JOHNSON MATTHEY TECHNOLOGY REVIEW

www.technology.matthey.com

Johnson Matthey's international journal of research exploring science and technology in industrial applications

**********Accepted Manuscript**********

This article is an accepted manuscript

It has been peer reviewed and accepted for publication but has not yet been copyedited, house styled, proofread or typeset. The final published version may contain differences as a result of the above procedures

It will be published in the **OCTOBER 2020** issue of the *Johnson Matthey Technology Review*

Please visit the website https://www.technology.matthey.com/ for Open Access to the article and the full issue once published

Editorial team

Johnson Matthey Technology Review
Johnson Matthey Plc
Orchard Road
Royston
SG8 5HE
UK
Tel +44 (0)1763 253 000
Email tech.review@matthey.com

Unlocking the Full Evolutionary Potential of Artificial Metalloenzymes Through Direct Metal-Protein Coordination

George S. Biggs,† Oskar James Klein,† Sally. R. Boss,* Paul D. Barker.*

Chemistry Department, University of Cambridge, Lensfield Road, Cambridge CB2 1EW

*srb39@cam.ac.uk, *pdb30@cam.ac.uk

Abstract

Generation of artificial metalloenzymes has gained much inspiration from the general understanding of natural metalloenzymes. Over the last decade, a multitude of methods generating transition metal-protein hybrids have been developed and many of these new-to-nature constructs catalyse reactions previously reserved for the realm of synthetic chemistry. This perspective will focus on artificial metalloenzymes incorporating 4d and 5d transition metals. It aims to summarise the significant advances made to date and asks whether there are chemical strategies, used in nature to optimise metal catalysts, that have yet to be fully recognised in the synthetic enzyme world, particularly whether artificial enzymes produced to date fully take advantage of the structural and energetic context provided by the protein.

Further, the argument is put forward that, based on precedence, in the majority of naturally evolved metalloenzymes the direct coordination bonding between the metal and the protein scaffold is integral to catalysis. Therefore, the protein can attenuate metal activity by positioning ligand atoms in the form of amino acids, as well as making non-covalent contributions to catalysis, through intermolecular interactions that pre-organise substrates and stabilise transition states. This highlights the often neglected but crucial element of natural systems that is the energetic contribution towards activating metal centres through protein fold energy. Finally, general principles needed for a different approach to the formation of artificial metalloenzymes are set out, utilizing direct coordination inspired by the activation of an organometallic cofactor upon protein binding. This methodology, observed in nature, delivers true interdependence between metal and protein. When combined with the ability to efficiently evolve enzymes, new problems in catalysis could be addressed in a faster and more specific manner than with simpler small molecule catalysts.

Doi: 10.1595/205651320X15928204097766

Introduction

Metalloenzymes have been prominent in the field of enzyme engineering since its emergence some 40 years ago, at the birth of protein/enzyme engineering.(1, 2) Metal ions or co-factors in solution have an intrinsic chemistry that can be catalytic and these are accessible to detailed mechanistic study. These properties mean that co-localization of substrate and metal within a peptidic scaffold can be sufficient in forming an artificial metalloenzyme (ArM), without further influence from the protein on the catalytic mechanism. With the advent of modern protein engineering and design technologies, ArMs were developed by incorporating metal binding sites in or adjacent to hydrophobic pockets. While the resulting ArMs were active, they often displayed low efficiency and specificity. Therefore, directed evolution (i.e. iterative rounds of mutagenesis and selection for activity, Figure 1) has become a key step in creating enzymes with new and useful properties. The choice of starting point for such a forced evolution campaign, in this case the metal-protein complex formed initially, is of great importance. Since any particular enzyme follows a unique evolutionary trajectory as new mutations move it along the fitness landscape towards (potentially local) maxima, choice of the starting point may directly predetermine the result. By nature of the selection process, it is further possible, that trajectories leading to the global maximum fall beneath the cut-off limit for further evolution, becoming inaccessible. For instance, a mutation introduced in the first round of mutagenesis may lead to a destabilisation of the protein at assay conditions, causing that initial variant to be discarded through selection. However, a compensating mutation to that variant in a subsequent round of mutagenesis could result in an enzyme which is stable, active and closer to a global fitness maximum. Finally, not every method of generating ArMs may be compatible with current methods for directed evolution and therefore limit the extent of evolution that can be achieved.

Doi: 10.1595/205651320X15928204097766

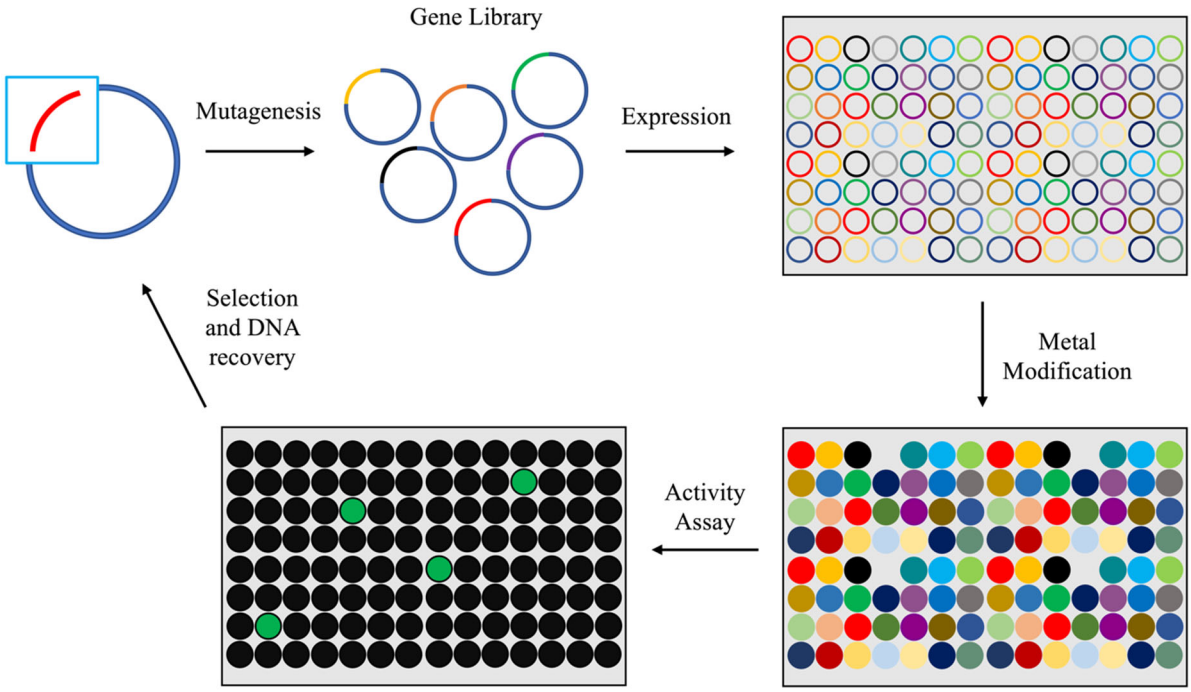


Figure 1 | The general overview of a directed evolution campaign for artificial metalloenzymes. | The Darwinian algorithm can be reproduced in the laboratory, greatly increasing the speed of evolution. Mutagenesis methods introduce mutations with various levels of randomness, depending on the method used, to the starting point gene, forming a gene library. This library can then be expressed in a manner that couples expression products and genetic sequence information to yield the different proteins. Upon addition of the metal cofactor, the artificial metalloenzymes are formed and can be selected for improved variants in regard to desired parameters (reaction rates, yield, stereoselectivity, stability etc.). The metal modification step must itself clearly be efficient and high yielding to avoid limiting the library size at that stage. The sequence information of the improved candidates is recovered and can be subjected to further rounds of directed evolution.

In this perspective, different routes towards artificial metalloenzymes are considered in the context of the starting protein scaffold as well as the type of catalytic centre and reactions involved. Advances in artificial metalloenzymes have recently been reviewed and the reader is referred to these for further details of the strategies used to find new systems.(3–5) This article aims to provide an overview of the strengths and weakness of these different approaches and to provide a perspective of some challenges that remain.

Why Do We Want New Artificial Metalloenzymes?

One particular area that will greatly impact chemical production on this planet is synthetic biology. Replacing synthetic catalysts, acting on petrochemical feedstocks in non-aqueous solvents, with biocatalytic systems working in water with simple carbon neutral feedstocks (CO_2 even?) is clearly highly desirable. But why engineer new enzymes, particularly using expensive and relatively scarce transition metals, when the ability to find new catalysts amongst gene products from all corners of the biological world has developed at staggering pace?(6–8) As a consequence of the latter, any target chemical can conceivably be obtained by

Doi: 10.1595/205651320X15928204097766

recombining pre-existing metabolic pathways.(9) What will new and unnatural metalloenzymes provide?

One clear feature is orthogonality – the objective of introducing functionality into a cell that has no counterpart in the natural world could provide chemistry that biology cannot currently catalyse – alkene metathesis for example. As there is a limit to the number of additional transformations a viable cell will perform, these orthogonal reactions may allow access to much shorter, and therefore more efficient, pathways. If not for a synthetic purpose, one could also imagine orthogonal catalytic chemistry providing a diagnostic or reporter output without interference from the host endogenous processes. For it to be truly orthogonal, it is difficult to imagine evolving a new enzyme based around metals already abundant in nature and already used as catalysts in biology. The transition metals used by nature are very carefully controlled by acquisition and regulatory networks that ensure catalytic metal ions are not free to operate outside the endogenous metabolism. Therefore, there is significant advantage in trying to introduce metals that biology currently has no evolved means of metabolising. This work therefore focuses primarily on non-biological transition metal co-factors as a route to introducing novel orthogonal activity into a biologically viable system.

Evolutionary Routes to Optimised Artificial Metalloenzymes

Natural evolution has provided numerous examples of metal ions used by enzymes for a plethora of different catalytic purposes. Rigorous mechanistic and structural biochemistry has advanced understanding of the mechanistic detail of metalloenzyme activity significantly, to the point that a few underpinning principles can be identified, linking protein structure and thermodynamics to catalytic activity of metal centres. Together with the knowledge garnered from extensive research on transition metal catalysts, it is possible to establish key properties desirable for novel ArMs.

Considerations on Protein-Substrate Interactions

As mentioned above, the ability of enzymes to organise reactants cooperatively can in itself give rise to enhanced activity over background rates in solution and in highly evolved systems this may even be the greatest factor driving increased reaction rates. It is important to realise that while metal-substrate proximity may be enough to confer reactivity, directional metal-substrate orbital overlap also plays a crucial role in activating the substrate to react. Indeed, it is via the formation of metal-ligand/substrate molecular orbitals that the substrate chemistry is attenuated by the presence of the metal and that catalytic reactivity can be achieved. Significant

Doi: 10.1595/205651320X15928204097766

computational advancements have been made in the *in-silico* design of catalytic metal binding sites(10, 11) and the mechanistic understanding of reported ArMs.(12–15) However, given the lack of reliable parameters for defining transition metal bonding, and the immense complexity of the many low energy interactions that determine the coupling of protein folding to the binding of small molecules, it is beyond current computational capabilities to predict what primary sequence and cofactors are necessary to achieve the optimal arrangement for metal catalysis. It therefore becomes important to have a malleable, promiscuous starting system that can be used to sample a large space of different structures.(16) Hence, while choosing proteins with well-defined properties and unique structures has some advantages from a design point of view, starting points that do not fold into one specific structure may be desirable, since they are not as closely constrained by any one particular energy well. For similar reasons, in choosing a particular chemical strategy for introducing a metal cofactor into the protein, it becomes essential to use a method that allows for high throughput selection or screening.(17)

Considerations on Metal Chemistry in Proteins

In addition to sampling sequence space to optimize the geometrical factor, protein evolution offers the unique possibility of sampling transition metal chemistry by poising the metal in energized states. In small molecule transition metal catalysis, ligands will arrange around the metal centre to maximise bonding interactions and reach a thermodynamic minimum. In order to maintain the ligand exchange necessary for catalysis, some ligands tend to be weakly bonding, with the presence of strongly bonding ligands (for instance water or hydroxide) being a major factor in catalyst poisoning. In enzymes however, the intramolecular bonds generated within the whole protein scaffold can be used to place and maintain coordinating atoms from amino acids. These interactions can be seen as the second coordination sphere, shaping the metal complex and potentially leaving the first sphere ligand atoms in a suboptimal configuration around the metal centre so that the energy of the resulting complex is not at a minimum on the coordination energy landscape. The stabilisation of this complex is made possible by the favourable intramolecular peptidic interactions (i.e. protein fold energy) offsetting the steric and or/electronic distortion of the optimum geometry.(18) These energized, or entatic, states have a reactivity that is not easily realized in conventional, synthetic metal catalysts, if it is possible at all.(19) This effect is most easily visualised by considering the common biological process of activation of inert co-factors by alteration of coordination upon binding to their respective apoenzymes. For instance, on their own the cobalt metallo-organic cofactor, vitamin B12, and methionine synthase are catalytically inert; upon protein-

Doi: 10.1595/205651320X15928204097766

cofactor binding and coordination of the cobalt centre to a specific histidine, methyl transfer activity is unmasked with great control and substrate specificity.(20–22) Applying this principle, it can be envisioned that even with the limited donor atoms available to proteins, a vast number of different complexes with different chemistries can be accessed, because the exact positioning as well as characteristics of the ligands dictate metal properties such as electron density, redox potential, Lewis acidity and ligand exchange rates. Further, the metal cofactor does not need to be a bare metal ion but could be incorporated with other ligands already attached. Interaction between these ligands (for instance π-π stacking with an arene ligand) and the protein can be relayed to the metal centre and allow for an even finer tuning of the metal centre. Again, current possibilities for design are insufficient to predict these effects which can be very subtle, highlighting the need for biochemical high throughput screening methods.

The Optimal Method of ArM Formation

The above considerations define a range of requirements for potential methods of forming ArMs. Primarily, there needs to be a direct connection between the protein scaffold and the metal ion in the form of at least one coordination bond, not only for localisation but also for poising the metal reactivity. As will be detailed below, most of the successful methods of generating ArMs published to date are efficient but rely on fully saturated, catalytically active co-factors such as commercial transition metal catalysts decorated with a linker moiety. These cannot make use of the protein fold energy to optimize the chemical process of catalysis, a potential factor in why directed evolution campaigns of ArMs have been of limited success. Whereas improvements in enantioselectivity and turnover number have been reported, which can be traced to substrate binding and the hydrophobic micro-environment respectively,(23–26) significant increases in the chemical turnover rate (in many systems characterised by the initial k_{cat}) from the free cofactor to the formed ArM have so far been limited. Small changes in k_{cat} can be explained by organisational effects and indirect interactions with the substrate orbitals, such as charge compensation. As demonstrated by Hilvert *et al*., significant increases in k_{cat} have been shown to be possible by fine tuning the actual centre of reaction, which is the first coordination sphere of the metal complex.(27) From the perspective of the protein scaffold, the formation of an entatic state requires the peptide to be at least partially folded before binding the metal, the more defined the fold, the greater the ability of the fold to energise the metal complex. This is in contrast to the desirable dynamic system for the evolutionary process. A potential compromise can be struck by using a starting scaffold that is partially

Doi: 10.1595/205651320X15928204097766

folded as the apoprotein and upon cofactor binding rigidifies to a completely folded form. The initial folding energy can be used to poise the metal in an activated state, while the folding process occurring during cofactor binding allows for the system to adapt during directed evolution. Once the ArM becomes more specialised after rounds of evolution, the apoprotein will probably approach a more fully folded form, yielding an ArM after cofactor addition that is less promiscuous but contains a more energized and active metal centre.

To summarize, the number of different complex chemical factors required of ArMs demand the use of directed evolution in order to form enzymes with industrially and medically relevant properties. In order to ensure a high level of engineerability, an optimal methodology for combining 4d and 5d metals starts with a highly promiscuous and malleable holoprotein that further has dative bonds between the metal ion and the peptidic moieties. A further point considering the cofactor attachment point is that the cofactor should be in a deep cleft within the protein topology rather than at the surface. This is to allow the protein to maximise substrate binding and secondary transition state stabilising effects, as well as second sphere interactions influencing the metal complex.

Strategies for Generating Artificial Metalloenzymes

Artificial metalloenzymes are generated either from the combination of an unnatural transition metal cofactor being introduced into a protein scaffold or a natural metalloprotein being evolved in a laboratory to enhance or alter its natural catalytic reactivity. A detailed review of the field of the directed evolution of natural metalloproteins is out of the scope of this perspective, however, the engineering and evolutionary approaches developed by Frances Arnold, and applied to haem metalloproteins (e.g. Cytochrome P450) is particularly noteworthy and applicable when evolving unnatural metal-protein hybrid catalysts.(28–30)

Four successful strategies have been employed to localise an unnatural metal to a well-defined location within a protein matrix.

Metal Ion Substitution in Natural Enzymes

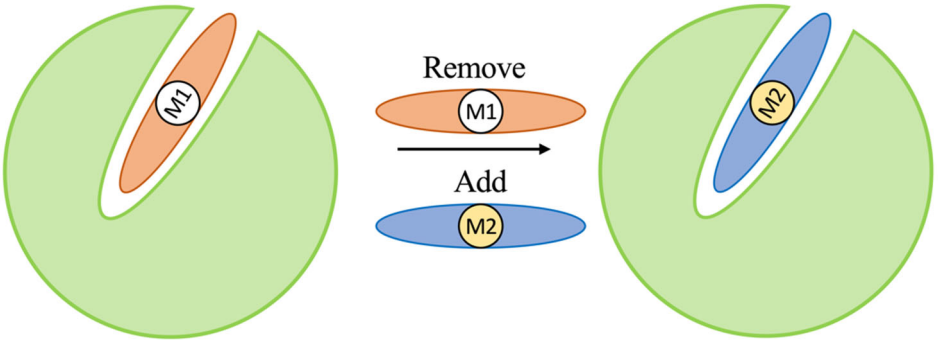


Figure 2 | Schematic representation of metal substitution | The natural cofactor (red) can be substituted with a suitable unnatural cofactor (blue). This may include the bare metal ion or larger cofactors such as heme.

Natural metal cofactors can be found in proteins encapsulated by ligands supplied by the protein or with non-protein ligands also coordinated. This enables two different methods of metal substitution: (i) substituting the metal ion in a protein defined coordination site or (ii) substituting the metal ion in a natural metal-organic cofactor (e.g. haem) (Figure 2).

Many ArMs have been generated by substituting the catalytic Zn(II) ion located in a His_3 binding site of carbonic anhydrase with different metals, for example, Coleman *et al.* reported esterase activity of a Co(II) substituted carbonic anhydrase.(31) Replacement with different Rh(I) species has also been explored, with catalytic hydrogenation(32) and hydroformylation(33) demonstrated. However, these rhodium metalloenzymes have a much slower activity than commercial small molecule rhodium catalysts alone. Although in these examples it is demonstrated that unnatural metal complexes can coordinate to the natural Zn(II) binding site, relatively low catalytic activity is observed. The highly evolved Zn binding site contains a complex secondary sphere architecture, in order to modulate the Lewis acidity of Zn. The chemically different demands for rhodium catalysed hydrogenation/hydroformylation reactions will therefore not be met in this system. Further, evolution of such a specialised system may be difficult.

Hartwig *et al*. reported taking the metal-organic cofactor haem and substituting Fe for a range of different 4d and 5d metals (including Rh, Ru, Ir and Ag).(34) In one particularly comprehensive example, an Ir(Me) porphyrin was incorporated into the cytochrome P450 enzyme CYP119 and catalytic functionalisation of C-H bonds to C-C bonds by carbene insertion was demonstrated, capable of high stereo-specificity.(26) Evolutionary campaigns on this artificial iridium metalloenzyme generated variants with an impressive 4000 fold increase in catalytic efficiency (defined by the k_{cat}/K_M), with kinetic parameters and selectivities

Doi: 10.1595/205651320X15928204097766

matching those of native enzymes. These parameters highlight the potential of this attachment method, and in particular the advantages of introducing exogenous metal cofactors with non-protein ligands remaining coordinated upon ArM formation.

In this case, the mutations made to this iridium CYP119 metalloenzyme have greatly optimised the binding and pre-organisation of the substrate for catalysis, lowering the value for K_M, (Figure 3). In this system there is no direct iridium-protein coordination; the iridium metal is coordinatively saturated by four haem nitrogens, one methyl ligand and coordination to the substrate. Therefore, the moderate increase in k_{cat} cannot have come through an electronic (through bond) contribution to catalysis from amino acid side chain ligands and protein fold energy but must arise from other minor contributors as discussed in the previous section. Another limitation of such a system is that it does not allow for the metal to interact with more than one substrate at a time, an essential feature of many interesting organometallic transformations e.g. metathesis.

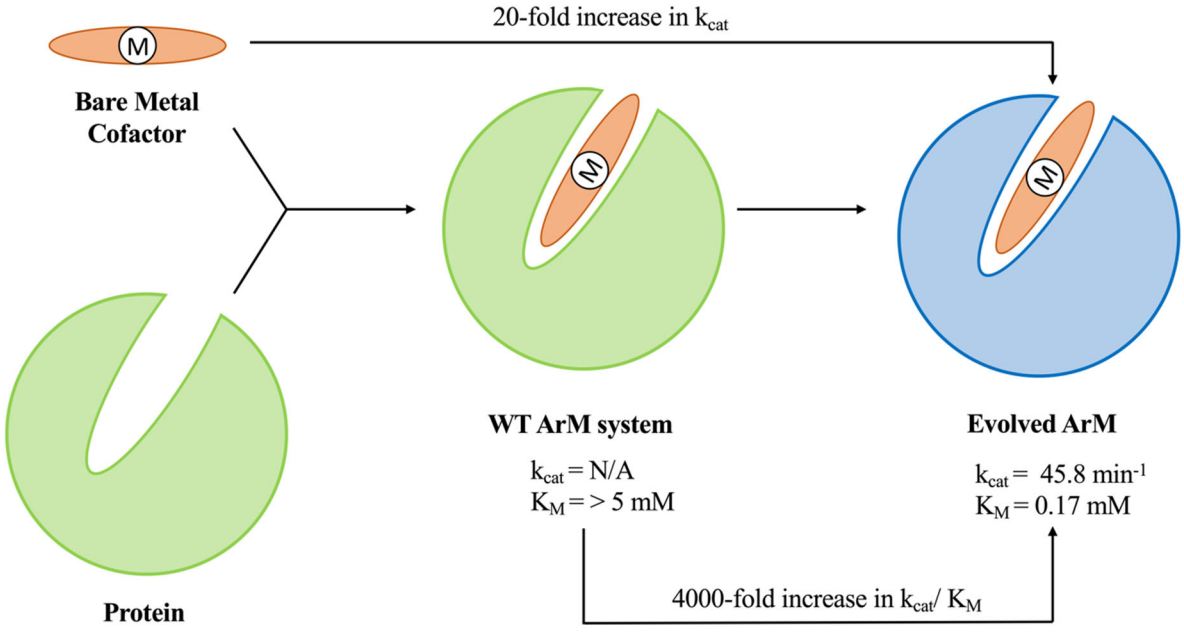


Figure 3 | Comparisons between the activities of a bare cofactor and artificial metalloenzyme before and after directed evolution | The data in this figure are taken from the work from Hartwig *et.al.*(26) This elegant study is a good example of the issues encountered when using fully substituted artificial cofactors, even in highly optimized systems. Whereas directed evolution was able to achieve an impressive 4000-fold increase in k_{cat}/K_M, the actual chemical turnover rate (k_{cat}) was only moderately enhanced when compared to the cofactor in solution. This can be explained by the enzyme evolving to more strongly bind the substrate and optimize the orientation of the substrate-metal complex. However, as there is no direct metal-protein interface, directed evolution cannot influence the metal chemistry, capping the chemical potential at that observed for the free cofactor in solution.

Doi: 10.1595/205651320X15928204097766

Supramolecular, Non-Covalent Binding of Tagged Complexes

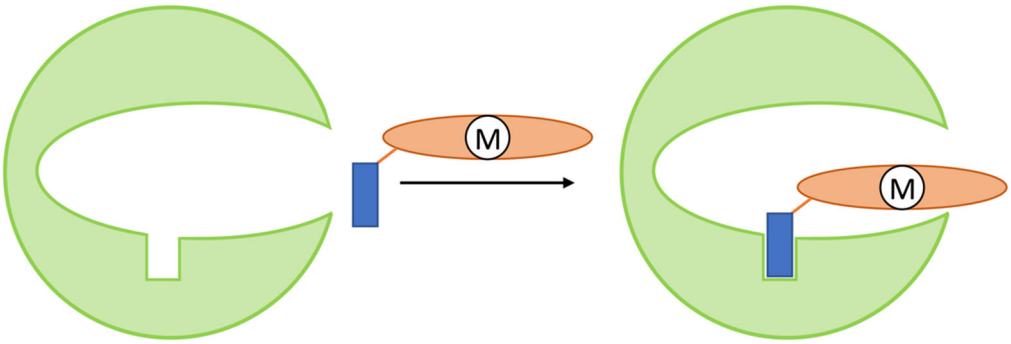


Figure 4 | Schematic representation of supramolecular assembly | The metal cofactor (red) is localised by non-covalent interaction between a ligand bound recognition group (blue) and the protein.

There are many specific complexes between proteins and small molecules which are well understood and have very high affinity. Artificial metalloenzymes have therefore been generated where a catalytic metal complex has been attached to a small molecule with high affinity for a protein target. (Figure 4) This means of localising the new cofactor into a protein scaffold has been widely explored. Building on the work of Wilson and Whitesides in the 1970's,(35) Thomas Ward and co-workers have assembled ArMs based on the high supramolecular affinity of small molecule biotinylated metal catalysts for the protein streptavidin. As many as 12 different catalytic transformations have been performed by these metal-streptavidin hybrids, including ruthenium catalysed olefin metathesis,(36) ruthenium-catalysed deallylation,(37) iridium-catalysed transfer hydrogenation(25) and dirhodium-catalysed cyclopropanation,(38) all *in vivo*.

This strategy has also been employed in ArMs that were reported by Tanaka *et al.* for potential therapeutic application. In this example, a coumarin derivative tagged with a ruthenium metathesis catalyst was localised to a hydrophobic binding site in human serum albumin. The metalloenzyme was directed to cancerous tissue (through specific glycosylation) and a pro-drug was administered which upon metathesis induced cellular death.(39)

One key benefit to supramolecular assembly is apparent in the examples described above, and that is that the conjugation between metal and protein is robust enough to be performed in complex cellular environments. Furthermore, unlike covalent attachment, supramolecular assembly can be a reversible process, which allows for component recycling. In a recent report of Duhme-Klair *et al*. catalytic transfer hydrogenation is demonstrated from a siderophore–protein combination that enables strong but redox-reversible catalyst anchoring.(40) All

Doi: 10.1595/205651320X15928204097766

current examples of ArMs generated by supramolecular assembly do, however, rely on the assembly of proteins with known, highly catalytically active metal complexes. As discussed previously, using complexes which maintain their ligand set during ArM formation does not allow the metal complex to be subjected to evolutionary pressures limiting the evolutionary potential.

Covalent Anchoring Through Metal Ligands

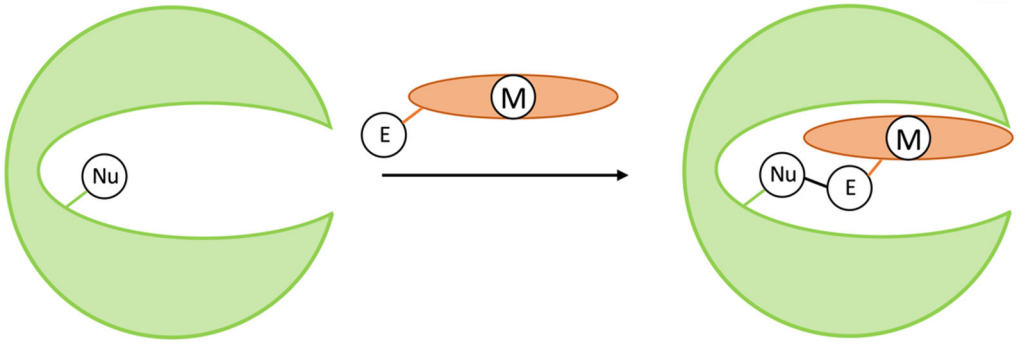


Figure 5 | Schematic representation of covalent anchoring | The metal cofactor (red) is attached to the protein by a reaction forming a covalent bond, for instance nucleophile (Nu) attacking an electrophile (E).

Covalent anchoring relies on using a chemical reaction to covalently link a protein side chain to a strong ligand for a metal. (Figure 5) Covalent anchoring methods can be split into two broad categories: (i) modification of a natural amino acid side chain (e.g. cysteine, lysine or tyrosine), via a nucleophilic– electrophilic reaction and (ii) coupling through a genetically encoded unnatural amino acid (UAA).

There is a resurgence in research for developing novel bioconjugation and protein modification techniques of natural amino acids (e.g. cysteine, lysine or tyrosine).(41, 42) Generating ArMs through cysteine modification is attractive due to the high nucleophilicity and rarity of free cysteines allowing for greater control of reactivity. Salmain and co-workers have modified the free Cys25 in the cysteine protease papain, using a variety of Ru, Re and Rh complexes all functionalised with either a maleimide or chloroacetamide group.(43–45)

The pioneering work of the Schultz laboratory enabled incorporation of UAAs into protein scaffolds.(46) Since then, the most successful generation of ArMs involving a covalent linkage to an UAA were reported by Lewis *et al.* and involve a reaction between an alkyne-substituted dirhodium catalyst and a genetically encoded L-4-azidophenylalanine residue through strain-promoted azide-alkyne cycloaddition (SPAAC) (47–49) Hypothetically, UAAs could be encoded into a specific residue of most proteins; here, the protein scaffold selected was a β-

barrel prolyl oligopeptidase and the resulting metalloenzymes generated catalysed olefin cyclopropanation.

The effectiveness of introducing UAA via stop codon methodology is that theoretically the same conjugation technology is applicable to many different proteins to generate diverse artificial metalloenzymes through a specific, fast and irreversible covalent conjugation. Beside commonly relying on pre-formed metal complexes, an overarching issue of covalent attachment and supramolecular assembly is that the protein scaffold is used predominantly as an auxiliary providing a chiral and hydrophobic micro-environment. Further, many reported methods utilise a long flexible linker between the point of attachment and the metal complex which could remove the catalytic centre from the very interactions needed for the protein to exert an influence on transition states.

Direct Activation by Metal Coordination to Protein Side Chains

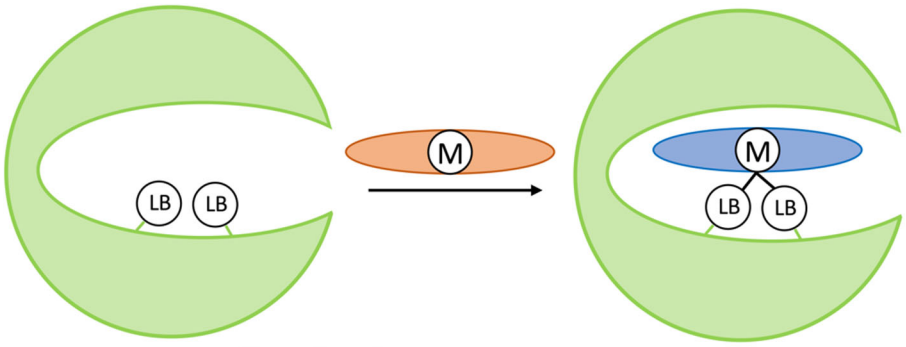


Figure 6 | Schematic representation of cofactor attachment via direct coordination | The free metal cofactor (red) attaches to Lewis basic residues on the protein (LB) via ligand substitution reactions, forming a new metal-protein complex (blue).

Dative ArMs have one or more coordination bonds directly from the metal to a Lewis basic amino acid residue (His, Cys, Ser, Glu, Asp, etc.) on the protein scaffold (Figure 6). The protein therefore has a direct electronic influence on the reactivity at the metal centre. The active hybrid molecule is formed by substitution reactions from a precursor metal species and the apoprotein. This allows for potentially very clean reaction conditions for assembly of the metal-protein complex. Although advances have been made, the complexity of these metal-protein binding processes remain a major challenge for the design of competently folded and catalytically active metalloproteins from scratch. It is important to distinguish between metalloenzymes where coordination to the metal is provided only by amino acid sidechains, substrates and solvents, and those in which the metal brings its own specific ligands with it. The latter, metal cofactors would be artificial versions of commonly encountered natural examples such as

Doi: 10.1595/205651320X15928204097766

heme, vitamin B12 and molybdopterin which are (bio)synthesised separately and bind to the protein through both non-covalent interactions and coordination. As pointed out above, their activity is defined by the other ligands they carry to an active site as well as the coordination by the protein.

Degrado and co-workers have pioneered the design of a number of synthetic proteins which directly coordinate bare metal atoms or metal cofactors.(50, 51) For example, in some of the earliest work, the His_3-Zn(II) binding motif found in carbonic anhydrase was introduced into a designed four helical bundle protein, and hydrolytic activity was observed.(52) More recently, de-novo design has been coupled with directed evolutionary approaches to generate an artificial zinc metalloenzyme capable of accelerating ester cleavage with un-paralleled catalytic efficiency (k_{cat} / K_M – 10^6 $M^{-1}s^{-1}$).(27)

In a range of studies,(53–56) Roelfes and co-workers use amber stop codon technology to introduce the UAA (2,2′-bipyridin-5yl)alanine into a range of protein scaffolds. Upon addition of different bare metal ions, they were able to obtain ArMs catalysing the Friedel-Craft alkylation of indoles, enantioselective metallohydration and the stabilisation of a semiquinone radical. By the use of sophisticated computational design, the group was able to introduce beneficial point mutations in many of the novel hybrid molecules, improving both enantioselectivity and yield. The advances in stop codon technology to introduce UAAs, especially in the context of directed evolution, make their use a promising option and provides an enticing method for expanding the ligand set available to the protein scaffold.(57–59) In another study, Reetz and co-workers computationally designed a copper (II) ion binding site into the thermostable protein imidazole glycerol phosphate synthase.(60) The resulting ArM was able to catalyse the Diels-Alder cycloaddition of an azachalcone and cyclopentadiene with medium selectivity, however, to our knowledge no subsequent directed evolution experiments have been reported.

In contrast to these examples of forming the complete coordination sphere by binding a bare metal to the apoprotein state of the ArM, to the best of our knowledge there are no examples of adding exogenous metal complexes (particularly 4d and 5d metal complexes) as precursor cofactors which then show catalytic activity upon direct coordination to a protein. This is a particularly attractive methodology as the challenges of taking unnatural ligands such as arenes, carbenes and phosphanes into biology become opportunities for expanding the repertoire of chemistries available for catalysis. Controlling the ligand exchange behaviour of 4d and 5d metal complexes with protein side chain ligands is challenging, not least because the

Doi: 10.1595/205651320X15928204097766

bonds are often stronger and hence exchange rates are slow. This, however, remains an exciting area of research due to the catalytic diversity demonstrated by many 4d and 5d metal complexes. In this specific area our own work has focused upon ruthenium complexes and their ligand exchange behaviour with biological systems, laying the foundation for future work into artificial metalloenzymes with direct metal-protein coordination.(61, 62)

Summary and Outlook

Significant advances in the incorporation of organometallic complexes into proteins in order to generate artificial metalloenzymes have been made. The studies highlighted above reliably create hybrid molecules where the stability and turnover number of the metal centre is higher than the comparable small molecule organometallic complex in aqueous solution. Maybe unsurprisingly, the propensity for side reactions and catalyst decomposition is lowered once the complex is in a hydrophobic protein environment, already showcasing the usefulness of these hybrid systems. However, the question remains, as to whether or not these strategies make full use of the protein component. The unique and numerous demands of ArMs call for a highly integrated approach. To date, most of the work described in the literature attempts to exploit the chemistry of metal ions and their complexes in a protein scaffold but with limited influence from the protein on any catalytic activity because metal-protein coordination is largely indirect and so co-operativity is limited.

The potential for synthetic organometallic chemistry to deliver cofactors which utilise ligand chemistry not available to naturally evolved systems can vastly expand the orthogonal catalysis available in synthetic biological applications. Using such molecules to embed novel metal-peptide hybrid complexes in protein scaffolds allows for three-dimensional and electronic control around the metal centre that reduces the need for intricate synthetic catalyst generation. Instead, control of the steric and electronic environment around the metal ion can be delivered via the protein coordination sphere, particularly where a direct coordination bond is used to anchor the metal ion to the protein. When combined with the ability to efficiently evolve enzymes, a sophisticated organometallic precursor complex together with a suitable apoprotein could potentially give rise to a number of diverse reactivities. Therefore, new problems in catalysis could be addressed in a faster and more specific manner than with small molecule catalysts. Together with non-covalent contributions to catalysis and the intermolecular interactions that pre-organise substrates and stabilise transition states, such a system contains many readily evolvable components.

Doi: 10.1595/205651320X15928204097766

The majority of protein scaffolds selected for ArM construction have been chosen because of their apparent engineerability. However, in most cases the focus seems to lie solely on the peptidic component with little consideration for evolution of the metal complex. Although methods of selection and directed evolution have been applied, these are often operating on already well-defined protein scaffolds that carry an abiotic cofactor but not a direct protein-metal complex, which inevitably limits the scope for evolution. Arguably it is desirable, therefore, to select for a promiscuous and versatile protein starting point which is not constrained by one energy minima but instead can potentially offer numerous distinct metal-binding environments, both in terms of direct coordination and through secondary, intramolecular spheres of influence, ultimately generating differential catalytic ArM activity.

Performing catalysis with exogenous metal complexes within cellular environments has enormous potential applications in medicinal chemistry and synthetic biology. Given the potential difficulties associated with cell-uptake, minimising deactivation, overcoming toxicity of exogenous metal ions and precise localisation of metal cofactors in cells, the idea of using traditionally inert organometallic complexes has obvious advantages in that reactive promiscuity is reduced. As pointed out above, such complexes would be designed to have a latent catalytic activity which emerges once the metal complex is bound to a protein. The design challenges raised by this approach are not just as a result of a need to control the electronic and three-dimensional steric coordination sphere of the metal ion, but also to limit ligand exchange processes, restricting lability of a precursor complex (in the cellular milieu) until it reaches a specific protein target. Since the metal-ligand exchange processes for 4d and 5d metal complexes are typically slow, they are particularly attractive from this point of view but are hard to predict *ab initio*.

In conclusion, in order to optimise the chemistry and biochemistry of artificial metalloenzymes, directed evolutionary campaigns coupled with high throughput screening methods rather than individually-designed synthetic strategies are much more likely to generate optimised orthogonal catalysts for new and efficient metabolic processes. Direct coordination between metal ions and enzymes is essential in order to deliver truly interdependent systems, ideally where entatic states deliver enhanced reactivity, efficiency and selectivity that cannot easily be replicated in conventional, synthetic metal catalysis. Going forward, methods of generating artificial metalloenzymes should be evaluated and developed for both their ability to be used in directed evolution procedures and the extent to which the protein scaffold participates in the activity of the metal complex.

Doi: 10.1595/205651320X15928204097766

Acknowledgements

GSB is supported by the EPSRC (EP/N509620/1) and Peterhouse, Cambridge. OJK is supported by the EPSRC (EP/R513180/1). SRB and PDB thank the Chemistry Department. We thank Florian Hollfelder for deep discussions.

Bibliography

1. G. Winter, A. R. Fersht, A. J. Wilkinson, M. Zoller and M. Smith, *Nature*, 1982, **299**, (5885), 756 LINK https://doi.org/10.1038/299756a0

2. A. R. Fersht, J. P. Shi, J. Knill-Jones, D. M. Lowe, A. J. Wilkinson, D. M. Blow, P. Brick, P. Carter, M. M. Y. Waye and G. Winter, *Nature*, 1985, **314**, (6008), 235 LINK https://doi.org/10.1038/314235a0

3. F. Schwizer, Y. Okamoto, T. Heinisch, Y. Gu, M. M. Pellizzoni, V. Lebrun, R. Reuter, V. Köhler, J. C. Lewis and T. R. Ward, *Chem. Rev.*, 2018, **118**, (1)142 LINK https://doi.org/10.1021/acs.chemrev.7b00014

4. M. Jeschek, S. Panke and T. R. Ward, *Trends Biotechnol.*, 2018, **36**, (1) 60 LINK https://doi.org/10.1016/j.tibtech.2017.10.003

5. J. G. Rebelein and T. R. Ward, *Curr. Opin. Biotechnol.*, 2018, **53**, 106 LINK https://doi.org/10.1016/j.copbio.2017.12.008

6. J. D. Tyzack, A. J. M. Ribeiro, N. Borkakoti and J. M. Thornton, *ACS Synth. Biol.*, 2019, **8**, (11), 2494 LINK https://doi.org/10.1021/acssynbio.9b00105

7. A. Bhushan, P. J. Egli, E. E. Peters, M. F. Freeman and J. Piel, *Nat. Chem.*, 2019, **11**, (10), 931 LINK https://doi.org/10.1038/s41557-019-0323-9

8. M. C. Wilson, T. Mori, C. Rückert, A. R. Uria, M. J. Helf, K. Takada, C. Gernert, U. A. E. Steffens, N. Heycke, S. Schmitt, C. Rinke, E. J. N. Helfrich, A. O. Brachmann, C. Gurgui, T. Wakimoto, M. Kracht, M. Crüsemann, U. Hentschel, I. Abe, S. Matsunaga, J. Kalinowski, H. Takeyama and J. Piel, *Nature*, 2014, **506**, (7486), 58 LINK https://doi.org/10.1038/nature12959

9. A. Casini, F. Y. Chang, R. Eluere, A. M. King, E. M. Young, Q. M. Dudley, A. Karim, K. Pratt, C. Bristol, A. Forget, A. Ghodasara, R. Warden-Rothman, R. Gan, A. Cristofaro, A. E. Borujeni, M. H. Ryu, J. Li, Y. C. Kwon, H. Wang, E. Tatsis, C. Rodriguez-Lopez, S. O'Connor, M. H. Medema, M. A. Fischbach, M. C. Jewett, C. Voigt and D. B.

Gordon, *J. Am. Chem. Soc.*, 2018, **140**, (12), 4302 LINK https://doi.org/10.1021/jacs.7b13292

10. W. F. DeGrado, C. M. Summa, V. Pavone, F. Nastri and A. Lombardi, *Annu. Rev. Biochem.*, 1999, **68**, (1), 779 LINK https://doi.org/10.1146/annurev.biochem.68.1.779

11. L. Falivene, Z. Cao, A. Petta, L. Serra, A. Poater, R. Oliva, V. Scarano and L. Cavallo, *Nat. Chem.*, 2019, **11**, (10), 872 LINK https://doi.org/10.1038/s41557-019-0319-5

12. L. Alonso-Cotchico, J. Rodríguez-Guerra, A. Lledós and J.-D. Maréchal, *Acc. Chem. Res.*, 2020, **53**, (4), 896 LINK https://doi.org/10.1021/acs.accounts.0c00031

13. I. Drienovská, L. Alonso-Cotchico, P. Vidossich, A. Lledós, J.-D. Maréchal and G. Roelfes, *Chem. Sci.*, 2017, **8**, (10), 7228 LINK https://doi.org/10.1039/C7SC03477F

14. L. Alonso-Cotchico, G. Sciortino, P. Vidossich, J. Rodríguez-Guerra Pedregal, I. Drienovská, G. Roelfes, A. Lledós and J.-D. Maréchal, *ACS Catal.*, 2019, **9**, (5), 4616 LINK https://doi.org/10.1021/acscatal.8b04919

15. H. Feng, X. Guo, H. Zhang, L. Chen, P. Yin, C. Chen, X. Duan, X. Zhang and M. Wei, *Phys. Chem. Chem. Phys.*, 2019, **21**, (42), 23408 LINK https://doi.org/10.1039/C9CP04473F

16. R. B. Leveson-Gower, C. Mayer and G. Roelfes, *Nat. Rev. Chem.*, 2019, **3**, (12) 687 LINK https://doi.org/10.1038/s41570-019-0143-x

17. M. Jeschek, R. Reuter, T. Heinisch, C. Trindler, J. Klehr, S. Panke and T. R. Ward, *Nature*, 2016, **537**, (7622), 661 LINK https://doi.org/10.1038/nature19114

18. W. R. Hagen, *Metallomics*, 2019, **11**, (11), 1768 LINK https://doi.org/10.1039/C9MT00208A

19. B. L. Vallee and R. J. Williams, *Proc. Natl. Acad. Sci. U. S. A.*, 1968, **59**, (2), 498 LINK https://doi.org/10.1073/pnas.59.2.498

20. C. L. Drennan, S. Huang, J. T. Drummond, R. G. Matthews and M. L. Lidwig, *Science*, 1994, **266**, (5191), 1669 LINK https://doi.org/10.1126/science.7992050

21. M. L. Ludwig, C. L. Drennan and R. G. Matthews, *Structure*, 1996, **4**, (5), 505 LINK https://doi.org/https://doi.org/10.1016/S0969-2126(96)00056-1

22. R. Banerjee and S. W. Ragsdale, *Annu. Rev. Biochem.*, 2003, **72**, (1), 209 LINK https://doi.org/10.1146/annurev.biochem.72.121801.161828

23. M. T. Reetz, *Acc. Chem. Res.*, 2019, **52**, (2), 336 LINK https://doi.org/10.1021/acs.accounts.8b00582

24. M. Jeschek, R. Reuter, T. Heinisch, C. Trindler, J. Klehr, S. Panke and T. R. Ward, *Nature*, 2016, **537**, (7622), 661 LINK https://doi.org/10.1038/nature19114

25. J. Zhao, J. G. Rebelein, H. Mallin, C. Trindler, M. M. Pellizzoni and T. R. Ward, *J. Am. Chem. Soc.*, 2018, **140**, (41), 13171 LINK https://doi.org/10.1021/jacs.8b07189

26. P. Dydio, H. M. Key, A. Nazarenko, J. Y.-E. Rha, V. Seyedkazemi, D. S. Clark and J. F. Hartwig, *Science*, 2016, **354**, (6308), 102 LINK https://doi.org/10.1126/science.aah4427

27. S. Studer, D. A. Hansen, Z. L. Pianowski, P. R. E. Mittl, A. Debon, S. L. Guffy, B. S. Der, B. Kuhlman and D. Hilvert, *Science*, 2018, **362**, (6420), 1285 LINK https://doi.org/10.1126/science.aau3744

28. P. S. Coelho, E. M. Brustad, A. Kannan and F. H. Arnold, *Science*, 2013, **339**, (6117), 307 LINK https://doi.org/10.1126/science.1231434

29. F. H. Arnold, *Acc. Chem. Res.*, 1998, **31**, (3), 125 LINK https://doi.org/10.1021/ar960017f

30. F. H. Arnold, *Angew. Chemie - Int. Ed.*, 2018, **57**, (16), 4143 LINK https://doi.org/10.1002/anie.201708408

31. J. E. Coleman, *Nature*, 1967, **214**, (5084) 193 LINK https://doi.org/10.1038/214193a0

32. Q. Jing, K. Okrasa and R. J. Kazlauskas, *Chem. - A Eur. J.*, 2009, **15**, (6), 1370 LINK https://doi.org/10.1002/chem.200801673

33. Q. Jing and R. J. Kazlauskas, *ChemCatChem*, 2010, **2**, (8), 953 LINK https://doi.org/10.1002/cctc.201000159

34. H. M. Key, P. Dydio, D. S. Clark and J. F. Hartwig, *Nature*, 2016, **534**, (7608), 534 LINK https://doi.org/10.1038/nature17968

35. M. E. Wilson and G. M. Whitesides, *J. Am. Chem. Soc.*, 1978, **100**, (1) 306 LINK https://doi.org/10.1021/ja00469a064

36. M. Jeschek, R. Reuter, T. Heinisch, C. Trindler, J. Klehr, S. Panke and T. R. Ward, *Nature*, 2016, **537**, (7622), 661 LINK https://doi.org/10.1038/nature19114

37. T. Heinisch, F. Schwizer, B. Garabedian, E. Csibra, M. Jeschek, J. Vallapurackal, V. B.

Pinheiro, P. Marlière, S. Panke and T. R. Ward, *Chem. Sci.*, 2018, **9**, (24), 5383 LINK https://doi.org/10.1039/c8sc00484f

38. A. D. Liang, J. Serrano-Plana, R. L. Peterson and T. R. Ward, *Acc. Chem. Res.*, 2019, **52**, (3), 585 LINK https://doi.org/10.1021/acs.accounts.8b00618

39. S. Eda, I. Nasibullin, K. Vong, N. Kudo, M. Yoshida, A. Kurbangalieva and K. Tanaka, *Nat. Catal.*, 2019, **2**, (9), 780 LINK https://doi.org/10.1038/s41929-019-0317-4

40. D. J. Raines, J. E. Clarke, E. V. Blagova, E. J. Dodson, K. S. Wilson and A. K. Duhme-Klair, *Nat. Catal.*, 2018, **1**, (9), 680 LINK https://doi.org/10.1038/s41929-018-0124-3

41. J. N. DeGruyter, L. R. Malins and P. S. Baran, *Biochemistry*, 2017, **56**, (30), 3863 LINK https://doi.org/10.1021/acs.biochem.7b00536

42. E. A. Hoyt, P. M. S. D. Cal, B. L. Oliveira and G. J. L. Bernardes, *Nat. Rev. Chem.*, 2019, **3**, (3) 147 LINK https://doi.org/10.1038/s41570-019-0079-1

43. P. Haquette, B. Talbi, S. Canaguier, S. Dagorne, C. Fosse, A. Martel, G. Jaouen and M. Salmain, *Tetrahedron Lett.*, 2008, **49**, (31), 4670 LINK https://doi.org/10.1016/j.tetlet.2008.05.043

44. P. Haquette, M. Salmain, K. Svedlunq, A. Martel, B. Rudolf, J. Zakrzewski, S. Cordier, T. Roisnel, C. Fosse and G. Jaouen, *ChemBioChem*, 2007, **8**, (2), 224 LINK https://doi.org/10.1002/cbic.200600387

45. N. Madern, N. Queyriaux, A. Chevalley, M. Ghasemi, O. Nicolotti, I. Ciofini, G. F. Mangiatordi and M. Salmain, *J. Mol. Catal. B Enzym.*, 2015, **122**, 314 LINK https://doi.org/10.1016/j.molcatb.2015.10.007

46. C. J. Noren, S. J. Anthony-Cahill, M. C. Griffith and P. G. Schultz, *Science*, 1989, **244**, (4901), 182 LINK https://doi.org/10.1126/science.2649980

47. J. C. Lewis, *Curr. Opin. Chem. Biol.*, 2015, **25**, 27 LINK https://doi.org/10.1016/j.cbpa.2014.12.016

48. H. Yang, P. Srivastava, C. Zhang and J. C. Lewis, *ChemBioChem*, 2014, **15**, (2), 223 LINK https://doi.org/10.1002/cbic.201300661

49. P. Srivastava, H. Yang, K. Ellis-Guardiola and J. C. Lewis, *Nat. Commun.*, 2015, **6**, (1), 7789 LINK https://doi.org/10.1038/ncomms8789

50. W. F. DeGrado, C. M. Summa, V. Pavone, F. Nastri and A. Lombardi, *Annu. Rev.*

Biochem., 1999, **68**, (1), 779 LINK https://doi.org/10.1146/annurev.biochem.68.1.779

51. A. Lombardi, F. Pirro, O. Maglio, M. Chino and W. F. DeGrado, *Acc. Chem. Res.*, 2019, **52**, (5), 1148 LINK https://doi.org/10.1021/acs.accounts.8b00674

52. T. Handel and W. F. DeGrado, *J. Am. Chem. Soc.*, 1990, **112**, (18), 6710 LINK https://doi.org/10.1021/ja00174a039

53. I. Drienovská, A. Rioz-Martínez, A. Draksharapu and G. Roelfes, *Chem. Sci.*, 2015, **6**, (1), 770 LINK https://doi.org/10.1039/C4SC01525H

54. I. Drienovská, L. Alonso-Cotchico, P. Vidossich, A. Lledós, J.-D. Maréchal and G. Roelfes, *Chem. Sci.*, 2017, **8**, (10), 7228 LINK https://doi.org/10.1039/C7SC03477F

55. N. Ségaud, I. Drienovská, J. Chen, W. R. Browne and G. Roelfes, *Inorg. Chem.*, 2017, **56**, (21), 13293 LINK https://doi.org/10.1021/acs.inorgchem.7b02073

56. M. Bersellini and G. Roelfes, *Org. Biomol. Chem.*, 2017, **15**, (14), 3069 LINK https://doi.org/10.1039/C7OB00390K

57. Q. Wang, A. R. Parrish and L. Wang, *Chem. Biol.*, 2009, **16**, (3), 323 LINK https://doi.org/https://doi.org/10.1016/j.chembiol.2009.03.001

58. C. Mayer, C. Dulson, E. Reddem, A.-M. W. H. Thunnissen and G. Roelfes, *Angew. Chem. Int. Ed. Engl.*, 2019, **58**, (7), 2083 LINK https://doi.org/10.1002/anie.201813499

59. A. Fallah-Araghi, J.-C. Baret, M. Ryckelynck and A. D. Griffiths, *Lab Chip*, 2012, **12**, (5), 882 LINK https://doi.org/10.1039/C2LC21035E

60. J. Podtetenieff, A. Taglieber, E. Bill, E. J. Reijerse and M. T. Reetz, *Angew. Chemie Int. Ed.*, 2010, **49**, (30), 5151 LINK https://doi.org/10.1002/anie.201002106

61. T. G. Scrase, M. J. Oneill, A. J. Peel, P. W. Senior, P. D. Matthews, H. Shi, S. R. Boss and P. D. Barker, *Inorg. Chem.*, 2015, **54**, (7), 3118 LINK https://doi.org/10.1021/ic502051y

62. G. S. Biggs, M. J. O'Neill, P. Carames Mendez, T. G. Scrase, Y. Lin, A. M. Bin-Maarof, A. D. Bond, S. R. Boss and P. D. Barker, *Dalt. Trans.*, 2019, **48**, (20), 6910 LINK https://doi.org/10.1039/c8dt05159c

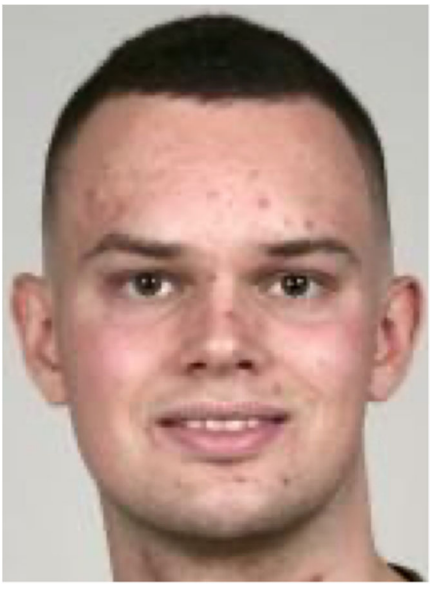
Authors

George Biggs

George Biggs completed an MChem in Chemistry at the University of Bath in 2016. He is now a PhD student in the Department of Chemistry at the University of Cambridge. Supervised by Dr Paul Barker and Dr Sally Boss, his project is focused on understanding the reactivity of Ru(II) arene complexes with proteins for the development of novel artificial metalloenzymes.

Oskar James Klein

Oskar James Klein obtained a MSc in Chemistry from the University of Cambridge in 2019, where he remains as a PhD student in the Department of Chemistry. Supervised by Dr Paul Barker and Dr Sally Boss and in collaboration with Prof. Florian Hollfelder his project tries to develop a high throughout methodology for the formation and evolution of novel artificial metalloenzymes.



Sally Boss

Sally Boss studied for an M.Sci. in Chemistry at the Universities of Bristol and Heidelberg, before moving to Cambridge University to begin a PhD on the synthesis and reactivity of Lewis acidic, heterobimetallic main group compounds. Shortly after obtaining her PhD in 2005, she was appointed to a joint College Lectureship in the Department of Chemistry and at Churchill College, Cambridge. Her time is split between teaching and research and her specific interest is in improving the utility of heavy metals in biology by careful design of complexes, targeted direction of metal-cofactors to protein targets and using spectroscopy to understand how they behave in situ.

Paul Barker

Paul Barker is Senior Lecturer at the University of Cambridge, Department of Chemistry, and a Fellow of Downing College. His research has always been at the interface between inorganic chemistry and biology. It started in the field of electron transfer proteins studied by biophysical methods and mutagenesis, in the early days of protein engineering. After two, independent MRC and BBSRC Fellowships in Cambridge he joined the Chemistry faculty and has been combining protein engineering with synthesis and self-assembly for the purposes of generating novel protein based electronic and catalytic systems. His current interests span protein design and evolution, self-assembling materials and synthesis of organometallic complexes.

Doi: 10.1595/205651320X15928204097766

Doi: 10.1595/205651320X15928204097766