the enantioselectivity obtained with SCP-Q111BpyAla experimentally (Table 2, entry 2), no clear role was observed in the computational work.

The only amino acid shown to make direct contact with the substrate during the course of the reaction was M112: the methyl makes a weak hydrophobic interaction with **3** during substrate binding in the *proS* pathway (Figure 6), whilst the amide backbone of M112 makes a weak hydrogen bond with **3** in the *proR* transition state (Figure 5A). Mutating M112 to alanine gave a small reduction in enantioselectivity (Table 2, entry 5). Whilst alanine can also make hydrophobic interactions, its reduced chain length would preclude interactions with **3** in this instance and thus M112 could be helping to stabilise the substrates to a small extent. No difference would be expected for the *proR* pathway as alanine can still participate in amide backbone bonding.

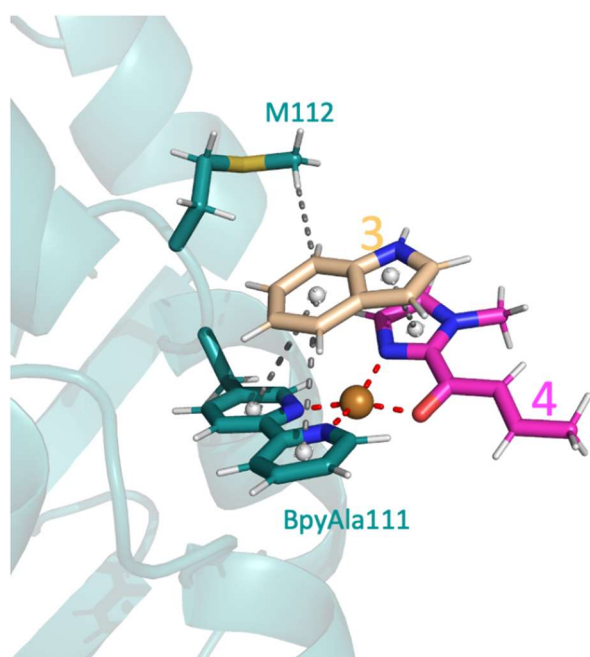


Figure 6: Docking pose of **3** and **4** with the SCP_Q111BpyAla ArM, highlighting the key interactions of **3** including one hydrophobic interaction with M112 and three π - π interactions, two with BpyAla111 and one with **4** (represent by grey spheres).

The lack of side chain interactions with the substrates and no clear results from the alanine scanning suggests that a rational approach to designing the active site via mutagenesis may not lead to an improved enzyme. Indeed, the exposed nature of the active site on the side of the protein suggests that extensive backbone engineering to build up the protein bulk around the substrates may lead to more promising candidates for enantioselective catalysis.

Conclusions

Using two different approaches to forming artificial metalloenzymes, a library of successful catalysts for the Cu-catalysed Friedel Crafts alkylation were obtained. Moderate activities and good enantiomeric ratios were observed with SCP_Q111BpyAla giving the highest e.r. of 14:86 in favour of the *S* enantiomer. In general, lower enantioselectivities were observed for the cysteine-modified catalysts and this is attributed to increased flexibility around the active site due to the additional S-CH₂ linker. Changing from genetic incorporation of BpyAla at Q111 to bioconjugation of bipyridine through a cysteine, resulted in an unexpected change in the major isomer observed from *S* to *R*, albeit with a reduction in enantiomeric excess. This result is intriguing as a major challenge in enzyme catalysis is accessing the opposite product enantiomer. Crystallographic and computational studies confirm that changes in the structure of the active site led to the stabilisation of indole **3** on different sides of the bound enone **4**, with the enantiomeric ratios observed correlating with the difference in the calculated transition state energies for the pro*R* and pro*S* SCP_Q111BpyAla and SCP_Q111CBpy Cu catalyst models. Based on these findings, future work should prioritize determining the most suitable modification method for the desired application of the ArM before proceeding with reaction screening and optimization, as subtle differences in protein structure can substantially change the stabilised transition state.

The crystal structure obtained of SCP-Q111BpyAla_Cu is a rare example of a protein structure containing an unnatural amino acid. Our work shows that increasing our understanding of artificial metalloenzyme structures facilitates their development as catalysts, and moving forwards will be vital to enable future *de novo* design of metalloproteins.

Acknowledgements

AGJ, RCB and HSA acknowledge the support of the UKRI through the award of a UKRI FLF fellowship to AGJ (MR/S017402/1). In addition, this work was supported by the Wellcome Trust [204804/Z/16/Z] and The Carnegie Trust for the Universities of Scotland. EK was funded through an EaSiCAT studentship. We would like to thank Dr Antonia Mey for computational discussions; Diamond Light Source for beamtime (proposals mx-1815 and mx-25233) and the staff of beamlines I03, I04, I04-1 and I24; SIRCAMS for mass spectrometry support. We thank Prof. Peter Schultz of The Scripps Research Institute for the gift of the pEVOL plasmid. This work has made use of the resources provided by the Edinburgh Compute and Data Facility (ECDF) (<http://www.ecdf.ed.ac.uk/>). RCB is currently an employee of AstraZeneca.

Author Contributions

AGJ conceived, supervised, and obtained funding for the project. All authors contributed to data analysis and preparing the manuscript. JR obtained and analysed the X-ray Crystal Diffraction data, and uploaded the data to the PDB. EV and RB designed the experiments and collected the data for the protein production, modification, and catalysis. HA carried out the computational aspects of the project.

Supporting information

ESIa: Additional experimental details, materials, methods and data.

ESIb: Additional information on the computational studies.

References

- 1 S. Hanreich, E. Bonandi and I. Drienovská, Design of Artificial Enzymes: Insights into Protein Scaffolds, *ChemBioChem*, 2022, e202200566.
- 2 F. Schwizer, Y. Okamoto, T. Heinisch, Y. Gu, M. M. Pellizzoni, V. Lebrun, R. Reuter, V. Köhler, J. C. Lewis and T. R. Ward, Artificial Metalloenzymes: Reaction Scope and Optimization Strategies. *Chem. Rev.*, 2018, **118**, 142–231.
- 3 H. J. Davis and T. R. Ward, Artificial Metalloenzymes: Challenges and Opportunities, *ACS Cent. Sci.*, 2019, **5**, 1120–1136.
- 4 A. G. Jarvis, Designer Metalloenzymes for Synthetic Biology: Enzyme Hybrids for Catalysis. *Curr. Opin. Chem. Biol.*, 2020, **58**, 63–71.
- 5 M. Dürrenberger, T. Heinisch, Y. M. Wilson, T. Rossel, E. Nogueira, L. Knörr, A. Mutschler, K. Kersten, M. J. Zimbron, J. Pierron, T. Schirmer and T. R. Ward, Artificial Transfer Hydrogenases for the Enantioselective Reduction of Cyclic Imines, *Angew. Chemie Int. Ed.*, 2011, **50**, 3026–3029.
- 6 A. G. Jarvis, L. Obrecht, P. J. Deuss, W. Laan, E. K. Gibson, P. P. Wells and P. C. J.

- Kamer, Enzyme Activity by Design: An Artificial Rhodium Hydroformylase for Linear Aldehydes, *Angew. Chem.*, 2017, **129**, 13784–13788.
- 7 M. V. Doble, L. Obrecht, H.-J. Joosten, M. Lee, H. J. Rozeboom, E. Branigan, J. H. Naismith, D. B. Janssen, A. G. Jarvis and P. C. J. Kamer, Engineering Thermostability in Artificial Metalloenzymes to Increase Catalytic Activity, *ACS Catal.*, **11**, 2021, 3620–3627.
 - 8 M. Jeschek, R. Reuter, T. Heinisch, C. Trindler, J. Klehr, S. Panke and T. R. Ward, Directed Evolution of artificial metalloenzymes for in vivo metathesis, *Nature*, 2016, **537**, 661–665.
 - 9 M. V. Doble, A. G. Jarvis, A. C. C. Ward, J. D. Colburn, J. P. Götze, M. Bühl and P. C. J. Kamer, Artificial Metalloenzymes as Catalysts for Oxidative Lignin Degradation, *ACS Sustain. Chem. Eng.*, 2018, **6**, 15100–15107.
 - 10 R. B. Leveson-Gower and G. Roelfes, Biocatalytic Friedel-Crafts Reactions, *ChemCatChem*, 2022, **14**, e202200636.
 - 11 P. J. Deuss, M. V. Doble, A. G. Jarvis and P. C. J. Kamer, in *Artificial Metalloenzymes and MetalloDNAzymes in Catalysis*, Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, Germany, 2018, pp. 285–319.
 - 12 C. Van Stappen, Y. Deng, Y. Liu, H. Heidari, J.-X. Wang, Y. Zhou, A. P. Ledray and Y. Lu, Designing Artificial Metalloenzymes by Tuning of the Environment beyond the Primary Coordination Sphere, *Chem. Rev.*, 2022, **122**, 11974–12045.
 - 13 B. J. Bloomer, D. S. Clark and J. F. Hartwig, Progress, Challenges, and Opportunities with Artificial Metalloenzymes in Biosynthesis, *Biochemistry*, 2023, **62**, 221-228.
 - 14 A. Pordea, Metal-binding promiscuity in artificial metalloenzyme design, *Curr. Opin. Chem. Biol.*, 2015, **25**, 124–132.
 - 15 S. C. Hammer, A. M. Knight and F. H. Arnold, Design and evolution of enzymes for non-natural chemistry, *Curr. Opin. Green Sustain. Chem.*, 2017, **7**, 23–30.
 - 16 X. Ren and R. Fasan, Engineered and artificial metalloenzymes for selective C–H functionalization, *Curr. Opin. Green Sustain. Chem.*, 2021, **31**, 100494.
 - 17 J. Klehr, J. Zhao, A. S. Kron, T. R. Ward and V. Köhler, Streptavidin (Sav)-Based Artificial Metalloenzymes: Cofactor Design Considerations and Large-Scale Expression of Host Protein Variants, 2020, 213–235.
 - 18 N. Stephanopoulos and M. B. Francis, Choosing an effective protein bioconjugation strategy, *Nat. Chem. Biol.*, 2011, **7**, 876–884.
 - 19 S.-M. Jung, M. Yang and W. J. Song, Symmetry-Adapted Synthesis of Dicopper Oxidases with Divergent Dioxygen Reactivity, *Inorg. Chem.*, 2022, **61**, 12433–12441.
 - 20 I. Drienovská, A. Rioz-Martínez, A. Draksharapu and G. Roelfes, Novel artificial metalloenzymes by in vivo incorporation of metal-binding unnatural amino acids, *Chem. Sci.*, 2015, **6**, 770–776.
 - 21 G. Roelfes, LmrR: A Privileged Scaffold for Artificial Metalloenzymes, *Acc. Chem. Res.*,

- 2019, **52**, 545–556.
- 22 K. R. Miller, S. Biswas, A. Jasiewicz, A. H. Follmer, A. Biswas, T. Albert, S. Sabuncu, E. L. Bominaar, M. P. Hendrich, P. Moënne-Loccoz and A. S. Borovik, Artificial Metalloproteins with Dinuclear Iron-Hydroxido Centers, *J. Am. Chem. Soc.*, 2021, **143**, 2384–2393.
- 23 A. M. Haapalainen, D. M. F. Van Aalten, G. Meriläinen, J. E. Jalonen, P. Piriälä, R. K. Wierenga, J. K. Hiltunen and T. Glumoff, Crystal structure of the liganded SCP-2-like domain of human peroxisomal multifunctional enzyme type 2 at 1.75 Å resolution, *J. Mol. Biol.*, 2001, **313**, 1127–1138.
- 24 P. J. Deuss, G. Popa, A. M. Z. Slawin, W. Laan and P. C. J. Kamer, Artificial Copper Enzymes for Asymmetric Diels–Alder Reactions, *ChemCatChem*, 2013, **5**, 1184–1191.
- 25 T. Kuckhoff, R. C. Brewster, J. Calum, T. J. Ferguson and A. G. Jarvis, Reactivity Tuning of Metal-Free Artificial Photoenzymes through Binding Site Specific Bioconjugation, *European J. Org. Chem.*, 2023, e202201412.
- 26 R. R. Davies, H. Kuang, D. Qi, A. Mazhary, E. Mayaan and M. D. Distefano, Artificial metalloenzymes based on protein cavities: Exploring the effect of altering the metal ligand attachment position by site directed mutagenesis, *Bioorg. Med. Chem. Lett.*, 1999, **9**, 79–84.
- 27 H. Osseili, D. F. Sauer, K. Beckerle, M. Arlt, T. Himiyama, T. Polen, A. Onoda, U. Schwaneberg, T. Hayashi and J. Okuda, Artificial Diels–Alderase based on the transmembrane protein FhuA, *Beilstein J. Org. Chem.*, 2016, **12**, 1314–1321.
- 28 S. Kim, J. Yun, H. Yoo, S. Kim, H. M. Kim and H. S. Lee, Metal-Mediated Protein Assembly Using a Genetically Incorporated Metal-Chelating Amino Acid, *Biomacromolecules*, 2020, **21**, 5021–5028.
- 29 C. Kaes, A. Katz and M. W. Hosseini, Bipyridine: The most widely used ligand. A review of molecules comprising at least two 2,2'-bipyridine units, *Chem. Rev.*, 2000, **100**, 3553–3590.
- 30 P. J. Deuss, G. Popa, C. H. Botting, W. Laan and 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.
- 31 J. Xie, W. Liu and P. G. Schultz, A Genetically Encoded Bidentate, Metal-Binding Amino Acid, *Angew. Chemie Int. Ed.*, 2007, **46**, 9239–9242.
- 32 V. Kumar, W. B. Turnbull and A. Kumar, Review on Recent Developments in Biocatalysts for Friedel-Crafts Reactions, *ACS Catal.*, 2022, **12**, 10742–10763.
- 33 D. A. Evans, K. R. Fandrick and H. J. Song, Enantioselective Friedel-Crafts alkylations of α,β -unsaturated 2-acyl imidazoles catalyzed by bis(oxazoliny)pyridine-scandium(III) triflate complexes, *J. Am. Chem. Soc.*, 2005, **127**, 8942–8943.
- 34 H. S. Ali, R. H. Henchman and S. P. de Visser, Lignin Biodegradation by a Cytochrome P450 Enzyme: A Computational Study into Syringol Activation by GcoA, *Chem. – A Eur.*

- J.*, 2020, **26**, 13093–13102.
- 35 S. Louka, S. M. Barry, D. J. Heyes, M. Q. E. Mubarak, H. S. Ali, L. M. Alkhalaf, A. W. Munro, N. S. Scrutton, G. L. Challis and S. P. De Visser, Catalytic Mechanism of Aromatic Nitration by Cytochrome P450 TxtE: Involvement of a Ferric-Peroxynitrite Intermediate, *J. Am. Chem. Soc.*, 2020, **142**, 15764–15779.
- 36 L. Villarino, S. Chordia, L. Alonso-Cotchico, E. Reddem, Z. Zhou, A. M. W. H. Thunnissen, J. D. Maréchal and G. Roelfes, Cofactor Binding Dynamics Influence the Catalytic Activity and Selectivity of an Artificial Metalloenzyme, *ACS Catal.*, 2020, **10**, 11783–11790.
- 37 O. Trott and A. J. Olson, AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization and multithreading, *J. Comput. Chem.*, 2010, **31**, 455.
- 38 S. M. Simonsen, L. Sando, K. J. Rosengren, C. K. Wang, M. L. Colgrave, N. L. Daly and D. J. Craik, Alanine Scanning Mutagenesis of the Prototypic Cyclotide Reveals a Cluster of Residues Essential for Bioactivity, *J. Biol. Chem.*, 2008, **283**, 9805–9813.
- 39 C. G. de Souza, L. Alonso-Cotchico, M. Bersellini and G. Roelfes, Unexpected Catalytic Activity of the Regulatory Protein QacR, doi:10.26434/chemrxiv.12662003.v1.