Expanding the secondary coordination sphere of streptavidin-based artificial metalloenzymes and their characterization

Inauguraldissertation

zur

Erlangung der Würde eines Doktors der Philosophie

vorgelegt der

Philosophisch-Naturwissenschaftlicher Fakultät

der Universität Basel

von

Nico Valerio Igareta

Basel 2023

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät auf Antrag von

Prof. Dr. Thomas R. Ward

Prof. Dr. Florian P. Seebeck

Prof. Dr. Francesca Paradisi

Basel, den 20.09.2022

Prof. Dr. Marcel Mayor

Dekan

Acknowledgements

Abstract

Over billions of years, enzymes evolved to be nature's catalyst. Under (aqueous) mild conditions they promote a vast number of chemical reactions, achieving remarkable accelerations and exquisite selectivities. In addition, scientists designed and optimized organometallic catalysts in order to perform new-to-nature chemical transformations. Organometallic catalysts and enzymes have complementary properties that are combined and exploited in artificial metalloenzymes. In this approach, the reactivity of the catalyst can be tuned not only by chemical optimization of the catalyst, but furthermore by genetic optimization of the protein scaffold. These systems can be especially useful in bringing new-to-nature reactions in living systems. Various artificial metalloenzymes are designed, assembled and evolved in the Ward group. Although several host proteins have been exploited, the biotin-streptavidin technology occupies a place of choice in this context.

This thesis aims to create a chimeric streptavidin variant and to study its utility as a scaffold for the creation of ArMs. Furthermore, native MS and protein crystallography were used to characterize ArMs.

Table of Contents

Chapter 1. Introduction

1.1 Artificial Metalloenzymes

Life wouldn't be possible without the multitude of biochemical reactions spontaneously occurring in living cells. Enzymes catalyze the majority of these reactions. These evolved over time to be exquisitely effective catalysts. But what do enzymes consist of and how do they achieve their reactivity and selectivity? Enzymes consist of a combination of 20 amino acids, growing to a polypeptide chain with a molecular weight in a range of 10'000–100'000 Da. It is important to note that the polypeptide chains can also self-assemble into larger complexes. The amino acids have different side chains with different chemical and electronic properties (Figure 1). By combining these properties within the active site of an enzyme, perhaps with a metal ion or a coenzyme, it can confer the ability to catalyze a wide range of chemical reactions, including carbon-carbon bond formation, reductions, oxidations and hydrolyses.
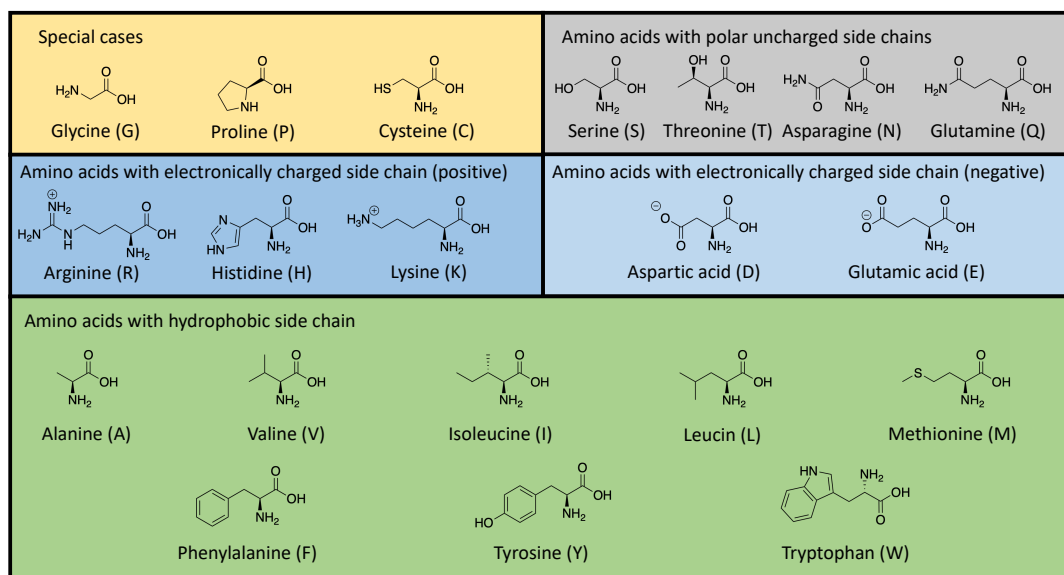
Figure 1 The 20 natural amino acids categorized according to the chemical properties of their side chains.

The combination of the amino acids leading to a protein, give rise to several levels of hierarchical structure, starting with the primary structure consisting of the amino acid

sequence. This amino acid chain then folds into a secondary structure, which defines the local topology of the polypeptide chain with structured units, such as α-helixes and β-strands. These secondary structure motifs are then further folded, leading to a globular structure: the tertiary structure. Finally, if more than one polypeptide chain associates to a functional enzyme, the resulting polypeptide assembly is coined the quaternary structure.

Within the active site, enzymes contribute to stabilize the transition state of a reaction[1] and furthermore enzymes may serve as an "entropic trap".[2] This leads to outstanding properties, including high catalytic activity at ambient conditions, substrate specificity and high selectivity with respect to chemo-, regio- and enantioselectivity, often outperforming the catalytic potential of most of the chemical catalysts. By means of example, the rate of an enzyme-catalyzed reaction can be increased by a factor of up to 10^{17} compared to its non-catalyzed reaction. In addition, enantioselectivities > 99 ee % are achieved routinely in case of asymmetric transformations.[3] The basic mechanism of transition state stabilization by enzymes was first formulated by Haldane[4] and expanded by Pauling[5] and Wolfenden.[6] According to their concept, the Gibbs energy ($\Delta G^{\ddagger}$) of the transition state of a given reaction catalyzed by enzymes is lowered by binding of the corresponding transition state with higher affinity than the ground state of the substrate(s). This model is further supported by the fact that some enzymes are strongly inhibited by transition state analogs.[7]

Along with the enzymes, also chiral transition metal complexes are used as homogenous catalysts in asymmetric synthesis. Their use in catalysis relies on the tuneability and versatile reactivity of these compounds. The transition metals can switch between different oxidation states; therefore, they are able to recombine certain fragments of molecules. This can be seen in the catalytic cycle of reactions catalyzed by transition metal complexes, where they can go through processes such as an oxidative addition or reductive elimination. The electron density and steric environment of the metal can furthermore be fine-tuned by the choice of ligand with tailored chemical properties. With this approach so called fine-tuning of the reactivity and selectivity (e.g. regio- and chemoselectivity) of the catalyst can be achieved, affording an opportunity to perform asymmetric catalysis relying on transition metals in combination with enantiopure ligands. The traditional way of preparing such enantiopure ligands is by fractional recrystallisation of a diastereomeric salt.[8] The pioneering achievements of W.S. Knowles[9] and R. Noyori[10] and B. Sharpless[11] revealed the potential of homogenous catalysts in asymmetric

synthesis. Their work on the asymmetric hydrogenation of functionalized alkenes, using chiral phosphine ligands and the application of enantiopure ligands derived from natural products for the asymmetric epoxidation and dihydroxylation of olefins was honored with the 2001 Nobel Prize in Chemistry. Inspired by their work, the scope of enantiopure ligands was continuously expanded and many other enantioselective transformations were developed.

Despite the great catalytic potential of enzymes, they suffer some major drawbacks. For organic synthetic purpose, thes include: limited stability towards organic solvents, elevated temperatures, extreme pH values and limited substrate scope. Inspection of the pros and cons of homogeneous catalysts and enzymes reveal a striking degree of complementarity (Table 1). In order to combine these advantages, the idea of the incorporation of a catalytically-active transition metal complex within a biomolecular scaffold (protein, DNA, RNA) to form an artificial metalloenzyme (ArM) arose in the 1970s. In 1976, Kaiser and Yamamura reported that the hydrolytic enzyme carboxypeptidase A (CPA) could be repurposed into an oxidase upon substitution of the native Zn(II) by Cu(II) for the oxidation of ascorbic acid.[12] In 1978, Whitesides and Wilson exploited the high affinity of biotinylated probes for avidin to anchor a Rh(diphosphine) within avidin.[13] Actually, the first report of an artificial metalloenzymes was used, dates back to 1956, when Fujii and co-workers described the use of reduced palladium chloride adsorbed on silk fibers for the asymmetric reduction of dehydro-amino acid derivatives.[14] By the creation of artificial metalloenzymes, next to the high selectivity of natural enzymes and the wide substrate scope of homogeneous catalysts, new-to-nature reactions can be implemented.[15,16] The full potential of ArMs was only be realized with the advent of recombinant protein production.

Table 1 Properties of homogenous - and enzymatic catalysis. Labeled in blue are shown the properties of ArMs.

	Homogenous Catalysis	Enzymatic Catalysis
Enantiomers	**both**	Single
Solvent	mostly organic	**mostly water**
Specificity	**broad**	**narrow**
Optimization	**chemical**	**genetic**
Metal	**precious metals**	**base metals**
1st coord. sphere	**vast choice**	limited choice
2nd coord. sphere	ill-defined	**well defined**
In silico design	easier	**challenging**
Lifetime	limited	**extended**
In vivo compat.	limited	**broad**

Taking a closer look at ArMs, reveals that the high activity and selectivity is by-and-large governed by the ligand forming the first coordination sphere in the case of the metal complexes. In metalloenzymes, the activity and selectivity relies to a large extent on the secondary coordination sphere, which is provided by hydrophobic interactions and hydrogen bonds between the cofactor and the protein. The secondary coordination sphere further contributes to the stabilization of the transition state of the reaction, the chiral discrimination and/or the activation of the substrate by secondary interactions. Therefore, the activity and especially the selectivity of the artificial metalloenzyme can be engineered by genetic modification of the host protein. This can either be addressed by rational design or directed evolution.[17,18] This ultimately leads to kinetics and catalytic efficiencies approaching those of native enzymes, as reported by Hartwig *et al.* for a reconstituted cytochrome P450 performing C-H insertion reactions[19] or by Basler *et al.* for an abiological hetero-Diels-Alderase.[20]

To convert proteins into ArMs, three important parameters must be carefully considered: i) the transition metal catalyst, ii) the protein scaffold and iii) the anchoring mode of the complex to the scaffold.

To compete with natural enzymes and introduce new-to-nature reactions, usually second or third row transition metal catalyst are used for the formation of ArMs. This catalyst can be

fine-tuned by the selection of the ligands. Therefore, the catalyst can be chemically optimized and adapted to the protein scaffold. Concerning the protein host, next to the de novo design of peptides and proteins, the creation of new or enhanced actives sites in native proteins and protein self-assembly were developed by rational design.[21–23] DeGrado and coworkers introduced the de-novo-designed peptides for the creation of ArMs in the 1980s, using single-chain peptides that fold into three- or four-helix bundles that can model the naturally-occurring structural motifs seen in native metalloproteins and bind metal ions via amino acid residues in the interior of the peptide fold.[24] Next to it, the DeGrado group also developed de novo peptides that model several important active sites in nature, such as porphyrins and FeS clusters.[25] Furthermore, Tezcan *et al.* used the de-novo approach to create an artificial β-lactamase[26] and Pecoraro *et al.* for an artificial hydrolase based on the TRI-peptide family.[27] Since the beginnings of the de-novo design of peptides, the computational methods have improved remarkably[28] cuminating with the AlphaFold protein structure prediction software, among other powerful prediction tools.[29] In addition to the natural and de-novo proteins, also DNA and small peptides have been utilized as hosts for transition metal catalysts.[30] Lastly, the incorporation of unnatural amino acids[31] into host proteins completes the toolbox for the design of artificial metalloenzymes.

There are four strategies that are usually applied for the attachment of the catalyst to the protein to assemble an artificial metalloenzymes (Figure 2).

(I)   **Dative Anchoring**: An interaction between a coordinatively unsaturated metal (cofactor) within a cavity and a Lewis-basic amino acid (e.g. histidine) is formed.[32]

(II) **Supramolecular Anchoring**: Either a high-affinity inhibitor or substrate is used to anchor a metal cofactor within a host protein (by supramolecular interactions).[13]

(III) **Metal Substitution**: In a native metalloenzyme the metal is substituted for another metal. It may be part of a prosthetic group (e.g. heme) or bound solely to amino acids.[33]

(IV) **Covalent Anchoring**: The covalent immobilization of a metal cofactor is achieved by an irreversible reaction between complementary functional groups on a ligand and on the host protein, respectively.[34,35]
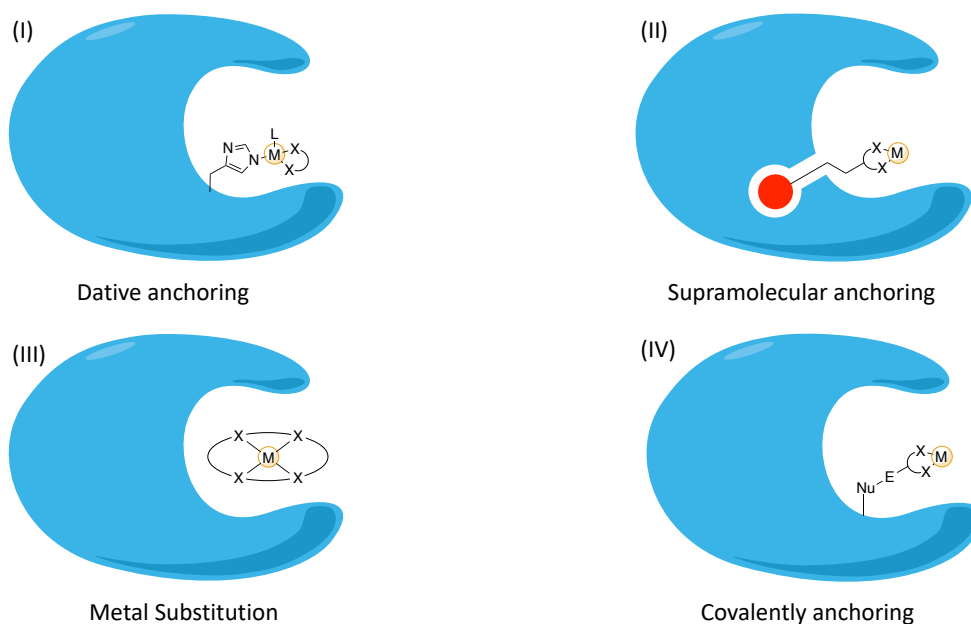
Figure 2 The four strategies for the assembly of ArMs. (I) Dative anchoring (II) Supramolecular anchoring (III) Metal substitution and (IV) Covalent anchoring.

Inspired by the seminal paper of Whitesides *et al.*[13], where they used the biotin-avidin technology to generate ArMs, the Ward group has capitalized on this technology and designed several biotinylated organometallic complexes, that catalyze a variety of reactions, including sulfoxidation, hydrogenation, dihydroxylation, allylic alkylation, alcohol oxidation, transfer hydrogenation, Suzuki cross-coupling, C–H bond activation, metathesis and hydroamination, etc.[36–43]

1.2 The Biotin-(Strept)avidin Technology

The relation of avidin and biotin was first reported by Boaz in the 1920s, while he was in investigating chicken egg white.[44] In the 1940s Williams *et. al.* showed that animals fed a diet that contains a large proportion of fresh or commercial egg white and is devoid of a specific protective factor (biotin, vitamin B7, vitamin H) an egg-white injury is produced.[45] In the case of streptavidin (Sav), during an antibiotic screening in the 1960s, it was shown that the antibacterial effects, of the medium containing the culture filtrate of *Streptomyces* (*avidiini*), could be reversed by high concentrations of biotin in the medium.[46] Since then,

(strept)avidin is widely used in a number of applications in life sciences, ranging from purification, labeling techniques, diagnostics, targeted drug delivery to nanotechnology and artificial metalloenzymes.[47] Both avidin and streptavidin are well known for their high biotin (Vitamin B7, Vitamin H) binding capacity. The affinity for biotin is ~10^{15} M^{-1} for avidin and ~10^{13} - 10^{14} M^{-1} for streptavidin respectively and is the strongest noncovalent molecule-protein interaction known.[48] In their natural hosts, avidin and streptavidin are secreted.[49] During secretion, in case of the streptavidin (159 aa), the so-called core-streptavidin (125-127 aa) is formed as a result of processing from both N and C termini simultaneously. Contrary to that, only the signal sequence of avidin is cleaved off during the secretion process. The core streptavidin version and avidin have only a 33 % sequence homology. However, the two protein structures are very similar and both are homotetrameric, which can be better described as a dimer of dimers with a D2-symmetry.[50] All four monomers consist of eight-stranded β-barrel with seven interconnecting loops, with each monomer binding one biotin molecule within his core. The main differences between streptavidin and avidin are in the interconnecting loop regions. The four biotin binding pockets can be categorized into two sets, cis (two biotins facing each other) and trans (long distance between two biotin molecules) (Figure 3). Noteworthy, no cooperativity was observed for the individual biotin binding events.[51] In both proteins, the quaternary structure is stabilized predominately by a network of van der Waals interactions and hydrogen bonds between each individual monomer.[52] Additionally, avidin (M_w= 62'400 Da for the tetramer) contains two cysteines forming a disulfide-bridge and two methionine residues as well as a glycosylation at asparagine 17.[53] Streptavidin (M_w= 65'700 Da) in contrast has none of these modifications. Both proteins are extremely stable to high temperatures , over a wide pH range and in the presence of denaturing agents as well as organic solvents.[54] The high affinity towards biotin and the high stability of the protein, makes (strept)avidin an ideal host for ArMs. Especially streptavidin, as its lack of glycosylation and disulfide bond and the more favorable isoelectric point (pI_{Avidin}= 10, $pI_{Streptavidin}$= 6.8-7.5) makes it easier to work with.
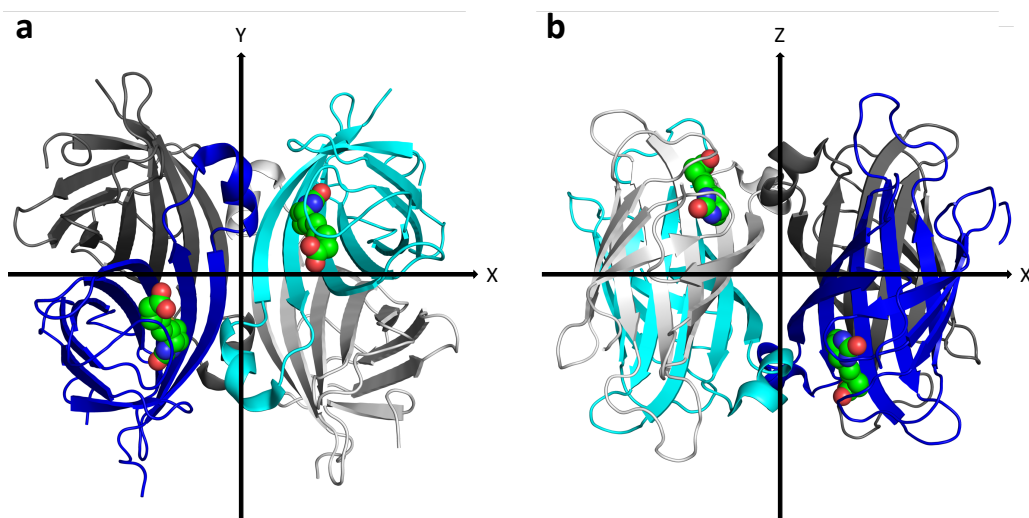
*Figure 3 Streptavidin homotetramer as cartoon representation (pdb: 3RY2) with biotin represented as cpk model. The two facing biotins (cis) are shown in **a**, whereas in **b** the two distant biotins (trans) are highlighted. The 2-fold rotation axis in x,y,z direction are also displayed to highlight the symmetry relation of the monomers in the tetramer.*

In order to improve the soluble protein fraction of streptavidin, the Ward lab relies on a genetically optimized variant.[55] In this variant, the protein sequence consists of the full-length streptavidin with the first 13 residues of streptavidin being replaced by a T7 epitope tag. By introducing this tag, the production levels could be improved to up to 200 mg/L in shake flasks and even TO up to 8 g/L in a fed batch mode.[56] As streptavidin itself, the tagged streptavidin tolerates a wide pH range (pH 3-11), elevated temperatures (75°C without biotin, up to 110°C with biotin bound), organic solvents (e.g. DMSO, ethyl acetate), high concentrations of guanidinium chloride (6M) or urea (8M), as well as the presence of chaotropic agents (e.g. sodium dodecyl sulfate).[54,57–59] The most important feature of streptavidin as ideal host for the assembly of ArMs may be its remarkable affinity towards biotin as well as biotinylated probes. Upon biotin binding, the loop connecting the strands 3 and 4 of the β-barrel of streptavidin, changes its conformation and thereby closes the binding site such that the ligand is buried almost completely.[52,60] However, the valeric acid group is located at the entrance of the β-barrel. Therefore, covalent modification of this carboxylic moiety does not significantly affect the biotin-Sav affinity. As the 3,4-loop is critical in the biotin binding, several modified variant of streptavidin were developed with affinities for peptides (as streptactin)[61] and circularly permutated streptavidin[62], among others.

The tight biotin binding originates from an extended hydrogen bond network and several hydrophobic interactions.[63] Interestingly, the ureido oxygen forms 3 hydrogen bonds with the residues Asn23, Ser27, Tyr43, suggesting that the groups involved stabilize an sp^3 oxyanion.[64] Furthermore, the NH-groups of the ureido moiety form H-bonds with Ser45 and Asp128, the sulfur of the thiolane moiety with Thr90 and the carboxyl group of the valeric acid chain with Asn49 and Ser88 (Figure 4). Additionally, the tryptophan residues Trp79, Trp92, Trp108 and Trp120 (from the adjacent monomer) form a hydrophobic binding pocket.
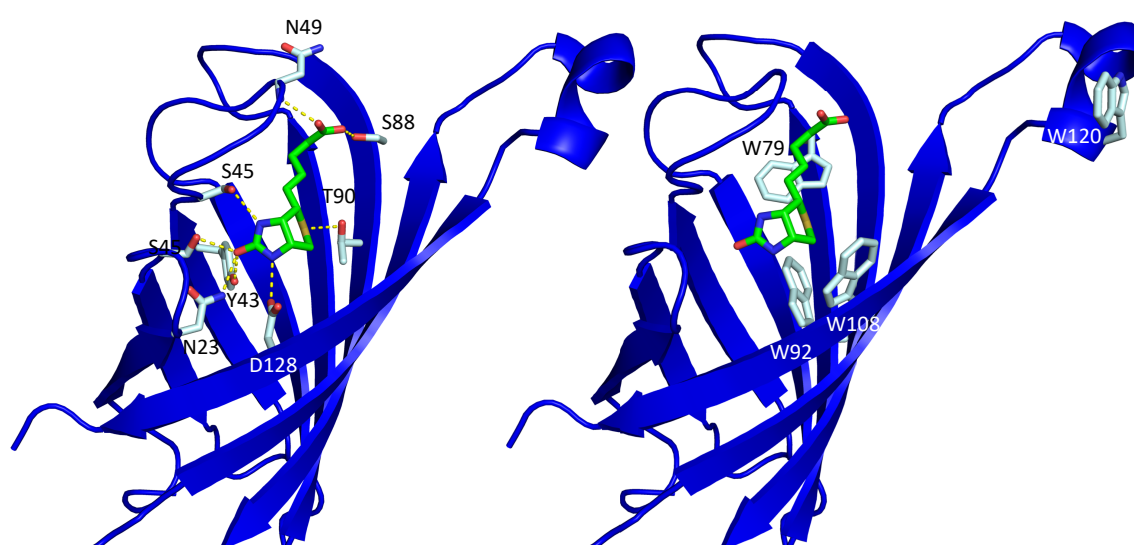
Figure 4 X ray structure of the streptavidin-biotin complex (pdb 3RY2, only one monomer of Sav displayed). The protein is represented as blue cartoon, whereas the residues involved in biotin binding are represented as sticks (carbon: cyan, oxygen: red and nitrogen: blue). The biotin is displayed as sticks (carbon: green, oxygen: red, sulfur: yellow and nitrogen: blue). On the left panel, the hydrogen bond network is highlighted, with the H-bond interactions being displayed as yellow dotted lines. On the right panel, the hydrophobic interaction by the tryptophan residues are shown, W120 is involved in an interaction with the adjacent Sav monomer (not displayed).

As previously mentioned, streptavidin's quaternary structure consists of a tetramer, which is stabilized by interactions of the residues in the subunit interfaces (including Val55, Thr76, Thr90, Leu109, Trp120, Val125, His127 and Asp128) as well as an important inter-subunit salt-bridge between Asp61 and His87 (at physiological pH). Crystallographic studies on the T7-taged full length streptavidin revealed that, in the apo-protein form, the C-terminus is folded into the biotin binding pocket.[65] The binding of the C-terminus can be described as

follows: "C-terminal extension stretches away from the β-barrel, forms a section of α-helix, and then folds against the β-barrel with several residues located within the biotin-binding site".[65] Nevertheless, due to the high production yield and stability the Ward lab relies on the T7-tagged full length streptavidin for its use as scaffold for ArMs.

Chapter 2. Chimeric Streptavidin

2.1 Why engineer (strept)avidin?

In 2008, the Ward group reported the first X-ray structure of an ArM based on streptavidin.[66] Since then, the group published more than 30 structures of Sav-based ArMs.[37,43,66–71] As previously mentioned, streptavidin consists of a dimer of dimers, where two neighboring subunits are related by a C2 axis. Inspection of the crystal structures, a narrow distribution of the metals, close to the C2 axis can be found (Figure 5).
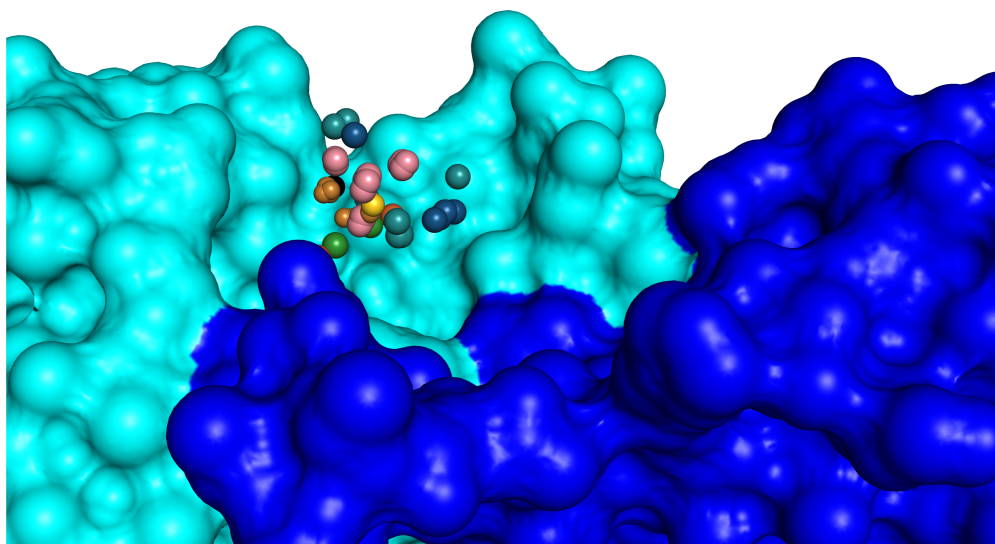
Figure 5 Close-up view of the metal distribution of ArMs based on Sav. The protein is represented as surface, where the two monomers are color-coded in cyan and blue, respectively. The metal atoms of biotinalyted metal cofactors are represented by spheres: Ir (Blue, pdb 6GMI, 6ESS, 6ESU, 4OKA, and 3PK2), Ru (Green, pdb 6FH8, 5F2B, 5IRA, 2QCB, and 2WPU), Cu (Brown, pdb 5VKX, 5VL5, 5VL8, 5WBA, 5WBB, 5WBD, 6ANX, 5WBC, 5K67, 5K68, and 5L3Y), Pd (black, pdb 5CSE), Rh (turquoise, pdb 4GJV and 4GJS) and the averaged metal position is represented as a yellow sphere. For clarity, only one metal is displayed for the homodimeric Sav making-up the biotin-binding vestibule.

As displayed in Figure 5, the metals are rather solvent-exposed. This characteristic surely contributes towards a broad substrate scope of the corresponding ArMs. As a consequence, however, the influence of the secondary coordination sphere interactions between the host protein and the catalyst is rather limited. Furthermore, the catalyst is less shielded against external factors than in traditional enzymatic cavities, where the active site is often well

shielded. The concept of introducing additional structural features which might offer the opportunity to tailor further and confine the active site, was initially addressed by the Ward group in 2008[72] and further expanded in 2018.[73] In their first report they implemented avidins larger 3,4-loop to streptavidin, whereas in the second, they introduced well-structured, naturally-occurring motifs, displaying close-lying N and C termini, to the 3,4-, 4,5-, 5,6-, 7,8-loop and were able to show that their incorporation to the 3,4-loop and C-terminus were the most promising ones. Furthermore, the biotin-binding of those chimeras was not significantly affected by the introduction of these motifs. Nevertheless, the newly introduced motifs needed to be further engineered to form a protective lid over the active site, as the lids did not cover the whole biotin-binding vestibule. Most importantly, these initial designs did not take into account the opportunity to fully shield the biotin-binding vestibule by capitalizing the dimeric nature of Sav.

One way of improving the properties of the ArMs is to genetically modify the protein's primary sequence. By doing so the properties of the ArMs can be "fine-tuned". In the field of biochemistry, an approach of so-called directed evolution is pursued, where a protein with a pre-existing activity is evolved to perform better by random mutagenesis.[74] In this approach, the protein sequence is changed by iterative cycles of mutagenesis and the beneficial mutations are taken as a starting point for the next round of evolution. By doing so, the natural evolution is simulated in a lab environment in a much faster time-scale. A different approach, which relies on the structural information of proteins is called targeted mutagenesis, where structural information about the protein and its active site are used in order to identify amino acid residues to target. As there are several structures for streptavidin and their ArMs, we inspected them and determined the average metal position (Figure 6). By doing so, we were able to identify potential targets for mutagenesis, especially by inspecting the distances to the C-β atom (from the amine and carboxyl group). This group was selected, as all amino acids (except glycine) have this position in common, whereas the remainder of the side chains is different in size and direction. In Sav-WT, there are no C-β atoms of residues within 5 Å of the averaged metal center. "Zooming out" a bit further to 8 Å, several residues were identified that might interact with the transition state of the metal-bound substrate, namely the residues at position N49, S112, and K121' (' = of the adjacent streptavidin monomer). Finally, at a distance of 10 Å more residues could be identified. In total, there are

11 residues within 10 Å of the averaged metal center (N49, A86, H87, S88, L110, S112, T114, N118, K121, S122, L124, and K121' (Figure 6). Unsurprisingly, most of them were already identified and genetically modified in previous studies of the Ward lab.[37,75]
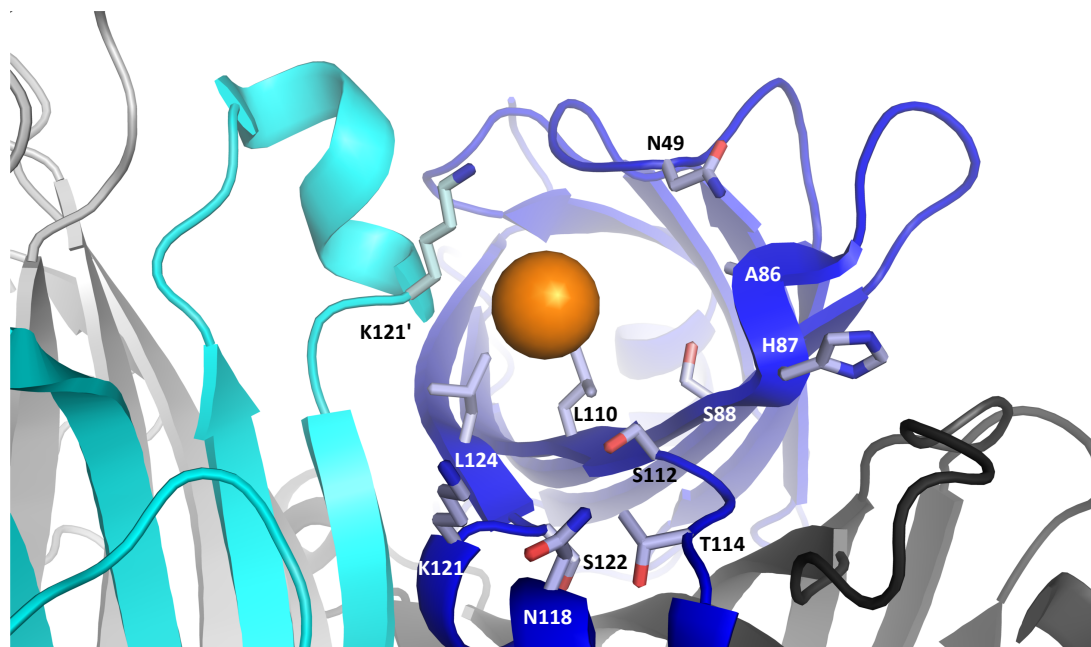
Figure 6 Close-up view of the biotin binding site with the averaged metal position highlighted. The monomers of streptavidin are color coded (black, white, cyan and blue, respectively). For simplification, only one metal (orange sphere) is displayed. The residues with C-βatoms at < 10 Å from the averaged metal center are displayed as sticks (carbon: green, oxygen: red and nitrogen: blue) and labeled.

One drawback of the quaternary structure of streptavidin is that the monomers cannot be genetically modified independently, as one gene encodes all four monomers. This leads to an uncertainty as to which of the two symmetry-related mutations (for positions S112 and K121 most prominently) affects most the catalytic performance. An additional challenge arises when using an ArMs where the cofactor:streptavidin ratio influences the catalytic properties. This effect is most prominently in the case for the artificial transfer hydrogenase consisting of [Cp*Ir(biot-*p*-L)Cl]·Sav-S112A. In this ArM, the enantioselectivity for the reduction of the salsolidine precursor erodes from 93 ee % to 45 ee % for the *(R)*-enantiomer upon increasing the [Cp*Ir(biot-*p*-L)Cl]:Sav-S112A from one to four.[76] One way of addressing this challenge is to assemble two separate streptavidin forms, one with the natural biotin-binding affinity and the other with ablated biotin affinity. Chilkoti *et al.* demonstrated that this could be achieved by denaturing and mixing monomers with high affinity and low affinity and renaturing a mixture thereof.[77] However, following refolding, a wide variety of combinations were

produced. Some contained four high-affinity binding sites, whereas other tetramers had a decreasing number of high affinity binding sites. To produce a more controlled means of combination, Reznik *et al.* engineered the subunit interfaces by modifying the residue at position 127 of streptavidin.[78] One more effective way of modifying each monomer individually would be to fuse them together. This was what Nordlund *et al.* reported for avidin.[79] In their study, they connected two circularly-permutated avidins over the old N- and C-terminus of avidin via a glycine/serine-rich linker and created new termini. In solution, this scaffold spontaneously formed dimers (pseudo tetramers, i.e., four biotin-binding sites per dimeric quaternary structure) with equivalent characteristics to wild-type avidin. However, when forming this dimer, there were two different ways how they assembled, which lead to the problem that the exact position of the applied modification in the quaternary structure could not be controlled (Figure 7). One year later, they addressed this problem by introducing the tetravalent single-chain avidin, where they joined all four subunits of avidin into one polypeptide chain.[80] In the same year, Hytönen et al. published a different strategy to control the assembly of the quaternary structure.[81] By engineering a disulfide bridge within the protein, they could control the quaternary structure of the dual chain avidin into a single conformation.
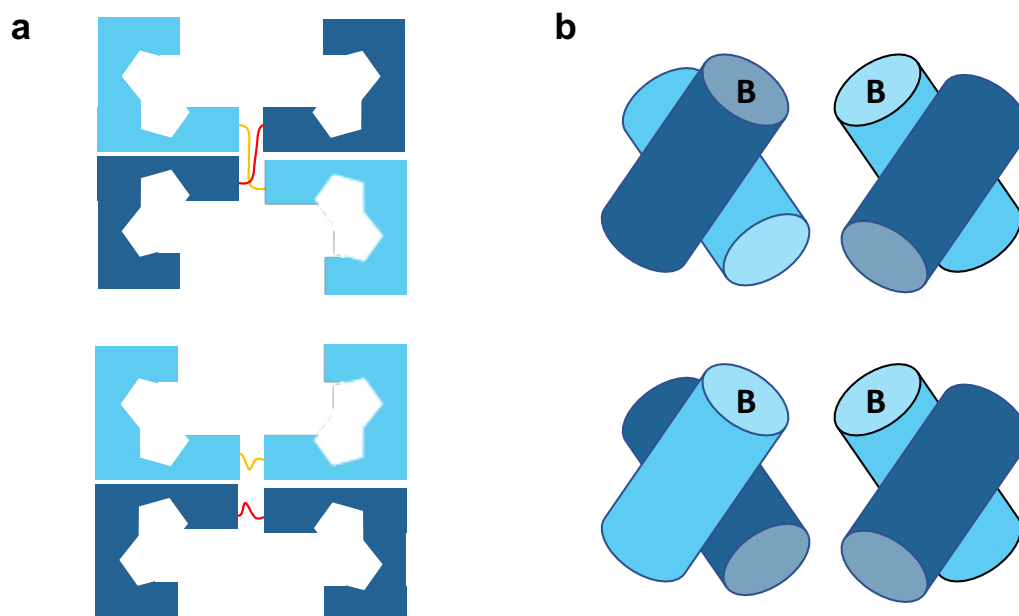
*Figure 7 Schematic representation of the fused avidin from Nordlund et al..[79] The fused dimers are color-coded in different shades of blue. **a** Two possible dimerization topologies to afford a the pseudotetrameric single chain dimer. The connection loop is represented as a red or orange line. **b** simplified representation of the β-barrels from each avidin monomer. B marks the biotin binding site.*

In 2019 the Ward group combined the previous approaches and introduced a single-chain dimeric streptavidin (scdSav).[69] In contrast to the single-chain avidin, where a circular permutated version of the protein was used, the C-terminus from the first monomer (Sav^A) of streptavidin was fused to the N-terminus of the second (Sav^B) via a 26 amino acid linker. Furthermore, the residue H127 which forms a π-stacking interaction with its neighboring H127 residue was mutated to cysteine in the Sav^B to favor the formation of a disulfide bond with an adjacent scdSav and lock the conformation (Figure 8). Additionally, the mutations N23A/S27D/D128A were introduced in Sav^B, in order to significantly reduce the biotin-binding affinity of this triple mutant.[64,82] This variant was called scdSav(SARK)mv2, whereas the mv2 stands for monovalent (N23A/S27D/D128A) and the letters in the parenthesis stand for the mutations in the Sav^A S112S and K121A, followed by the mutations in the Sav^B S112R and K121K. In the crystal packing of the scdSav, each tetramer can occupy two orientations, which leads to a mixed electron density. However, a preferential orientation of scdSav in the crystals resulted in differences in the electron density of the residues N23A, S27D, H127C, and D128A, and the cofactor binding of scdSav(SARK)mv2 could be observed. With this chimeric streptavidin, the Ward group expanded the genetic optimization potential and resolved the issue of multiple cofactor binding within the biotin-binding vestibule in ArMs based on the biotin–streptavidin technology. Based on the previous studies on the modifications of streptavidin, we set out to engineer a protecting lid on the biotin-binding vestibule, to more closely mimic a shielded enzymatic active site.
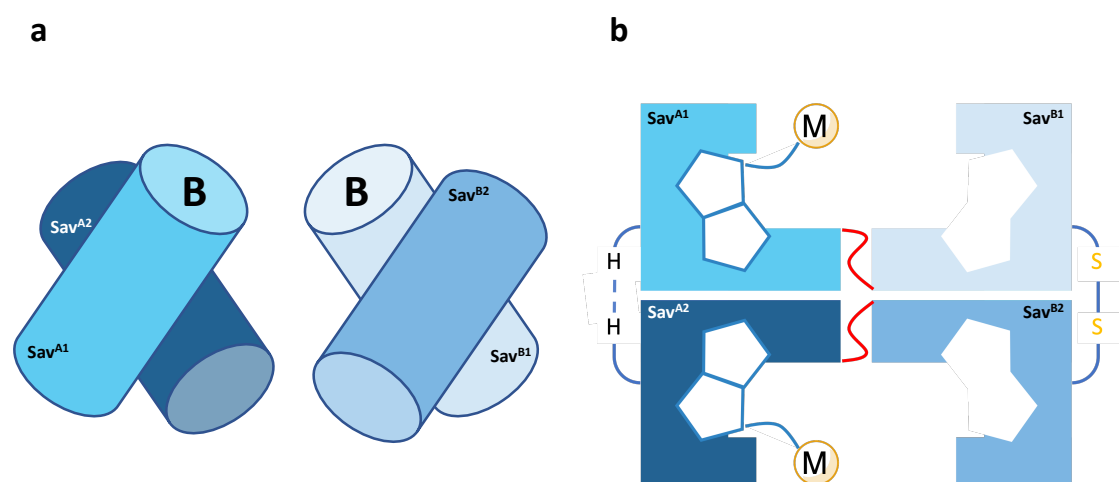
*Figure 8 Illustration of the scdSav from Ward et al.[69] **a** Simplified representation of the β-barrels from each scdSav monomer. **b** Schematic representation of scdSav mv2 with two bound cofactors, each in one biotin binding vestibule. Sav$^{B1 \& B2}$ contain the mutations N23A, S27D, H127C, and D128A.*

In the following section, I outline the group's effort to design a streptavidin chimera whereby the active site of Sav is shielded by a well-structured lid borrowed from a superoxide dismutase.

2.2 Streptavidin SODc Chimera

2.2.1 Outline of the Authors Contribution

TRW, RLP, and NVI conceived and designed the study. NVI, DCS, RLP contributed to mutagenesis, protein expression, protein purification and protein characterization. NVI performed the crystallization, X-ray structure determinations and the native MS experiments.

2.2.3 Results and Discussion

A careful pdb-search for naturally occurring dimerization domains present in β-sheet-rich proteins, the dimerization domain of the superoxide dismutase C (SODc) from *M. tuberculosis* (pdb 1pzs)[83] was identified as a possible candidate to close the"active site" of streptavidin and cover it with a lid. The SODc is a 20.7 kDa homodimeric enzyme comprised of a Greek key β-barrel domain and possesses a 34 amino acid dimerization domain (DD) that forms an elaborate dimerization interface spanning across the ~ 29 Å of the two Greek Key β-barrel subunits (Figure 10). Furthermore, its 2-fold symmetry matches with the symmetry of streptavidin, as well as the C- and N-terminus of the DD have a similar distance to each other as the b-sheets connected via a loop region in streptavidin (~5-7 Å). In order to engineer the chimeric Sav-SODs, we started by designing the genes containing the 34 aa DD from SODc on the loops of streptavidin and a circular permutated version of Sav (cpSav).[62] To further decrease the aggregation propensity and increase the protein's solubility, all of the genes were designed to include an N-terminal SUMO (small ubiquitin-related modifier) tag.[84] This protein tag was shown to circumvent challenges with "difficult to express" proteins. Importantly, the SUMO tag can be cleaved off by the ULP1 protease, originally identified in *Saccharomyces cerevisiae*.[84] Contrary to other proteases, ULP1 exclusively cuts after the SUMO-tag, recognizing its fold as the target rather than an amino acid sequence.[84–86] In this

way, five different capped Sav (K121A; L124H) mutants were created and cloned into the pET28a-vector. Initial small-scale expression tests of the constructs in *E.coli* BL21 (DE3) at 30°C for 24h, using the modified autoinduction media reported by Studier *et al.*[87] revealed that the incorporation of the SODc domain between the third and fourth β-sheet was yielding as only construct tetrameric soluble protein, which was capable of binding B4F (Biotin-4-fluorescein)(Figure 13).[88] Refolding experiments, as well as positioning of the SUMO tag on either N- or C- terminus of the protein, did not yield soluble proteins, therefore we focused on the SUMO-Sav-SOD (3,4 loop). The constructs that were evaluated are collected in Table 2. In order to obtain more comparable constructs to Sav-WT, the L124H of the SUMO-Sav-SOD was mutated back to L124. Furthermore, to streamline the process, by avoiding the UPL1 cleaving step, we investigated the influence of the SUMO tag on the protein yields and cloned a Sav-SOD version without the tag. The Sav-SOD production yields were very good with ±200 mg/L of culture, and the protein could be lyophilized for storage after purification and dialysis.

Table 2 Constructs prepared for the initial studies of the Sav-SOD variants.

Gene
SUMO–SODc–Circularly Permuted SAV (K121A, L124H)
SUMO–Sav (SODc between 5. and 6. β-sheet) (K121A, L124H)
SUMO–Sav (SODc between 3. and 4. β-sheet) (K121A, L124H)
SUMO–Sav (SODc between 7. and 8. β-sheet) (K121A, L124H)
SUMO–Sav (SODc between 4. and 5. β-sheet) (K121A, L124H)
SUMO–Circularly Permuted Sav–SODc (K121A, L124H)
Sav (SODc between 3. and 4. β-sheet) (K121A)

In order to verify the protein structure and topology, there are different methods, ranging from *in vitro* assays to X-ray crystallography. Each method has its advantages, and some are complimentary. By SDS-PAGE analysis, only an estimation for the size and no information of the binding stoichiometry can be obtained. Therefore, we used mass spectral analysis to verify the precise mass of the chimeric proteins. By mass spectral analysis, different information about the proteins can be gained, ranging from the protein sequence (e.g., tryptic digest) to the quaternary protein structure (native MS). We could show by native MS that the Sav-SOD formed tetramers (~80 kDa) and binds up to four biotins (Figure 19). Furthermore, initial crystallographic experiments revealed that the SOD loop of the Sav-SOD could not be resolved. This might be due to the high flexibility of the loop. In order to stabilize the structure of the DD in Sav-SOD, we introduced cysteines at the dimer/dimer interface of the SOD cap, in order to potentially form a disulfide bridge with the cysteine at the same position in the adjacent monomer, thereby rigidifying the cap. In order to find potential targets, a homology search of the SODc was performed and we searched for positions with natural diversity and amino acid residues of similar size to cysteine. We could identify in a related structure from a different *mycobacterium* a serine, which usually is considered to be an excellent candidate to be substituted by cysteine, at the equivalent position of SODc V15 (being position V62 in Sav-SOD). Furthermore, to keep the α-helical structure at the position selected for mutation, the V17 residue of SOD (V64 Sav-SOD) was mutated to a threonine. Having two cysteines in close proximity, in order to facilitate the formation of the disulfide bridge, the construct was expressed in *E.coli* SHuffel T7 cells, as they have been shown to have a more oxidative cytoplasmic environment, as well as special chaperones designed to "shuffle" S–H and S–S bonds, thereby favoring the creation of the bridge between the monomers of the construct.[89] By doing so, the protein yield was significantly lowered. Nevertheless, we obtained enough protein to run initial crystallographic experiments. Unfortunately, the introduction of the disulfide bridge did not help to stabilize the cap and the SOD-loop could not be resolved. Therefore, in order to gain some information about the structure of Sav-SOD, we teamed up with computational scientists, which modeled the structure based on a crystal structure.[36] A detailed discussion can be found in Section 3.2 Design and Evolution of Chimeric Streptavidin for Protein-Enabled Dual Gold Catalysis. With the model in hand, we inspected the protein for potential targets for saturation mutagenesis.
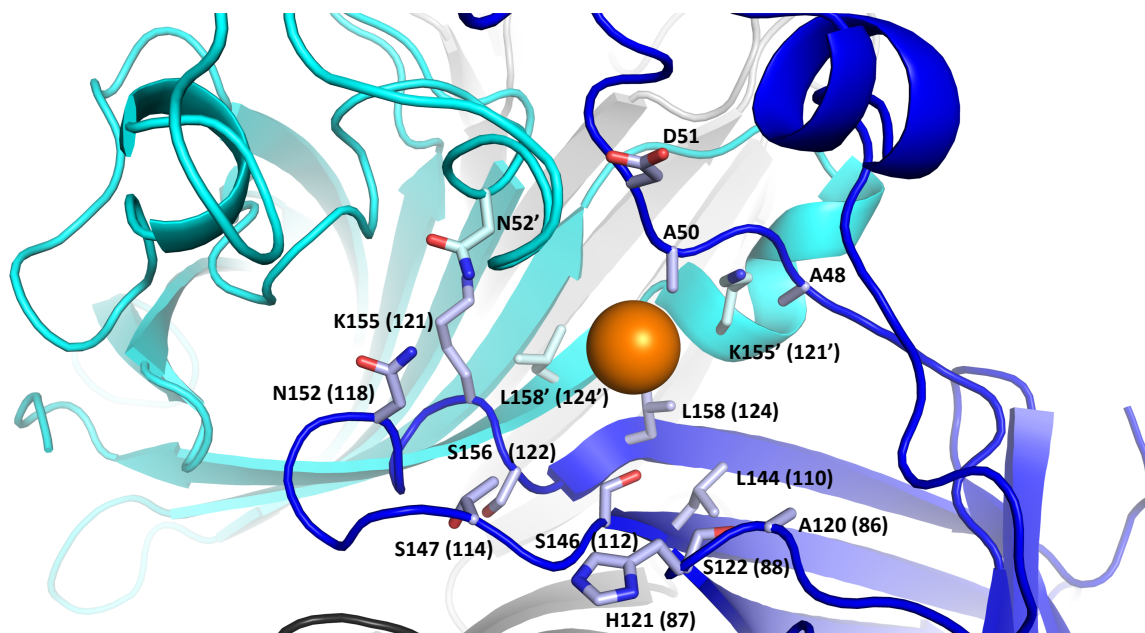
Figure 9 Close-up view of the biotin binding site with the averaged metal position. The monomeric chains of Sav-SOD are color coded (black, white, cyan and blue). For simplification, only one metal (orange sphere) is displayed. The residues with C-β atoms < 10 Å distance to the metal center are shown as sticks (carbon: green, oxygen: red and nitrogen: blue) and labeled according to their type of amino acid sidechain and number in sequence (the numbers in parenthesis correspond to the aa numbering scheme for the Sav devoid of the SOD lid).

We searched for C-β atoms of residues within 5 Å of an averaged metal center and could find one residue (A48), the first amino acid of the SOD-loop. "Zooming out" to 8 Å, we identified more residues that might interact with the metal center: N52' and K155' of the adjacent monomer and A50. At 10 Å, sixteen residues could be identified (Figure 9). Of these, 4 amino acid residues are part of the SOD-loop, which confirms that we gained at least four positions for targeted mutagenesis compared to the Sav-WT. In summary, we introduced the SODc DD to the 3,4-loop of Sav in order to shield its "active site" and created a new chimera which can be used as scaffold for ArMs (Figure 10).
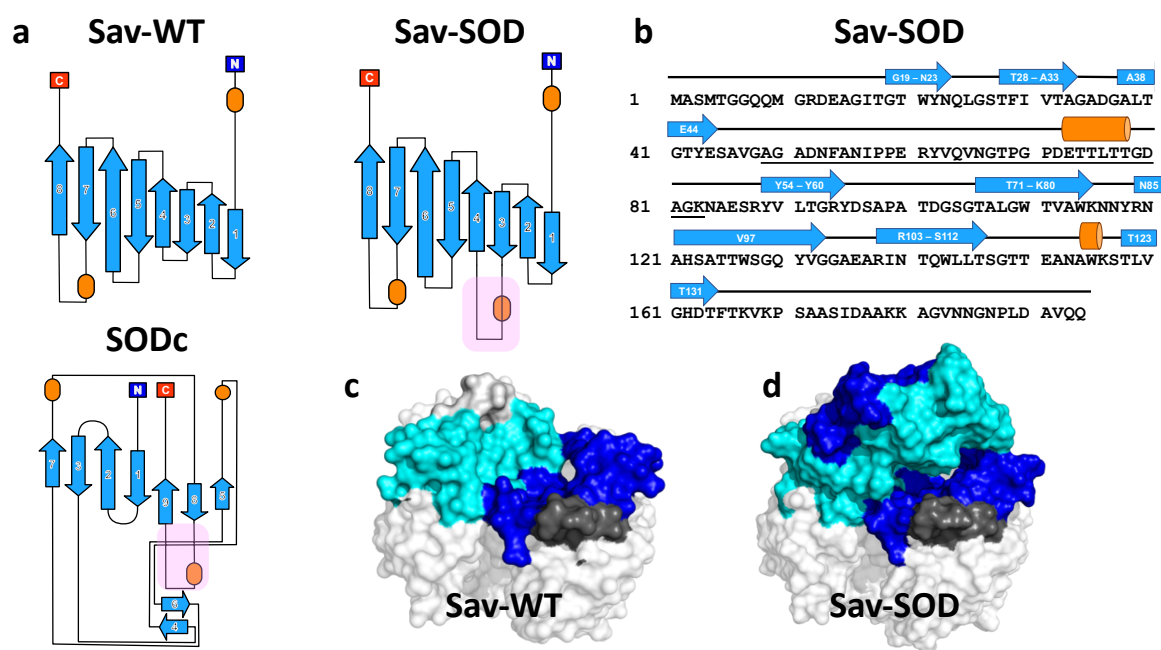
*Figure 10 **a** Origami scheme of Sav-WT, SODc and the chimeric Sav-SOD used throughout this thesis. The β-sheets are represented by arrows, the α-helixes by an orange oval. The dimerization domain of SODc is highlighted in pink. The origami were generated with http://munk.csse.unimelb.edu.au/pro-origami/. **b** Sequence of the Sav-SOD. The sequence of the SOD-lid part is underlined. **c** surface representation of the Sav tetramer (pdb 3pk2) with the biotin binding vestibule colored (each chain has another color, blue, cyan, dark grey and white). **d** surface representation of the Sav-SOD tetramer from the predicted model from Section 3.2 Design and Evolution of Chimeric Streptavidin for Protein-Enabled Dual Gold Catalysis with the same color code as **c**, highlighting the shielding of the "active site".*

One of the main reasons, next to expanding the secondary coordination sphere of the cofactor, to create the Sav-SOD, was to shield the metal cofactor from cellular components. We hypothesized that such a shielded environment may protect the cofactor from cellular components, thus enabling the screening directly from cell lysates.

It has been shown that the incorporation of the catalysts with reaction conditions in the cytoplasm (*in vivo*) (or in cell-free extracts) is challenging. Only a handful of reports on ArMs within living cells or in the presence of cell lysates have been reported to date.[90–93] However, the catalytic efficiency remains modest in most cases. Furthermore, there are several challenges to overcome like:

(I)   The inherent cytotoxicity of the abiotic catalyst (for *in vivo* catalysis only)

(II)   Ensuring that the metal cofactor is efficiently uptaken by the cell (for *in vivo* only)

(III)   Circumventing the deactivation of the catalyst by cellular components like thiols.

The uptake of the cofactor can be facilitated by engineering the host cells with a transporter protein. This has been shown to be very effective in case of a Ir(Me)-porphyrin IX CYP119 ArM

20

for carbene insertion.[19,33,94] Once the ArM is assembled, it needs to withstand the cellular components, especially glutathione (GSH), the most abundant low-molecular thiol in living aerobic cells. This thiol has been shown to be a major source of inhibition when performing transition metal catalysis *in vivo* or in cell-free extracts.[95] Nevertheless, Tanaka *et al.*[96], Hartwig *et al.*[94], Okamoto *et al.*[97] and Schwaneberg *et al.*[98] among others have shown that ArMs can be used *in vivo* for catalysis. Some of them by shielding the cofactor and some by localizing the ArM to different cellular compartments.

To overcome these challenges, our group relied mostly on three different strategies. Either we added suitable GSH-oxidizing agents to protect the cofactor[99], or secreted the protein to the periplasm[100–102] or displayed it on the surface[103] of *E. coli* cells. As the periplasmic protein accumulation has been shown to speed up the turnover of the directed evolution campaigns, we engineered the Sav-SOD to be secreted to the periplasm. Our group has previously shown that Sav can be exported efficiently by the SEC pathway, by fusion of the N-terminal signal peptide (MKKTAIAIAVALAGFATVAQA) of the outer membrane protein A (OmpA) to Sav and using modified Top10 (DE3) cells as expression host (Figure 11).[37] A big advantage of this strategy is that, after secretion of the unfolded protein to the periplasm, the signal peptide is cleaved off and the remaining Sav folds, so that the Sav construct is identical to the one in the cytosolic fraction. We therefore cloned the OmpA signal peptide to the N-terminus of the Sav-SOD gene and checked for protein production in the periplasm. By SDS-PAGE analysis we could show that the periplasmic fractionation contained Sav-SOD and maintained its biotin-binding activity (Figure 11).
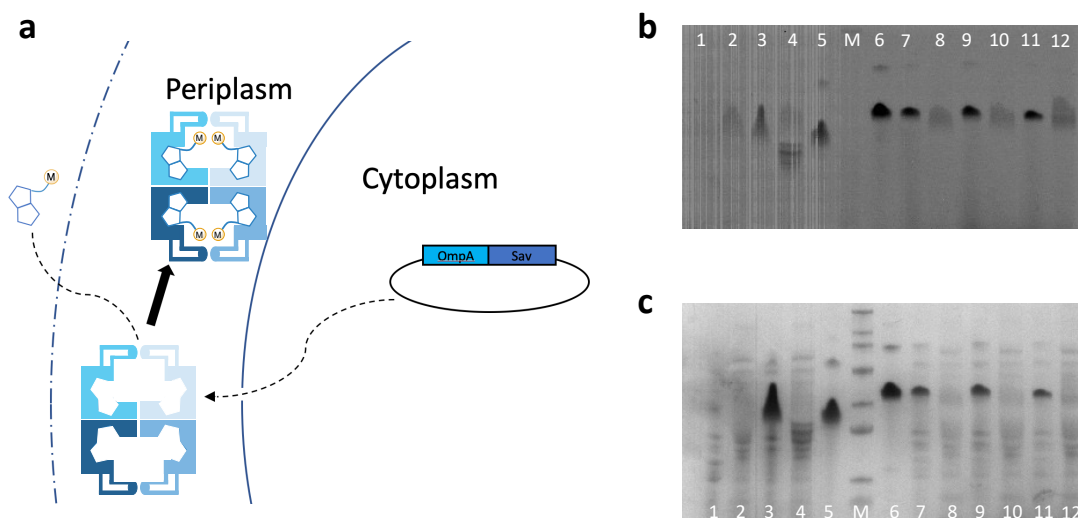
*Figure 11 **A** Sav-SOD is secreted to the periplasm by fusion to an N-terminal signal peptide from the outer membrane protein A (OmpA). **B** UV-light exposed SDS-PAGE gel in presence of B4F prepared according to Zhao et al.[101] **C** Same SDS-PAGE gel under white light. 1) periplasmic fraction of the empty vector cells 2) cytosolic fraction of the empty vector 3) periplasmic fraction Sav-WT 4) cytosolic fraction Sav-WT 5) Sav-WT refrence M) Marker (ID-MWBPS1-250) 6) Sav-SOD K121A reference 7) periplasmic fraction Sav-SOD WT 8) cytosolic fraction Sav-SOD WT 9) periplasmic fraction Sav-SOD K121A 10) cytosolic fraction Sav-SOD K121A 11) periplasmic fraction Sav-SOD S112A 12) cytosolic fraction Sav-SOD S112A.*

As mentioned in the introduction, the catalytic performance of some ArMs based on Sav highly depend on the biotinylated-cofactor:Sav ratio. In order to overcome this challenge, our group engineered the scdSav in which the biotin-binding capacity of two fo the four binding sites has been ablated.[69] Encouraged by the previous work with scdSav and the robustness of Sav-SOD against modifications, we envisioned that, by using the same linker and using different genes for each monomeric unit of the linked dimer, we could generate scdSav-SOD. Therefore, we ordered the gene construct and tested the protein production in different *E. coli* strains (Top10 (DE3), BL21 (DE3), ArcticExpress (DE3) and Lemo21 (DE3)) with manual induction, as well as autoinduction media, and different temperatures. Only Lemo21 (DE3) cells led to the production of scdSav-SOD in a reasonable yield (14 mg/L) after 48h at 25°C. By mass spectral analysis, the exact mass of the tetramer could be verified and by SDS-PAGE analysis the B4F binding was observed. We conclude that the construct maintains its ability to bind biotin (Figure 20-21). To further reduce the flexibility of the SOD lid, we teamed up

22

with the Correia Lab from the École polytechnique fédérale de Lausanne (EPFL). They suggested ten slightly modifed SOD sequences, which are presented in Figure 12.

		variation		variation		
SavSOD	30	VTAGADGALTGTY**ESAVGAGADNF**ANIPPERYVQVNGTPGPDETTLT**TGDAGKNAESR**YVLTGRY				95
SODL1	30	VTAGADGALTGTY**QPKDVAKSAST**ANIPPERYVQVNGTPGPDETTLT**TNQDDKKQTSGP**FVLTGRY				95
SODL2	30	VTAGADGALTGTY**QPKPVADYKFT**ANIPPERYVQVNGTPGPDETTLT**TDQDDKKQSSGP**YVLTGRY				95
SODL3	30	VTAGADGALTGTY**QPKPVADYKES**ANIPPERYVQVNGTPGPDETTLT**TRQDDKKQSSGP**YVLTGRY				95
SODL4	30	VTAGADGALTGTY**QPKPVADYKFT**ANIPPERYVQVNGTPGPDETTLT**TNTDDKNQSSGP**YVLTGRY				95
SODL5	30	VTAGADGALTGTY**QPKNVAQPAAT**ANIPPERYVQVNGTPGPDETTLT**TNTDDKNQHPGP**FVLTGRY				95
SODL6	30	VTAGADGALTGTY**QPKDVADYKWA**ANIPPERYVQVNGTPGPDETTLT**TNQDDKKQSSGP**YVLTGRY				95
SODL7	30	VTAGADGALTGTY**QPKPVADYKES**ANIPPERYVQVNGTPGPDETTLT**TRQDDKNQSSGP**YVLTGRY				95
SODL8	30	VTAGADGALTGTY**ASTGTEDPRDS**ANIPPERYVQVNGTPGPDETTLT**TGHPPTSPDQSR**FVLTGRY				95
SODL9	30	VTAGADGALTGTY**ASTGTEDPRDS**ANIPPERYVQVNGTPGPDETTLT**TGHPPVPPDQSR**FVLTGRY				95
SODL10	30	VTAGADGALTGTY**ASTGTEDPRDS**ANIPPERYVQVNGTPGPDETTLT**TGHPPTSPDQTR**FVLTGRY				95
		************* ********************** :******				

Figure 12 Loop linker variations designed to limit the flexibility of the SOD-loop. The sequence of the starting and ending amino acids of the DD are varied and are highlighted in the orange boxes.

These constructs were cloned to enable periplasmic localization of the protein by fusing the OmpA sequence in the same way as for Sav-SOD. Having the new Sav-SOD chimeras at hand, we selected the ATHase based on [Cp*Ir(biot-*p*-L)Cl], as well as the a gold (I) catalyzed hydroamination. The results of the studies thereof are summarized in the following chapters.

2.2.4 Experimental

General information. Water used for molecular biology and in the catalytic reactions was purified by Milli-Q Advantage system. High-resolution mass spectrometry (HR-MS) was performed on a Bruker maXis II QTOF ESI mass spectrometer coupled to a Shimadzu LC. Native-MS was performed via direct injection (aqueous buffer) to a Bruker maXis II QTOF ESI mass spectrometer. Molecular biology reagents were purchased from New England Biolabs (NEB), Integrated DNA Technologies (IDT), and Macherey-Nagel and were used as described in the accompanying protocols. PCR reactions were performed with an Eppendorf

23

Mastercycler Gradient. Mutations were verified by Sanger sequencing performed by Microsynth (Balgach, Switzerland).

Cloning. Q5® Hot Start High-Fidelity DNA Polymerase (NEB, cat. no. M0493) was used with the standard protocol and the corresponding T_m for the desired primers.

Table 3 Q5® Hot Start High-Fidelity DNA Polymerase protocol.

Compound	Volume (µl)	Final concentration
mQ-H$_2$O	22.5	
Template plasmid (100 ng/ µl)	1	100 ng/µl
Q5 polymerase 5x buffer	10	1x
Q5 5x GC enhancer	10	1x
dNTPs (10 mM)	1	200 µM each
Primers (10 µM)	2.5	0.5 µM
Q5® Hot Start polymerase	0.5	0.02 U/µl

Table 4 Thermocycling Conditions for routine PCR

Step	Temperature (°C)	Time (sec.)
Initial denaturation	98	30
	98	10
25-30 cycles	50-72[*]	30
	72	30
Final extension	72	120
Hold	4	Until next step

To minimize the background template DNA, DPNI digestion was performed followed by either a heat-shock transformation into chemically competent Top10 (DE3) cells, or further digestion and ligation reactions were performed. Mini-prep (Sigma, GenElute™ Plasmid Miniprep Kit) was performed for each sample according to the protocol. The sequence of the modified plasmids was verified by Sanger Sequencing (Microsynth, Balgach).

***E.Coli* Strains.** The following strains were used throughout the thesis, depending on their application; I) *E. coli* TOP10 DE3: F⁻ mcrA Δ(mrr-hsdRMS-mcrBC) φ80lacZΔM15 Δlac^X74 nupG recA1 araD139 Δ(ara-leu)7697 galE15 galK16 rpsL(Str^R) endA1 λ⁻, λ(DE3))[37] for cloning approaches and periplasmic protein production II) *E. coli* BL21 DE3 (NEB, cat. no. C2527): B F⁻dcm ompT hsdS(rB⁻mB⁻) gal λ(DE3), III) *E. coli* ArcticExpress DE3 (Agilent, cat. no. #230192) B F⁻ ompT hsdS(r_B⁻ m_B⁻) dcm⁺ Tet^r gal λ(DE3) endA Hte [cpn10 cpn60 Gent^r], IV) *E. coli* SHuffle® T7 (NEB, cat. no. C3026J) fhuA2 lacZ::T7 gene1 [lon] ompT ahpC gal λatt::pNEB3-r1-cDsbC (Spec^R, lacI^q) ΔtrxB sulA11 R(mcr-73::miniTn10--Tet^S)2 [dcm] R(zgb-210::Tn10 --Tet^S) endA1 Δgor Δ(mcrC-mrr)114::IS10, V) Lemo21 DE3 (NEB, cat. no. C2528J), fhuA2 [lon] ompT gal (λ DE3) [dcm] ΔhsdS/pLemo(Cam^R) λ DE3 = λ sBamHIo ΔEcoRI-B int::(lacI::PlacUV5::T7 gene1) i21 Δnin5 pLemo = pACYC184-PrhaBAD-lysY for (large scale) cytosolic protein production.

Lysogeny Broth (LB) medium. 10 g of tryptone (1 % w/v), 5 g of yeast extract (0.5 % w/v), and 5 g of NaCl (0.5 % w/v) are dissolved in mQ-H2O (final volume of 1 L), autoclaved (20 min at 121 °C, 1.5 bar)

LB agar plates: In addition to the LB components 15 g of agar (1.5 % w/v) is added to 1 L medium and autoclaved.

20X ZYP 5052 salts. 136 g of KH$_2$PO$_4$ (50 mM), 142 g of solved in mQ-H$_2$O (final volume of 1 L) and autoclaved.

20X ZYP 5052 sugars. 100 g of glycerol (10 % v/v), 11 g of glucose- monohydrate (4 % w/v) are dissolved in mQ-H2O (final volume of1 L) and autoclaved.

200 mM MgCl$_2$ stock solution. 3.81 g of MgCl$_2$ is dissolved in 200 ml mQ-H$_2$O and autoclaved.

200 mM MgSO$_4$. 4.8 g of anhydrous MgSO$_4$ is dissolved in mQ-H$_2$O (final volume of 200 ml) and autoclaved.

Iminobiotin binding (IBB) buffer, pH 10.8. 438 g of NaCl (final concentration 0.5 M) and 63 g of NaHCO$_3$ (final concentration 50 mM) are dissolved in 15 L mQ-H$_2$O. The pH is adjusted to 10.8 with 5M NaOH and the solution is stored at 4 °C.

Elution buffer. 20 ml of glacial acetic acid (1 % v/v) are mixed with mQ- H$_2$O.

B4F stock solution. 10 mg of B4F (CAS: 1032732–74-3, MW: 644.7 g/mol) were dissolved in 25.85 ml of DMSO (0.6 mM).

Protein production tests. *E. coli* cells were transformed with the plasmid and selection on an agarose plate with Kanamycin (50 mg/L) was undertaken. A single colony of a fresh transformant was used for an overnight culture and was grown in Luria–Bertani (LB) medium (37 °C, 300 rpm), supplemented with glucose (1 %) and kanamycin (50 µg/mL). For small scale production, a 96-well plate containing 500 ml modified ZYP-5052 medium[104] supplemented with kanamycin (50 µg/mL) in each well, was inoculated with a single colony per well. The plate was shaken at 30°C for 24h. The cells were harvested by centrifugation (4 °C, 12000 g, 10 min), the supernatant was discarded and the cell pellet was frozen in liquid nitrogen. After 3 cycles of freezing and thawing, the cell pellet was resuspended in lysis mix (100 µl/pellet, composed of lysis buffer (1 ml, 50 mM Na phosphate, 300 mM NaCl, 0.5 % DM, pH = 7-8), enzyme mix (20 µl, 100:10:10 mg/mL lysozyme:DNase:RNase) and Mg acetate solution (5 µl, 500 mM)). The soluble and insoluble parts of the cell lysate were subsequently analyzed by SDS-PAGE.

Protein refolding. After cell lysis, the samples were centrifuged and the pellet was resuspended in 8 M urea. Tris-HCl (pH 8.0) was added to the solution until the end concentration of 50 mM was reached. The mixtures were dialyzed in 50 mM Tris-HCl (pH 8.0) at 4 °C overnight with gentle stirring. The mixture was centrifuged and the supernatant was used for SDS-PAGE analysis

SDS sample preparation. To 19 µl of the sample 1µl of the B4F solution was added, followed by the addition of 4 µl 6X SDS sample buffer solution. Load to SDS-Gel

Removal of the SUMO-tag by ULP1. To 19 µl of the sample ULP1 protease was added to a final concentration of 0.13 mg/mL. After 30 min at room temperature the sample was prepared for SDS analysis as described above.

Large scale protein production and purification. The expression and purification of Sav-SOD proteins were achieved using auto induction media (ZYP-5052) followed by cell lysis and purification using iminobiotin agarose affinity chromatography as previously described.[105] In short, after chemical cell lysis, the supernatant was equilibrated in IBB overnight, followed by

an affinity purification with an 2-iminobiotin agarose column coupled to an ÄKTA prime Plus. After elution with 1 % v/v acetic acid, the collected fractions were pooled and dialyzed at least 3 times in 15 L of mQ-H$_2$O, followed by a freezing step and consequently lyophilized.

SDS-PAGE. SDS-PAGE analysis was performed according to a published procedure.[87] Hand-casted 12 % acrylamide gels were used. After B4F visualization the gels were stained with Coomassie Brilliant blue R-250 dye in MeOH and acetic acid, followed by de-staining.

Western-Blot. The gel from SDS-PAGE analysis was transferred (prior to staining and without addition of B4F) to a nitrocellulose membrane at 45 V overnight. After completion of the transfer, the membrane was rinsed with water and blocked with TBST (Tris-buffered saline + 0.5 % Tween 20) containing 4 % BSA (25 ml) at RT for 1.5 h. The membrane was rinsed twice with TBST and incubated in fresh TBST (20 ml) containing the primary antibody (100 µL, Streptavidin Antibody, pAb, Rabbit, GenScript, Cat. no. A00621, 1:2500). After thorough washing with TBST (4 x 10 min), the membrane was incubated in fresh TBST (20 ml) containing the secondary antibody fused with HRP (10 µL, Abcam, Cat. no. ab7403, 1:10'000). After washing the membrane with TBST (4 x 10 min) the detection was carried out according to the manufacturer's protocol (1-Step™ TMB-Blotting Substrate Solution, Thermo Scientific, Cat. no. 34018).

Sequence of the Sav-SOD and cpSav-SOD constructs. All the gene of interest were cloned into the pET-28a expression vector, except the scdSav-SOD gene, which was cloned into the pRSFduet-1 vector.

SUMO_Sav_SOD (3,4-loop)

```
ATGGGCTCGGACTCAGAAGTCAATCAAGAAGCTAAGCCAGAGGTCAAGCCAGAAGTCAAGCCTGAGACTCACATCAATTTAAA
GGTGTCCGATGGATCTTCAGAGATCTTCTTCAAGATCAAAAAGACCACTCCTTTAAGAAGGCTGATGGAAGCGTTCGCTAAAA
GACAGGGTAAGGAAATGGACTCCTTAAGATTCTTGTACGACGGTATTAGAATTCAAGCTGATCAGACCCCTGAAGATTTGGAC
ATGGAGGATAACGATATTATTGAGGCTCACAGAGAACAGATTGGTGGTGCAAGCATGACCGGTGGCCAGCAGATGGGTCGTGA
TCAGGCAGGTATTACCGGCACCTGGTATAATCAGCTGGGTAGCACCTTTATTGTTACCGCAGGCGCAGATGGTGCACTGACCG
GTACGTATGAAAGCGCAGTTGGTGCCGGCGCCGACAACTTTGCCAACATTCCGCCAGAACGCTACGTCCAGGTCAATGGGACT
CCGGGTCCCGACGAGACGACGTTGACCACCGGCGACGCCGGCAAGAATGCAGAAAGCCGTTATGTTCTGACAGGTCGTTATGA
TAGCGCACCGGCAACCGATGGTAGCGGCACCGCACTGGGTTGGCAAGAATGCAGAAAGCCGTTATGTTCTGACAGGTCGTTAT
GATAGCGCACCGGCAACCGATGGTAGCGGCACCGCACTGGGTTGGACCGTTGCATGGAAAAATAACTATCGTAATGCACATAG
```

CGCAACCACCTGGTCAGGTCAGTATGTTGGTGGTGCAGAAGCACGCATTAATACCCAGTGGCTGCTGACCAGCGGCACtactG

AAGCAAATGCCTGGaaAAGCACCCTGGTTGGTCATGATACCTTTACCAAAGTTAAACCGAGCGCAGCAAGCATTGATGCAGCA

AAAAAAGCCGGTGTGAATAATGGTAATCCGCTGGATGCAGTTCAGCAGTAA

Legend for SUMO_SOD_cpSav: SUMO_Tag, Sav_Domain; sodC_Dimerization Domain.

SUMO_SOD_cpSav

ATGGGCTCGGACTCAGAAGTCAATCAAGAAGCTAAGCCAGAGGTCAAGCCAGAAGTCAAGCCTGAGACTCACATCAATTTAAA

GGTGTCCGATGGATCTTCAGAGATCTTCTTCAAGATCAAAAAGACCACTCCTTTAAGAAGGCTGATGGAAGCGTTCGCTAAAA

GACAGGGTAAGGAAATGGACTCCTTAAGATTCTTGTACGACGGTATTAGAATTCAAGCTGATCAGACCCCTGAAGATTTGGAC

ATGGAGGATAACGATATTATTGAGGCTCACAGAGAACAGATTGGTGGTACTAGTGCCGGCGCCGACAACTTTGCCAACATTCC

GCCAGAACGCTACGTCCAGGTCAATGGGACTCCGGGTCCCGACGAGACGACGTTGACCACCGGCGACGCCGGCAAGGGCAGCG

AAAGCCGTTATGTTCTGACAGGTCGTTATGATAGCGCACCGGCAACCGATGGTAGCGGCACCGCACTGGGTTGGACCGTTGCA

TGGAAAAATAACTATCGTAATGCACATAGCGCAACCACCTGGTCAGGTCAGTATGTTGGTGGTGCAGAAGCACGCATTAATAC

CCAGTGGCTGCTGACCAGCGGCACCACCGAAGCAAATGCCTGGGCAAGCACCCTGGTTGGTCATGATACCTTTACCAAAGTT

Legend for SUMO_SOD_cpSav: SUMO_Tag, cpSav_Domain; sodC_Dimerization_Domain.

Sav_SOD (3,4-loop)

CGTGATCAGGCAGGTATTACCGGCACCTGGTATAATCAGCTGGGTAGCACCTTTATTGTTACCGCAGGCGCAGATGGTGCACT

GACCGGTACGTATGAAAGCGCAGTTGGTGCCGGCGCCGACAACTTTGCCAACATTCCGCCAGAACGCTACGTCCAGGTCAATG

GGACTCCGGGTCCCGACGAGACGACGTTGACCACCGGCGACGCCGGCAAGAATGCAGAAAGCCGTTATGTTCTGACAGGTCGT

TATGATAGCGCACCGGCAACCGATGGTAGCGGCACCGCACTGGGTTGGACCGTTGCATGGAAAAATAACTATCGTAATGCACA

TAGCGCAACCACCTGGTCAGGTCAGTATGTTGGTGGTGCAGAAGCACGCATTAATACCCAGTGGCTGCTGACCAGCGGCACTA

CTGAAGCAAATGCCTGGAAAAGCACCCTGGTTGGTCATGATACCTTTACCAAAGTTAAACCGAGCGCAGCAAGCATTGATGCA

GCAAAAAAAGCCGGTGTGAATAATGGTAATCCGCTGGATGCAGTTCAGCAGTAA

Legend for Sav_SOD: Sav_Domain; sodC_Dimerization_Domain

scdSav-SOD (SavA S112S K121A, SavB S112S K121K)

ATGGCAAGCATGACCGGTGGCCAGCAGATGGGTCGTGATCAGGCAGGTATTACCGGCACCTGGTATAATCAGCTGGGTAGCAC

CTTTATTGTTACCGCAGGCGCAGATGGTGCACTGACCGGTACGTATGAAAGCGCAGTTGGTGCCGGCGCCGACAACTTTGCCA

ACATTCCGCCAGAACGCTACGTCCAGGTCAATGGGACTCCGGGTCCCGACGAGACGACGTTGACCACCGGCGATGCCGGCAAG

AATGCAGAAAGCCGTTATGTTCTGACAGGTCGTTATGATAGCGCACCGGCAACCGATGGTAGCGGCACCGCACTGGGTTGGAC

CGTTGCATGGAAAAATAACTATCGTAATGCACATAGCGCAACCACCTGGTCAGGTCAGTATGTTGGTGGTGCAGAAGCACGCA

TTAATACCCAGTGGCTGCTGACCAGCGGCACCACCGAAGCAAATGCCTGGGCAAGCACCCTGGTTGGTCATGATACCTTTACC

AAAGTTAAACCGAGCGCAGCAAGCATTGATGCAGCAAAAAAAGCCGGTGTGAATAATGGTAATCCGCTGGATGCAGTTCAGCA

GGGATCCGGTGGCGGTAACGGTGGGGGAAACGGTGGCGGAAATGGCGGAGGGAACATTGATGGTCGCGGTGGTGGTAATGCTA

GCATGACTGGTGGACAGCAAATGGGTCGGGATCAGGCCGGCATAACCGGCACCTGGTACGCCCAGCTCGGCGATACCTTCATC

GTGACCGCGGGCGCCGACGGCGCCCTGACCGGAACCTATGAAAGCGCAGTTGGTGCCGGCGCCGACAACTTTGCCAACATTCC

GCCAGAACGCTACGTCCAGGTCAATGGGACTCCGGGTCCCGACGAGACGACGTTGACCACCGGCGACGCCGGCAGAATGCAG

AAAGCCGTTATGTCCTGACCGGTCGTTACGACAGCGCCCCAGCCACCGACGGCTCTGGCACCGCCCTCGGTTGGACGGTGGCC

TGGAAGAACAATTACAGAAACGCCCACTCCGCGACCACGTGGAGCGGCCAATACGTCGGCGGCGCCGAGGCGAGGATCAACAC

ACAATGGTTATTAACATCTGGAACTACTGAGGCCAACGCATGGAAGTCCACGCTGGTCGGCTGCGcCACCTTCACCAAGGTGA

AGCCTTCCGCCGCCTCAATCGACGCGGCGAAGAAGGCTGGCGTCAACAACGGCAACCCTCTCGACGCCGTACAACAATAA

Legend for Sav_SOD: scdSav_1_Domain; sodC_Dimerization_Domain 1; linker; scdSav_2_Domain; sodC_2_Dimerization_Domain

Primers.

SUMOless_Rev	GGTATATCTCCTTCTTAAAGTTAAACAAAATTATTTCTAGAGGG
SUMOless_Fwd	ATGGCAAGCATGACCGGTG
SODc_V15C_V17T_Fwd	AGAACGCTACtgcCAGACCAATGG
SODc_Rev	GGCGGAATGTTGGCAAAG
SUMO_cpSav_SOD_1_Fwd	tataACTAGTGAAAGCCGTTATGTTCTGACAG
SUMO_cpSav_SOD_1_Rev	tataACTAGTTGCGCTTTCATACGTACCGG
SUMO_cpSav_SOD_2_Fwd	tattAGGCCTttaGCTTCCCTTCCCCGCATCACCGGTGGTCAACGTCG
SUMO_cpSav_SOD_2_Rev	ataAGGCCTGAGCACCACCACCACCACCACTG
L124L_Fwd	CCAACCAGGGTGCTTGCCCAGGCATTTGC
L124L_Rev	CAAGCACCCTGGTTGGTCATGATACCTTTACC

Crystallization and data processing. Purified proteins were dissolved with a concentration of 20 mg/mL in mQ-H_2O for initial crystallization trials. Commercial screens were used to identify suitable crystallization conditions (Crystal Screen HT (Hampton Research, HR2-130), PEG Ion Screen HT (Hampton Research, HR2-139), Grid Screen MPD (Hampton Research, HR2-215), Index HT (Hampton Research HR2-134), PACT premier™ HT (Molecular Dimensions, MD1-36) and JCSG-plus™ (Molecular Dimensions, MD1-37). The crystallization plates were either set up manually or by a pipetting robot (Gryphon, Art Robbins Instruments). Manual set up for sitting drop vapor diffusion in Swissci 96-Well 3-Drop Plates (Molecular Dimensions, MD11-003) was done the following: The protein solution (1 µL) was mixed with the precipitation buffer (1 µL). The drop was equilibrated against a reservoir of the precipitation buffer (40 µL at 20 °C). The Gryphon was programmed to pipette the plates as follows: I) 40 µl of the precipitation solution was added to each reservoir of the plate, II) 0.3 µl of the precipitation buffer was added to each sitting drop, III) 0.3 µl of the protein solution was added to the sitting drop and IV) the plate was sealed and stored in the RockImager (Formulatrix), where pictures were taken in regular intervals.

Protein-crystal diffraction data were collected (at 100 K) at the Swiss Light Source beam line PSI and PSIII at a wavelength of 1.0 Å. Crystal indexing, integration and scaling[106,107] were carried out with the program XDS[108] and AIMLESS[109] using the graphical interface CCP4i2.[110] The structures was solved by molecular replacement using the program PHASER MR[111] and structure pdb: 3pk2, devoid of the Ir-cofactor and water molecules. For the structure

refinement, REFMAC[112,113] was used. For structure modelling, water picking and electron-density visualization the software COOT[114] was used. Figures were generated with PyMOL (the PyMOL Molecular Graphics System, Version 1.821, Schrödinger, LLC).

Protein ESI-MS characterization. Proteins were dissolved in Mili-Q water, 0.1% formic acid pH = 2.5 with a final concentration of 0.2 mg/mL and clarified by centrifugation. A HPLC (Shimadzu, equipped with a Jupiter® 5 μm C4 300 Å)-ESI QTof (Bruker maXis II QTOF ESI) system was used to record the data. The ESI-QTof mass spectrometer was calibrated with ESI-Tof TuneMix (Agilent). The charge envelope from 800-1200 m/z was de-convoluted using the Compass DataAnalysis software (Bruker Daltonics) with the Maximum Entropy set-up. The results are summarized in the HR-MS section.

Table 6. MS settings for protein mass spectral analysis.

Source Type	ESI	Ion Polarity	Positive	Set Nebulizer	1.4 Bar
Focus	Active	**Set Capillary**	4500 V	**Set Dry Heater**	220 °C
Scan Begin	500 m/z	**Set End Plate Offset**	-500 V	**Set Dry Gas**	9.0 l/min
Scan End	6000 m/z	**Set charging Voltage**	2000 V	**Set Divert Valve**	Source
		Set Corona	0 nA	**Set APCI Heater**	0 °C

Native MS analysis. The experiments were adapted from the studies of biotin binding of Sav described by Eckart *et al.*[115] To assemble the Sav-biotin complex, the protein was dissolved in ammonium acetate buffer (50 mM, pH 6.8) to a concentration of 2 mg/mL, supplemented with 0.1 mM biotin and incubated while shaking (1 h, 25°C, 600 rpm). The protein concentration was determined with a NanoDrop microvolume spectrophotometer (Thermo Fisher Scientific, USA). Electrospray ionization (ESI) low concentration tuning mix from Agilent was used as calibrant. The samples were directly injected into the high-resolution MS (HRMS) with a syringe pump using a flow rate of 5 μL/min. The ESI source parameters were optimized for the streptavidin analysis. The charge-state distribution from 2200-3000 m/z was deconvoluted using the Compass Data Analysis software (Bruker Daltonics) with the maximum entropy setup.

Table 7 MS settings for native MS analysis of the Sav-biotin, Sav-SOD-biotin complexes.

Source Type	ESI	Ion Polarity	Positive	Set Nebulizer	1.8 Bar
Focus	Active	Set Capillary	4000 V	Set Dry Heater	220 °C
Scan Begin	1000 m/z	Set End Plate Offset	-500 V	Set Dry Gas	3.0 l/min
Scan End	12000 m/z	Set charging Voltage	2000 V	Set Divert Valve	Source
		Set Corona	0 nA	Set APCI Heater	0 °C

Figure 13 SDS-PAGE gel images of the test protein production. L) Protein Ladder (NEB, *P7712S*) ***a*** *B4F biotin binding assay* ***b*** *Coomassie blue staining 1) total fraction SUMO-SODc-cpSav 2) soluble fraction SUMO-SODc-cpSav 3) insoluble fraction SUMO-SODc-cpSav 4) total fraction SUMO-Sav-SOD (5,6-loop) 5) soluble fraction SUMO-Sav-SOD (5,6-loop) 6) insoluble SUMO-Sav-SOD (5,6-loop) 7) total fraction SUMO-Sav-SOD (3,4-loop) 8) soluble fraction SUMO-Sav-SOD (3,4-loop) 9) insoluble fraction SUMO-Sav-SOD (3,4-loop) 10) total fraction SUMO-Sav-SOD (7,8-loop) 11) soluble fraction SUMO-Sav-SOD (7,8-loop) 12) insoluble fraction SUMO-Sav-SOD (7,8-loop) 13) total fraction SUMO-Sav-SOD (4,5-loop) 14) soluble fraction SUMO-Sav-SOD (4,5-loop) 15) insoluble fraction SUMO-Sav-SOD (4,5-loop). In yellow the corresponding bands of SUMO-Sav-SOD are marked. In red the corresponding bands for SUMO-Sav SODc variants are marked.*

*Figure 14 SDS-PAGE gel images of the test protein production. L) Protein Ladder (Eurogentec, ID-MWBRU1-250) **a** B4F biotin binding assay **b** Coomassie blue staining 1) SUMO-SODc-cpSav soluble after lysis 2) total fraction after lysis of SUMO-SODc-cpSav 3) SUMO-SODc-cpSav soluble after refolding 4) SUMO-SODc-cpSav insoluble fraction after refolding 5) SUMO-Sav-SOD (5,6-loop) soluble after lysis 6) total fraction after lysis of SUMO-Sav-SOD (5,6-loop) 7) SUMO-Sav-SOD (5,6-loop) soluble after refolding 8) SUMO-Sav-SOD (5,6-loop) insoluble fraction after refolding. 9) SUMO-Sav-SOD (7,8-loop) soluble after lysis 10) total fraction after lysis of SUMO-Sav-SOD (7,8-loop) 11) SUMO-Sav-SOD (7,8-loop) soluble after refolding 12) SUMO-Sav-SOD (7,8-loop) insoluble fraction after refolding. 13) SUMO-Sav-SOD (4,5-loop) soluble after lysis 14) total fraction after lysis of SUMO-Sav-SOD (4,5-loop) 15) SUMO-Sav-SOD (4,5-loop) soluble after refolding 16) SUMO-Sav-SOD (4,5-loop) insoluble fraction after refolding. In yellow, the corresponding bands of SUMO-Sav-SOD are marked. In red, the corresponding bands for SUMO-Sav SODc variants are marked.*
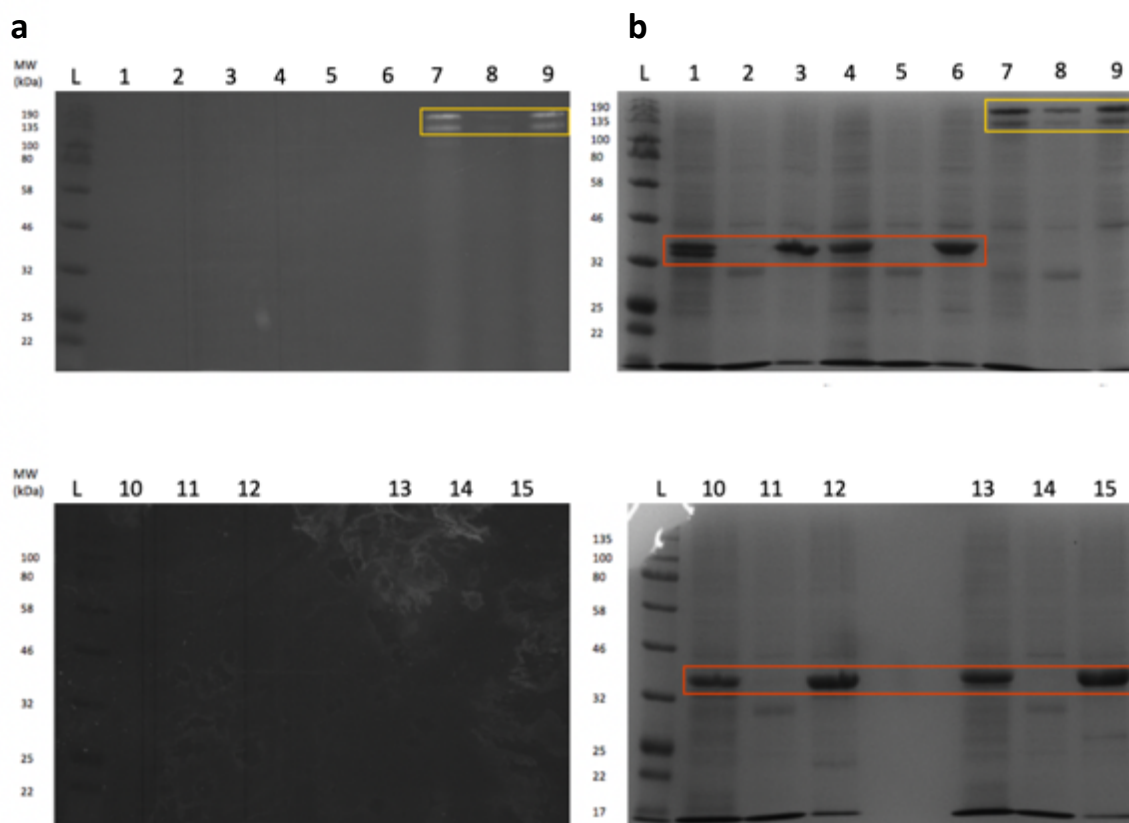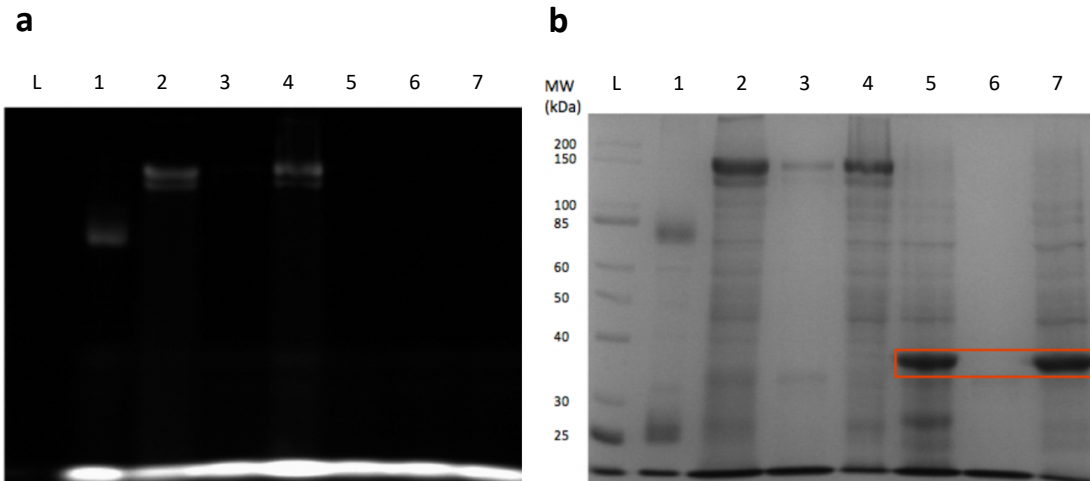
*Figure 15 SDS-PAGE gel images of the test protein production. L) Protein Ladder (Eurogentec, ID-MWBRU1-250) **a** B4F biotin binding assay **b** Coomassie blue staining 1) pellet of Sav-SOD (4,5-loop) after SUMO cleavage 2) total fraction of SUMO-SODc-cpSav 3) soluble fraction of SUMO-SODc-cpSav 4) insoluble fraction of SUMO-SODc-cpSav 5) total fraction of SUMO-cpSav-SODc 6) soluble fraction of SUMO-cpSav-SODc 7) insoluble fraction of SUMO-cpSav-SODc. In red, the corresponding bands for SUMO-Sav SODc variants are marked.*
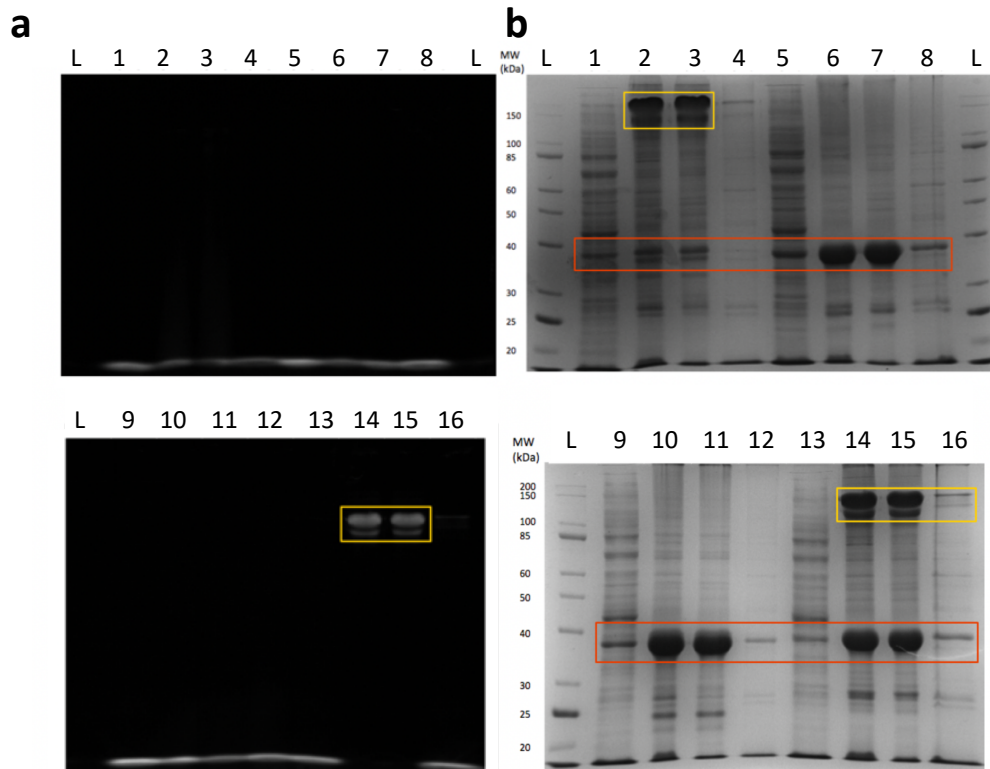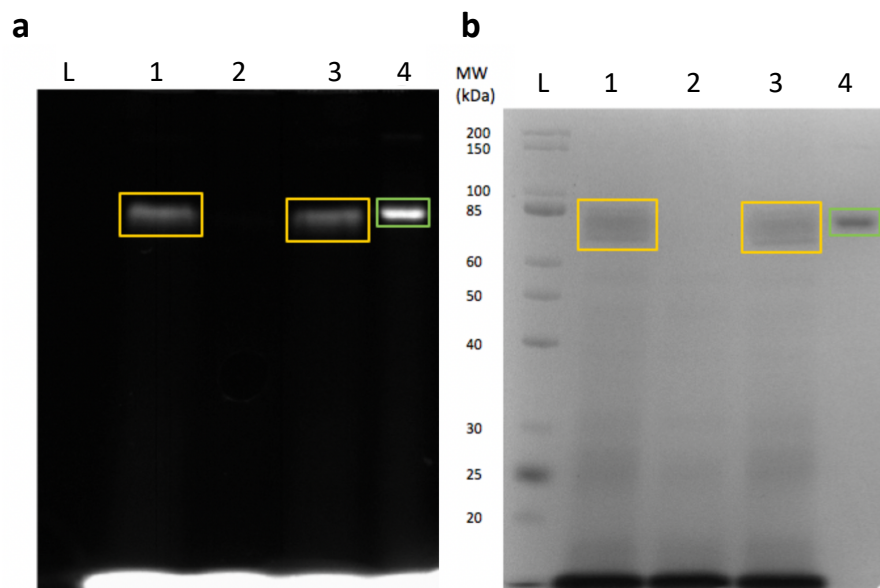

*Figure 16 SDS-PAGE gel images of the test protein production. L) Protein Ladder (Eurogentec, ID-MWBRU1-250) **a** B4F biotin binding assay **b** Coomassie blue staining 1) Sav-SOD (3,4-loop) total 2) Sav-SOD (3,4-loop) insoluble 3) Sav-SOD (3,4-loop) soluble 4) SUMO-Sav-SODc (3,4-loop) after SUMO cleavage. In yellow the corresponding bands of Sav-SOD are marked. In green the biotin binding band of the the Sav-SODc (3,4-loop) after SUMO cleavage is shown.*
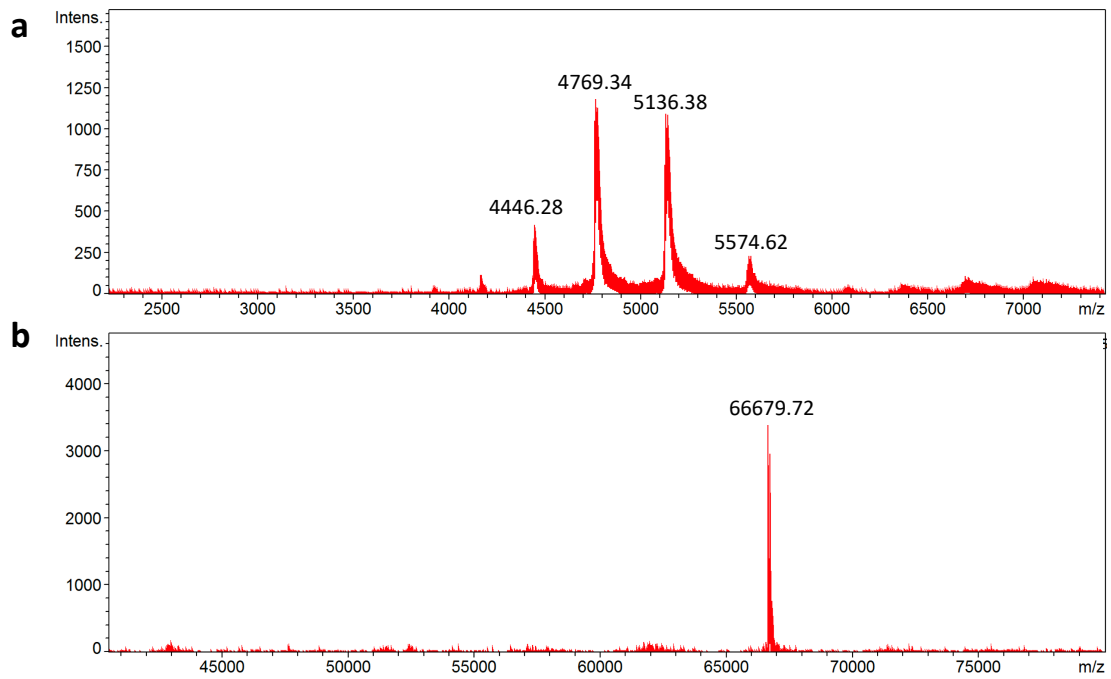
Figure 17 Native MS spectrum of SavWT with four biotins bound. **a** *Charge state envelope of the complex with the charges +15, +14, +13 and +12.* **b** *Deconvoluted spectrum. The calculated mass of Sav-WT + 4 Biotins is:* **66676.96 Da.**
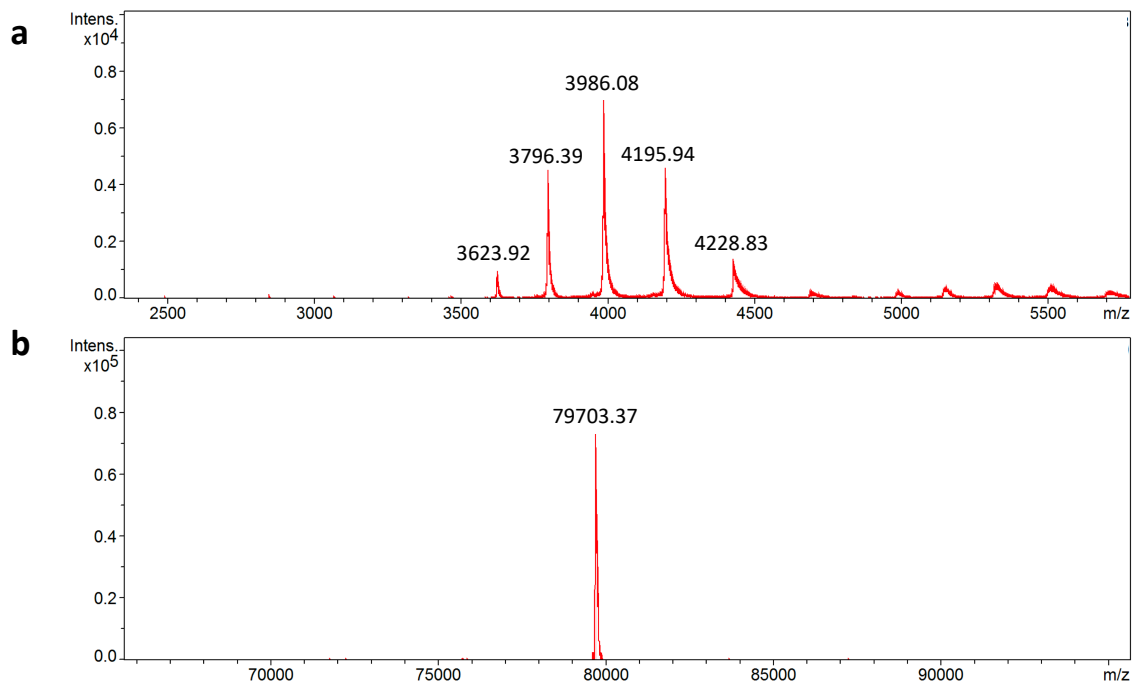
Figure 18 Native MS spectrum of Sav-SOD K121A. **a** *Charge state envelope of the complex with the charges +22, +21, +20, +19 and +18.* **b** *Deconvoluted spectrum. The calculated mass of Sav-SOD K121A is: 79698.56 Da.*
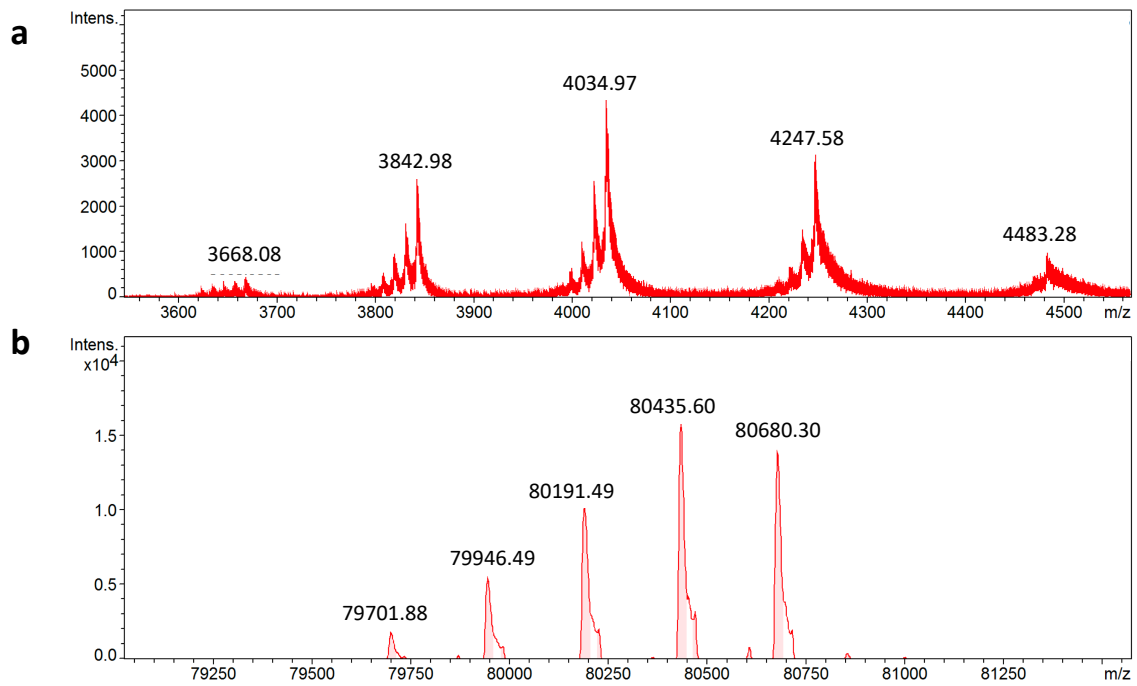
*Figure 19 Native MS spectrum of Sav-SOD K121A with (up to) four biotins bound. **a** Charge state envelope of the complexes. **b** Deconvoluted spectrum. The calculated mass of Sav-SOD K121A is: 79698.56 Da, for Sav-SOD K121A + 1 biotin: 79942.87 Da, for Sav-SOD K121A + 2 biotins: 80187.18 Da, for Sav-SOD K121A + 3 biotins: 80431.49 Da, for Sav-SOD K121A + 4 biotins: 80675.80 Da.*
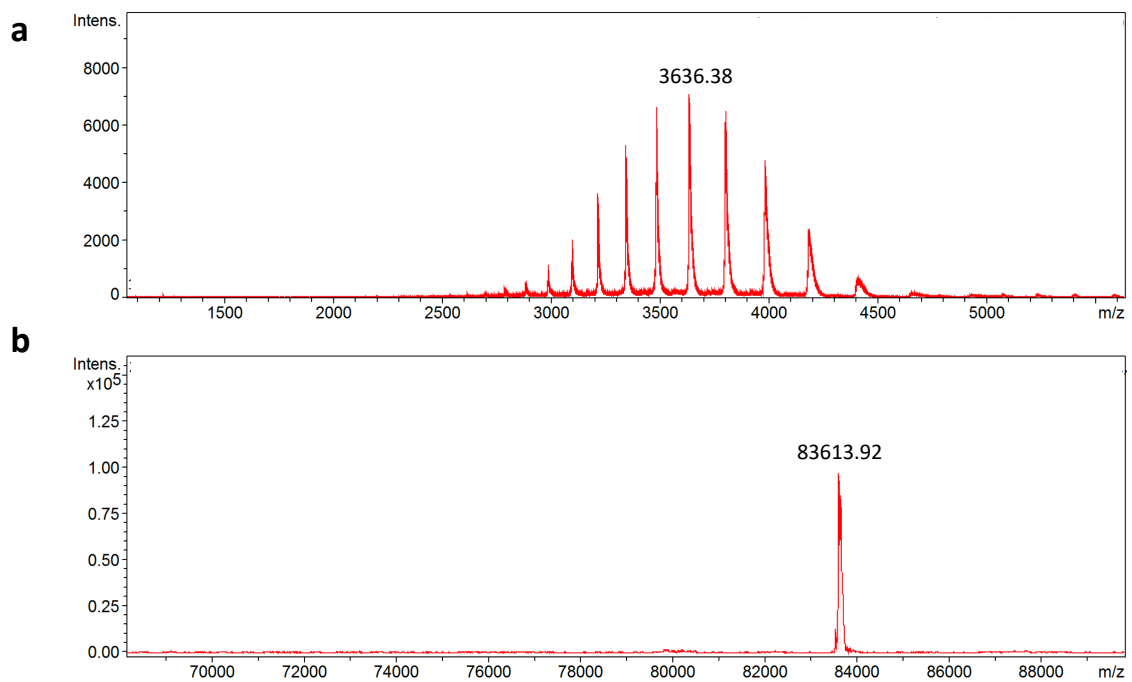
*Figure 20 MS spectrum of scdSav-SOD SASK. **a** Charge state envelope **b** deconvoluted stepctrum.*

36

Table 8 Mass spectral results for the proteins produced on a larger scale.

Construct	Mass observed (Da) for the monomer	Calculated Mass (Da) for the monomer
Sav-SOD K121A	19924.58	19924.64
Sav-SOD L1	20116.49	20114.97
Sav-SOD L2	20281.86	20280.16
Sav-SOD L3	20391.90	20390.28
Sav-SOD L4	20340.77	20339.18
Sav-SOD L5	20215.39	20214.06
Sav-SOD L6	20391.46	20389.29
Sav-SOD L7	20377.35	20376.20
Sav-SOD L8	20182.13	20180.90
Sav-SOD L9	20190.19	20188.42
Sav-SOD L10	20192.59	20194.93
	Mass observed (Da) for the tetramer	
scdSav-SOD SASK	83613.92	83614.50
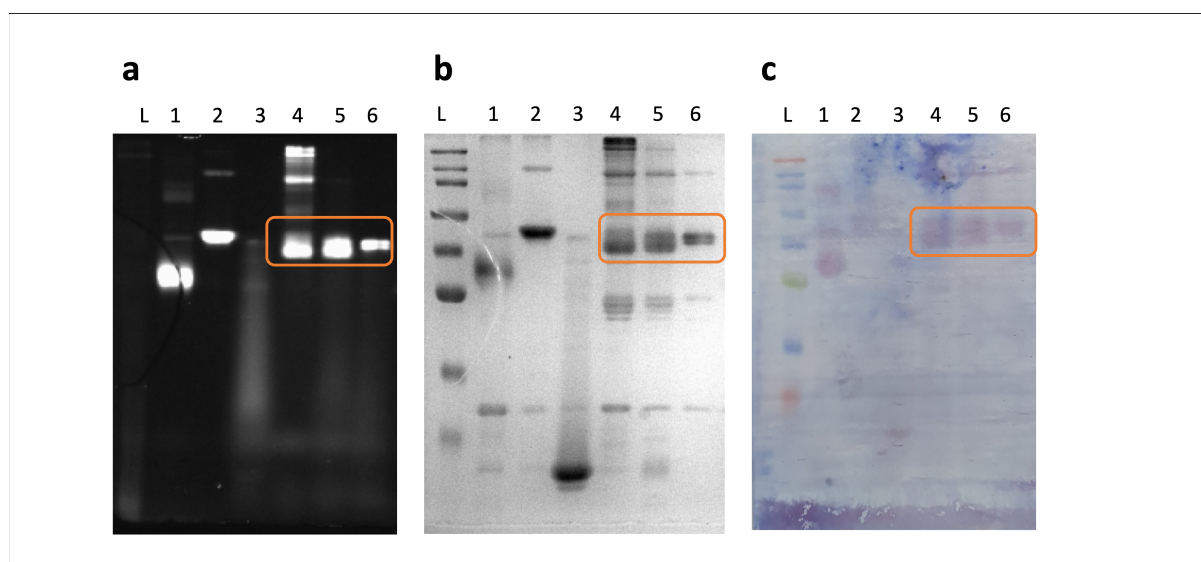
*Figure 21 SDS-PAGE and Western-Blot gel images of **a** B4F biotin binding assay **b** Coomassie blue staining. L) Protein Ladder (Eurogentec, ID-MWBRU1-250) 1) Sav-WT 2) Sav-SOD batch 1 3) Sav-SOD batch 2 4) scdSav-SOD batch 1 5) scdSav-SOD batch 2 6) scdSav-SOD batch 3. In orange, the corresponding bands of scdSav-SOD are marked.*

Chapter 3. Gold-(I) ArMs

3.1 Introduction

The human fascination for gold has persisted since ancient times. Attracted by its shiny appearance and rare existence, the human greed for gold was consistently high, and many wars were fought for its cause. Nevertheless, the desire for gold can be seen as a driving force for beneficial processes like the discovery of the new continent by Columbus or the development of early chemistry, better known as alchemy. Its name comes from the Arabic *al-kimia*, meaning "the art of transformation", with its highest goal being to find a transformation procedure for ordinary metals into gold (chrysopoeia). Nowadays, gold is used to acknowledge outstanding achievements, seen by the gold medal at the Olympics, and value something like the marriage with the golden wedding ring. Furthermore, terming particular "divine" principles like the golden rule are omnipresent in our language nowadays. The name for this "rare" element comes from aurum ("shining dawn"), which might come from the properties of the element itself, as it is resilient towards external influences. Gold can be termed as the noblest of all metals, as it is the least reactive metal towards atoms or molecules at the interface with a gas or a liquid.[116] Also, it is frequently used in dental/medicinal applications as it does not lead to problems of allergic reactions as other metals.

Early reports on the use of gold as a catalyst showed no superiority compared to other catalysts, even though its potential was already suspected.[117] In 1973, Bond *et al.* reported the hydrogenation of olefins over supported gold catalysts.[118] The following statement of Bond is representative of many discoveries in the gold catalysis field: "We are at a loss to understand why these catalytic properties of gold have not been reported before, especially since the preparative methods we have used are in no way remarkable." More than a decade later, the first reports, where gold was the best catalyst, appeared in the literature. In these, on the one side, Haruta *et al.* investigated the low-temperature oxidation of CO[119], and on the other side Hutchings *et al.* the hydrochlorination of ethyne to vinyl chloride.[120] Almost at the same time, Ito *et al.* introduced gold catalysis to the homogenous asymmetric catalysis

field.[121] Using an enantiomerically pure ferrocenylphosphine ligand complexed with gold (I), they could catalyze the asymmetric aldol reaction of an isocyanoacetate with aldehydes. From there on, the interest in gold catalysis increased exponentially, leading to reactions including nucleophilic additions, activation of carbonyl groups and alcohols, hydrogenation, reactions with carbon monoxide as nucleophiles and oxidation reactions.[122] As previously stated, elemental gold is inert; however, the examples above highlight that gold(I) or gold (III) complexes could be used for catalysis.

Due to its high stability against oxidation, and in contrast to other transition metals, gold(I) and gold (III) complexes do not readily cycle between oxidation states and thus rarely follow the typical oxidative addition/reductive elimination pattern.[123] The potential to stabilize cationic intermediates, the strong Lewis acidity, and the ability to coordinate unsaturated bonds combined with the exceptionally high stability towards oxygen and acidic protons imparts unique reactivity for catalysis.[124] These properties come from the relativistic contraction of the valence s orbitals of Au, leading to a relatively low-lying lowest unoccupied molecular orbital (LUMO) and, therefore, strong Lewis acidity. Furthermore, the relativistic contraction of the 6s orbital results in greatly strengthened Au-ligand bonds, as well as gold-gold interactions, a phenomenon referred to as "aurophilicity".[123]

In many cultures, gold has occupied a special place in medicine as a potential "cure-all" for diseases.[125] Especially, gold nanoparticles are heavily used in modern cancer research.[126] However, also various gold catalysts display remarkable biocompatibility, good reactivity at room temperature, low toxicity, and a propensity to coordinate alkynes, the most widely used biorthogonal functional group. Therefore, gold catalysis gained even more interest in recent years for its therapeutic potential.[91] In one approach Tanaka *et al.* used human serum albumin to anchor an NHC based Au(I) catalyst for the synthesis of fluorophores and drugs for anticancer treatments.[96,127–129] In this study, we first introduced a gold(I) based hydroamination, followed by a second study, where we were especially interested in utilizing the alkynophilic, Lewis acid-like properties of Au(I) complexes to convert alkynes in a biorthogonal and selective manner under mild reaction conditions.

3.2 Design and Evolution of Chimeric Streptavidin for Protein-Enabled Dual Gold Catalysis

3.2.1 Outline of the Authors Contribution

T.R.W., R.L.P. and F.C. conceived and designed the study. F.C., M.M.P. and B.L. contributed to the synthesis of the substrates, products and complexes. N.V.I., D.C.S., R.L.P. and F.C. contributed to mutagenesis, protein expression, protein purification and protein characterization. N.V.I. performed the crystallization, X-ray structure determinations and native MS experiments. F.C. performed the catalytic, preparative and deuterium-labelling experiments, designed the screening protocol and recorded the data. T.R.W., F.C. and N.V.I. analyzed the data. J.D.M., A.L. and L.T.S. contributed to the molecular modelling experiments. T.R.W., F.C. and N.V.I. wrote the manuscript, which was further supplemented through contributions from R.L.P. and J.-D.M. All authors have given approval to the final version of the manuscript.

3.2.2 Abstract

Artificial metalloenzymes result from anchoring an organometallic catalyst within an evolvable protein scaffold. Thanks to its dimer of dimers quaternary structure, streptavidin allows the precise positioning of two metal cofactors to activate a single substrate, thus expanding the reaction scope accessible to artificial metalloenzymes. To validate this concept, we report herein on our efforts to engineer and evolve an artificial hydroaminase based on dual gold activation of alkynes. Guided by modelling, we designed a chimeric streptavidin equipped with a hydrophobic lid shielding its active site, which enforces the advantageous positioning of two synergistic biotinylated gold cofactors. Three rounds of directed evolution using *Escherichia coli* cell-free extracts led to the identification of mutants favouring either the anti-Markovnikov product (an indole carboxamide with 96% regioselectivity, 51 turnover numbers), resulting from a dual gold σ,π-activation of an ethynylphenylurea substrate, or the Markovnikov product (a phenyl-dihydroquinazolinone with 99% regioselectivity, 333 turnovers), resulting from the π-activation of the alkyne by gold.

3.2.3 Introduction

Thanks to their unique affinity towards alkynes, allenes and alkenes, gold complexes have attracted considerable attention for their catalytic potential.[130–132] In addition to the activation of unsaturated substrates via π-coordination, terminal alkynes undergo dual gold activation via synergistic σ,π-coordination.[133–138] This mode of activation, which proceeds via a di-aurated transition state, affords distinct products/regioisomers, notably broadening the scope of gold-catalysed reactions. Such synergistic action of two metals in catalysis is reminiscent of polynuclear metallo-enzymes, whereby (at least) two metals act in concert to catalyse challenging reactions.[139]

In the context of in-vivo ligation and bioconjugation, alkynes occupy a place of choice, as this functional group was shown to be bio-orthogonal, thus finding widespread use in click chemistry.[140,141] Although Cu and Ru are privileged catalysts in this context,[142,143] recent reports suggest that gold complexes maintain catalytic activity in a cellular environment, albeit for a different type of reactivity.[91,144–147] To the best of our knowledge, however, these

biocompatible reactions rely on a π-activation of the alkyne, rather than the dual activation so distinctive of gold catalysis.

With the aim of complementing natural enzymes,[74] artificial metalloenzymes (ArMs) have experienced a renaissance in the past two decades.[13,26,148–150] For this purpose, an abiotic metal cofactor is compartmentalized within a protein scaffold that can be optimized by genetic means. Thus far, more than 40 reactions can be catalysed by ArMs.[18] Current challenges in the field include protein-accelerated catalysis, whereby a precatalyst is activated upon incorporation within the host protein,[70] dual catalysis[151,152] and compatibility of the ArM with a cytosolic environment.[153] Privileged scaffolds for ArMs include carbonic anhydrase,[154] hemoproteins,[21,155] prolyl oligopeptidase,[156] lactococcal multiresistance regulator,[150] four-helix bundles,[25,157] nitrobindin,[158] human serum albumin[129] and streptavidin.[13,159–161] The work presented herein capitalizes on the unique topology of streptavidin (Sav), enabling the localization of two close-lying biotinylated probes within a hydrophobic environment. This enables the engineering and evolution of a biocompatible artificial hydraminase (HAMase) based on either single or dual gold activation of an alkyne (Figure 22).
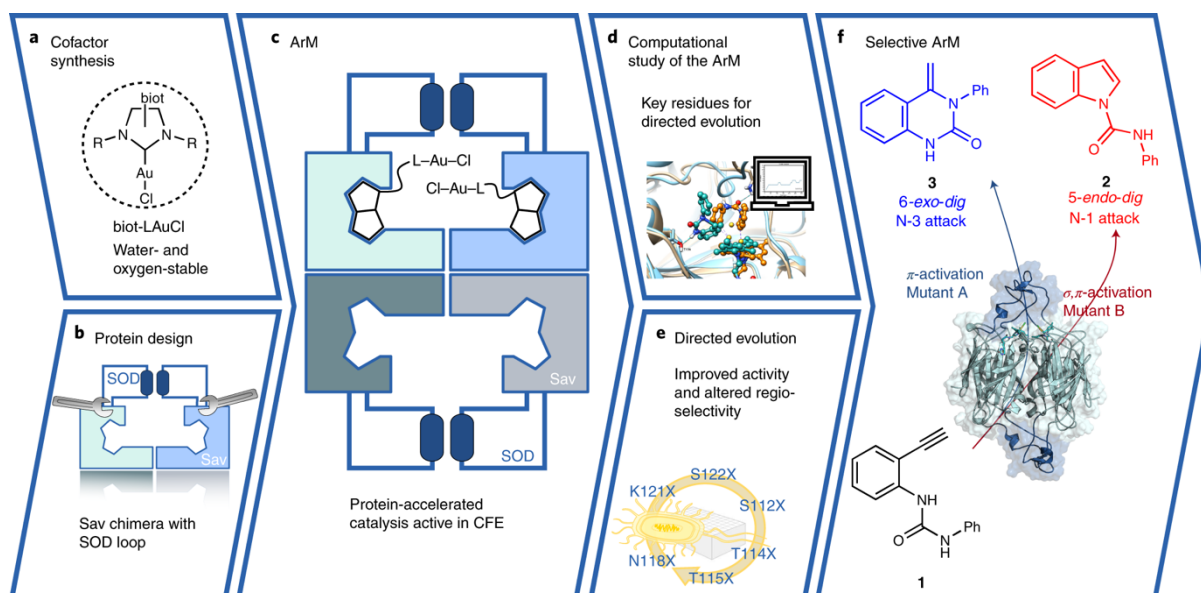
*Figure 22 Engineering and evolving an HAMase based on dual gold activation of alkynes. **a–c**, Chemo-genetic optimization of the catalytic performance relies on combining a biotinylated (biot) cofactor (LAuCl, where L corresponds to the N-heterocyclic carbene ligand) (**a**) and a tailored chimeric protein (**b**) to assemble an ArM equipped with two adjacent gold cofactors (**c**). **d–f**, genetic optimization is guided by modelling (**d**) to identify advantageous amino acids (**e**) for directed evolution to favour either σ,π-activation or π-activation of the alkyne (**f**) to afford either indole **2** or quinazolinone **3**, respectively.*

42

3.2.5 Results and Discussion

Design of the artificial HAMase. As reported by Asensio[134,162] and van der Vlugt[163], the regioselectivity for the hydroamination of ethynylurea **1** is by and large governed by the mode of activation of the alkyne by gold: the canonical π-activation favours the quinazolinone **3** (Markovnikov, *6-exo-dig* addition product), while the dual σ,π-gold activation affords preferentially the indole **2** (anti-Markovnikov, *5-endo-dig* addition product).[134,163,164] Upon π-coordination of the alkyne to gold, the acid dissociation constant (pK_a) of the terminal C–H bond is lowered, thus favouring its deprotonation and coordination by a second gold to afford the σ,π-activation mode.[162] Accordingly, the spatial arrangement of the two gold species is critical in determining the regioselectivity of the reaction. We thus selected the gold-catalysed cyclization of the ethynylurea **1** to engineer and evolve a dual-gold-catalysed HAMase based on the biotin–Sav technology.

Thanks to its dimer of dimers quaternary structure, which places the valeric acid side chains of two proximal biotins 19.8 Å apart (Protein Data Bank (pdb) 3ry2), we designed N-heterocyclic carbene ligands (L) equipped with a biotin anchor (biot) introduced at various positions (Figure 23a).[165] We hypothesized that the relative position of two gold moieties within the biotin-binding vestibule may influence the mode of alkyne activation as reflected by the indole **2** versus the quinazolinone **3** ratio. Sterically crowded imidazolium precursors were metalated using a one-step procedure,[166] and less-hindered carbenes were prepared through transmetalation of the silver–carbene complex, to afford the corresponding gold complexes: **biot-Au 1–5** (Figure 38).[167] These are air-stable and water-stable and can be stored for months as stock solutions in dimethylsulfoxide (DMSO) at 5 °C.
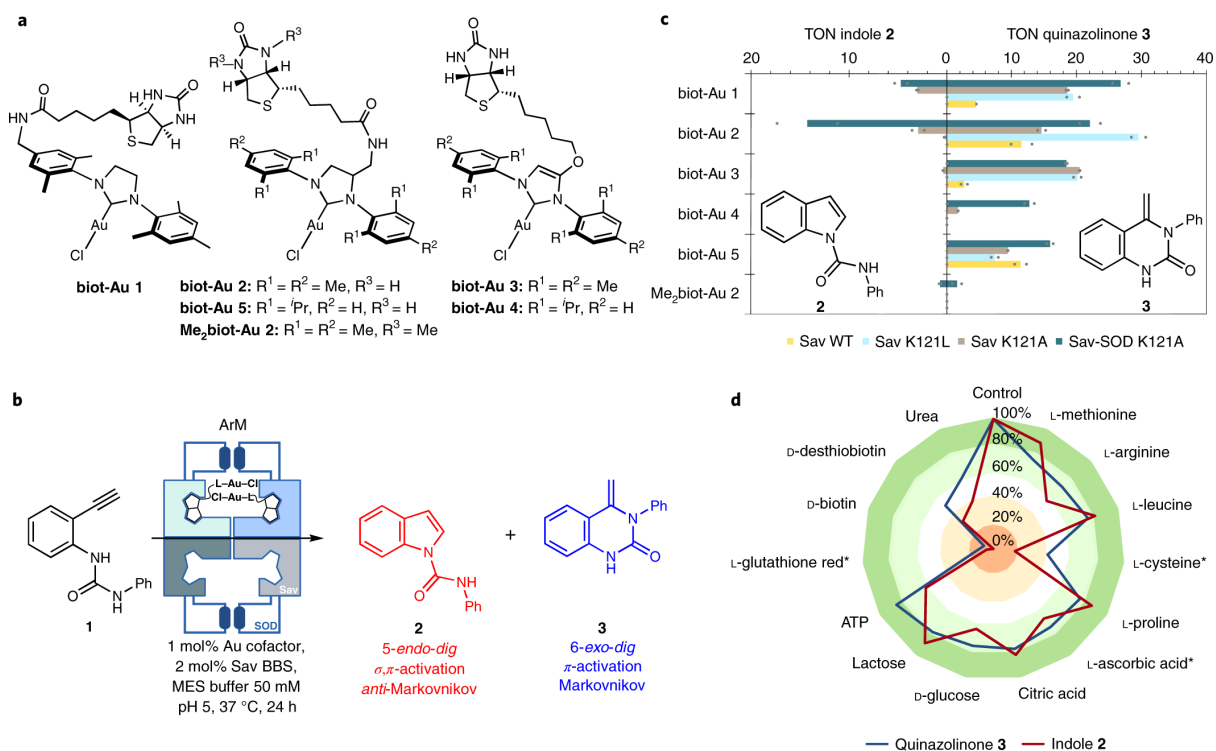
*Figure 23 Chemo-genetic optimization of HAMase activity. **a–c**, Biotinylated gold complexes **biot-Au 1–5** (**a**) tested in the presence of Sav isoforms (**b**) for the hydroamination of substrate **1** to afford indole **2** or quinazolinone **3** (**c**). Screening conditions using purified Sav samples: V_{tot}, 200 µl (V_{DMSO}, 15 µl); [Sub], 5 mM; [Au], 50 µM; [Sav], 100 µM; [MES buffer], 50 mM; pH, 5; 37 °C for 24 h, performed twice (grey dots). The turnover number (TON) was calculated in relation to the cofactor concentration [Au]. **d**, Bio-additive-based screening of **biot-Au 2**·Sav-SOD K121A allows identifying detrimental cellular components. Reaction conditions: V_{tot}, 200 µl (V_{DMSO}, 15 µl); [Sub], 2.5 mM; [bio-additive], 2.5 mM; [**biot-Au 2**], 50 µM; [Sav], 100 µM; [MES buffer], 50 mM; pH, 5; 37 °C for 24 h. BBS, biotin-binding sites. The * symbol indicates cross-reaction with the substrate.*

Engineering a chimeric protein with a hydrophobic biotin-binding vestibule. Initial studies using Au(i) complexes bearing either commercially available N-heterocyclic carbenes or **biot-Au 1–5** in buffered aqueous solutions afforded <1 turn-over number (TON; using 7.5% DMSO, [substrate **1**] = 5 mM and [Au] = 50 µM = 1 mol%, 24 h at room temperature; Table 18). Addition of wild-type Sav (Sav WT; 25 µM, corresponding to 100 µM biotin-binding sites) to a biotinylated cofactor **biot-Au 1–5** improved the catalytic activity leading to up to 12 TONs and affording the quinazolinone **3** exclusively (Figure 23b, c and Figure 39). This protein-acceleration phenomenon upon addition of Sav was not observed with cofactors devoid of biotin or with **Me₂biot-Au 2**, which bears a dimethylated biotin anchor and thus a markedly decreased affinity towards Sav as highlighted by a 4'-hydroxyazobenzene-2-carboxylic acid

(HABA) displacement assay (Figure 40). Fine-tuning the reaction conditions revealed that 2-(N-morpholino)ethanesulfonic acid (MES) buffer at pH = 5 affords the highest TONs, without considerably affecting the regioselectivity (Figure 41 - Figure 45). Next, we selected **biot-Au 2** and screened it in the presence of a focused library of Sav isoforms bearing mutations at S112 and/or K121. While the TON could be improved, especially in the presence of small/hydrophobic residues at position K121, the critical 3/2 ratio, that reflects the different gold-activation modes, remained heavily biased in favour of the *6-exo-dig* product **3** (Figure 46 and Table 19). The addition of various bio-additives in 50-fold excess versus catalyst was tolerated in most cases. Reduced glutathione and cysteine poisoned the catalytic system, fortunately only at higher concentrations (>20 × [**biot-Au 2**]). Strikingly, biotin inhibits the ArM already at low concentrations (>4× [**biot-Au 2**]). These results highlight the promising bio-robustness of the ArM, suggesting it may be used in a cellular medium (Table 20). With directed evolution in mind, this is a highly desirable feature as it allows the screening of cell-free extracts (CFEs), without the need to purify the Sav mutants.[168]

Inspection of the >30 X-ray structures of ArMs based on the biotin–Sav technology reveals a tendency for biotinylated metal cofactors to be poorly localized within the biotin-binding vestibule.[159] We attribute this to its shallow topology, thus allowing a biotinylated cofactor to adopt multiple poses, resulting in apparent reduced occupancy. Multiscale-modelling strategies on some of these systems illustrate that the pronounced flexibility of the cofactor may compromise the regiospecificity of the reaction.[76] With the aim of shielding the biotin-binding vestibule, consisting of two eight-stranded β-barrels facing each other, we surveyed the literature for naturally occurring dimerization domains present in β -sheet-rich proteins. We identified a potential candidate fitting this criterion: the superoxide dismutase C (sodC) from *Mycobacterium tuberculosis* (pdb 1pzs; ref. [83]), which includes an ~30 amino acid dimerization domain (referred to as SOD). This SOD lid forms an interface that spans across the ~29 Å of the two Greek key β-barrel subunits of sodC. We thus set out to engineer a chimera consisting of the dimerization domain of sodC inserted in the 3–4 loop of the Sav to yield a Sav-SOD (Figure 24a). To our delight, Sav-SOD could be expressed in the soluble fraction in high yield (typically >100 mg l^{-1}) in shake flasks using *E. coli* BL21 DE3 (Figure 27 and Figure 28).

To scrutinize the effect on the perturbation resulting from the addition of the sodC dimerization domain, we performed isothermal titration calorimetry measurements with biotin. Nanomolar binding affinity for biotin is retained over a wide range of temperatures (10 to 40 °C; Figure 29 and Table *12* and Table 13). At 25°C, the dissociation constant (K_d) is 4.2 nM with standard binding enthalpy $\Delta H^0 = -20.90$ kcal mol^{-1} and standard binding enthalpy $T\Delta S^0 = -9.49$ kcal mol^{-1} where T is temperature.

The standard free energy $\Delta G^0 = -11.4$ kcal mol^{-1} at 25 °C (ref. [169]). These parameters suggest that the biotin binding is primarily enthalpically driven. Comparison of the turn-on fluorescence upon incorporation of the biotinylated solvatochromic fluorescent reporter 4-N,N-dimethylamino-1,8-naphthalimide (**biot-4DMN**) in Sav and Sav-SOD K121A reveals a >2.5-fold and 20-fold, respectively, increase in fluorescence compared to the free **biot-4DMN**, accompanied by a blue shift (with the emission wavelength λ_{em} = 532 and 526 nm, respectively, versus 556 nm; Figure 30). Such increases in fluorescence accompanied by an ipsochromic shift have been attributed to increased hydrophobicity.[169] Furthermore, a thermal shift assay highlights an increased thermal stability of the apo chimeric protein compared to apo Sav WT (Figure 31). We thus surmise that the SOD lid stabilizes the protein and considerably contributes to provide a hydrophobic and shielded environment for organometallic catalysis.
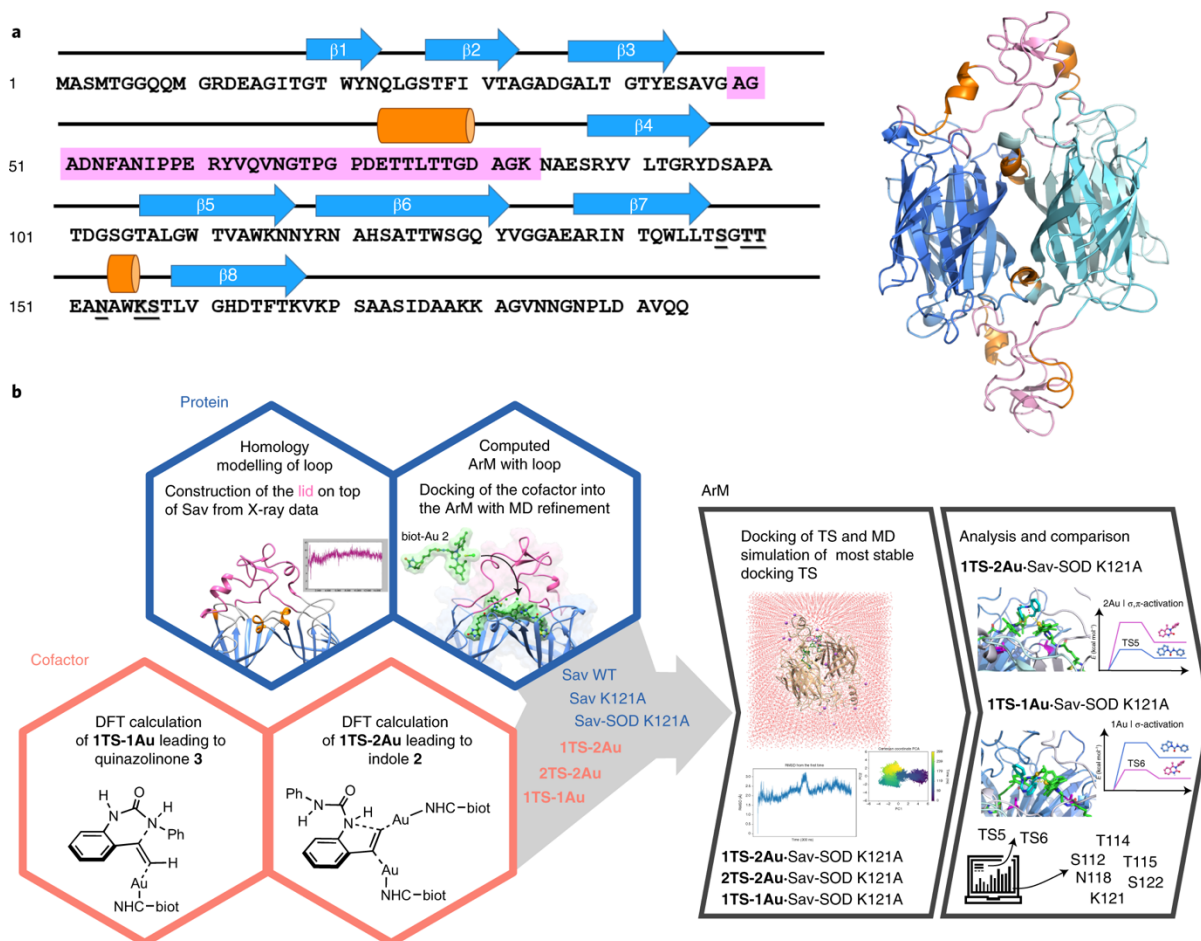
*Figure 24 Design and structural characterization of the chimeric ArM. **a**, Protein topology diagram of the sequence of one Sav-SOD monomer and the computed structure of chimeric Sav-SOD resulting from a 200 ns MD simulation; the SOD insert is highlighted in pink; residues subjected to saturation mutagenesis are underlined; blue arrows and orange cylinders represent β-sheets and α-helices, respectively. **b**, Workflow for the computational design and optimization of transition states (TSs) in chimeric Sav. Identification of the most promising amino acid residues to subject to mutagenesis. The red hexagonal frames depict the models used for the study of the solvated homogenous reaction (**1TS-1Au** on the left and **1TS-2Au** on the right) by quantum mechanical calculations. The blue hexagonal frames summarize the modelling steps carried out on the Sav-SOD prior to investigating the catalytic step (generation of a SOD model for the Sav-SOD (blue and pink ribbons, respectively) on the left, and insertion of the cofactor (ball and stick with green solvent-accessible surface) in the protein binding vestibule used for the molecular dynamic simulation on the right). The graphics inside the black-arrowed frames summarize the modelling steps for the ArMs' catalysed reaction. The left frame depicts the ArM once the transition state geometries have been introduced into the biotin-binding vestibule by protein-ligand docking, and hydrated for the molecular dynamics simulations. Exemplary graph used for analysis including root-mean-square deviation (RMSD) and principal component analysis (PCA) for the two first components (PC1 and PC2) are displayed. The black frame on the right contains refined 3D models of possible pseudo transition states and their analysis, including comparison with the energies (E) for the transition states of the homogeneous catalysts.*

47

Computational studies of the HAMase. To gain structural and mechanistic insight on the influence of the protein scaffold, we refined an integrative computational procedure that we have developed to model various ArMs (Figure 24b).[170] We analysed the formation of the *6-exo-dig* product **3** and *5-endo-dig* product **2** catalysed by **biot-Au 2**·Sav WT, **biot-Au 2**·Sav K121A and **biot-Au 2**·Sav-SOD K121A, where **biot-Au 2** corresponds to the cofactor incorporated (·) into the Sav or Sav-SOD protein. For this purpose, we applied the following workflow: (1) Density functional theory (DFT) calculations were carried out on the theozyme, the core catalytic centre (alkyne **1** coordinated to one or two biotinylated gold catalysts) for both π- and dual σ,π-activation modes. DFT calculations (B3LYP-D3 functional) were performed using continuum water solvent conditions (Table 21 and Figure 47 - Figure 50).[171,172] (2) With no X-ray structure of the Sav-SOD available, structural modelling (with Modeller[173]) was carried out on this system followed by classical molecular dynamics (MD) simulations (AMBER force field[174]) up to convergence (Figure 51). Then, after inclusion of **biot-Au 2** into the protein vestibule by protein–ligand docking, MD simulations on the three Sav scaffolds (**biot-Au 2**·Sav WT, **biot-Au 2**·Sav K121A and **biot-Au 2**·Sav-SOD K121A) were performed to determine the conformational space available for substrate binding (Figure 52). (3) Incorporation of the transition state structures for the *6-exo-dig* and *5-endo-dig* pathways in water (see step (1)) by protein–ligand docking approaches (including Gold 5.8.1 (ref. [175])) in the most representative structures of the MD simulations for the three Sav scaffolds (we term these 'pseudo-transition states'; Figure 53). (4) Further refinement was done by MD of the best results obtained for systems with reasonable predicted affinity (in step (3); Figure 54 - Figure 56). The final simulations were analysed focusing on (1) the complementarity of transition state structures for the *6-exo-dig* and *5-endo-dig* within the Sav vestibule, and (2) the number of gold cofactors biot-Au 2 involved (that is, one or two; Figure 57, Figure 58 and Table 22 and Table 23). The influence of the host protein on the transition state structures helped identify amino acid residues to randomize during directed evolution.

The DFT calculations revealed that the competition between π- and dual σ,π-activation modes also operates in water (Figure 47). The difference between the Gibbs energy barriers for both pathways is about 2 kcal mol^{-1} (19.5 versus 21.5 kcal mol^{-1} for *6-exo-dig* and *5-endo-dig* mechanisms, respectively), suggesting that subtle changes in the first or second coordination sphere of the metal may affect the ratio between both products. Calculations of the barriers

in solvents of different dielectric constant suggest that the regioselectivity is not considerably affected by the polarity of the medium (Figure 48 and Figure 49).

We surmised that a shift in regioselectivity may be promoted by the protein environment. DFT calculations carried out on small models were used to evaluate the impact on the ground and transition state structures of the gold complexes both in bulk water and in the confined, hydrophobic environment provided by the biotin-binding vestibule. Scrutiny of the active site led us to hypothesize that, following π-coordination, the alkyne's C–H may be deprotonated by close-lying amino acids or a water molecule to afford a di-aurated σ,π-acetylide species: Au–Cσ = 2.01 Å; Au–Cπ = 2.28, 2.33 Å; and the Au–Au distance is ~3 Å (Figure 50).

Having identified transition state structures in water for the isolated cofactor (in the absence of the protein scaffold), three different pseudo-transition states embedded in the protein were evaluated: **1TS-1Au**, **2TS-2Au** and **1TS-2Au**. The transition state **1TS-1Au** corresponds to the π-activated transition state, which occupies half of the biotin-binding vestibule, with a second, unligated **biot-Au 2** occupying the neighbouring biotin-binding site. The transition state **2TS-2Au** is similar to **1TS-1Au** but includes two gold complexes, each activating an alkyne via π-coordination. Finally, the transition state **1TS-2Au** contains two gold complexes interacting with a single alkyne substrate via σ,π-coordination (Figure 25a).

Molecular modelling of **biot-Au 2**·Sav WT provides the best fit for the pseudo-transition state **1TS-1Au** when inserted into the catalytic vestibule of Sav (Table 21 and Figure 53). Worse complementarities were computed for **2TS-2Au** and severe clashes with the amino acids on position 121 were predicted for **1TS-2Au**. All these pseudo-transition state structures are rather solvent exposed, with minimal impact of the second coordination sphere, resulting in similar energy barriers and similar predicted regioselectivity compared to the free cofactor. Both the free cofactor and **biot-Au 2**·Sav WT are thus predicted to favour π−activation to afford quinazolinone **3**. Gratifyingly, in vitro experiments support the model, revealing that **biot-Au 2**·Sav WT affords exclusively quinazolinone **3** (Figure 23c).

In the case of **biot-Au 2**·Sav K121A, similar or worse fitting scores compared to Sav WT are computed for all pseudo-transition states, except for **1TS-2Au**. In fact, the score associated per cofactor reveals that the binding affinity of **1TS-2Au** is close to the value obtained for **1TS-1Au** (that corresponds to a single transition state geometry with no geometric constraint from the adjacent Sav monomer; Table 21). These results suggest that a single-point mutation

K121A improves the docking score of **1TS-2Au** compared to **1TS-1Au**. This reflects a modest shift in favour of σ,π-alkyne coordination which should allow for the formation of the anti-Markovnikov product indole **2**. These findings are supported by experiments: the regioselectivity (3/2) varies from 100:0 (for **biot-Au 2**·Sav WT) to 83:17 for **biot-Au 2**·Sav K121A (Figure 23c and Table 9).

X-ray-quality crystals of **biot-Au 2**·Sav-SOD (including various mutants) were obtained through cocrystallization (Figure 37 and Table 17). Although a resolution down to 1.8 Å was achieved, the SOD lid could not be fully resolved due to its high flexibility (pdb: 7ALX). We thus set out to model the structure of Sav-SOD K121A, starting with homology modelling, followed by a long classical MD (200 ns; Figure 25b and Figure 51)[176]. The resulting models are stable with the SOD lid presenting the highest degree of flexibility (root-mean-square fluctuation 2.63 Å). Next, we docked **biot-Au 2** into the system. Good complementarities were obtained by collective motion of the entire SOD lid (Figure 52 and Figure 53). This hydrophobic lid contributes to shield both cofactors from the solvent. A second MD (300 ns) placed both **biot-Au 2** cofactors sufficiently close to synergistically engage in σ,π-activation of a terminal alkyne (Figure 54 - Figure 56). To our delight, **biot-Au 2**·Sav-SOD K121A indeed displayed the highest *5-endo-dig* regioselectivity (that is, 62:38 for 3/2) and TONs of up to 40 (Figure 23c and Table 9). We thus selected **biot-Au 2**·Sav-SOD K121A for directed evolution.

Docking of the three pseudo-transition state models into Sav-SOD K121A reveals that the best complementarity is obtained for **1TS-1Au** and **1TS-2Au** (especially for transition state **TS5**; Figure 25c,d). Scoring values for **2TS-2Au** were extremely low as there is limited space for such a large pseudo-transition state in the Sav-SOD's vestibule (Table 21). This also suggests that the possibility of another substrate approaching **1TS-2Au** to form **2TS-2Au** is unlikely in chimeric Sav-SOD. Depending on the transition state, the substrate or substrates occupy different positions within the active site (Figure 25c versus Figure 25d). The most notable difference is a 180 degree rotation of the substrate between **1TS-1Au** (*6-exo-dig*) compared to **1TS-2Au** (*5-endo-dig*). Further residue-contribution analysis of **1TS-2Au** and **1TS-1Au** (to afford *5-endo-dig* and *6-exo-dig* products, respectively) in the active ArM were performed. Calculations were carried out using Cytoscape[177] as implemented in UCSF Chimera[178] (for a qualitative analysis of the main interactions along the MD simulations) followed by MMGBSA (to extract indicative energetic values; Figure 57, Figure 58, Table 22 and Table 23). Direct

interactions of Au with close-lying amino acids are very weak (for example, purely van der Waals contacts). Both pseudo-transition state structures reveal common interactions between the coordinated substrate **1** and the residues SOD-N8, SOD-I9, SOD-A3 and Sav-N118. For **1TS-1Au**, additional contacts were identified: from Sav-S112 to S122, especially T114. This increased number of contacts is traced back to the high level of flexibility of the **1TS-1Au** versus the **1TS-2Au**. As the SOD lid is highly flexible (and disordered in the X-ray structure), we selected close-lying residues belonging to Sav rather than the SOD lid. Accordingly, the following amino acids were selected for the directed evolution campaign: S112, T114, T115, N118, K121 and S122 (Figure 25e).
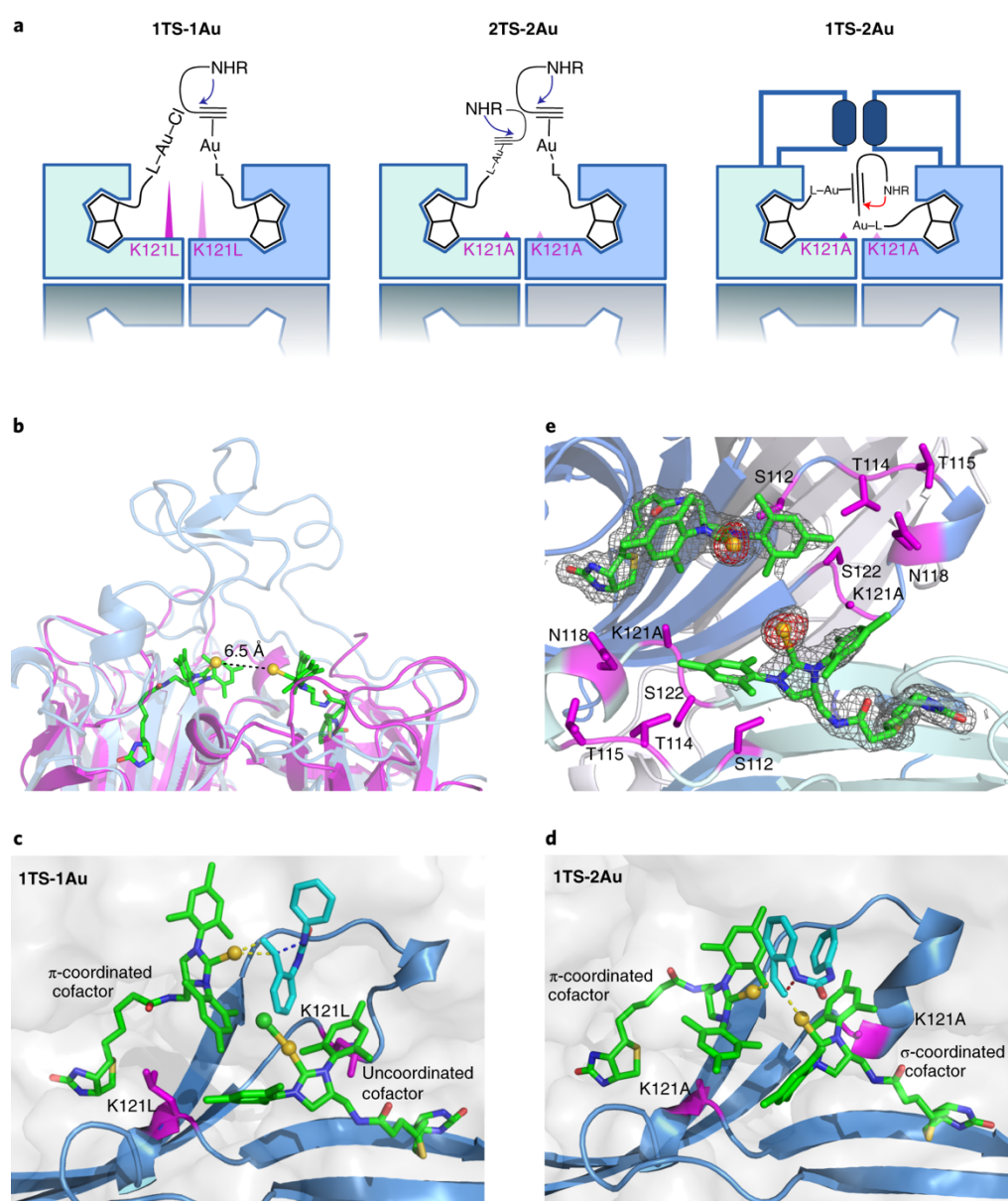
*Figure 25 Analysis of the transition state structure and close-lying amino acid residues in chimeric Sav. **b**, Schematic representation of postulated gold-catalysed hydroamination reaction mechanisms within the biotin-binding vestibule. **b**, X-ray crystal structure of **biot-Au 2**·Sav-SOD K121A (magenta) overlaid with the MD model (blue; Figure 37 for details). **c,d**, Computed transition states for the **biot-Au 2**-catalysed 6-exo-dig (**1TS-1Au**) and for the 5-endo-dig cyclization (**1TS-2Au**) docked within Sav-SOD K121A. e, Close-up view of the X-ray structure of **biot-Au 2** Sav-SOD K121A. Anomalous electron density (displayed as red mesh at 5σ (the cut-off point of the intensity of the electron density represented in standard deviation units) assigned to Au and modelled with a 50% occupancy; no electron density for Cl was detected. Electron density map (2Fo-Fc, corresponding to the electron-density difference map of the Fo: measured structure factors from the diffraction patterns and Fc: structure factor amplitudes calculated from the model) for the **biot-Au 2** (displayed as grey mesh at 1σ). **biot-Au 2** (green stick), close-lying amino acid residues (magenta stick representation), Au as golden sphere, Cl as green sphere and the protein as a cartoon representation.*

Directed evolution of the HAMase based on Sav-SOD. Having engineered an evolvable hydrophobic environment lining the biotin-binding vestibule and identified promising residues, we set out to optimize the HAMase by directed evolution. Building on the computational insight, we selected **biot-Au 2**·Sav-SOD K121A as the starting point for the iterative saturation mutagenesis. The free cofactor **biot-Au 2** did not display considerable HAMase activity in the absence of Sav, either using MES buffer or CFEs (Table 9 and Table 24). We hypothesize that as biot-Au 2 is insoluble in the aqueous CFE, the gold is shielded from poisoning by the cellular debris (in particular soluble thiols). Upon compartmentalization within Sav, a soluble and active ArM results, whereby the protein (partially) shields the cofactor from these detrimental metabolites, thus restoring catalytic activity, as previously observed in related studies.[99,129] We were delighted to observe HAMase activity upon addition of **biot-Au 2** to *E. coli* CFE (BL21 DE3) containing Sav and Sav-SOD (Table 24). This strategy allows bypassing the laborious protein purification step and complements our previous high-throughput screening platforms based on periplasmic and surface display.[18,37,68]

Table 9 Selected Results of evolved ArMs using purified HAMases[a]. The best hits for either regioisomer are highlighted in bold [b,c] and were tested on preparative scale[d,e].

Entry	Sav-Mutants	Total TON[a]	Selectivity (**3:2**)
1	no streptavidin	0.3 ± 0.1	100:0
2	*Sav K121A*	*21.2 ± 0.9*	*82:18*
3	Sav-SOD K121A	39.3 ± 1.9	62:38
4	Sav-SOD N118G K121G S122G	57.5 ± 4.7	45:55
5	Sav-SOD S112N N118G K121G S122G	64.6 ± 2.1	32:68
6	**Sav-SOD S112N T114S T115N N118G K121G S122G**	**51.2 ± 3.5[b]** **39[d]**	**4:96[b]** **8:92[d]**
7	Sav-SOD N118S K121F S122G	38.6 ± 1.0	99:1
8	Sav-SOD S112T N118S K121F S122G	42.0 ± 1.1	97:3
9	**Sav-SOD S112T T115A N118S K121F S122G**	**333 ± 57[c]** **94[e]**	**99:1[c]** **>99:1[e]**

[a]The analytical experiments were carried out in quadruplicates. The combined turnover for both products in relation to [**biot-Au 2**] is displayed as mean ± SD (n=4). The quinazolinone:indole ratio (**3:2**) with standard reaction conditions: V_{tot} 200 µL (V_{DMSO} 15 µL), [Sub] 5 mM, [**biot-Au 2**] 50 µM, [Sav] 100 µM, [MES-buffer] 50 mM, pH 5, 37 °C for 24 h.; [b]Optimized reaction conditions 5-endo-dig with the following changes: V_{tot} 100 µL (V_{MES} 45 µL, V_{DMSO} 20 µL), [Sub] 15 mM, [Diamide] 15 mM, [**biot-Au 2**] 100 µM, [Sav] 200 µM, 37 °C for 72 h; [c]Optimized reaction conditions 6-exo-dig with the following changes: V_{tot} 400 µL (365 µL of MES-buffer 50 mM, pH 5) [Sub] 2.5 mM, [**biot-Au 2**] 2.5 µM, [Sav] 5 µM, 39 °C for 72 h; [d]Preparative reaction (0.2 mmol): 72% yield of indole **2** and 6% yield of quinazolinone **3**; [e]Preparative reaction (0.1 mmol): 47% yield of quinazolinone **3** and no isolable trace of indole **2**.

First, positions N118, K121 and S122 were mutated simultaneously using a library consisting of NDT codons at these three positions (Figure 26a, Table 25 and Figure 59 and Figure 60).[179] For this purpose, thirteen 96-well plates, representing 24% coverage, were grown. Each well was subjected to a biotin-4-fluorescein (B4F) binding assay to identify biotin-binding Sav-SOD mutants that were then evaluated in catalysis. Twelve hits including Sav-SOD N118G K121G S122G (GGG hereafter) and Sav-SOD N118S K121F S122G (SFG hereafter) were quantified by ultra performance liquid chromatography and mass spectrometry (UPLC-MS) and displayed either the highest TON or the highest regioselectivity. They were sequenced, expressed, purified by affinity chromatography and tested in vitro. Next, the twelve hits were subjected to another round of directed evolution targeting position S112 with 18 possible amino acid combinations using VMA and NDT codons (Table 11 and Table 25). The best six hits were expressed, purified and evaluated. Mutants Sav-SOD S112N N118G K121G S122G (N-GGG) and Sav-SOD S112T N118S K121F S122G (T-SFG) were identified as most promising for *5-endo-dig* product **2** and *6-exo-dig* product **3**, respectively. These two quadruple mutants were subjected to another round of mutagenesis, targeting simultaneously positions T114 and T115.

From this screening campaign, the following trends emerge: large amino acids at position K121 favour the *6-exo-dig* product **3**, suggesting that the clashes observed via computational studies between K121 and the σ,π-gold intermediate are also valid for other mutants. Reaction conditions used for CFE experiments tend to favour the formation of *6-exo-dig* product **3**. After two rounds of directed evolution starting from Sav-SOD K121A, clear evolutionary trajectories favouring either regioisomer emerge (the scaffold GGG favours the indole **2** and SFG favours the quinazolinone **3** pathway; Figure 26b and Figure 59). To validate these trends, the most promising Sav chimeras were expressed in shake flasks and purified. Marked improvement of activity and selectivity over various generations was observed. Ultimately, we identified two chimeras Sav-SOD S112N T114S T115N N118G K121G S122G (N-SN-GGG) and Sav-SOD S112T T115A N118S K121F S122G (T-A-SFG) for the production of either regioisomer: **biot-Au 2**·Sav-SOD T-A-SFG and Sav-SOD N-SN-GGG afforded 333 and 51 TONs with a ratio 3/2 of 99:1 and 4:96, respectively (Figure 26c, Table 9 and Table 25). To further validate these results, the two best chimeras were scaled up (>500 mg) and the two HAMases were tested on a preparative scale (0.1 mmol and 0.2 mmol, respectively): **biot-Au 2**·Sav-SOD T-A-SFG afforded quinazolinone 3 with 47% isolated yield and no isolable amount of indole 2 (>99:1 regioselectivity) while **biot-Au 2**·Sav-SOD N-SN-GGG afforded quinazolinone 3 and indole 2 in 6% and 72% isolated yield, respectively (8:92 regioselectivity; Figure 61 - Figure 63).
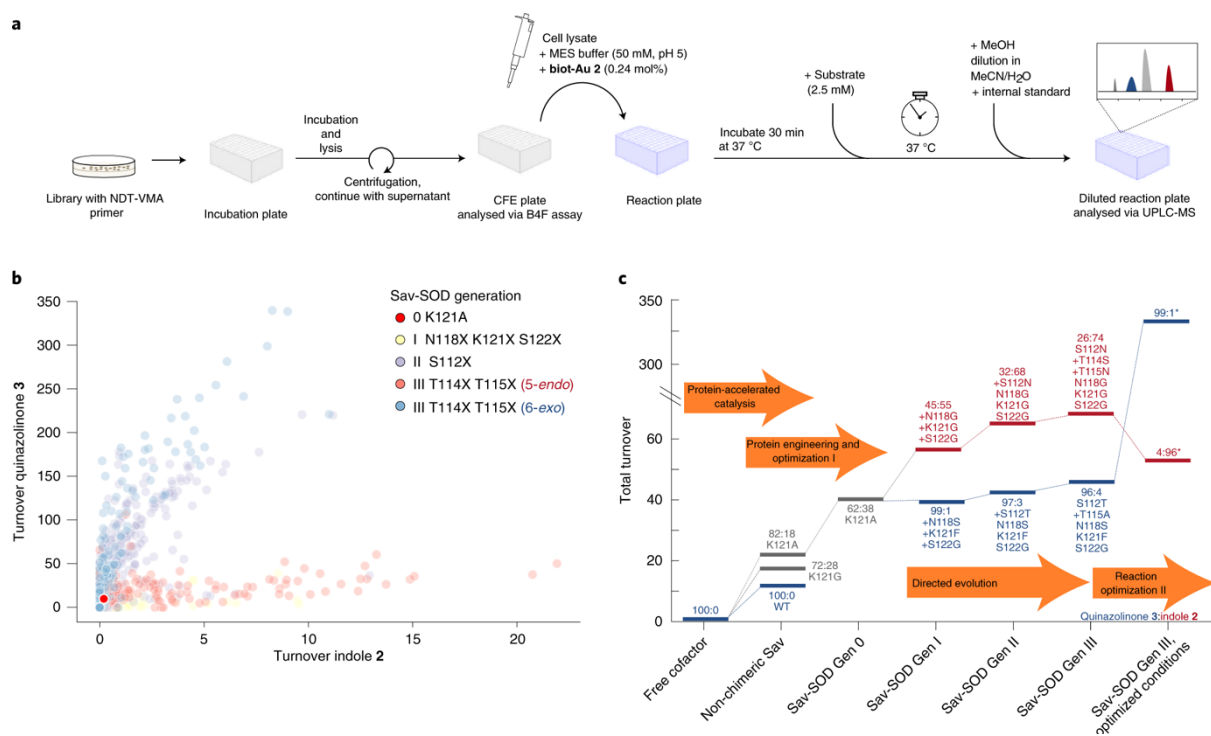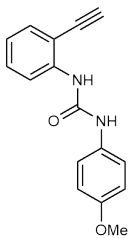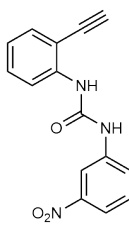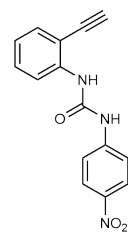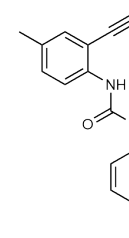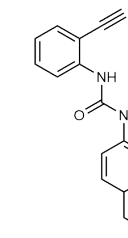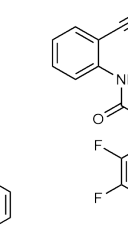
Figure 26 Directed evolution of a HAMase based on Sav-SOD. **a**, Streamlined protocol for the genetic optimization of **biot-Au 2**·Sav-SOD in CFE using a 96-well plate assay. **b,c**, Scatterplot of 2,500 CFE experiments (**b**) displaying the two evolutionary trajectories Sav-SOD ggg (red) and Sav-SOD SFg (blue), for amino acid positions selected for mutagenesis (**c**), ultimately leading to the identification of **biot-Au 2**·Sav-SOD T-A-SFg and **biot-Au 2**·Sav-SOD N-SN-ggg. Reaction conditions for 96-well CFE screening: V_{tot}, 400 µl (V_{DMSO}, 20 µl); [Sub], 2.5 mM; [cofactor], 6.25 µM; MES-buffer/Lysis-buffer, 1:1; pH, ~5.5; sealed at 37 °C for 4 days. Evolution trajectory (**c**) of HAMase using purified Sav isoforms of the evolution of **biot-Au 2**·Sav-SOD. Reaction conditions for catalysis with purified protein samples: V_{tot}, 200 µl (V_{DMSO}, 15 µl); [Sub], 5 mM; [cofactor], 50 µM; [Sav], 100 µM; [MES buffer], 50 mM; pH, 5; 37 °C for 24 h; except optimized reaction conditions (marked with *) for 6-exo-dig: V_{tot}, 400 µl (365 µl of MES buffer, 50 mM, pH 5); [Sub], 2.5 mM; [**biot-Au 2**], 2.5 µM; [Sav], 5 µM; 39 °C for 72 h; and for 5-endo-dig: V_{tot}, 100 µl (VMES, 45 µl); [Sub], 15 mM; [diamide], 15 mM; [**biot-Au 2**], 100 µM; [Sav], 200 µM; 37 °C for 72 h.

Dual gold mechanism and substrate scope. To confirm that the *5-endo-dig* product **2** isolated using the evolved HAMases Sav-SOD N-SN-GGG and Sav-SOD T-A-SFG indeed results from a dual gold mechanism—rather than a single gold mechanism—we carried out preparative-scale reactions in D_2O. The dual-activation mechanism proceeds via a di-aurated intermediate, which then undergoes deutero-deauration to regenerate the catalyst and release the (di-deuterated) product **2**-d_2 (Figure 65). Gratifyingly, both ^{1}H NMR and high-resolution MS studies confirm that the isolated indole **2**-d_2 is indeed equally di-deuterated at its C2 and C3 positions (>98% by ^{1}H NMR; Figure 65 and *Figure* 66). We thus conclude that the formation of the indole **2** through either evolved HAMase indeed proceeds via a dual gold

mechanism. Finally, six structurally related ethynylureas **1a–1f** were tested for their regioselective cyclization in the presence of the two evolved chimeric HAMases (Table 10 and Table 27). In some cases, the limited solubility and water stability of these aromatic ureas proved challenging. Nevertheless, both chimeras afforded preferentially either the quinazolinone (Sav-SOD T-A-SFG) or the indole (Sav-SOD N-SN-GGG) products.

*Table 10 Substrate scope using the fourth generation of **biot-Au 2** · Sav-SOD mutants obtained with the optimized reaction conditions for either 5-endo-dig[a] or 6-exo-dig[b] products.*

Substrate / Mutant	1a	1b	1c	1d	1e	1f
Sav-SOD S112N T114S T115N N118G K121G S122G[a]	45.9 (7:93)	35.9 (19:81)	22.0 (15:85)	15.3 (1:99)	5.2 (4:96)	6.9 (50:50)
Sav-SOD S112T T115A N118S K121F S122G[b]	104.7 (96:4)	35.3 (97:3)	41.7 (97:3)	26.0 (97:3)	16.7 (85:15)[c]	6.0 (96:4)[c]

The combined turnover for both products is displayed in relation to [**biot-Au 2**] as well as the quinazolinone:indole ratio (**3:2**). [a]Optimized reaction conditions 5-endo-dig: V_{tot} 100 µL (V_{MES} 45 µL, V_{DMSO} 20 µL), [Sub] 15 mM, [Diamide] 15 mM, [**biot-Au 2**] 100 µM, [Sav] 200 µM, [MES-buffer] 50 mM, pH 5, 37 °C for 24 h; [b]Optimized reaction conditions 6-exo-dig: V_{tot} 400 µL (V_{MES} 365 µL, V_{DMSO} 12 µL) [Sub] 2.5 mM, [**biot-Au 2**] 6.25 µM, [Sav] 12.5 µM, [MES-buffer] 50 mM, pH 5, 39 °C for 48 h; [c] with following changes: [**biot-Au 2**] 2.5 µM, [Sav] 5 µM, for 24 h. For a summary of results using the unoptimized reaction conditions used in Table 1, see *Table 27*.

3.2.6 Conclusions

Natural metalloenzymes often rely on dual catalysis to functionalize challenging substrates. Capitalizing on the unique topology of the biotin-binding vestibule of Sav, we designed an artificial HAMase that proceeds via a σ,π-activation of a terminal alkyne by two biotinylated gold cofactors. To ensure the positioning of the two gold moieties, Sav's biotin-binding vestibule was equipped with a hydrophobic lid, borrowed from sodC.[83] In silico modelling of the resulting chimeric HAMase provided insight into the two mechanistic manifolds and revealed close-lying amino acid residues to target by directed evolution, to favour the preferential formation of the anti-Markovnikov product indole **2** over the Markovnikov product quinazolinone **3**. These two products result from the dual gold σ,π-activation and the

single gold π-activation reaction manifolds, respectively. Thus far, optimization of ArMs' performance has mostly been focused on optimizing enantioselectivity.[18,149] In addition to displaying remarkable levels of enantiocontrol, enzymes excel at imposing catalyst control to address regioselectivity challenges. Herein, we combined protein engineering and directed evolution to fine-tune the second coordination sphere around the abiotic cofactor to control the regioselectivity of the hydroamination reaction. Despite its pronounced thiophylicity, the biotinylated cofactor **biot-Au 2** could be used in the presence of *E. coli* CFE, thus simplifying the directed evolution campaign. We tentatively assign this feature to the cofactor's insolubility in the reaction medium. Upon solubilization resulting from binding to Sav, the cofactor is partially shielded from thiols, thus affording biocompatible, active and selective artificial HAMases. Current efforts are aimed at integrating this versatile dual-activation reaction in vivo to complement the natural metabolism.

3.2.6 Experimental

General information. Chemicals were purchased from Sigma Aldrich, Acros Organics, Alfa Aesar or Fluorochem and used without further purification. Water used for molecular biology and in the catalytic reactions was purified by Milli-Q Advantage system. All catalytic reactions were carried out with non-degassed solvents under air. Temperature was maintained using Thermowatch-controlled heating blocks. ^{1}H NMR (500 MHz) and ^{13}C NMR (126 MHz) spectra were recorded at room temperature on a Bruker 500 MHz or 600 MHz spectrometer. Data are summarized with the coupling constants described as follows: chemical shift, multiplicity (s = singlet, d = doublet, t = triplet, q = quartet, m = multiplet, br = broad signal, bs = broad singlet, dd = doublet doublets, dt = doublet triplets, dq = doublet quartets, td = triplet doublets, ddd = doublet of doublet of doublets, ddt = doublet of doublet of triplets, dtd = doublet of triplet of doublets, dddd = doublet of doublet of doublet of doublets), coupling constants in Hertz (Hz). Chemical shifts (δ) are reported in ppm relative to the residual solvent peak. Peaks were fully assigned using 2D-NMR spectroscopy including COSY, NOESY, HMBC and HMQC. The NMR spectra were analysed using MestreNova$^©$ NMR data processing software. High-resolution mass spectrometry (HR-MS) was performed on a Bruker maXis II QTOF ESI mass spectrometer coupled to a Shimadzu LC. Prior to HR-MS submission, all

samples were characterized via direct injection or LC from solutions in MeOH (gold complex synthesis) or MeCN (substrate synthesis) on a Shimadzu LCMS-2020. Native-MS was performed via direct injection (aqueous buffer) to a Bruker maXis II QTOF ESI mass spectrometer. Fluorescence assays were recorded on a Tecan fluorimeter Infinite M1000Pro. Samples were prepared on a clear 96-well plate. Measurements settings for Fluorescence Top Reading: Excitation Wavelength 485 nm, Emission Wave length 520 nm, Excitation Bandwidth 5 nm, Emission Bandwidth 5 nm, Gain 94 (100%), Number of slashes 10, Flash frequency 400 Hz, Integration time 20 µs, Lag time 0 s, Settle Time 100 ms. Molecular biology reagents were purchased from New England Biolabs (NEB), Integrated DNA Technologies (IDT), and Macherey-Nagel and were used as described in the accompanying protocols. Mutations were verified by Sanger sequencing performed by Microsynth (Balgach, Switzerland).

Synthesis of biot-4DMN. The synthesis of **biot-4DMN** dye was achieved with slight modification of previously published protocols.[180]

6-(dimethylamino)-1H,3H-benzo[de]isochromene-1,3-dione: **C1**
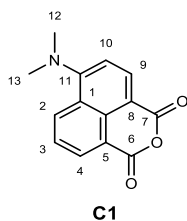
A round bottom flask was charged with 4-Bromo-1, 8-naphthalic anhydride (5.00 g, 18.1 mmol, 1 eq) and of $CuSO_4 * 5H_2O$ (315 mg, 1.26 mmol, 0.07 eq) and DMF (100 mL). The reaction was heated (1.5 h at 140 ºC). After cooling to room temperature, the brown crude was precipitated with water (1 L), dried and collected. The resulting solid was further purified by silica column chromatography using CH_2Cl_2/methanol as an eluent to yield C1 as an orange solid (3.0 g, 12.43 mmol, 69%).

HR-MS (ESI, pos): m/z calcd. for $C_{14}H_{11}NO_3H^+$ 242.0812 found 242.0808 $[M+H]^+$.

^1H NMR (500 MHz, DMSO-d_6) δ 8.58 (dd, J = 8.5, 1.1 Hz, H2), 8.46 (dd, J = 7.3, 1.1 Hz, H4), 8.32 (d, J = 8.4 Hz, H9), 7.76 (dd, J = 8.5, 7.3 Hz, H3), 7.19 (d, J = 8.4 Hz, H10), 3.16 (s, H12/H13).

^{13}C NMR (126 MHz, DMSO-d_6) δ 161.52 (C6), 160.35 (C7), 157.29 (C11), 134.28 (C9), 133.15 (C2), 132.50 (C4), 132.49 (C2), 125.00 (C3), 123.64 (C1), 118.84 (C5), 112.67 (C10), 107.94 (C8), 44.25 (C12/C13).

^{1}H NMR (500 MHz, DMSO-d_6) δ 8.58 (dd, J = 8.5, 1.1 Hz, H2), 8.46 (dd, J = 7.3, 1.1 Hz, H4), 8.32 (d, J = 8.4 Hz, H9), 7.76 (dd, J = 8.5, 7.3 Hz, H3), 7.19 (d, J = 8.4 Hz, H10), 3.16 (s, H12/H13).

N-(2-(6-(dimethylamino)-1,3-dioxo-1H-benzo[de]isoquinolin-2(3H)-yl)ethyl)pivalamide: **C2**

C2
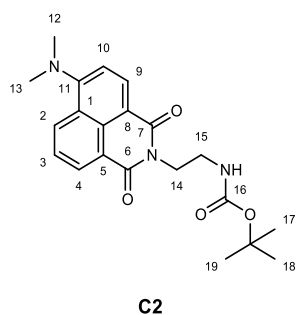
A round-bottom flask charged with **C1** (676 mg, 2.8 mmol, 1 eq), *tert*-butyl (2-aminoethyl)carbamate (492 mg, 3.08 mmol, 1.1 eq) and anhydrous potassium carbonate (774 mg, 5.6 mmol, 2 eq) was refluxed (24 h) in ethanol (25 mL). The resulting solution was concentrated under vacuum and dissolved in a minimal amount of CH_2Cl_2. The solution was then washed with aq. HCl (0.5 M, 100 mL), followed by water and saturated sodium bicarbonate. Removal of the CH_2Cl_2 yielded **C2** as an orange solid (995 mg, 2.60 mmol, 92%). TLC: R_f = 0.8 (Alumina, EtOAC).

HR-MS (ESI, pos): m/z calcd. for $C_{21}H_{25}N_3O_3H^+$ 384.1918 found 384.1912 $[M+H]^+$.

^{1}H NMR (500 MHz, DMSO-d_6) δ 8.47 (d, J = 8.5 Hz, 1H, C2*H*), 8.41 (d, J = 7.2 Hz, 1H, C4*H*), 8.30 (d, J = 8.2 Hz, 1H, C9*H*), 7.72 (t, J = 8.0 Hz, 1H, C3*H*), 7.19 (d, J = 8.4 Hz, 1H, C10*H*), 6.85 (t, J = 6.2 Hz, 1H, HN), 4.10 (t, J = 5.9 Hz, 2H, C14*H*), 3.23 (d, J = 6.1 Hz, 2H, C15*H*), 3.06 (s, 6H, C12*H*, C13*H*), 1.23 (s, 9H, C17*H*, C18*H*, C19*H*).

^{13}C NMR (126 MHz, DMSO-d_6) δ 163.81(C6), 163.15 (C7), 156.37 (C11), 155.68 (C16), 132.03 (C9), 131.19 (C2), 130.32 (C4), 129.69 (C, 124.92, 124.26, 122.60, 113.81, 112.96, 77.39, 44.38, 37.86, 28.08.

2-(2-aminoethyl)-6-(dimethylamino)-1H-benzo[de]isoquinoline-1,3(2H)-dione : **C3**
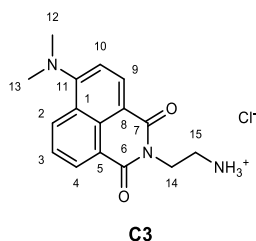
The Boc-protected dye **C2** (500 mg, 1.3 mmol, 1 eq) was dissolved in diethylether (10 mL) on ice and HCl in dioxane (4 M, 3.25 mL, 10 eq) was added. The solution was allowed to warm to RT and stirred for 4 hours. The solvent was then removed and the product was dried to yield **C3** · HCl as a red-orange salt (415 mg, 1.3 mmol, quant).

MS-Direct Injection: (ESI, pos) m/z: 284 [100 %, M+H]$^+$.

^{1}H NMR (500 MHz, DMSO-d_6) δ 8.53 (d, J = 8.5 Hz, 1H, C1*H*), 8.46 (d, J = 7.2 Hz, 1H, C5*H*), 8.35 (d, J = 8.3 Hz, 1H, C9*H*), 8.02 (s, 3H, C16*H*), 7.77 (t, J = 7.9 Hz, 1H, C6*H*), 7.23 (d, J = 8.3 Hz, 1H, C8*H*), 4.29 (t, J = 6.1 Hz, 2H, C14*H*), 3.17 − 3.12 (m, 2H, C15*H*), 3.10 (s, 6H, C13*H*).

^{13}C NMR (126 MHz, DMSO-d_6) δ 164.30 (C12), 163.54 (C11), 156.62 (C7), 132.28 (C9), 131.64 (C1), 130.56 (C5), 129.88 (C3), 125.01 (C6), 124.15 (C2), 122.58 (C4), 113.50 (C10), 112.95 (C8), 44.43 (C13), 37.54 (C15), 37.26 (C14).

N-(2-(6-(dimethylamino)-1,3-dioxo-1H-benzo[de]isoquinolin-2(3H)-yl)ethyl)-5-((3aS,4S,6aR)-2-oxohexahydro-1H-thieno[3,4-d]imidazol-4-yl)pentanamide: **biot-4DMN**

biot-4DMN
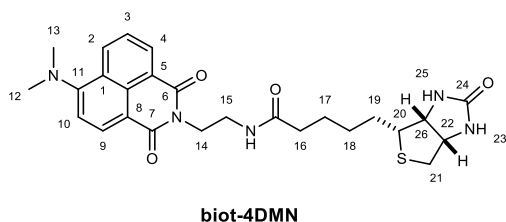
Biotin-pentafluorophenol (50 mg, 0.12 mmol, 1 eq) was dissolved in DMF (5 mL) and **C3 · HCl** (42.85 mg, 0.132 mmol, 1.1 eq) were added to the solution. The reaction was started by addition of triethylamine (50 uL, 0.36 mmol, 3 eq), and the reaction mixture was stirred (18 h). The volatiles were evaporated *in vacuo* and CH$_2$Cl$_2$ (2 mL) was added. The suspension

was filtered and the pure biotinylated fluorophore **biot-4DMN** was isolated by column chromatography (Al_2O_3, CH_2Cl_2:MeOH 9:1) as an orange solid (41 mg, 0.08 mmol, 67%).

HR-MS (ESI, pos): m/z calcd. for $C_{26}H_{31}N_5O_4SH^+$ 510.2170 found 510.2167 $[M+H]^+$.

^{1}H NMR (500 MHz, DMSO-d_6) δ 8.52 (dd, J = 8.5, 1.2 Hz, 1H, C1*H*), 8.46 (dd, J = 7.2, 1.1 Hz, 1H, C5*H*), 8.35 (d, J = 8.2 Hz, 1H, C9*H*), 7.87 (t, J = 6.1 Hz, 1H, C16*H*), 7.77 (dd, J = 8.5, 7.3 Hz, 1H, C6*H*), 7.23 (d, J = 8.3 Hz, 1H, *C8H*), 6.37 (d, J = 21.8 Hz, 2H, C25*H*, C23*H*), 4.30 (dd, J = 7.8, 4.9 Hz, 1H, C24*H*), 4.12 (t, J = 6.1 Hz, 2H, C14*H*), 4.06 (ddd, J = 7.4, 4.5, 1.6 Hz, 1H, C21*H*), 3.39 – 3.35 (m, 2H, C15*H*), 3.09 (s, 6H, C13*H*), 3.00 (ddd, J = 8.5, 6.3, 4.4 Hz, 1H, C26*H*), 2.80 (dd, J = 12.4, 5.1 Hz, 1H, C20*H*), 2.57 (d, J = 12.4 Hz, 1H, C21*H*), 1.94 (t, J = 7.4 Hz, 2H, C16*H*), 1.57 – 1.49 (m, 1H, C19*H*), 1.39 (qd, J = 8.4, 5.7, 4.5 Hz, 2H, C17*H*), 1.28 – 1.16 (m, 2H, C18*H*).

^{13}C NMR (126 MHz, DMSO-d_6) δ 172.64 (C17), 164.32 (C12), 163.67 (C11), 163.14 (C26), 156.92 (C7), 132.61 (C9), 131.79 (C1), 130.91 (C5), 130.20 (C3), 125.47 (C6), 124.75 (C2), 123.04 (C4), 114.17 (C10), 113.49 (C8), 61.40 (C28), 59.65 (C24), 55.80 (C22), 44.88 (C13), 36.88 (C15), 35.72 (C18), 28.54 (C19), 28.44 (C21), 25.63 (C25).

Cloning & expression of Sav-SOD K121A. The Sav-SOD chimera gene was synthesized and cloned in to the NcoI and EagII restriction sites of the pET28 vector by Gene Universal Inc. (Newark, DE). Site directed mutagenesis was achieved using the primers listed in Table 11, followed by BsaI/DpnI digestion and ligated using T4 ligase. Mutations were verified by Sanger sequencing performed by Microsynth (Balgach, Switzerland). The expression and purification of Sav and Sav-SOD proteins were achieved using auto induction media (ZYP-5052) followed by cell lysis and purification using iminobiotin agarose affinity chromatography as previously described.[105]
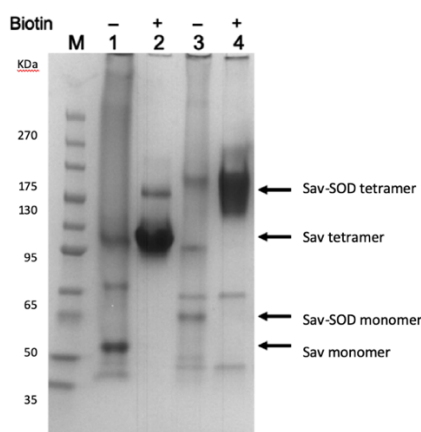
Figure 27 SDS-PAGE analysis of Sav and Sav-SOD soluble protein fractions after thermal shift assay (TSA) analysis in the presence and absence of excess biotin as cofactor. 1) Sav sample after TSA 2) Sav and biotin sample after TSA 3) Sav-SOD sample after TSA 4) Sav-SOD with biotin sample after TSA.
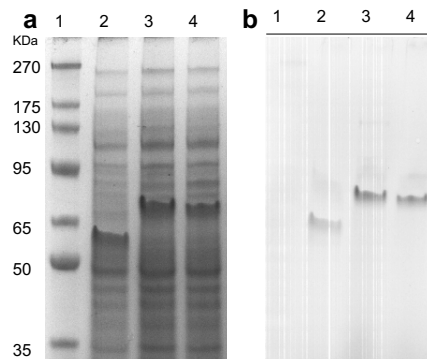
Figure 28 SDS-PAGE Analysis and B4F binding assay of E.coli lysate expressing Sav and Sav-SOD mutants. a Coomassie blue staining of total cell lysates. Lane: 1-Ladder, 2-Sav, 3-Sav-SOD and 4-Sav-SOD-K121A b. B4F biotin binding assay. Lane 1-Ladder, 2-Sav, 3-Sav-SOD, 4-Sav-SOD-K121A.

Sequence of Sav-SOD expression plasmid. *pET-28a-Sav-SOD VECTOR SEQUENCE*

Sav_SOD (3,4-loop)

CGTGATCAGGCAGGTATTACCGGCACCTGGTATAATCAGCTGGGTAGCACCTTTATTGTTACCGCAGGCGCAGATGGTGCACT
GACCGGTACGTATGAAAGCGCAGTTGGTGCCGGCGCCGACAACTTTGCCAACATTCCGCCAGAACGCTACGTCCAGGTCAATG
GGACTCCGGGTCCCGACGAGACGACGTTGACCACCGGCGACGCCGGCAAGAATGCAGAAAGCCGTTATGTTCTGACAGGTCGT
TATGATAGCGCACCGGCAACCGATGGTAGCGGCACCGCACTGGGTTGGACCGTTGCATGGAAAAATAACTATCGTAATGCACA
TAGCGCAACCACCTGGTCAGGTCAGTATGTTGGTGGTGCAGAAGCACGCATTAATACCCAGTGGCTGCTGACCAGCGGCACTA
CTGAAGCAAATGCCTGGAAAAGCACCCTGGTTGGTCATGATACCTTTACCAAAGTTAAACCGAGCGCAGCAAGCATTGATGCA
GCAAAAAAAGCCGGTGTGAATAATGGTAATCCGCTGGATGCAGTTCAGCAGTAA

Legend for Sav_SOD: Sav_Domain; sodC_Dimerization_Domain

Library generation. Mutant plasmid libraries were transformed into BL21 (DE3) cells and selected on LB-Agar plates containing kanamycin (50 mg/L). Individual colonies were used and inoculated in auto-induction media (500 µL) containing kanamycin (100 mg/L) in a deep-well plate and grown for 24 hours at 30 °C. Cells were harvested, frozen and lysed in Tris-buffer (20 mM, pH 7.4) containing DNAse (0.05 mg/mL) and lysozyme (1 mg/mL). After centrifugation, the resulting cell free extract (CFE) was used to determine the relative free biotin binding sites.[105] (BBS) and then used for catalysis trials. Wells with exceptionally low BBS (< 20%) were discarded.

Auto-induction media composition: tryptone (10 g/L), yeast extract (5 g/L), KH_2PO_4 (6.8 g/L), Na_2HPO_4 (7.1 g/L), $(NH_4)_2SO_4$ (3.3 g/L), glycerol (5 g/L), glucose-monohydrate (0.55 g/L) and lactose (2 g/L) dissolved in miliQ water.

Table 11 Primers used for saturation mutagenesis.

Name	SEQUENCE	
oRP265	TATAAGGTCTCAAGCANDTGCATGGNDTNDTACCCTGGTTGGTCATGATACC	NDT(118, 121, 122)
oRP266	CATATGGTCTCTTGCTTCAGTAGTGCCGGAGGTCAGCAGCCACTGGGTATT	Rev (118, 121, 122)
oRP299	TATAAGGTCTCAAGCATCTGCATGGTTTGGTACCCTGGTTGGTCATGATACC	pFC101
oRP300	TATAAGGTCTCAAGCAGTTGCATGGGATTGTACCCTGGTTGGTCATGATACC	pFC102
oRP301	TATAAGGTCTCAAGCAGTTGCATGGGGTTATACCCTGGTTGGTCATGATACC	pFC103
oRP302	TATAAGGTCTCAAGCAGGTGCATGGGGTGGTACCCTGGTTGGTCATGATACC	pFC104
oRP303	TATAAGGTCTCAAGCACATGCATGGTTTGGTACCCTGGTTGGTCATGATACC	pFC105
oRP304	TATAAGGTCTCAAGCAGGTGCATGGTCTGGTACCCTGGTTGGTCATGATACC	pFC106
oRP309	TATAAGGTCTCAAGCACTTGCATGGAGTCGTACCCTGGTTGGTCATGATACC	pFC107
oRP310	TATAAGGTCTCAAGCAGGTGCATGGTTTTGTACCCTGGTTGGTCATGATACC	pFC108
oRP313	TATAAGGTCTCAAGCAGGTGCATGGTTTGGTACCCTGGTTGGTCATGATACC	pFC109
oRP314	TATAAGGTCTCAAGCAAGTGCATGGTGTTGTACCCTGGTTGGTCATGATACC	pFC110
oRP315	TATAAGGTCTCAAGCAGGTGCATGGATTGGTACCCTGGTTGGTCATGATACC	pFC111
oRP316	TATAAGGTCTCAAGCAGGTGCATGGTATAGTACCCTGGTTGGTCATGATACC	pFC112
oNI84	TATAAGGTCTCAAGCAGGTGCATGGTATGGTACCCTGGTTGGTCATGATACC	pFC113
oRP306	CTATAAGGTCTCGCTGCTGACCNDTGGCACTACTGAAGCA	NDT S112
oRP307	CTATAAGGTCTCGCTGCTGACCVMAGGCACTACTGAAGCA	VMA S112
oRP308	CTATAAGGTCTCGCTGCTGACCVMAGGCACTACTGAAGCA	Rev S112
oNI65	CTAAATGGTCTCGTGGTTGTTGACCACAGGCACTACTGAAG	S112T
oNI66	CTAAATGGTCTCGTGGTTGTTGACCAATGGCACTACTGAAG	S112N
oRP319	CATTAGGTCTCAACCACTGGGTATTAATGCGTG	Rev S112 2
oNI98	GATTACTGGTCTCAATGGCNDTNDTGAAGCAGGTGCATGG	NDT (T114, T115) N-GGG
oNI99	GATTACTGGTCTCAATGGCVMAVMAGAAGCAGGTGCATGG	VMA (T114, T115) N-GGG
oNI124	CATATGGTCTCTTGCTTCATTGCTACCATTGGTCAGCAGCCACTGGG	T114S T115N
oNI125	TATAAGGTCTCAAGCAGGTGCATGGGGCGGCACCCTGGTTGGTCATGATACC	S112N N118G K121G S122G
oNI131	GATTACTGGTCTCGCCTGTTGTCAACAACCACTGGG	pFC101S112T-rev
oNI132	GATTACTGGTCTCACAGGCVMAVMAGAAGCATCTGCATGGTTTGG	pFC101S112T-T114VMA-T115VMA
oNI133	GATTACTGGTCTCACAGGCNDTNDTGAAGCATCTGCATGGTTTGG	pFC101S112T-T114NDT-T115NDT
oNI134	CATATGGTCTCTTGCTTCTGCAGTGCCATTGGTCAGCAGCCACTGGGTATT	T115A

Determination of the biotin-binding sites. First, a solution of B4F (1 μM) in MOPS (50 mM, pH = 7.4) was prepared. Then, 200 μL of this B4F solution was added to a clear micro well-

63

plate, followed by 10 μL of cell-free extract (cfe). The negative control contained 10 μL of lysis buffer (LB). The positive control (Sav$_{STD}$) contained 10 μL of 40 μM Sav as 100% saturation point. It should be noted that, the lysis buffer and protein components can alter the B4F fluorescence up to 10-20%. The fluorescence of each well was measured using a Tecan fluorimeter Infinite M1000Pro. The total concentration of Sav should be linear within this region and calculated using the following equation:

$$[SAV_{CFE}](\mu M) = [(F_{CFE} - F_{SAVSTD}) - (F_{LB} - F_{SAVSTD})]/(-F_{LB} - F_{SAVSTD}) * 20$$

F_{CFE} = Fluorescence intensity of Cell free extract

F_{SAVSTD} = Fluorescence intensity of Sav Standard to bind all free BF4

F_{LB} = Fluorescence intensity of Lysis buffer

ITC analysis. Biotin binding affinities were measured using a Microcal VP-ITC as previously described by Stayton and coworkers.[181] Sav-SOD solutions were prepared at 40 μM in 50 mM sodium phosphate buffer containing 100 mM sodium chloride at pH = 7.75 and titrated with 5 μL injections of buffer containing 250 μM biotin. Measurements were performed at 10, 25 and 40 °C. Biotin binding constant *Ka*, enthalpy *h*, and number of binding sites at each temperature were calculated using ITC data analysis origin software (MicroCal).

K_D (nM)	1.6	4.17	25.2
T (°C)	10	25	40
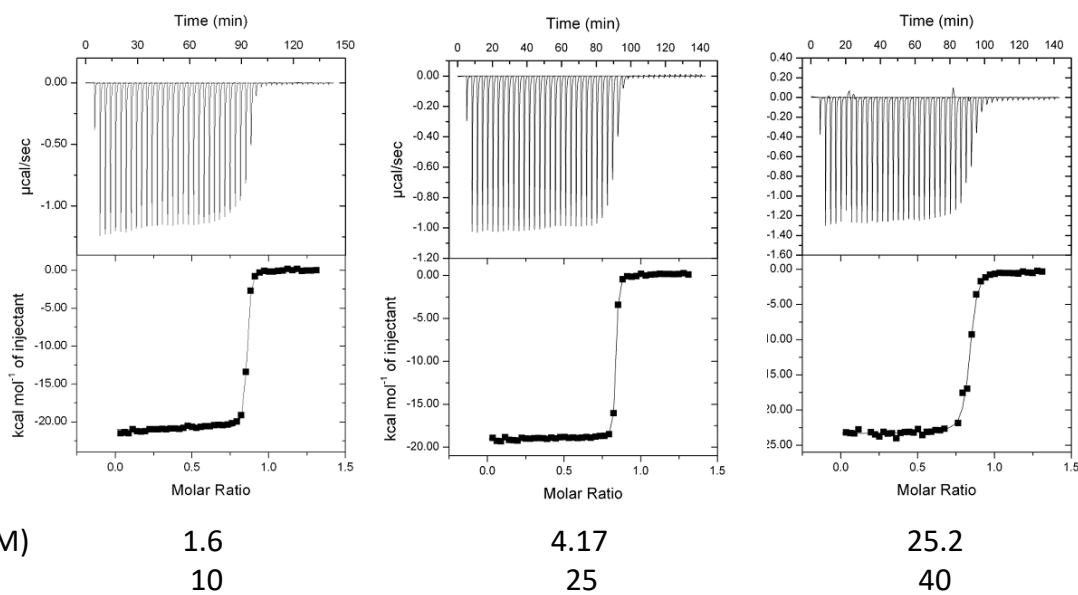
Figure 29 Variable temperature ITC analysis of biotin binding to Sav-SOD K121A.

Table 12 Thermodynamic parameters for biotin binding to Sav and Sav mutant proteins

Protein	K_a (M^{-1})	ΔG° (kcal/mol)	ΔH° (Kcal/mol)	TΔS° (Kcal/mol)	REF
Wild type[a]	$2.5 \cdot 10^{13}$	- 18	- 24.9 (0.4)	- 6.6 (0.4)	Chu[181]
CP51/46[a]	$2.3 (0.4) \cdot 10^7$	- 10.0	- 13.8 (0.8)	- 3.8 (0.8)	Chu[181]
Sav-SOD K121A[a]	$2.4 (0.3) \cdot 10^8$	- 11.4	- 20.9 (0.5)	- 9.49 (0.5)	This paper

[a]Collected at 25 ºC.

Table 13 Thermodynamic parameters for biotin binding to Sav-SOD K121A as a function of temperature.

Temp (Kelvin)	K_a (M^{-1})	ΔG (kcal/mol)	ΔH (kcal/mol)	TΔS (kcal/mol)
283.3	$6.1 (0.5) \cdot 10^8$	- 11.38	- 19.0 (0.3)	- 7.62
298.3	$2.4 (0.3) \cdot 10^8$	- 11.44	- 20.9 (0.5)	- 9.49
313.3	$4.0 (0.5) \cdot 10^8$	- 10.86	- 23.3 (1.1)	- 12.4

Fluorescent-probe binding assay of Sav mutants. Fluorescence emission spectra using the dielectric sensitive probe **biot-4DMN** was recorded on a Tecan M100 micro-plate reader using

40 mM sodium phosphate buffer pH = 7.0 employing 1.25 mM of Sav K121A or Sav-SOD K121A with 40 μM **biot-4DMN** (from a 2 mM DMSO stock solution). The excitation wavelength was 420 nm and the emission spectra were collected every 1 nm from 450 – 850 nm.
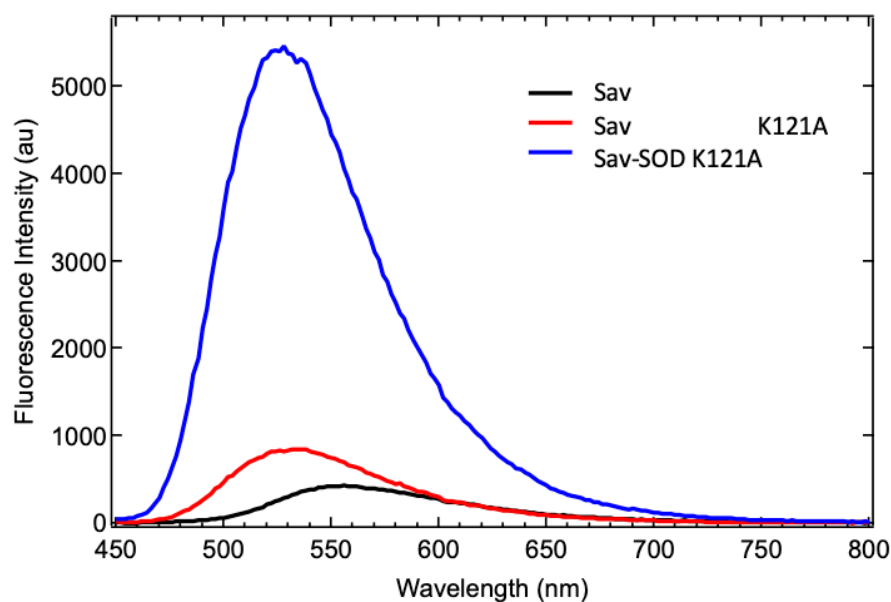
*Figure 30 Fluorescence intensity profiles of the solvatochromatic probe **biot-4DMN** irradiated at 420 nm in the presence and absence of Sav hosts. Shown in black is fluorescence of 40 μM **biot-4DMN** in the absence of a Sav protein host. Red is in the presence of 1.25 mM Sav K121A. Blue is in the presence of 1.25 mM Sav-SOD K121A.*

Thermal shift assays for protein melting-point determinations. Protein melting curves were determined using the protein thermal shift reagent kit (Applied Biosystems) as directed using a StepOne Real Time PCR system and buffer described for ITC analysis. A total sample volume of 20 μL/well was used with a final concentration of 5.25 mg/mL of protein and +/– (0.25 mg/mL) biotin. The melting curves of 3 independent wells were averaged and protein melting points estimated at the temperature with the minimum value on the dF/dT curve.
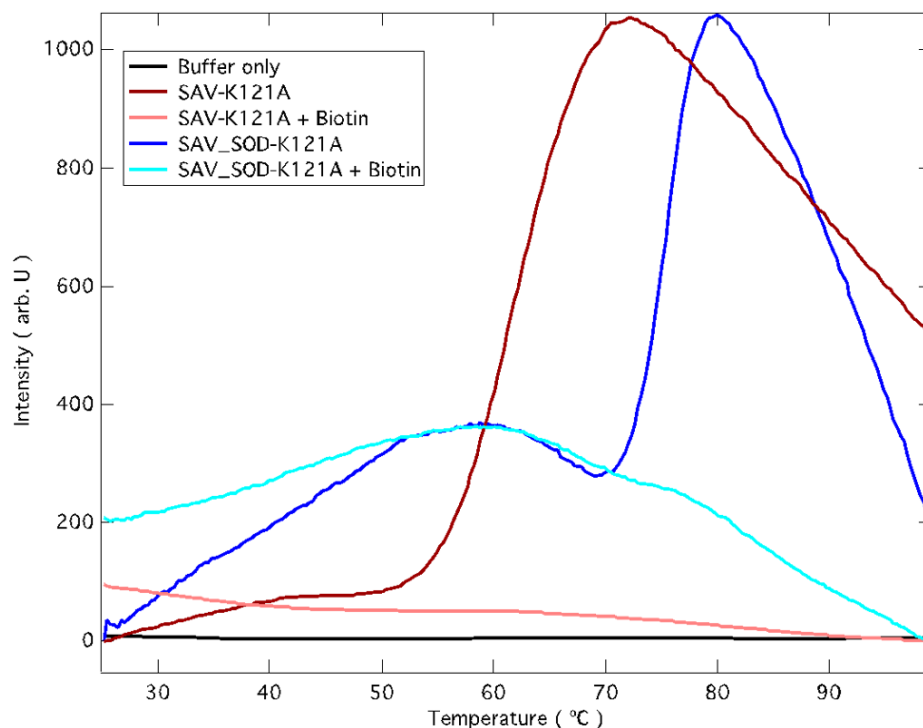
Figure 31 Protein melting curves for Sav-K121A and Sav-SOD K121A in the presence and absence of excess biotin cofactor. The curves corresponding to buffer only (black), Sav K121A(magenta), Sav K121A in the presence of 1 mM biotin (salmon), Sav-SOD K121A (blue), Sav-SOD K121A in the presence of 1 mM biotin (teal) are shown. Protein melting points of 60.9 and 74.6 ºC are estimated for Sav K121A and Sav-SOD K121A respectively.

Protein ESI-MS characterization. Proteins were dissolved in Mili-Q water, 0.1% formic acid pH = 2.5 with a final concentration of 0.2 mg/mL and clarified by centrifugation. A HPLC (Shimadzu, equipped with a Jupiter® 5 μm C4 300 Å)-ESI QTof (Bruker maXis II QTOF ESI) system was used to record the data. The ESI-QTof mass spectrometer was calibrated with ESI-Tof TuneMix (Agilent). The charge envelope from 800-1200 m/z was deconvoluted using the Compass DataAnalysis software (Bruker Daltonics) with the Maximum Entropy set-up. The results are summarized in the HR-MS section.
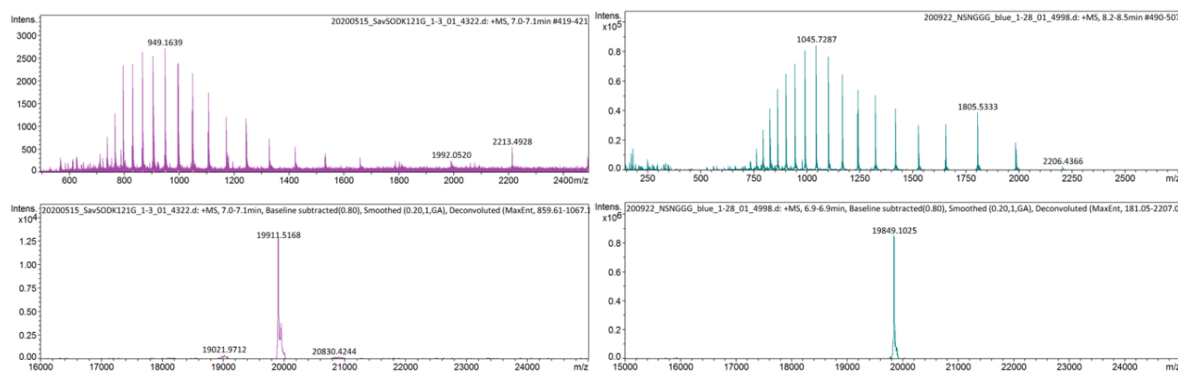
Figure 32 Charge state envelope of streptavidin (top) with the deconvoluted mass (bottom) observed by the HR-MS experiments. Sav-SOD K121G (left) and Sav-SOD S112N T114S T115N N118G K121G S122G (right) were taken as examples.

Table 14 Theoretical and observed masses for selected streptavidin mutants used in this study.

Mutant	Mass expected (Da)	Mass observed (D
Sav WT	16424.93	16425.56
Sav K121G	16353.81	16354.72
Sav K121A	16367.83	16339.64
Sav-SOD WT	19981.73	19981.17
Sav-SOD K121G	19910.61	19911.73
Sav-SOD K121A	19924.64	19925.20
Sav-SOD N118S K121F S122G	19943.68	19943.54
Sav-SOD N118G K121G S122G	19823.53	19825.15
Sav-SOD S112T N118S K121F S122G	19957.71	19959.22
Sav-SOD S112N N118G K121G S122G	19850.56	19852.48
Sav-SOD S112T T114T T115A N118S K121F S122G	19927.69	19927.95
Sav-SOD S112N T114S T115N N118G K121G S122G	19849.53	19849.10

Native-MS spectra of Sav mutants. Native mass spectrometry was performed using an adapted protocol based on Eckard and Spiess.[115] The lyophilized protein was dissolved in a 10 mM ammonium acetate pH 8.0 solution to a final concentration of 2 mg/mL. 1.2 eq of **biot-Au 2** in DMSO were added and the sample was incubated for 15 min at room temperature. The spectrum was then recorded by direct injection with a syringe pump at a flow rate of 5 μL/min using the maXis II Q-ToF (Bruker) ESI-MS. ESI Low Concentration Tuning Mix from Agilent was used as calibrant. The tetramer of streptavidin and the Sav-SOD mutants, as well as the complex with **biot-Au 2** could be detected between 3300-7000 m/z. The charge states were deconvoluted using the Compass DataAnalysis software (Bruker Daltonics) with the Maximum Entropy set-up. The results are summarized in Table 15 and Table 16.

*Table 15 Mass adducts of Sav WT with different equivalents of **biot-Au 2** per tetramer.*

|	Mass calculated (Da)	Mass measured (Da)	Deviation (ppm)
Sav WT	65699.72	65703.55	+ 58.30
Sav WT + **biot-Au 2**	66458.00	66461.36	- 50.56
Sav WT + 2 **biot-Au 2**	67216.28	67219.36	+ 45.82
Sav WT + 3 **biot-Au 2**	67974.57	67976.98	-35.45
Sav WT + 4 **biot-Au 2**	68732.86	68733.75	-12.95
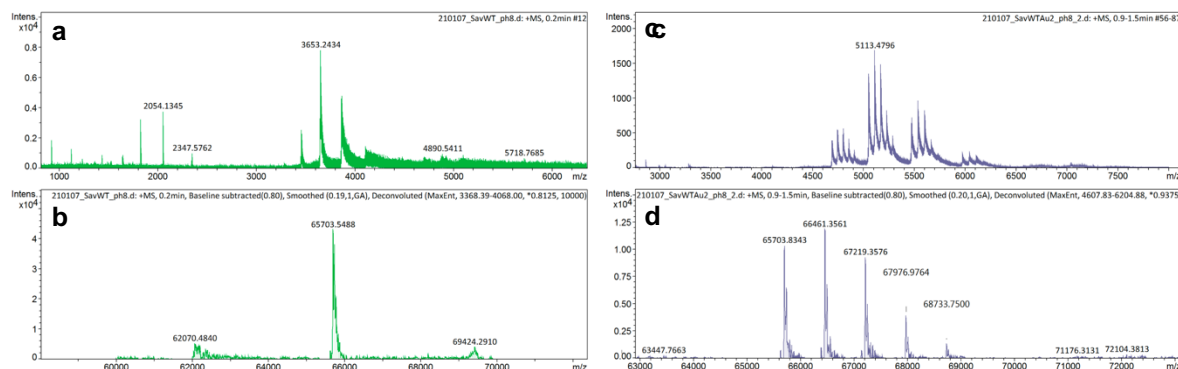
*Figure 33 Native MS analysis of HAMases. **a** Native MS spectrum the charge states of apo Sav WT and **b** deconvoluted spectrum of apo Sav WT; **c** Native MS spectrum the charge states of **biot-Au 2** · Sav WT, **d** deconvoluted spectrum of **biot-Au 2** · Sav WT, highlighting the presence of (up to) four **biot-Au 2** cofactors embedded in homotetrameric Sav WT (See Table 15 for calculated and measured masses).*

*Table 16 Summary of native-MS experiments confirming the incorporation of up to four **biot-Au 2** per tetramer for Sav-SOD K121A and Sav-SOD S112N T114S T115N N118G K121G S122G (N-SN-GGG).*

	Mass calculated (Da)	Mass measured (Da)	Deviation (ppm)
Sav_SOD K121A	79698.56	79738.13	+ 496.50
Sav_SOD K121A + **biot-Au 2**	80456.85	80501.57	+ 555.83
Sav_SOD K121A + 2 **biot-Au 2**	81215.14	81258.19	+ 530.07
Sav_SOD K121A + 3 **biot-Au 2**	81973.43	82040.43	+ 817.34
Sav_SOD K121A + 4 **biot-Au 2**	82731.72	82820.94	+ 1078.43
Sav_SOD N-SN-GGG	79398.12	79407.00	+ 111.84
Sav_SOD N-SN-GGG + **biot-Au 2**	80156.41	80162.00	+ 69.74
Sav_SOD N-SN-GGG + 2 **biot-Au 2**	80914.70	80922.00	+ 90.22
Sav_SOD N-SN-GGG + 3 **biot-Au 2**	81672.99	81678.00	+ 61.34
Sav_SOD N-SN-GGG + 4 **biot-Au 2**	82431.28	82436.00	+ 57.26
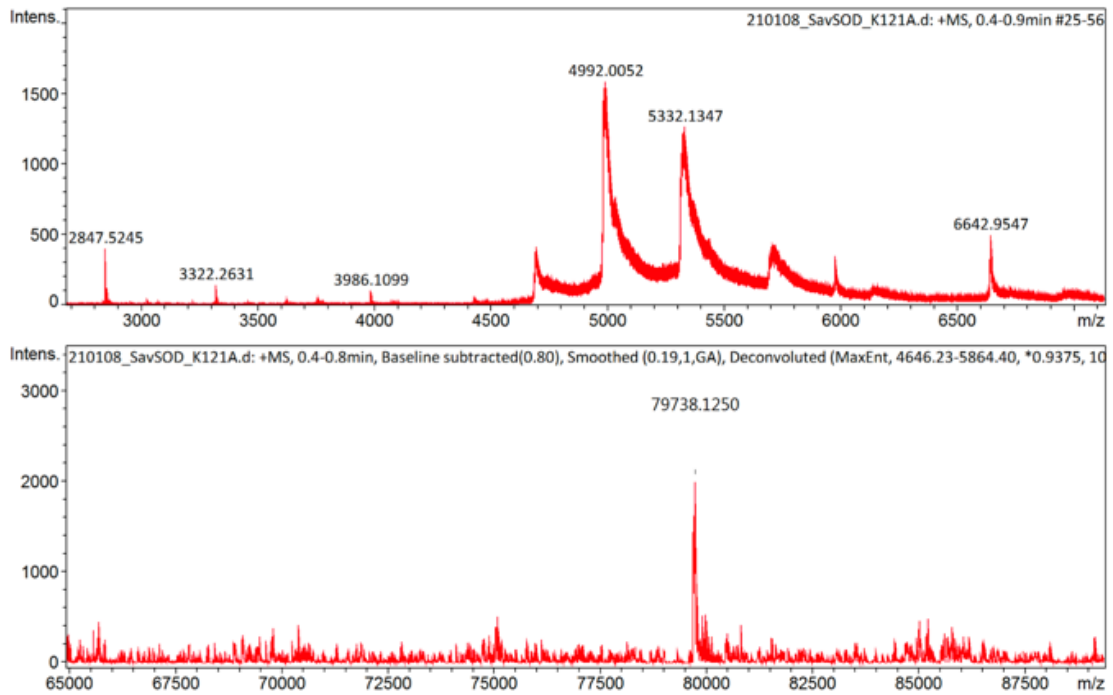
Figure 34 Native MS-Spectrum of Sav-SOD K121A. The upper window displays the charge states and the bottom window displays the deconvoluted spectrum (see Table 16 for calculated masses).
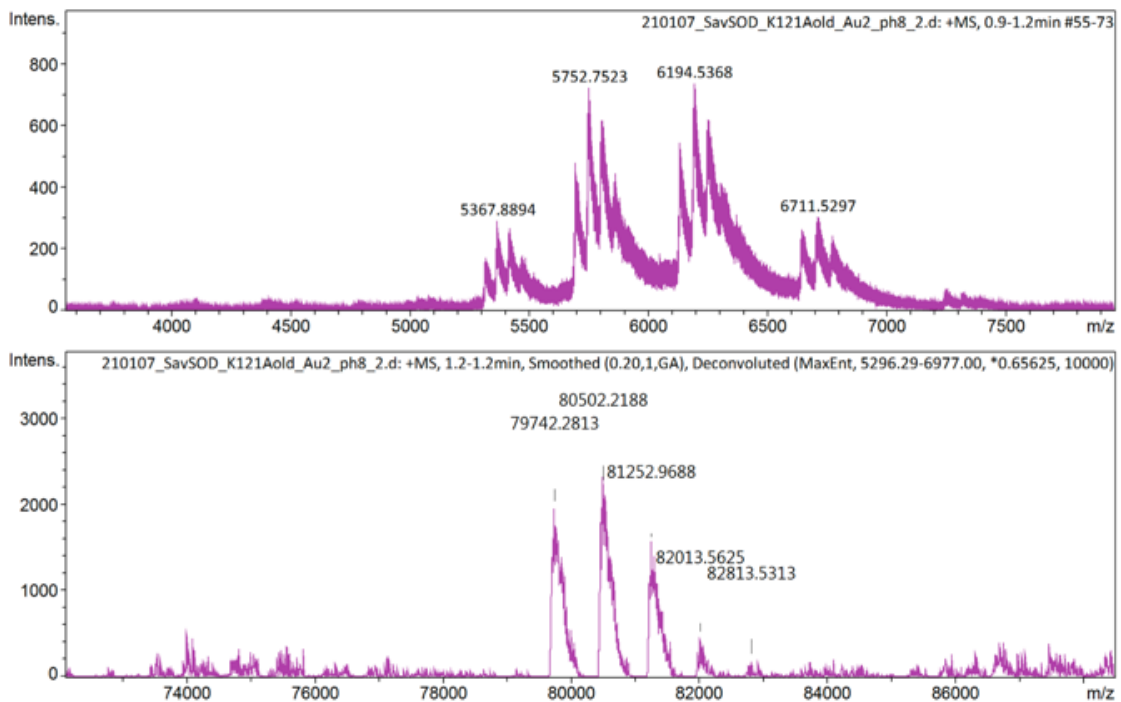
*Figure 35 Native MS-Spectrum of the **biot-Au 2** · Sav-SOD K121A complex. The upper window contains the charge states and the bottom window displays the deconvoluted spectrum (see Table 16 for calculated masses). The spectrum confirms the incorporation of up to four **biot-Au 2** per Sav-SOD tetramer.*
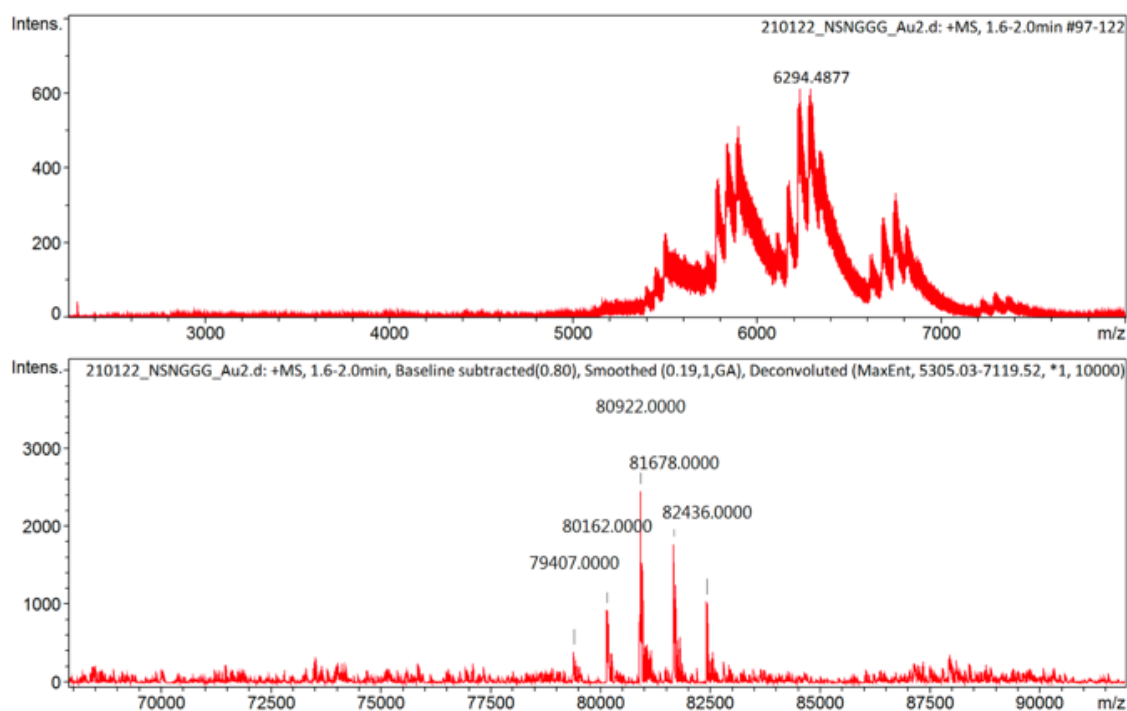
*Figure 36 Native MS-Spectrum of the **biot-Au 2** · Sav-SOD S112N T114S T115N N118G K121G S122G complex. The upper window contains the charge states while the bottom window displays the deconvoluted spectrum (seeTable 16 for calculated masses). The spectrum confirms the incorporation of up to four **biot-Au2** per Sav-SOD tetramer.*

Crystallization and data processing. For co-crystallization experiments, 10 μL of **biot-Au 2** (20 mM in DMSO) was added to a 2 mg/mL protein (in 1 mL 20 mM Tris-HCl, pH 7.4) solution. After 18 h of soaking at rt, the solution was subjected to concentration and buffer-exchange (water) using a Nanosep Centrifugal Device with an Omega Membrane (3K, Pall). Crystals grew within 21 days and were flash frozen. X-ray diffraction patterns were collected at 100 K at the Swiss Light Source beam line PXIII at a wavelength of 0.999 Å and analysed with CCP4i2 Suite.

The structure was resolved using the streptavidin structure pdb:4irw as model. Amino acid residues 1-12, 135-159 of the Sav core and the SOD loop are not resolved in the electron density, presumably due to the flexibility of these regions. In the biotin-binding pocket, residual electron density in the F_o-F_c map was observed (Figure 37b), as well as one anomalous density ellipsoid per Sav monomer (Figure 37a). The cofactor **biot-Au 2** could be modelled with an occupancy of 100%, with the exception of the gold and chloride atoms.

Their lowered occupancy (Au: 50%, Cl: 0%), is traced back to partial de-coordination of the gold and/or chloride during crystallization.

Despite the use of a "racemic" mixture of the biotinylated NHC-ligand, only the (*S*)-configuration of the NHC was detected in the biotin-binding site. Only three of the four mesityl groups could be unambiguously localized in the F_o-F_c electron density map. The point mutation K121A could be clearly distinguished. The structure was uploaded to the pdb database under the code: 7ALX.

Table 17 Data processing and crystal structure refinement statistics.

	Sav-SOD K121A
Data collection	
Space group	P 4_2 2_1 2
Cell dimension	
a, b, c(Å)	102.53, 102.53, 59.08
α, β, γ(°)	90, 90, 90
Resolution (Å)	45.85 – 1.80
R_{merge}	11.8 (61.6)
I/σ(I)	18.9 (5.4)
$CC_{1/2}$	0.99 (0.96)
Completeness (%)	100 (100)
Redundancy	24.4 (27.2)
Refinement	
Resolution (Å)	45.85 – 1.80
No. reflections	28298
R_{work}/R_{free}	18.2 / 22.1
No. atoms	
Protein	3591
biot-Au 2	164
Water	159
B factors	
Protein	29
biot-Au 2	36
Water	37
R.M.S. deviations	
Bond lengths (Å)	0.013
Bond angles (°)	2.17
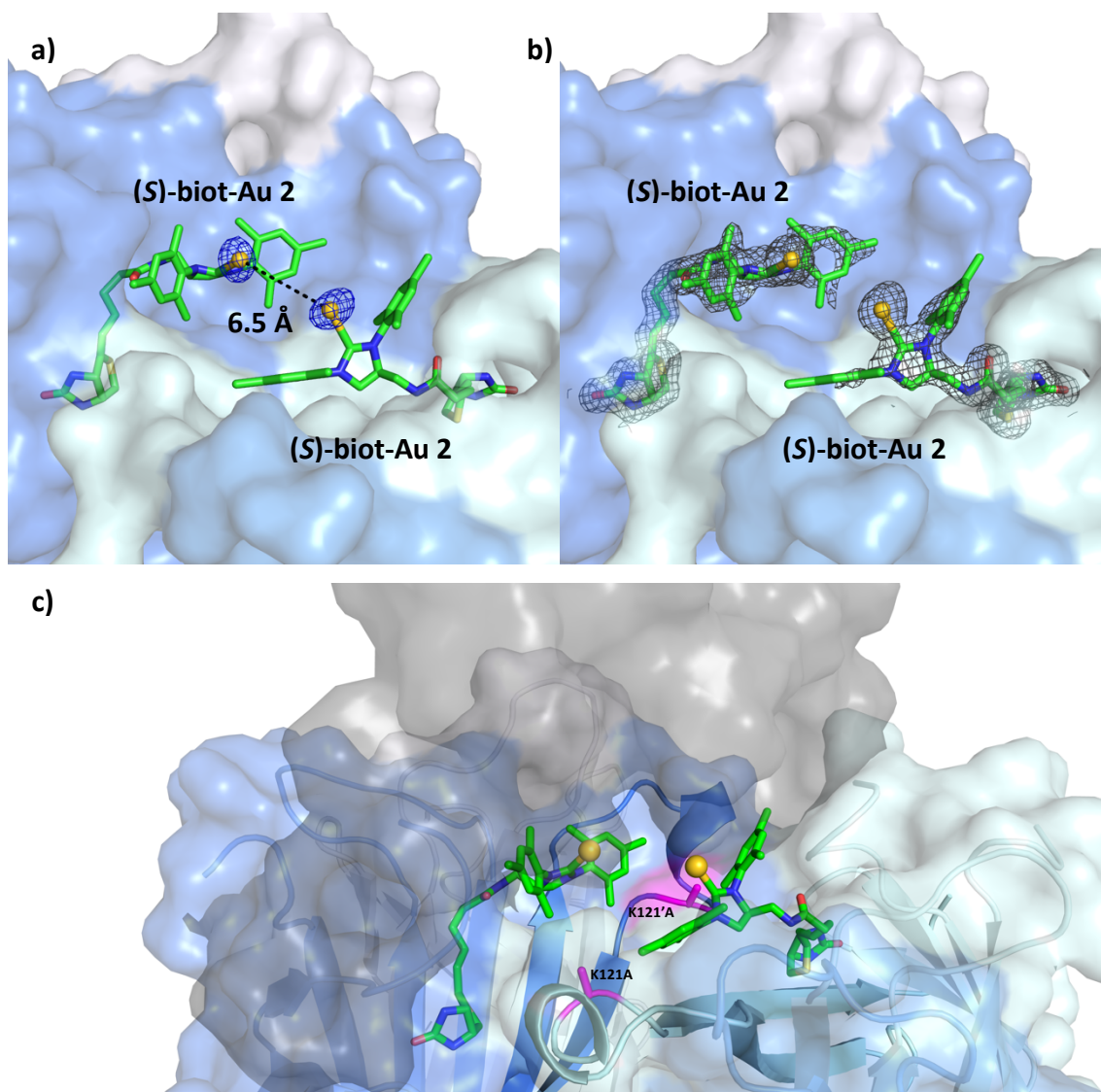
*Figure 37 Close-up view of the refined partial X-ray structure of **biot-Au 2** · Sav-SOD K121A. **a)** Anomalous electron density (displayed as blue mesh at 5σ) assigned to Au and modelled with a 50% occupancy; no electron density for Cl was detected. **b)** Electron density map (2Fo-Fc) for the **biot-Au 2** (displayed as grey mesh at 1σ). **c)** Superposition of the partial X-ray structure of **biot-Au 2** · Sav-SOD K121A with the SOD-lid obtained from the MD simulation. The **biot-Au 2** is represented as colour-coded stick, and Au as yellow sphere. Sav-SOD K121A is displayed as surface (the Sav monomers in different shades of blue and the SOD-lid in grey) and the mutations K121A and the symmetry related K121'A are highlighted as magenta sticks.*

3.2.7 Appendix
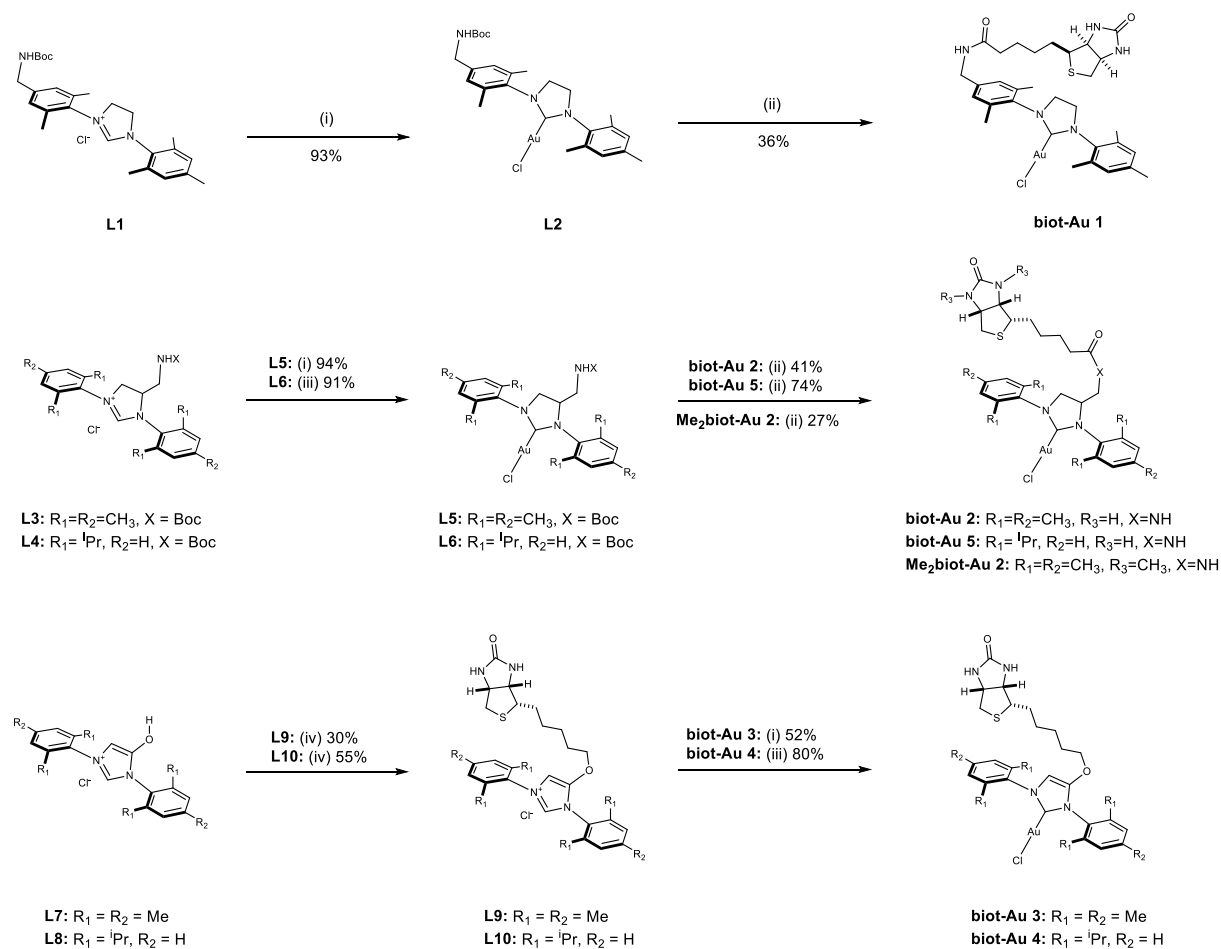
Figure 38 Synthesis of biotinylated gold complexes:(i) Ag₂O, CH₂Cl₂, rt, 4 h then (CH₃)₂SAuCl, CH₂Cl₂, rt, 16 h, (ii) TFA, CH₂Cl₂, rt, 2 h, then biotin pentafluorophenyl ester, NEt₃, DMF, rt, 16 h, (iii) (CH₃)₂SAuCl, K₂CO₃, acetone, 65 °C, 90 min, (iv) biotin mesylate, KI, K₂CO₃, acetone, reflux, 2 days, (v) HCl(g), CH₂Cl₂, rt, 4 h then Boc-Gly pentafluorophenyl ester, NEt₃, DMF, rt, 16 h.
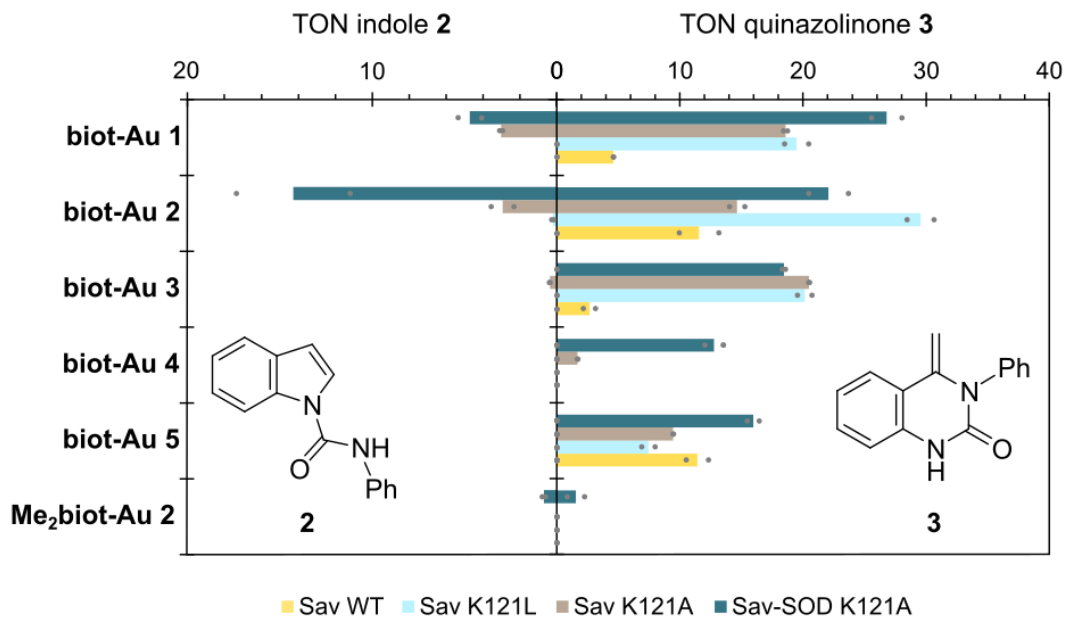
Figure 39 Cofactor screening using purified Sav samples: [biot-Au 2] = 1 mol%, [Sav] = 2 mol%, MES-buffer pH = 5, 37 °C and 24 h reaction time. All reactions were performed in duplicate.
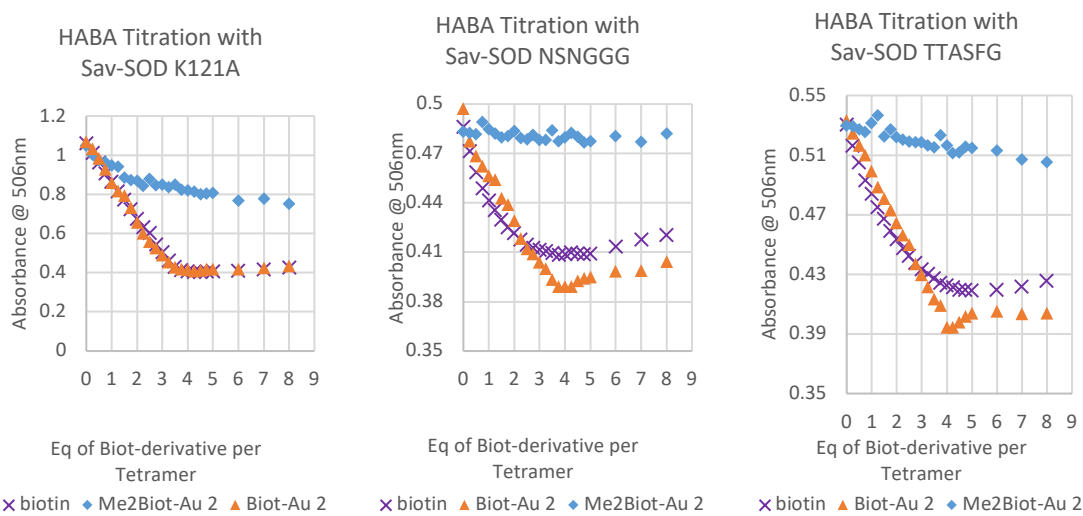
*Figure 40 HABA displacement titration of biotin and biotinylated cofactors **biot-Au 2** and **Me2biot-Au 2** in the presence of Sav-SOD K121A, Sav-SOD S112N T114S T115N N118G K121G S122G and Sav-SOD S112T T114T T115A N118S K121F S122G. For procedure see [12].*

Time (h)	TON quinazo-linone **3**	TON indole **2**
1	3.7	0.8
3	8.0	3.0
6	11.9	6.6
24	22.1	14.3
72	33.1	20.5
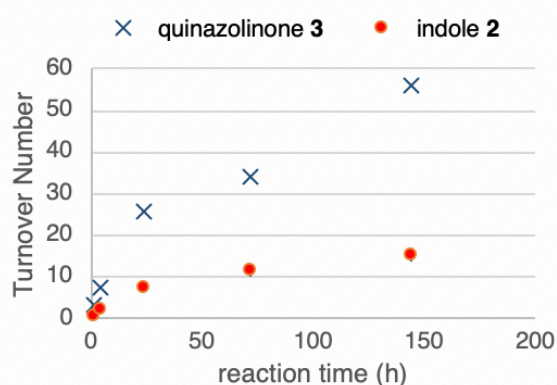
*Figure 41 Time course of the HAMase conversion using **biot-Au 2** · Sav-SOD K121A: [**biot-Au 2**] 1 mol%, [Sav] = 2 mol%, MES-buffer pH 5, 37 °C.*

Buffer	TON quinazolinone **3**	TON indole **2**
Acetate	15.5	12.0
Tris HCl	9.5	6.1
PBS	13.5	7.3
MQ H$_2$O	8.5	3.9
MES	21.7	10.4
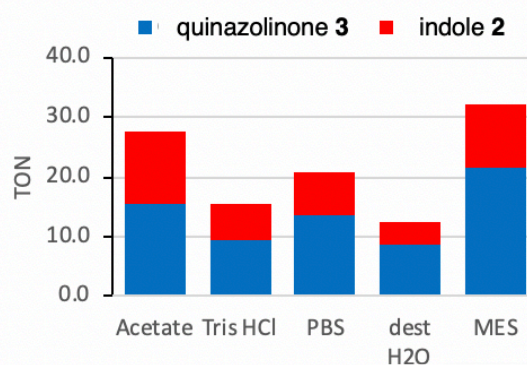
*Figure 42 Reactivity of **biot-Au 2** · Sav-SOD K121A in different aqueous buffer: [**biot-Au 2**] 1 mol%, [Sav] 2 mol%, pH 5, 37 °C, 24 h reaction time.*

pH	TON quinazo-linone **3**	TON indole **2**	TON dimer
2	17.8	1.0	1.4
3	12.5	8.6	7.1
4	18.3	10.9	5.3
5	32.8	11.0	7.1
6	28.4	20.7	16.7
7	16.0	5.8	6.8
8	4.1	0.7	0.5
9	1.2	0.2	0.2
10	0.5	0.1	0.2
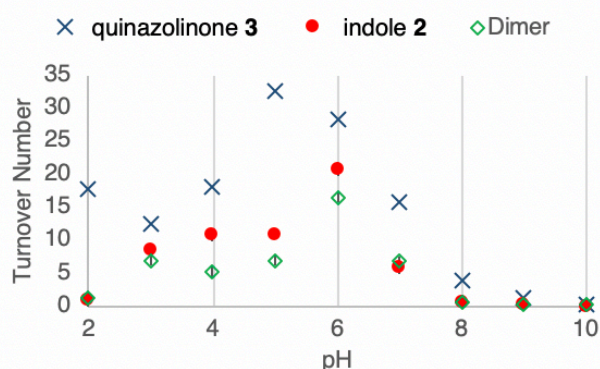
*Figure 43 Reactivity of **biot-Au 2** · Sav-SOD K121A at different pH in MES-Buffer: [**biot-Au 2**] 1 mol%, [Sav] 2 mol%, 37 °C and 72 h reaction time.*

77

[Sub 1] = mM	TON quinazo-linone 3	TON indole 2
5	25.5	20.3
2.5	20.7	19.0
1.5	16.6	8.1
0.5	10.0	1.8
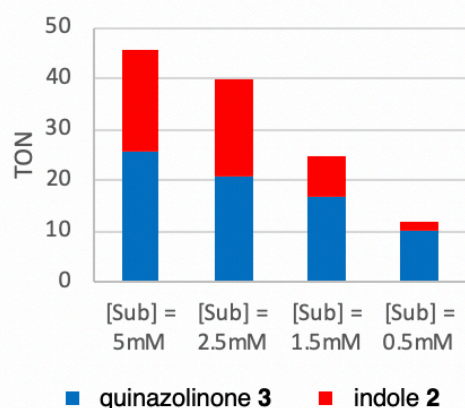
*Figure 44 Performance of chimeric streptavidin with different substrate concentrations. Reaction conditions: [**biot-Au 2**] = 1 mol%, [Sav] = 2 mol%, MES-buffer pH = 5, 37 °C and 28 h reaction time.*

c(**biot-Au 2**) in mol%	TON quinazolinone 3	TON indole 2
1.00	22.1	14.3
0.50	34.0	8.6
0.25	10.9	1.0
0.10	8.4	0
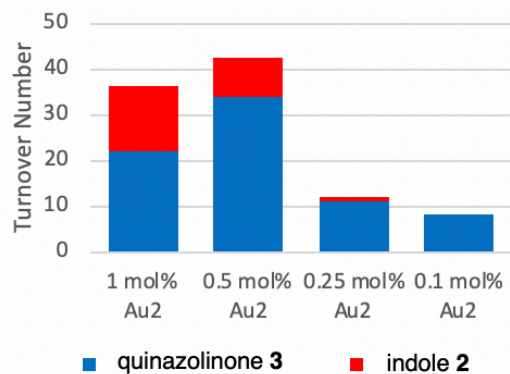
*Figure 45 Reactivity of **biot-Au 2** · Sav-SOD K121A with varying cofactor concentrations. Reaction conditions: [Sav] = 2x[**biot-Au 2**], MES-buffer pH = 5, 37 °C and 24 h reaction time.*
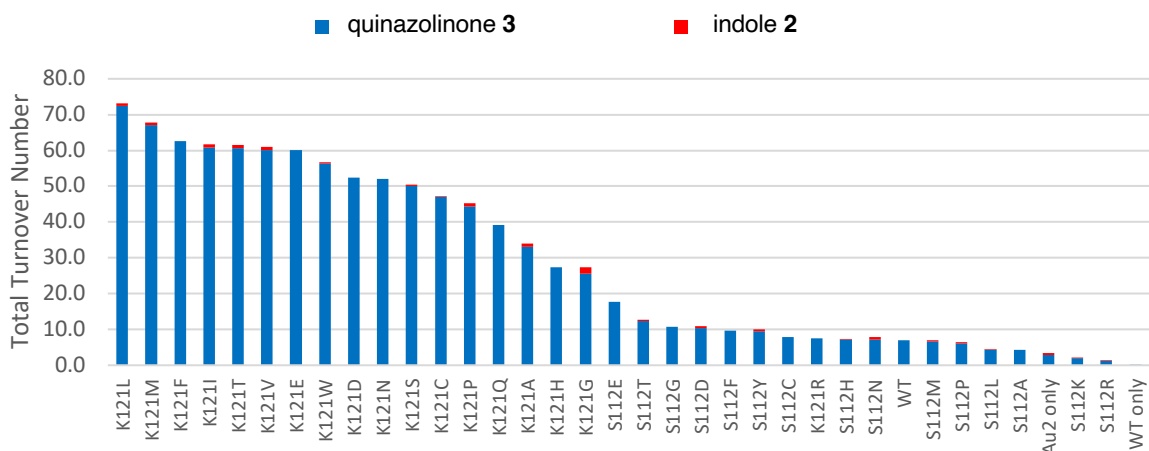
*Figure 46 Screening Sav mutants bearing mutations on position K121 or S112: [**biot-Au 2**] 1 mol%, [Sav] 2 mol%, MES-buffer pH 5, 37 °C and 5 days reaction time.*

78
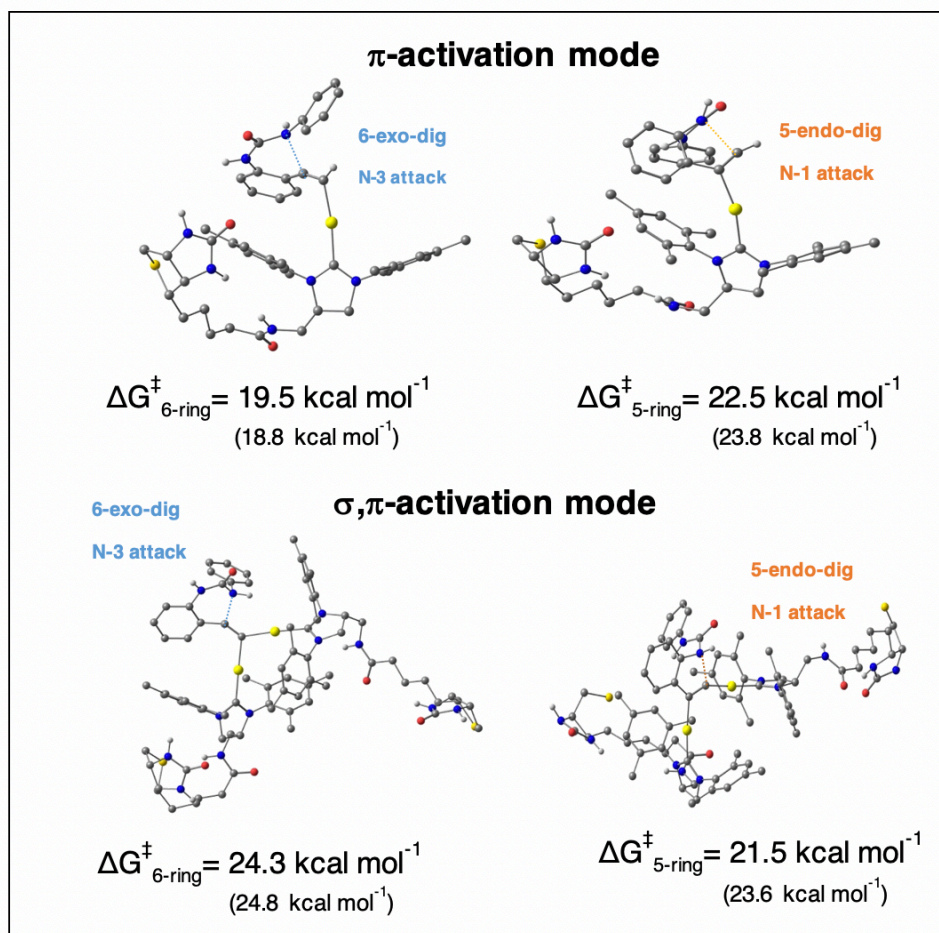
Figure 47 DFT B3LYP-D3 optimized structures of transition states in water for 5-endo-dig and 6-exo-dig additions by $\pi-$ and σ,π–activation modes of **biot-Au 2** with their computed barriers. The Gibbs energies of activation for the non-biotinylated IMes ligand are displayed in parenthesis. Atomic colouring code: yellow, gold; blue, nitrogen; red, oxygen; grey, carbon; hydrogen atoms omitted for clarity.
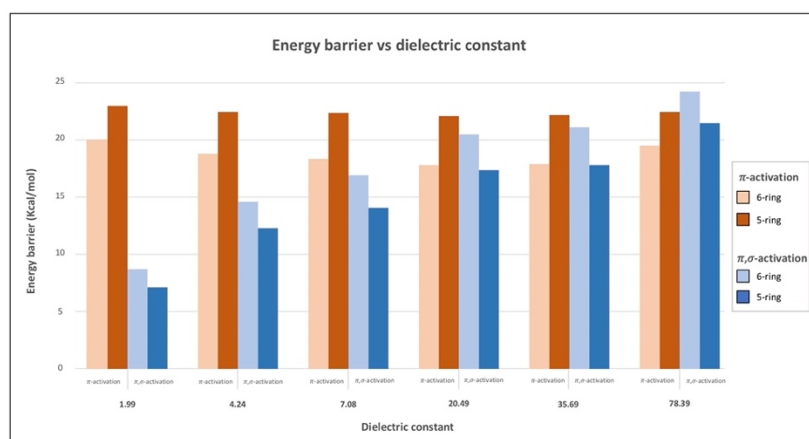
Figure 48 Comparison of the computed barriers for the reactions yielding the 6-exo dig (quinazolinone **3**) and the 5-endo-dig (indole **2**) addition products by both π- and σ,π-activation modes in water (ε = 78.39) and in solvents of varying dielectric constant, using the transition states depicted in *Figure 47*.
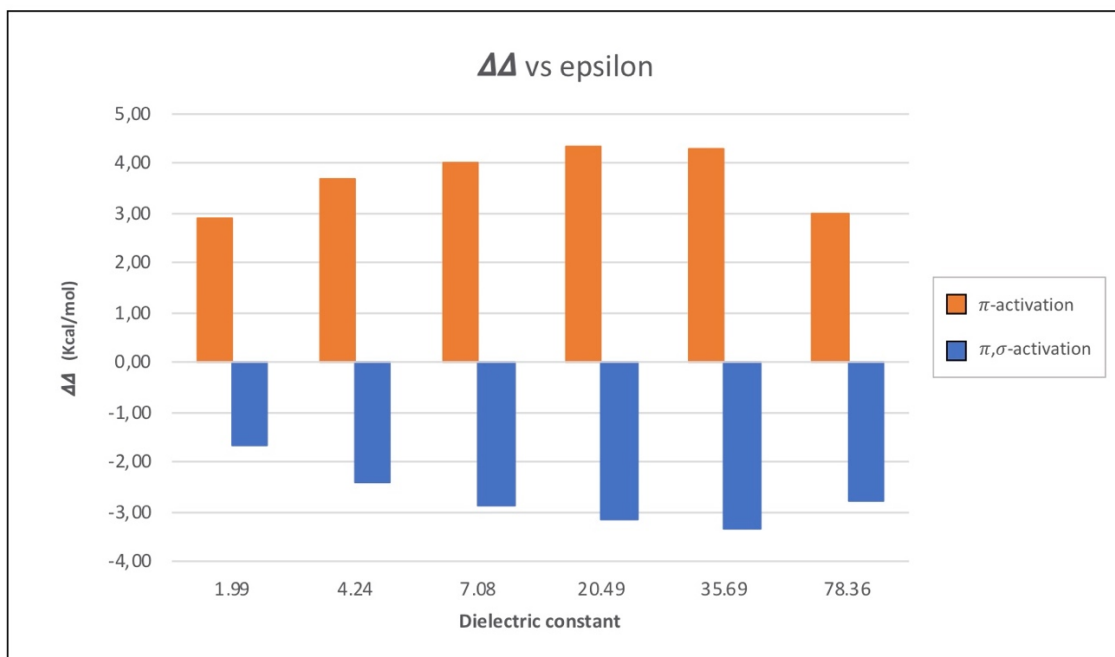
Figure 49 Difference between the barriers ($\Delta\Delta G^{\ddagger}$) for the 6-exo-dig and the 5-endo-dig additions for both π and σ,π-activation modes in water (ε = 78.39) and in solvents of different dielectric constant, computed using the transition states depicted in Figure 47.
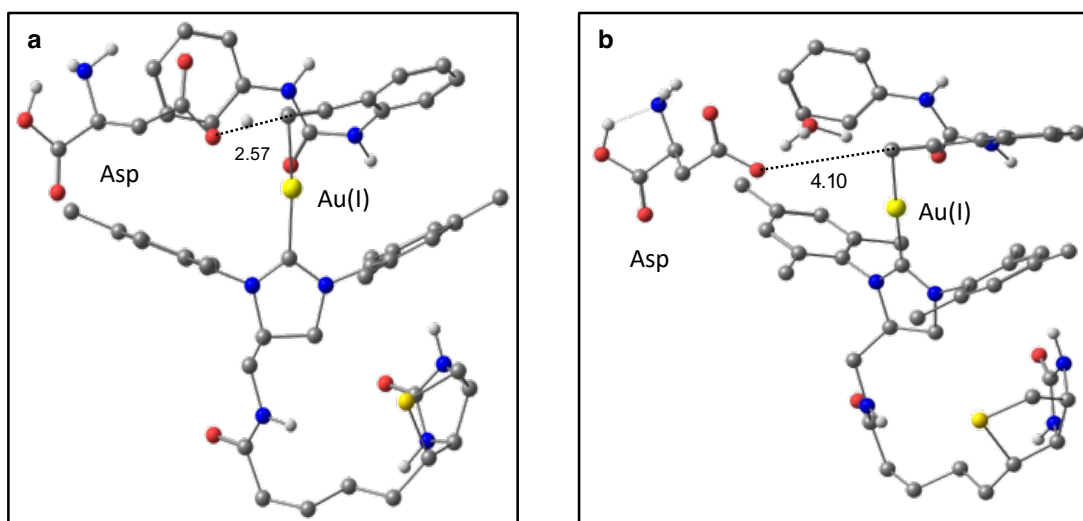
*Figure 50 **a** DFT B3LYP-D3 computed transition states in water for the alkyne's C–H deprotonation, either by an aspartate residue a of through a water molecule **b**. Atomic coloring code: yellow, gold; blue, nitrogen; red, oxygen; grey, carbon and white, hydrogen.*
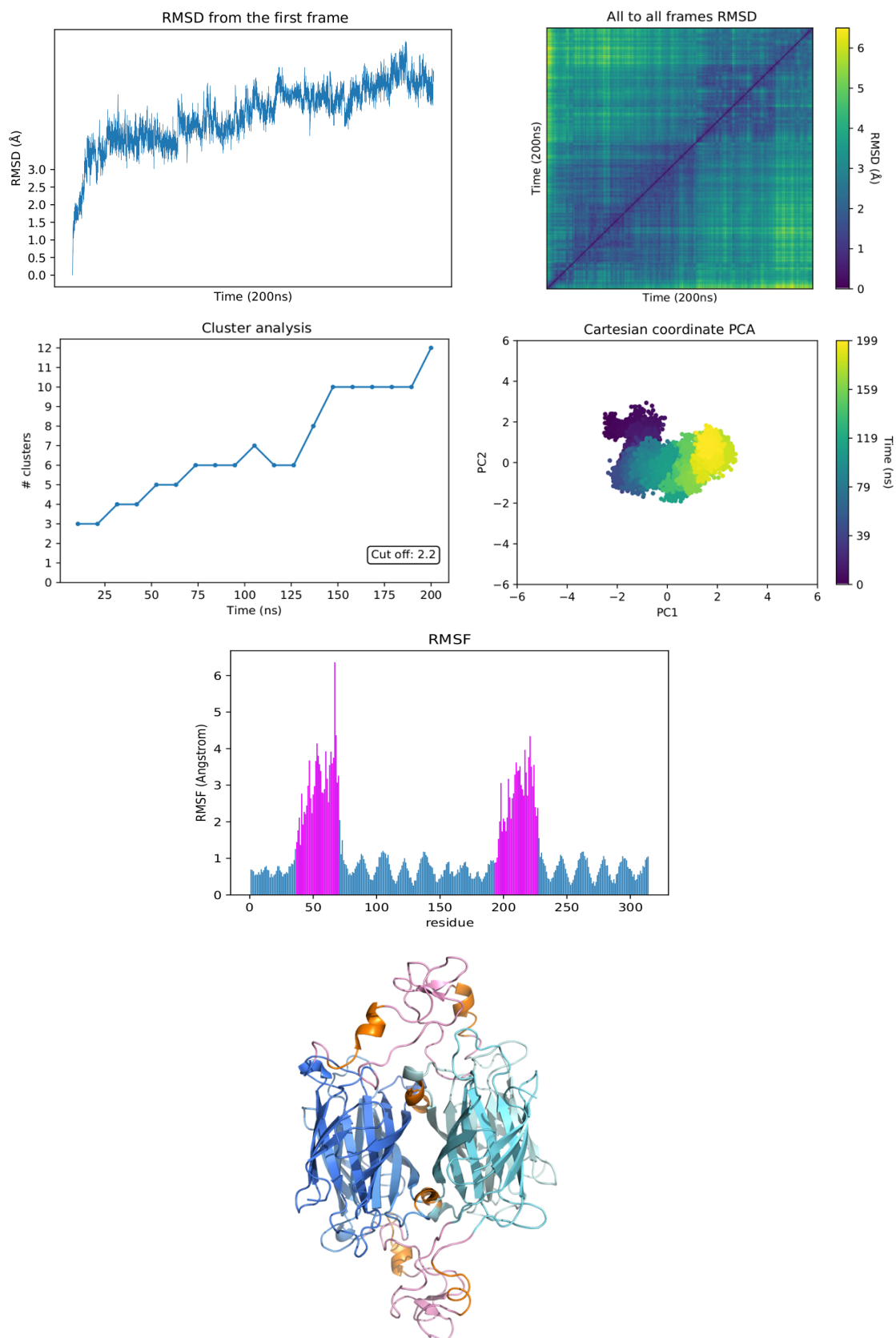
Figure 51 MD convergence analysis corresponding only to the SOD regions of initial MD of SAV-SOD (200 ns): RMSD, all-to-all RMSD, PCA, cluster counting and RMSF (pink corresponds to SOD regions). The representative of the most populated cluster is represented with SOD region in pink and its loop 1-8 in green.
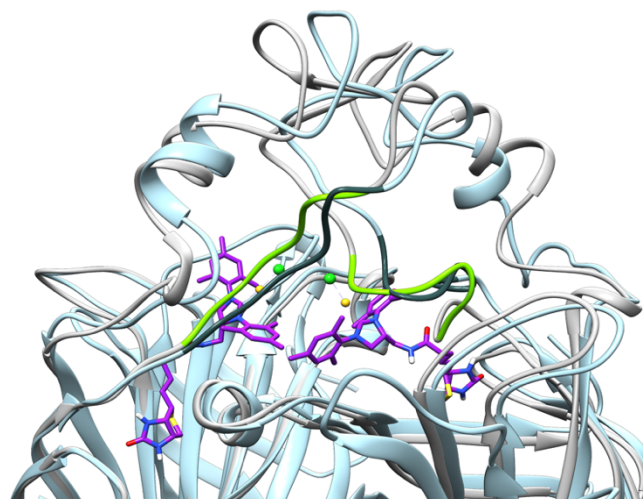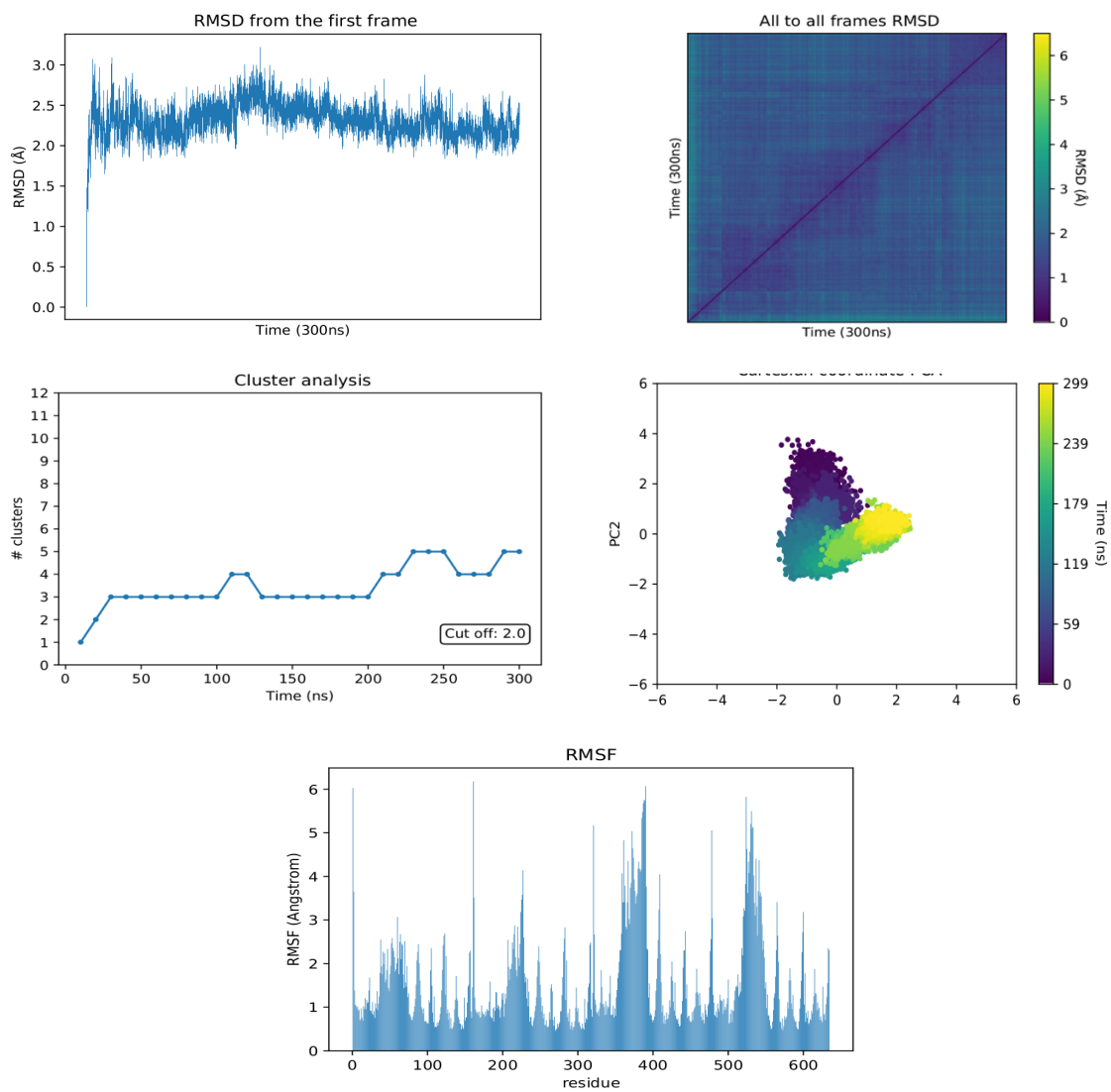
*Figure 52 MD convergence analysis of MD of SAV-SOD with cofactors **biot-Au 2** (300 ns): RMSD, all-to-all RMSD, PCA, cluster counting and RMSF. Superposition of most representative cluster of this MD (MD of SAV-SOD with cofactors - light blue) and without them (initial MD of SAV-SOD - gray) with SOD relevant loops SOD1-8 highlighted in green.*
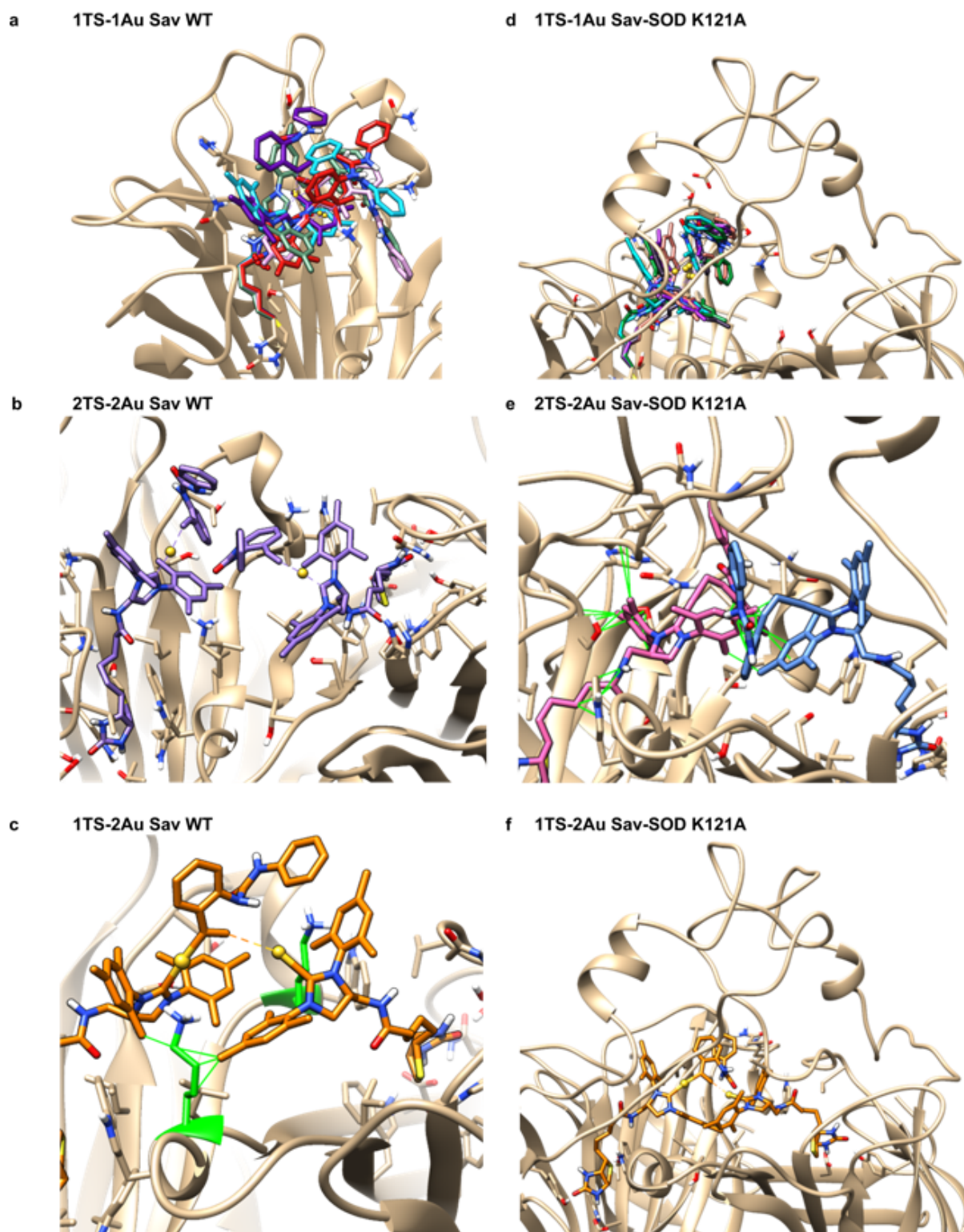
*Figure 53 Docking Study of **biot-Au 2** into Sav WT (**a-c**) and chimeric Sav (**d-f**). Clashes of the transition state with residual amino acids are highlighted with green lines. While docking **1TS-1Au** in Sav WT allows for multiple poses, Sav-SOD clearly favors one single cofactor-substrate orientation.*
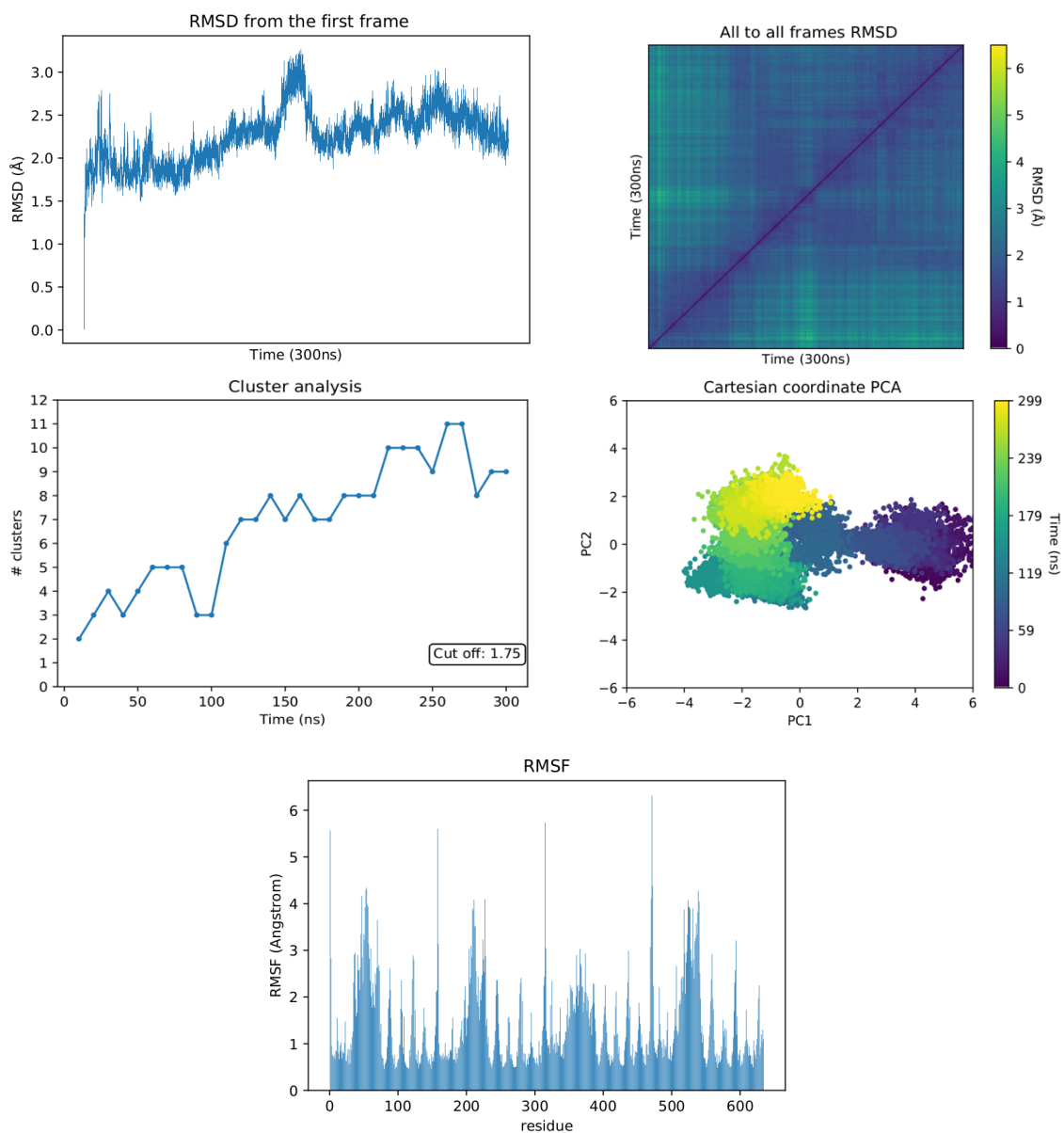
*Figure 54 MD convergence analysis of MD of Sav-SOD with **1TS-1Au** (5-endo-dig) (300 ns): RMSD, all-to-all RMSD, PCA, cluster counting and RMSF.*
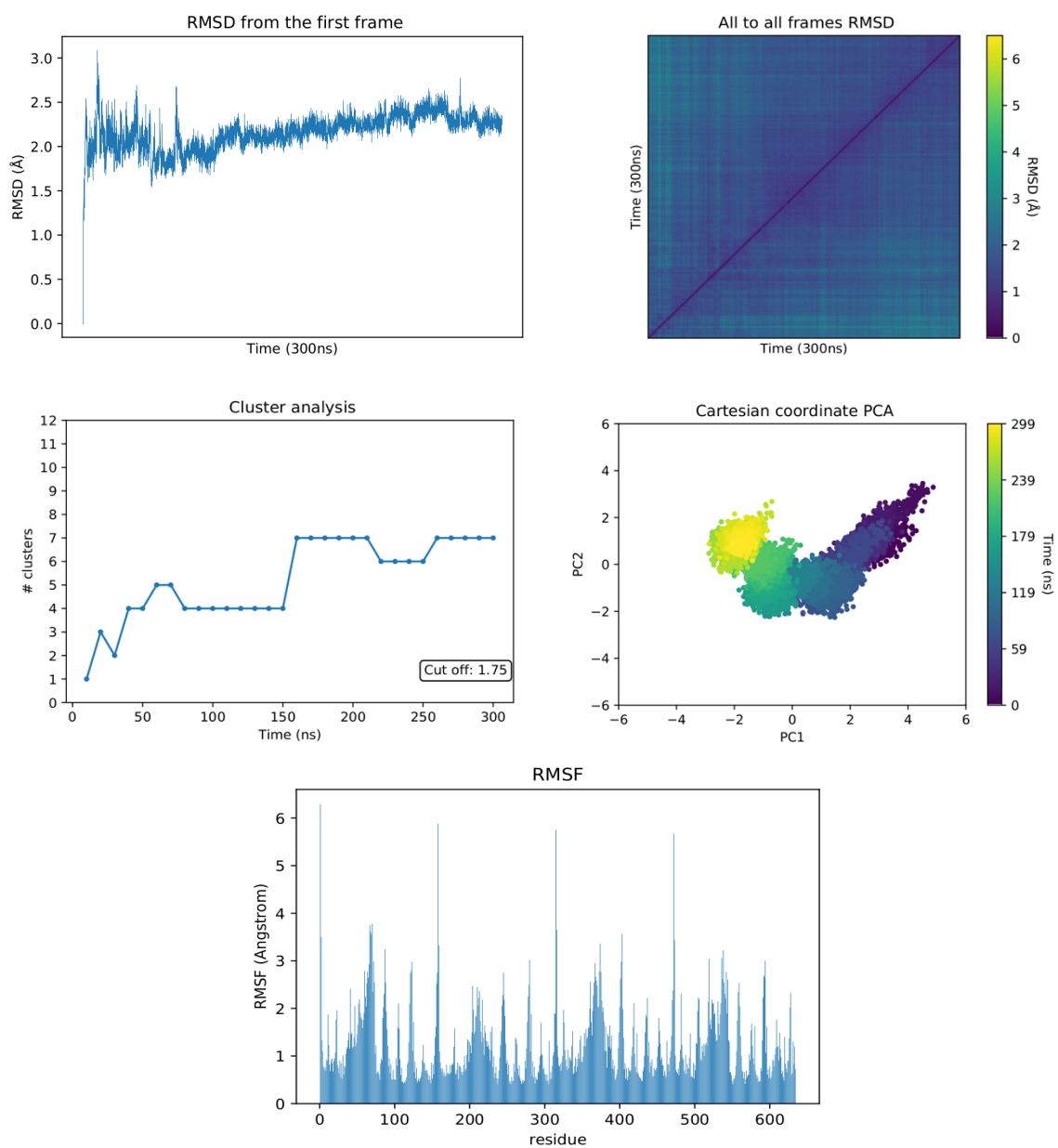
*Figure 55 MD convergence analysis of MD of Sav-SOD with **1TS-1Au** (6-exo-dig) (300 ns): RMSD, all-to-all RMSD, PCA, cluster counting and RMSF.*
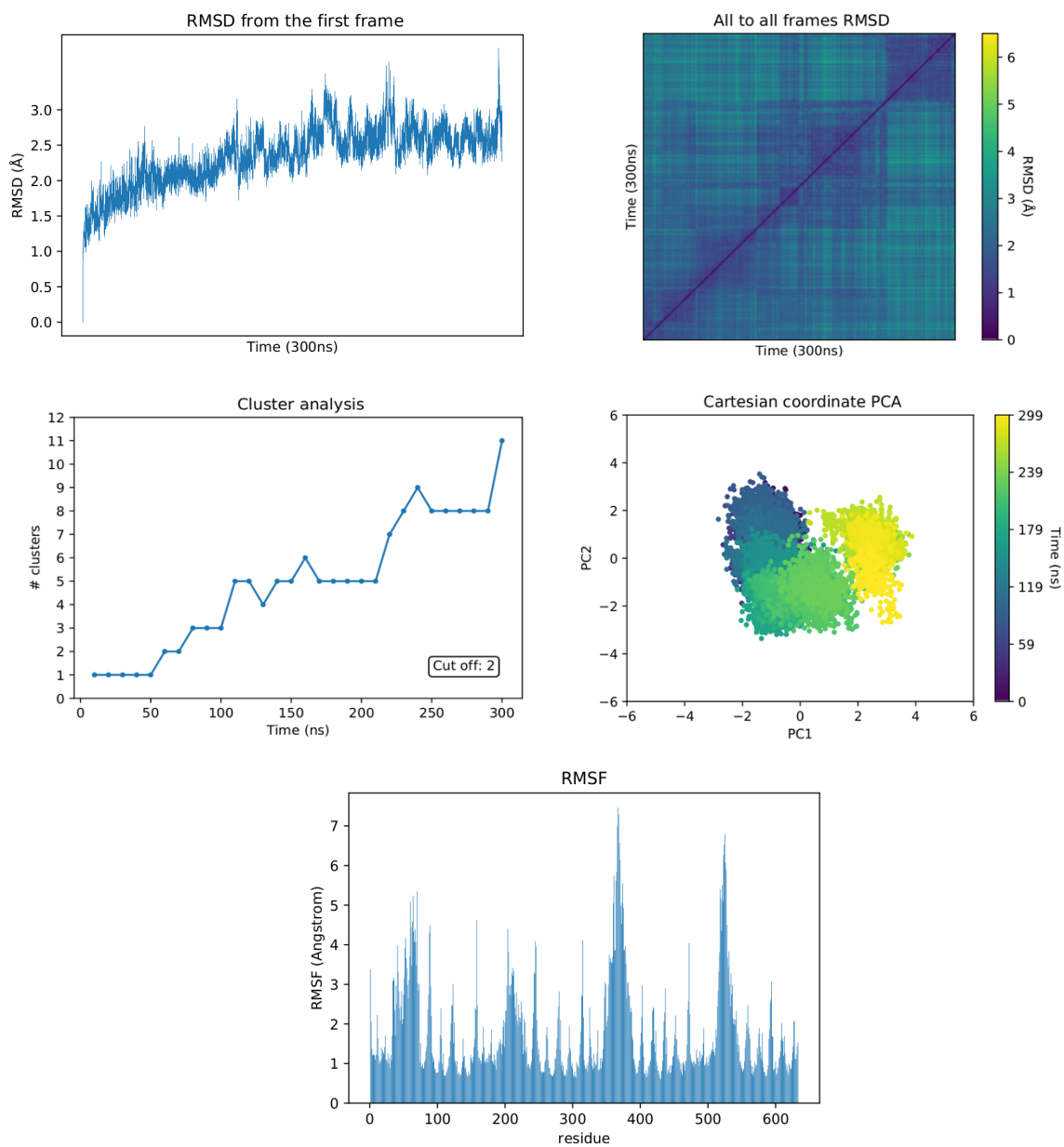
*Figure 56 MD convergence analysis of MD of Sav-SOD with **1TS-2Au** (5-endo-dig) (300 ns): RMSD, all-to-all RMSD, PCA, cluster counting and RMSF.*
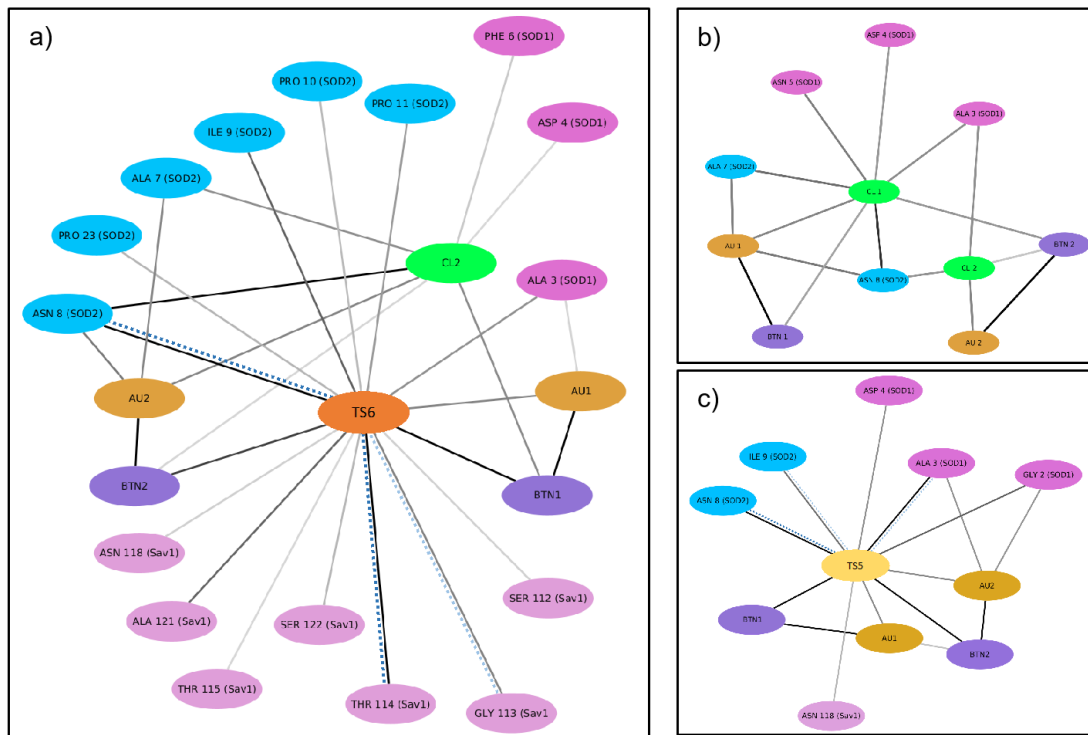
Figure 57 *Residues contribution network extracted from MD simulations of Sav-SOD K121A with:* ***a) 1TS-1Au*** *(6-exo-dig),* ***b)*** *cofactors (**biot-Au 2**) and* ***c) 1TS-2Au*** *(5-endo-dig). Network nodes represent protein residues, biotin (BTN), TS5/6 or atoms (Cl, Au) and edges represent VdW interactions (black-grey straight lines) or hydrogen bonds (blue dashed lines). Edges are colour-weighted by the number of interactions in each frame.*
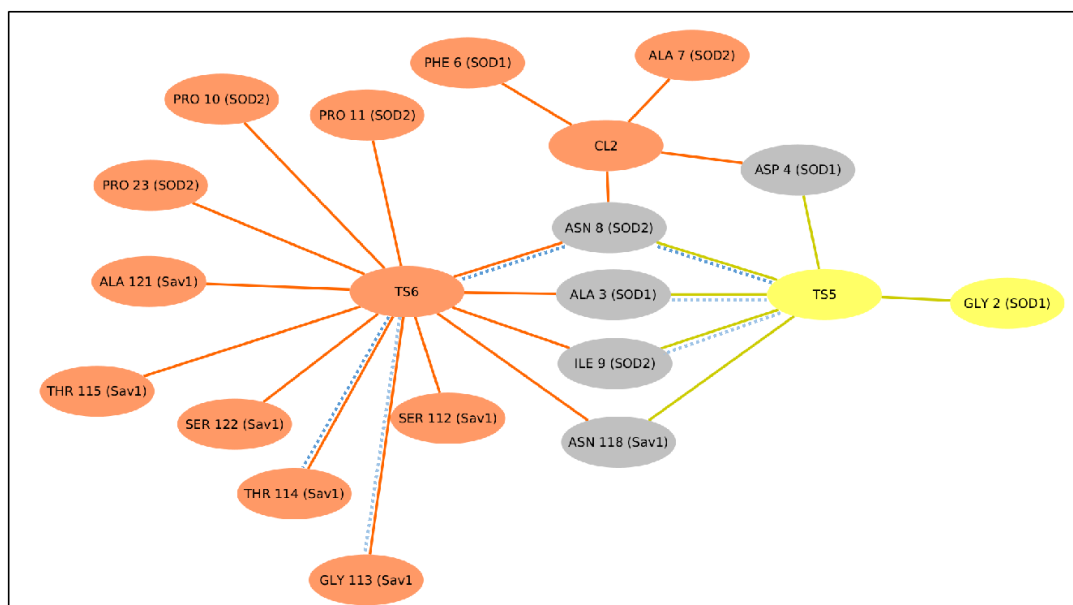
Figure 58 *Residues contribution network extracted from MD simulations of Sav-SOD K121A comparing **1TS-2Au** (5-endo-dig) and **1TS-1Au** (6-exo-dig) interactions. Common interacting residues coloured in grey, **1TS-1Au** interactions in orange and **1TS-2Au** interactions in yellow. Edges represent VdW interactions (black-grey straight lines) or hydrogen bonds (blue dashed lines).*
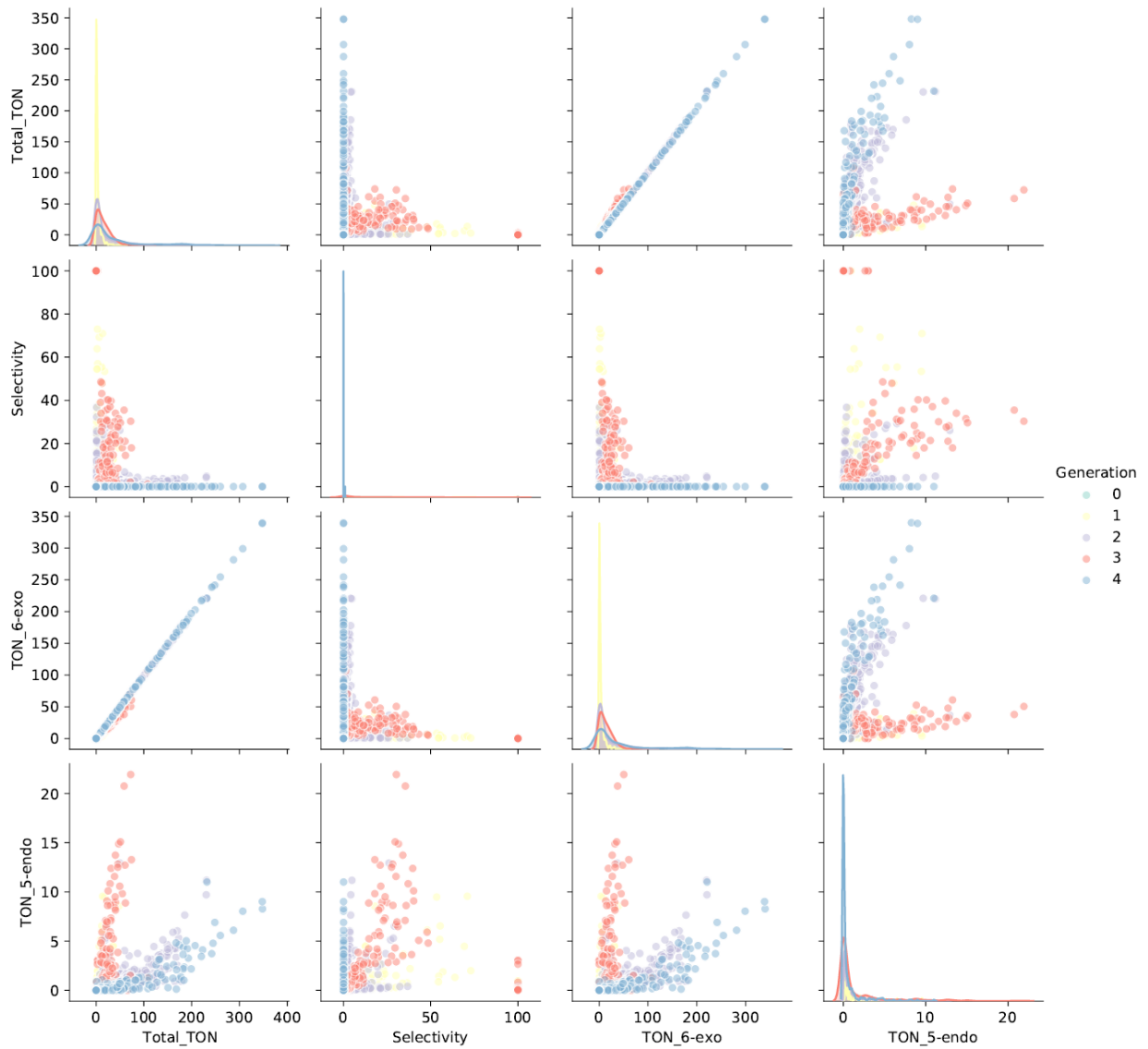
Figure 59 Scatterplot of 2500 cell-free-extract experiments covering 4 Generations of chimeric streptavidin analysed via UPLC-MS after preselection via B4F-assay displaying the two evolution trajectories of HAMase.
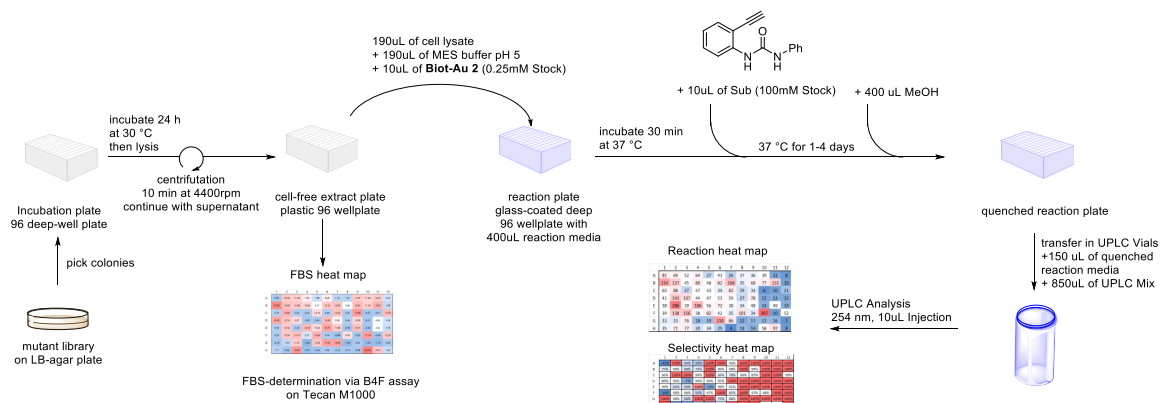
Figure 60 Work-flow for the directed evolution of Sav-SOD K121A using cell-free extracts for screening.
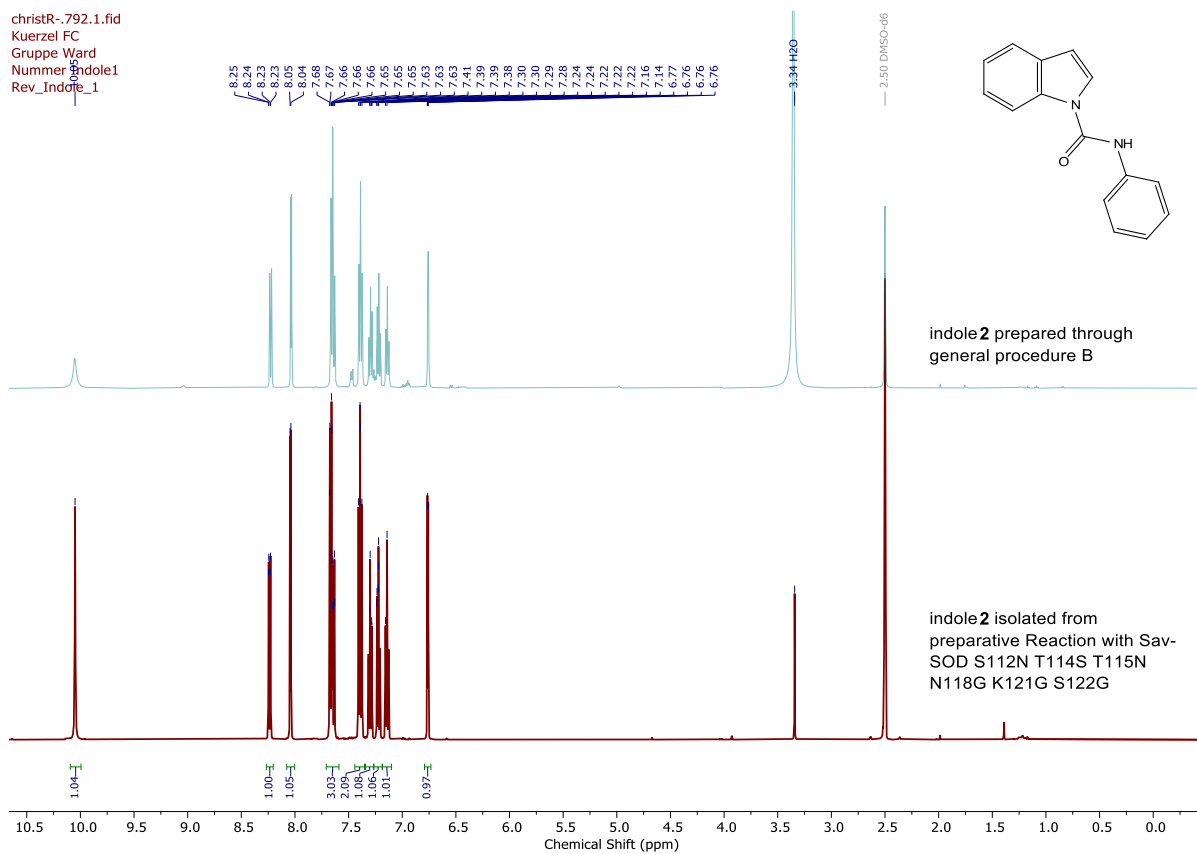
*Figure 61 Stacked ¹H NMR spectra of isolated and the reference indole **2** in DMSO-d₆.*
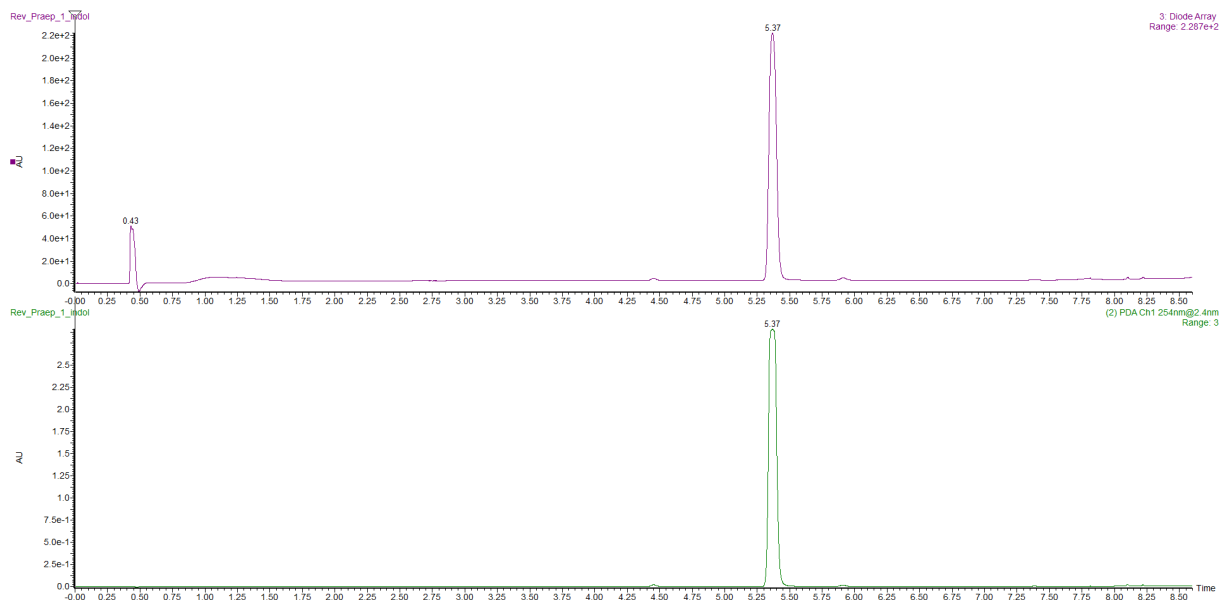
*Figure 62 UPLC-MS DAD (100-500 nm) and PDA (254 nm) Chromatogram of isolated indole **2** (5.37 min).*
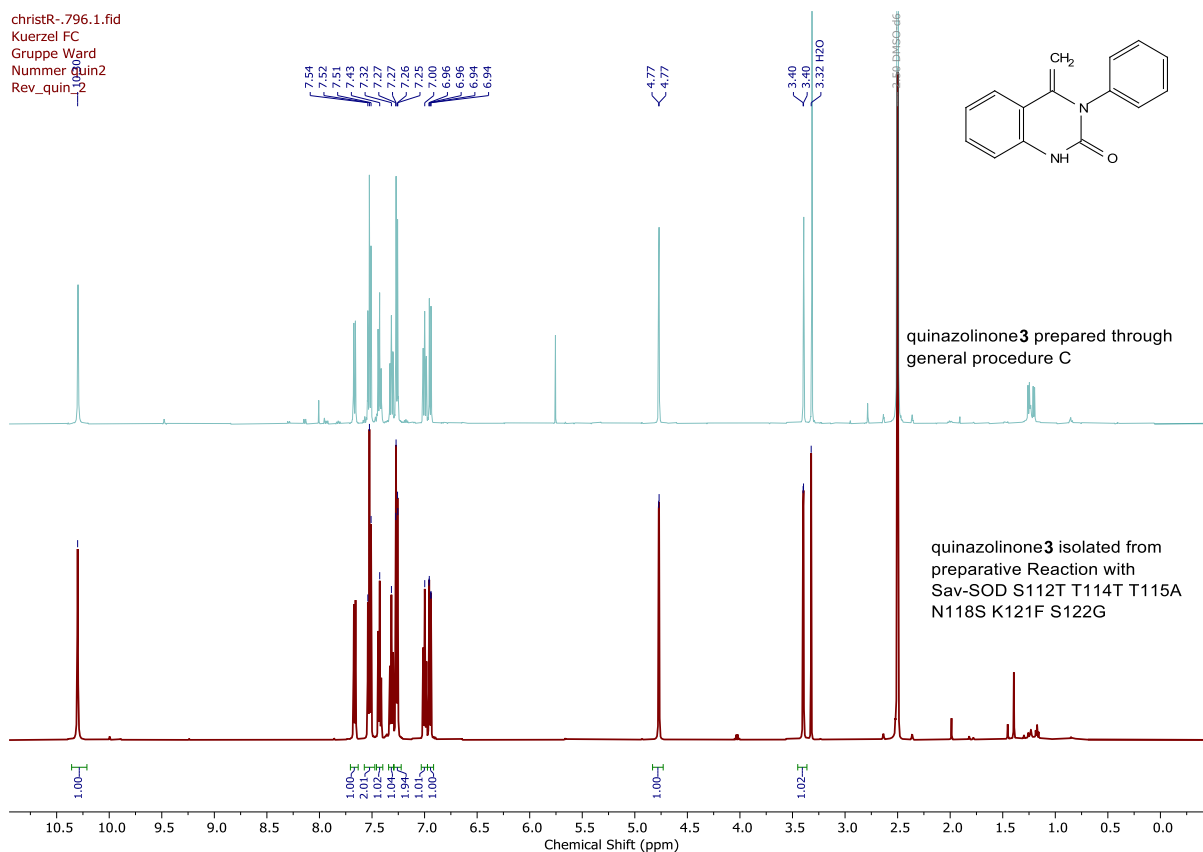
*Figure 63 Stacked ^1H NMR spectra of isolated and the reference quinazolinone **3** in DMSO-d$_6$.*
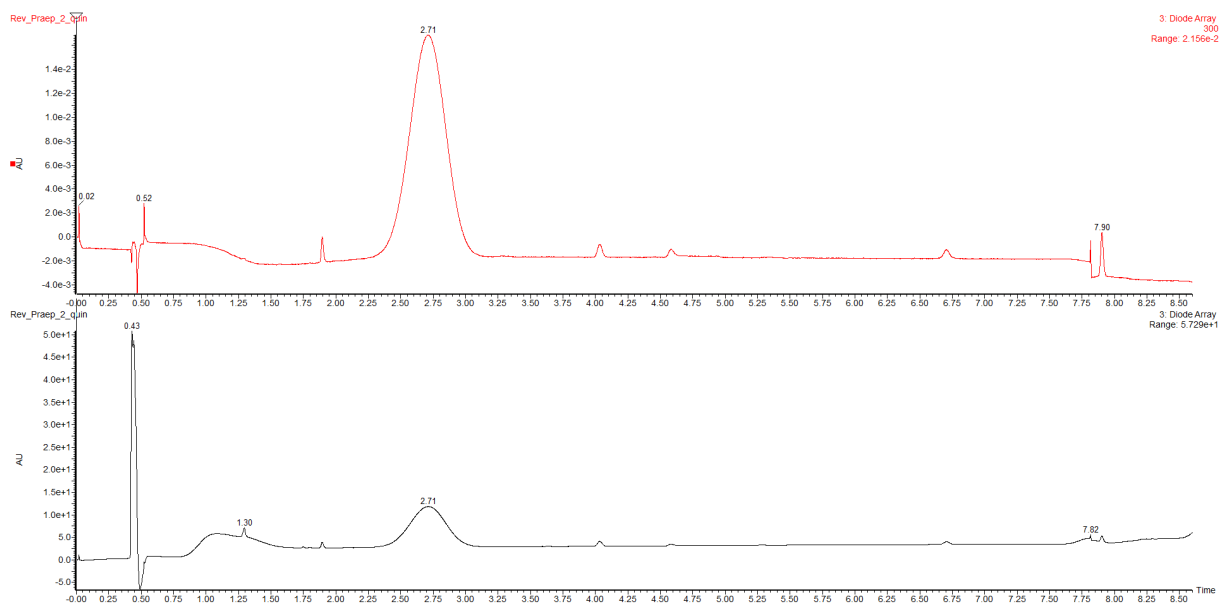
*Figure 64 UPLC-MS DAD (100-500 nm) and DAD (300 nm) Chromatogram of isolated quinazolinone **3** (2.71 min)*.
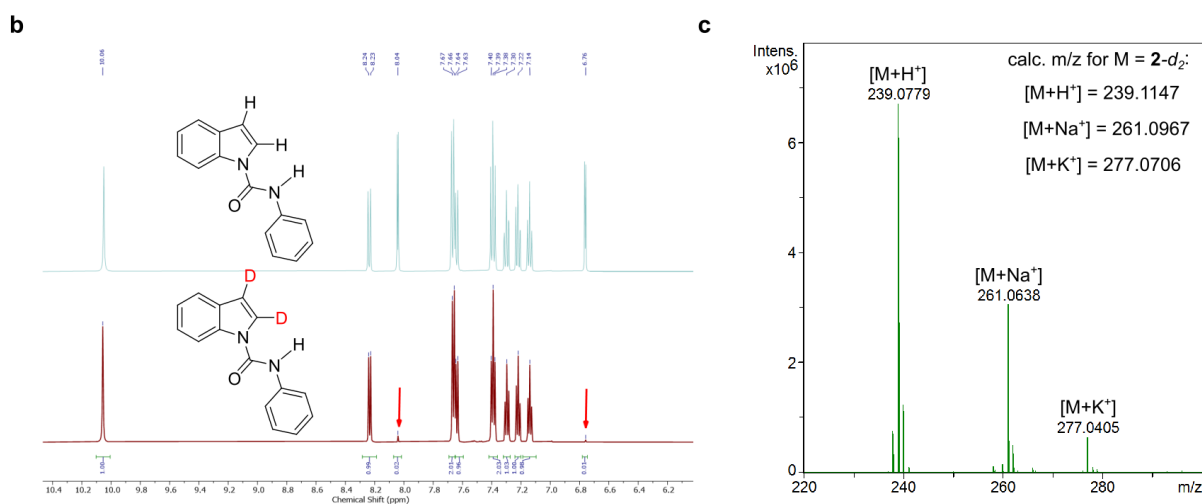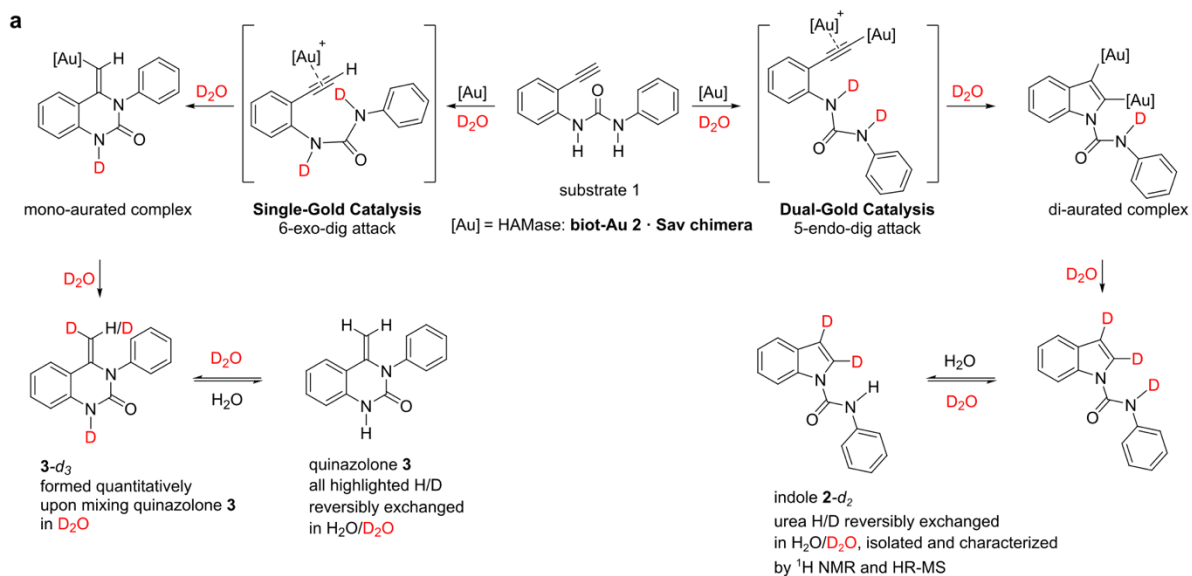
Figure 65 **Isotopic product-distribution of the HAMase biot-Au 2 · Sav-SOD N-SN-GGG carried out in D₂O. a** *Deuteration profile of quinazolinone* **3** *and indole* **2** *resulting from the gold-catalysed hydroamination. For the quinazolinone* **3**, *all highlighted protons/deuterons are readily exchanged in D₂O/H₂O;* **b** *¹H NMR analysis of the indole* **2**-d₂ *isolated from preparative reactions in H₂O (top) and D₂O (bottom, the urea N–H is readily exchanged in the presence of traces of H₂O) and* **c** *HR-MS analysis of* **2**-d₂ *resulting from the reaction performed in D₂O.*

91
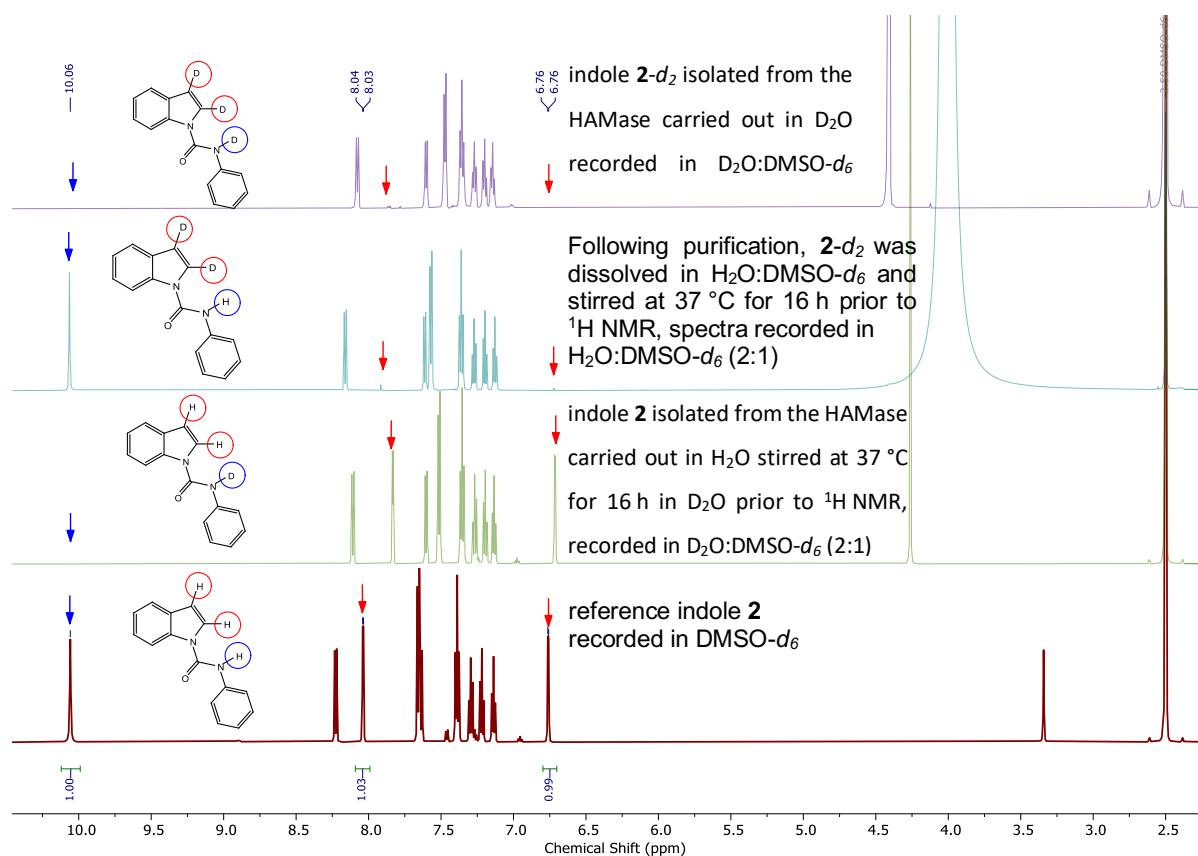
Figure 66 Stacked ¹H NMR spectrum confirming the inertness of indole 2 and 2-d₂ towards deuterium-proton exchange at the indole's C2 and C3 positions (highlighted in red). In contrast, the urea proton (blue) is readily exchanged.

Table 18 Tabulation of data displayed in Figure 39.

	Cofactor only		Sav WT		Sav K121L		Sav K121A		Sav-SOD K121A	
	TON quinaz-olinone 3	TON indole 2	TON quinaz-olinone 3	TON indole 2	TON quinaz-olinone 3	TON indole 2	TON quinaz-olinone 3	TON indole 2	TON quinaz-olinone 3	TON indole 2
biot-Au 1	0	0	4.6	0	19.5	0	18.6	3.0	26.8	4.7
biot-Au 2	0	0	11.6	0	29.5	0.2	14.7	2.9	22.1	14.3
biot-Au 3	0	0	2.7	0	20.1	0	20.5	0.4	18.5	0
biot-Au 4	0	0	0	0	0	0	1.7	0	12.8	0
biot-Au 5	0	0	11.4	0	7.4	0	9.5	0	16.0	0
Me₂biot-Au 2	0	0	0	0	0	0	0	0	1.55	0.7

*Table 19 Summary of the HAMAse activity screening using Sav mutants: [**biot-Au 2**] 1 mol%, [Sav] 2 mol%, MES-buffer pH 5, 37 °C and 5 days reaction time.*

Sav Mutants	TON quinazolinone 3	TON indole 2
K121L	72.4	0.8
K121M	67.0	0.8
K121F	62.5	0
K121I	60.8	0.8
K121T	60.6	0.9

92

K121V	60.1	0.8
K121E	60.1	0
K121W	56.4	0.4
K121D	52.3	0
K121N	52.1	0
K121S	50.1	0.4
K121C	47.0	0.2
K121P	44.3	0.9
K121Q	39.2	0
K121A	33.1	0.9
K121H	27.3	0
K121G	25.5	1.9
S112E	17.7	0
S112T	12.3	0.4
S112G	10.6	0
S112D	10.5	0.5
S112F	9.6	0
S112Y	9.5	0.5
S112C	8.0	0
K121R	7.6	0
S112H	7.2	0.2
S112N	7.2	0.6
WT	6.9	0
S112M	6.7	0.3
S112P	6.0	0.4
S112L	4.2	0.2
S112A	4.2	0
biot-Au 2 only	2.8	0.6
S112K	2.0	0.1
S112R	1.3	0.1
WT without **biot-Au 2**	0.1	0

Table 20 Evaluation of the reactivity and stability of various gold-complexes and gold-based ArM.

Au-Complex	Mutant	Additives	TON quinazolinone **3**	TON indole **2**
(CH$_3$)$_2$SAuCl	-	none	0	0.2
AuSIMes	-	none	1.3	0.6
AuIMes	-	none	0.3	0.2
AuSIPr	-	none	0.2	0.0
AuIPr	-	none	0.3	0.0
biot-Au 2	-	none	1.3	0.2
AuSIMes	Sav K121L	none	0	0.3
biot-Au 2	Sav K121L	none	47.5	2.1
biot-Au 2	Sav-SOD K121A	none	33.1	20.5
biot-Au 2	Sav-SOD K121A	10% cell lysate	14.7	5.1
biot-Au 2	Sav-SOD K121A	50% cell lysate	12.9	2.3
biot-Au 2	Sav-SOD K121A	2 mol% GSH red[c]	28.7	9.5
biot-Au 2	Sav-SOD K121A	20 mol% GSH red[c]	22.3	6.2
biot-Au 2	Sav-SOD K121A	100 mol% GSH red[c]	7.4	1.2
biot-Au 2	Sav-SOD K121A	100 mol% GSH red[d]	7.4	1.4
biot-Au 2	Sav-SOD K121A	100 mol% GSH ox[d]	33.6	12.8
biot-Au 2	Sav-SOD K121A	1 mol% biotin[d]	19.0	10.2
biot-Au 2	Sav-SOD K121A	2 mol% biotin[d]	12.0	3.8
biot-Au 2	Sav-SOD K121A	2 mol% biotin[c]	15.9	3.7

biot-Au 2	Sav-SOD K121A	4 mol% biotin[d]	6.7	1.8
biot-Au 2	Sav-SOD K121A	100 mol% Diamide[d]	20.2	34.2
biot-Au 2	-	100 mol% Diamide[d]	10.0	3.4

[a]*Reaction conditions: V_{tot} 200 μL (V_{DMSO} 15 μL), [Sub] 5 mM, [**biot-Au 2**] 50 μM, [Sav] 100 μM, MES-buffer pH 5 + cell lysate, 37 °C for 72 h.* [b]*Turnover Number* [c]*Additive added before cofactor incubation* [d]*Additive added after cofactor incubation*

*Table 21 GaudiMM scoring values using scoring function Ligscore for TS5 (5-endo-dig mechanism) and TS6 (6-exo-dig mechanism) models. Three different systems were calculated, lower scores highlight a better complementarity: 1) 1TS-1Au: Docking of π-activated TS (one **biot-Au 2**) into one monomer of Sav using Ligscore and clashes in GaudiMM. 2) 2TS-2Au: Docking of two π-activated TS (two **biot-Au 2**) into two adjacent monomers of Sav. Simultaneous docking calculations with GaudiMM using Ligscore and clashes. 3) 1TS-2Au: Docking of σ,π-activated TS (two **biot-Au 2**) into two adjacent monomers. Simultaneous docking calculations with GaudiMM and distance restriction for Au atoms (3 Å). To compare scoring values for systems with one and two Au centres, the energy of interaction by gold complex are presented (bold values).*

	Sav WT		Sav K121A		Sav-SOD K121A	
	TS5	**TS6**	**TS5**	**TS6**	**TS5**	**TS6**
1TS-1Au	**-71.83**	**-72.22**	**-62.74**	**-61.17**	**-65.5**	**-62.54**
2TS-2Au	-118.04 **-59.02**	-119.78 **-59.89**	-94.99 **-47.50**	-80.55 **-40.28**	-31.44 **-15.72**	-19.99 **-10.00**
1TS-2Au	-83.25 **-41.63**	-82.01 **-41.00**	-109.94 **-54.97**	-81.68 **-40.84**	-98.99 **-49.50**	-60.70 **-30.35**

*Table 22 Decomposition of total energy (kcal/mol) for most relevant residues of SOD chains using MMGBSA and based on cytoscape analysis for encapsulated pseudo-transition states TS5 (5-endo-dig) **1TS-2Au** (σ,π activation mode) and TS6 (6-exo-dig) **1TS-1Au** (π–activation mode) as well as the cofactor only (last column).*

	1TS-2Au (5-endo-dig)	**1TS-1Au (6-exo-dig)**	**Cofactor (biot-Au 2)**
Gly 2 (SOD1)	-1.887	-0.083	-0.027
Ala 3 (SOD1)	-4.053	-1.671	-1.624
Asp 4 (SOD1)	-0.026	0.237	-0.018
Ala 7 (SOD2)	-0.102	-2.108	-1.644
Asn 8 (SOD2)	-2.19	-4.930	0.224
Ile 9 (SOD2)	-1.045	-0.577	-2.248
Pro 10 (SOD2)	-0.023	-0.099	-0.04
Pro 11 (SOD2)	-0.014	-0.384	-0.087
Pro 23 (SOD2)	-0.015	-0.369	-0.002

*Table 23 Decomposition of total energy (kcal/mol) for most relevant amino acid residues of Sav chains using MMGBSA and based on cytoscape analysis for embedded pseudo-transition states TS5 (5-endo-dig) **1TS-2Au** (σ,π activation mode) and TS6 (6-exo-dig) **1TS-1Au** (π activation mode) as well as the **biot-Au 2** cofactor only (last column).*

	1TS-2Au (5-endo-dig)	1TS-1Au (6-exo-dig)	Cofactor (biot-Au 2)
Ser 112 (Sav1)	-0.615	-1.001	-1.415
Gly 113 (Sav1)	0.034	-0.011	-0.158
Thr 114 (Sav1)	-0.329	-3.019	-0.115
Thr 115 (Sav1)	-0.004	-0.369	-0.001
Glu 116 (Sav1)	0.007	-0.020	0.051
Ala 117 (Sav1)	0.048	0.008	0.047
Asn 118 (Sav1)	-0.186	-0.380	-0.028
Ala 119 (Sav1)	-0.121	-0.162	-0.101
Trp120 (Sav1)	-2.482	-3.003	-3.079
Ala 121 (Sav1)	-1.019	-0.937	-0.673
Asn122 (Sav1)	-0.296	-0.383	-0.172

Table 24 Investigation of E. coli strains for cell-free extract catalysis with various ratios of MES buffer (pH 5, 50 mM) to cell-free extract (cfe) before directed evolution.

E.Coli Strain	Mutant	Amount of cfe (%)	TON quinazo-linone **3**	TON indole **2**
BL21(DE3)	emptyV	100	0	0
BL21(DE3)	Sav_WT	100	0	0
BL21(DE3)	Sav_K121L	100	0	0
BL21(DE3)	emptyV	50	0	0
BL21(DE3)	Sav_WT	50	1	0
BL21(DE3)	Sav_K121L	50	11	0
BL21(DE3)	emptyV	30	0	0
BL21(DE3)	Sav_WT	30	2	0
BL21(DE3)	Sav_K121L	30	15	0
SHuffle-T7	emptyV	50	0	0
SHuffle-T7	Sav_WT	50	0	0
SHuffle-T7	Sav_K121L	50	0	0
SHuffle-T7	emptyV	30	0	0
SHuffle-T7	Sav_WT	30	0	0
SHuffle-T7	Sav_K121L	30	0	0

Table 25 Selected Results of evolved ArMs with and without SOD loop. [a] [b]

Mutant	TON quinaz-olinone 3	TON indole 2	TON Total	Selectivity for indole 2
Sav K121A	18.0	3.5	21.5	16.4%
Sav K121A	17.0	3.9	20.9	18.7%
Sav K121A	18.7	4.0	22.7	17.6%
Sav K121A	17.6	3.3	20.9	15.8%
Sav-SOD K121A	25.5	16.2	41.7	38.8%
Sav-SOD K121A	24.1	16.1	40.2	40.1%
Sav-SOD K121A	24.7	12.4	37.1	33.4%
Sav-SOD K121A	23.6	15.5	39.1	39.6%
Sav-SOD N118G K121G S122G	27.5	33.9	61.4	55.2%
Sav-SOD N118G K121G S122G	25.5	30.2	55.7	54.2%
Sav-SOD N118G K121G S122G	27.0	34.7	61.7	56.3%
Sav-SOD N118G K121G S122G	24.0	28.0	51.9	53.9%
Sav-SOD N118S K121F S122G	38.6	0.4	39.0	1.0%
Sav-SOD N118S K121F S122G	37.9	0.5	38.3	1.2%
Sav-SOD N118S K121F S122G	37.3	0.4	37.7	1.0%
Sav-SOD N118S K121F S122G	39.6	0.5	40.1	1.1%
Sav_SOD S112N N118G K121G S122G	21.9	44.8	66.7	67.2%
Sav_SOD S112N N118G K121G S122G	20.9	44.7	65.6	68.1%
Sav_SOD S112N N118G K121G S122G	21.1	44.5	65.6	67.8%
Sav_SOD S112N N118G K121G S122G	20.2	41.8	62.0	67.4%
Sav_SOD S112T N118S K121F S122G	40.4	1.1	41.5	2.6%
Sav_SOD S112T N118S K121F S122G	42.5	1.2	43.7	2.8%
Sav_SOD S112T N118S K121F S122G	40.2	1.0	41.3	2.5%
Sav_SOD S112T N118S K121F S122G	40.8	1.1	41.9	2.7%
Sav_SOD S112N T114S T115N N118G K121G S122G	18.1	55.1	73.2	75.2%
Sav_SOD S112N T114S T115N N118G K121G S122G	17.6	46.0	63.6	72.3%
Sav_SOD S112N T114S T115N N118G K121G S122G	18.5	52.3	70.9	73.8%
Sav_SOD S112N T114S T115N N118G K121G S122G	17.6	46.5	64.1	72.5%
Sav_SOD S112T T115A N118S K121F S122G	41.5	2.9	44.4	6.5%
Sav_SOD S112T T115A N118S K121F S122G	41.7	3.1	44.8	6.9%
Sav_SOD S112T T115A N118S K121F S122G	43.0	3.5	46.5	7.6%

Mutant				
Sav_SOD S112T T114S N118S K121F S122G	41.0	3.0	44.0	6.8%

Table 26 Optimized Reaction Screening with top Hits in quadruplets.

Mutant	TON quinaz-olinone 3	TON indole 2	TON Total	Selectivity for indole 2
Sav-SOD S112N T114S T115N N118G K121G S122G – OPTIMIZED 1	1.8	46.5	48.3	96.3%
Sav-SOD S112N T114S T115N N118G K121G S122G – OPTIMIZED 2	1.7	47.2	48.8	96.6%
Sav-SOD S112N T114S T115N N118G K121G S122G – OPTIMIZED 3	2.1	53.9	55.9	96.3%
Sav-SOD S112N T114S T115N N118G K121G S122G – OPTIMIZED 4	1.9	49.8	51.8	96.3%
Sav-SOD S112T T115A N118S K121F S122G – OPTIMIZED 1	265.0	2.4	267.4	0.9%
Sav-SOD S112T T115A N118S K121F S122G – OPTIMIZED 2	308.3	2.7	311.0	0.9%
Sav-SOD S112T T115A N118S K121F S122G – OPTIMIZED 3	349.2	2.2	351.4	1.0%
Sav-SOD S112T T115A N118S K121F S122G – OPTIMIZED 4	397.3	3.8	401.1	0.6%

*Optimized reaction conditions for quinazolinone 3: V_{tot} 400 μL (365 μL MES-buffer pH 5) [Sub] 2.5 mM, [**biot-Au2**] 2.5 μM, [Sav] 5 μM, 39 °C for 72 h and for indole 2: V_{tot} 100 μL (V_{MES} 45 μL, V_{DMSO} 20 μL), [Sub] 15 mM, [Diamide] 15 mM, [**biot-Au 2**] 100 μM, [SAV] 200 μM, 37 °C for 72 h. Both sets of experiments were quenched with MeOH to a total volume of 800 μL before UPLC MS analysis.*

*Table 27 Substrate scope using the evolved **biot-Au 2** · Sav-SOD chimeras under standard reaction conditions[a] used for directed evolution and under reaction conditions optimized for 5-endo-dig[b] and 6-exo-dig[c] products.*

Substrate / Mutant	1a	1b	1c	1d	1e	1f
Sav-SOD S112N T114S T115N	16.7 (36:64)[a]	21.5 (16:84)[a]	2.3 (44:56)[a]	6.3 (32:68)[a]	6.2 (34:66)[a]	0[a]
N118G K121G S122G	45.9 (7:93)[b]	35.9 (19:81)[b]	22.0 (15:85)[b]	15.3 (1:99)[b]	5.2 (4:96)[b]	6.9 (50:50)[b]
Sav-SOD S112T T115A N118S	13.2 (93:7)[a]	10.2 (86:14)[a]	4.0 (97:3)[a]	3.9 (90:10)[a]	7.3 (91:9)[a]	1.5 (91:9)[a]
K121F S122G	104.7 (96:4)[c]	35.3 (97:3)[c]	41.7 (97:3)[c]	26.0 (97:3)[c]	16.7 (85:15)[c,d]	6.0 (96:4)[c,d]

*The combined turnover for both products is displayed as well as the quinazolinone:indole ratio (**3:2**). [a]Standard Reaction conditions for directed evolution see Table 24; [b] Reaction conditions optimized for 5-endo-dig product: V_{tot} 100 µL (V_{MES} 45 µL, V_{DMSO} 20 µL), [Sub] 15 mM, [Diamide] 15 mM, [**biot-Au 2**] 100 µM, [Sav] 200 µM, [MES-buffer] 50 mM, pH 5, 37 °C for 24 h; [c] Reaction conditions optimized for 6-exo-dig product: V_{tot} 400 µL (V_{MES} 365 µL, V_{DMSO} 12 µL) [Sub] 2.5 mM, [**biot-Au 2**] 6.25 µM, [Sav] 12.5 µM, [MES-buffer] 50 mM, pH 5, 39 °C for 48 h; [d] with following changes: [**biot-Au 2**] 2.5 µM, [Sav] 5 µM, for 24 h.*
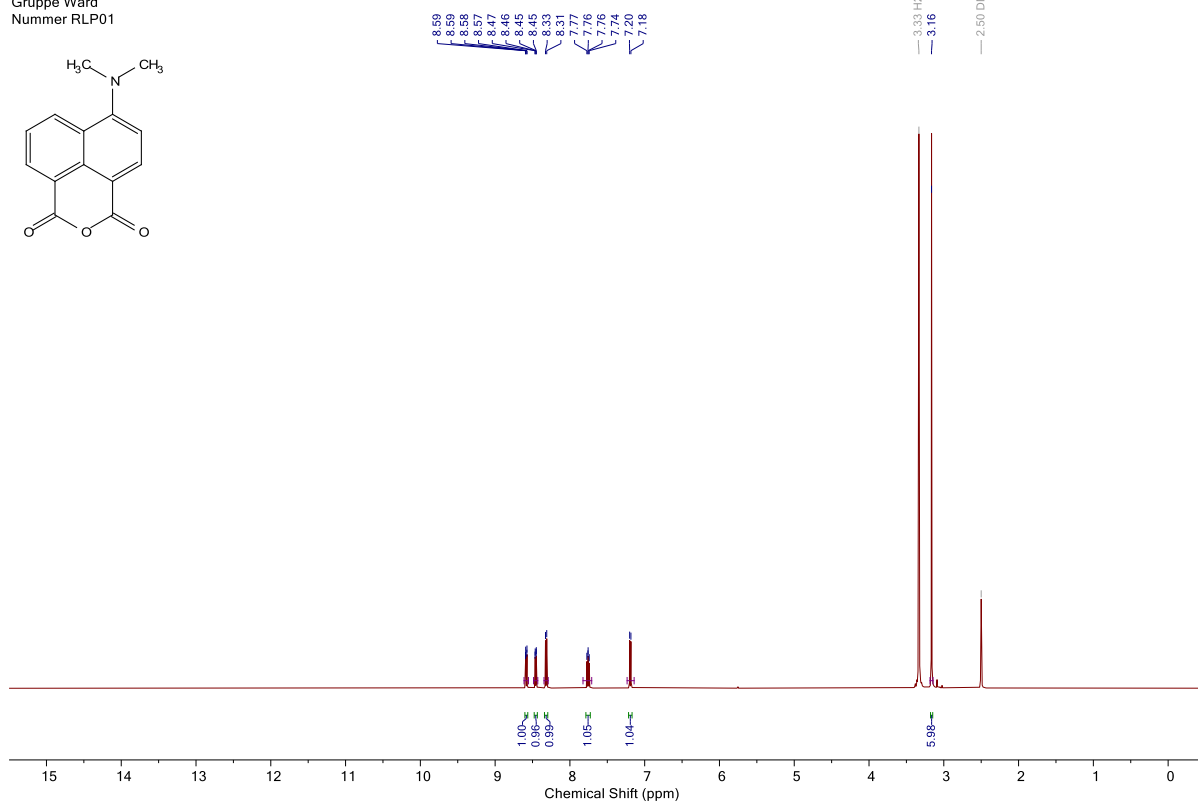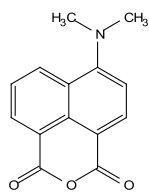
*Figure 67 ¹H NMR Spectrum of **C1** in DMSO-d₆.*
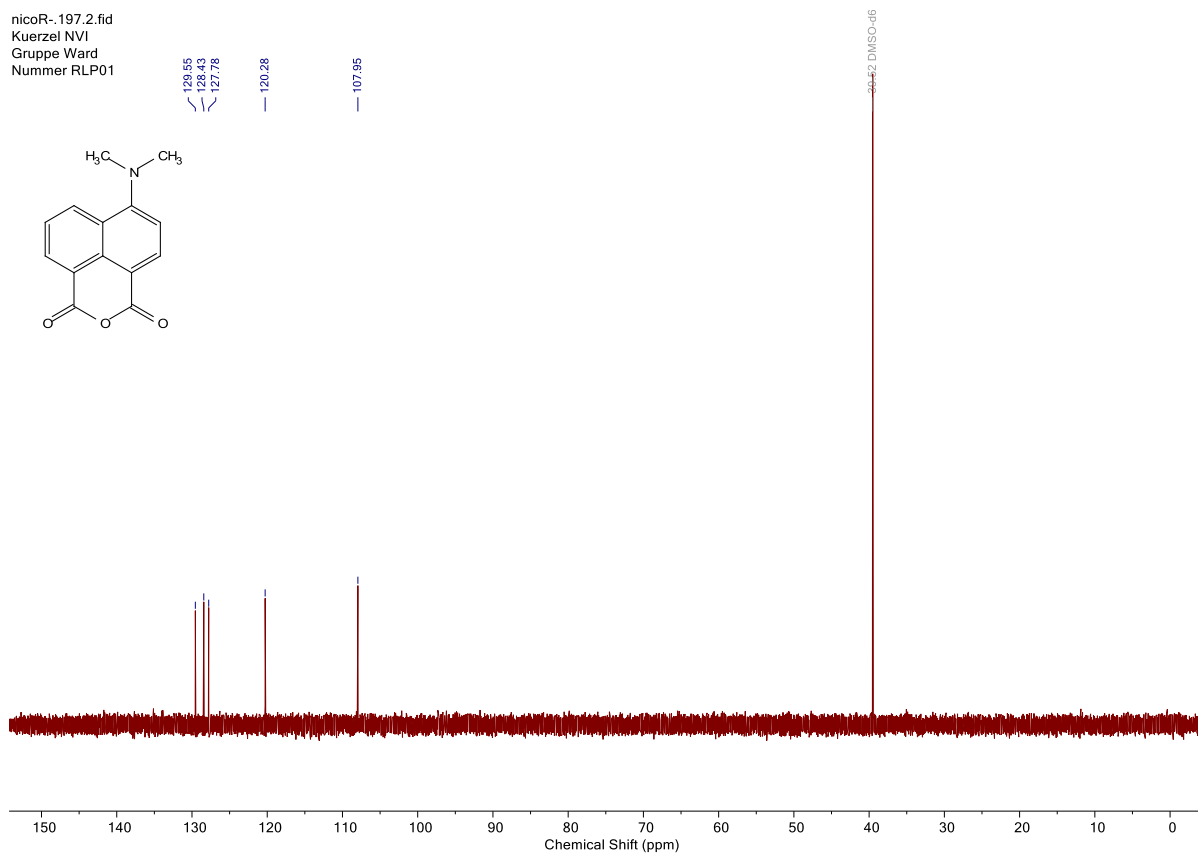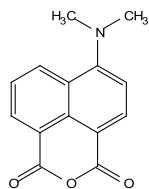
Figure 68 ¹³C NMR Spectrum of C1 in DMSO-d₆.

99
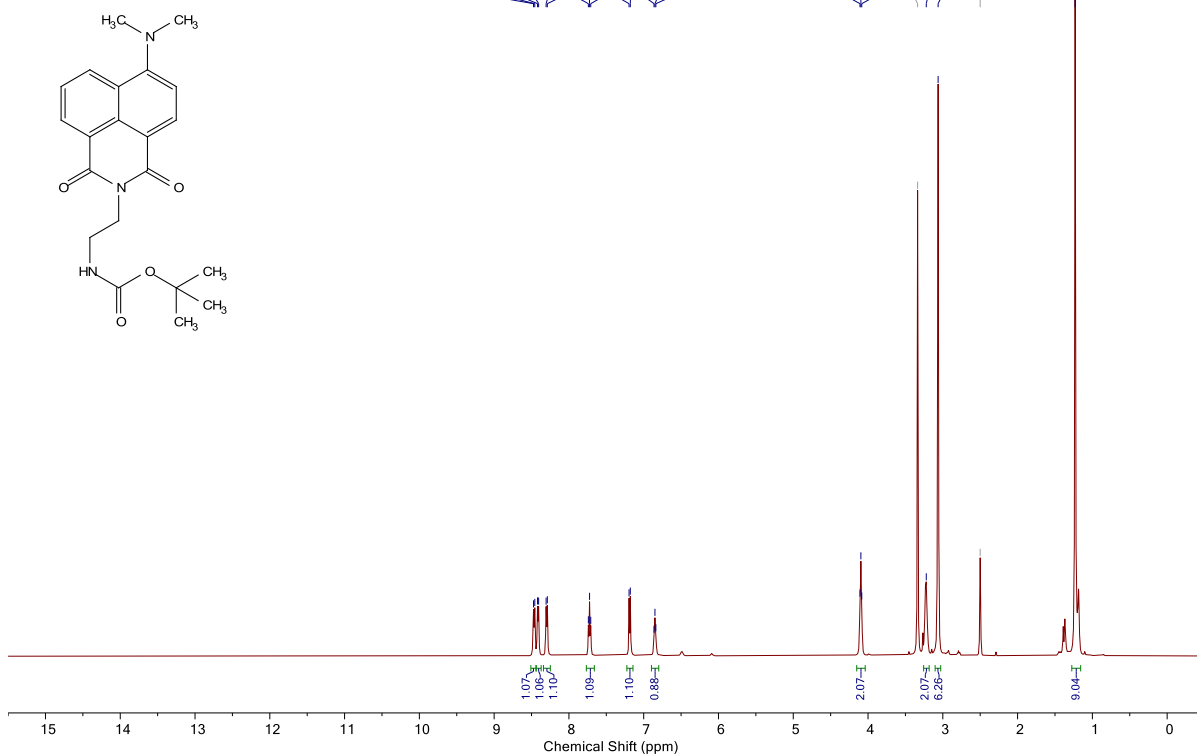
Figure 69 ¹H NMR Spectrum of C2 in DMSO-d₆.
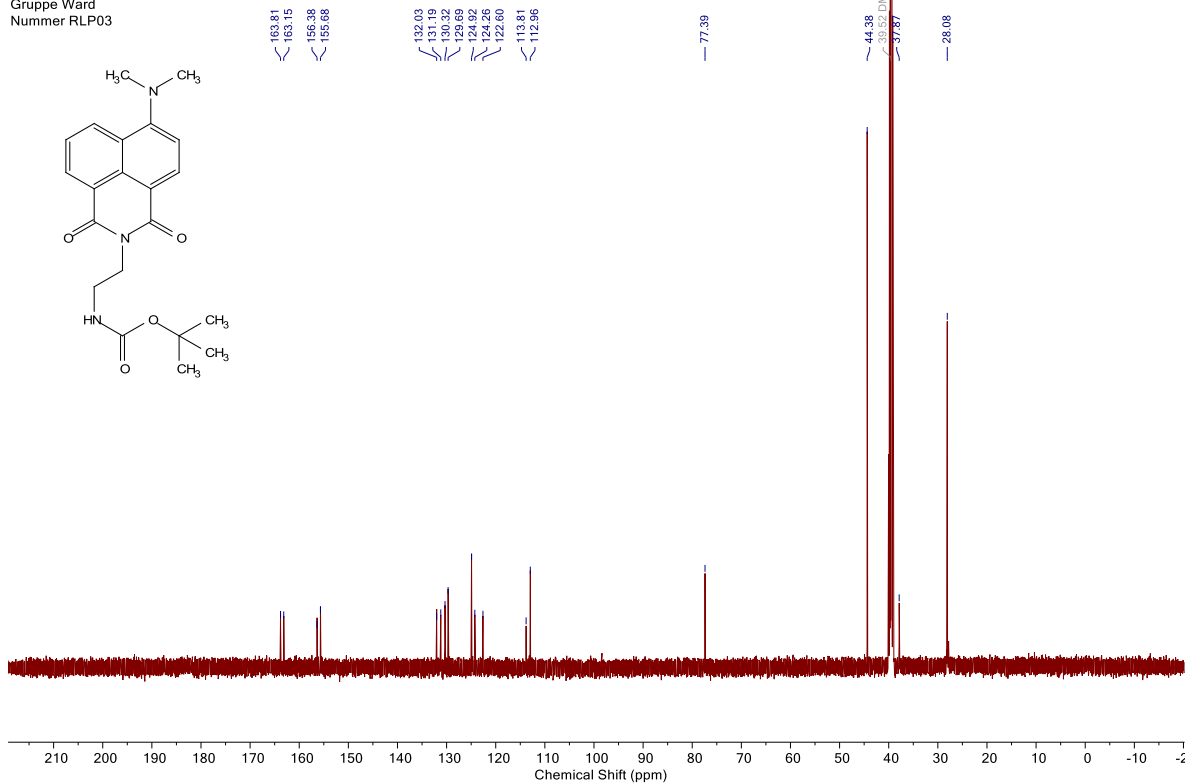
Figure 70 ¹³C NMR Spectrum of C2 in DMSO-d₆.
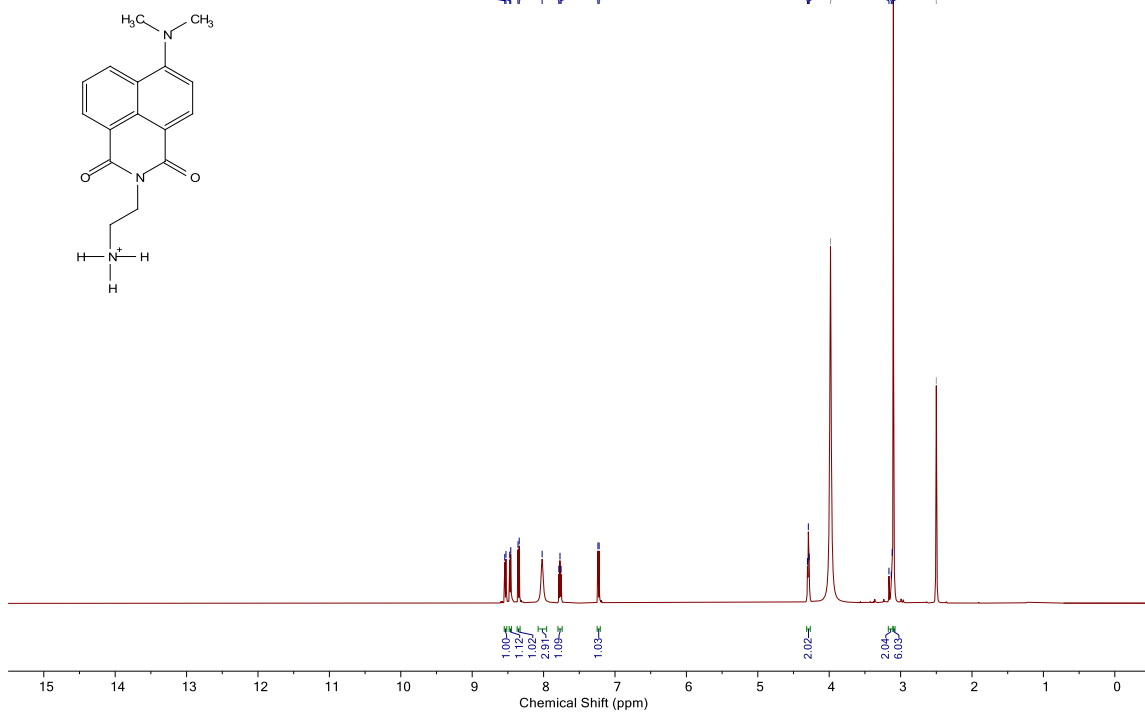
Figure 71 ^{1}H NMR Spectrum of C3 in DMSO-d$_6$.
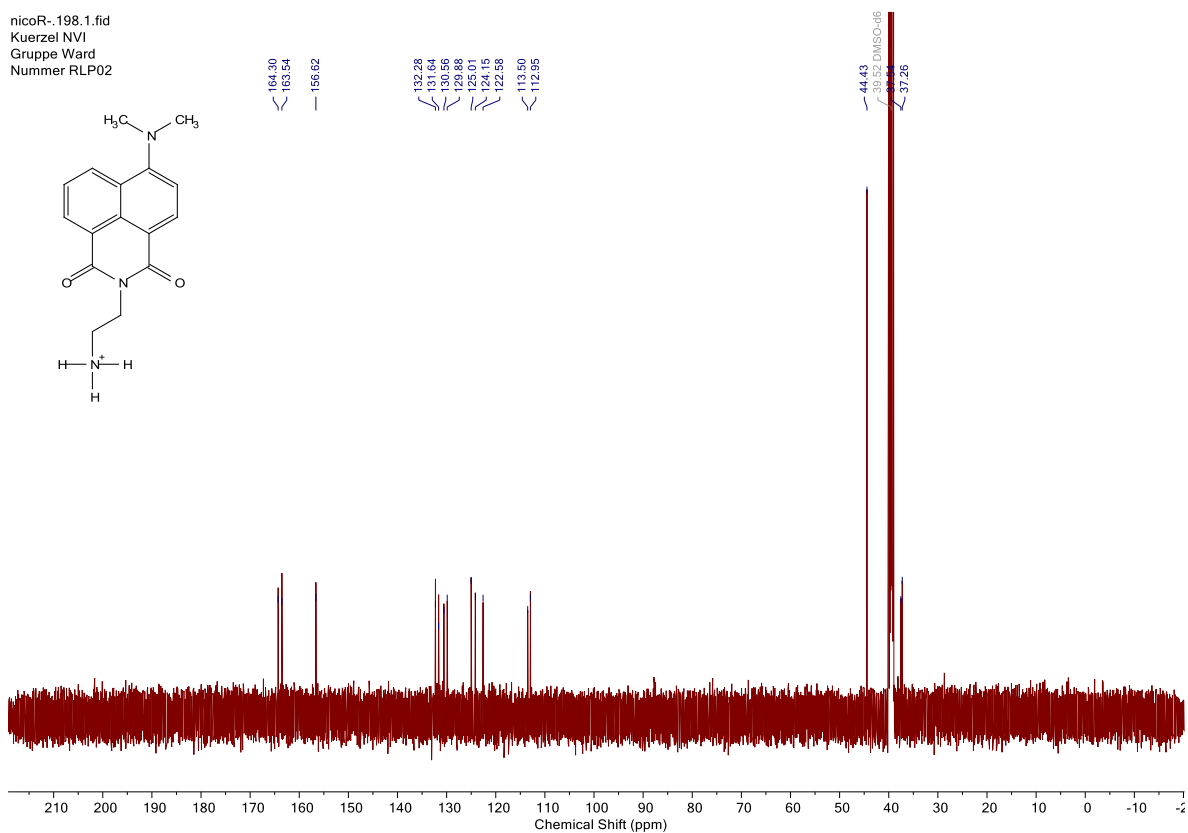
Figure 72 ^{13}C NMR Spectrum of C3 in DMSO-d$_6$.
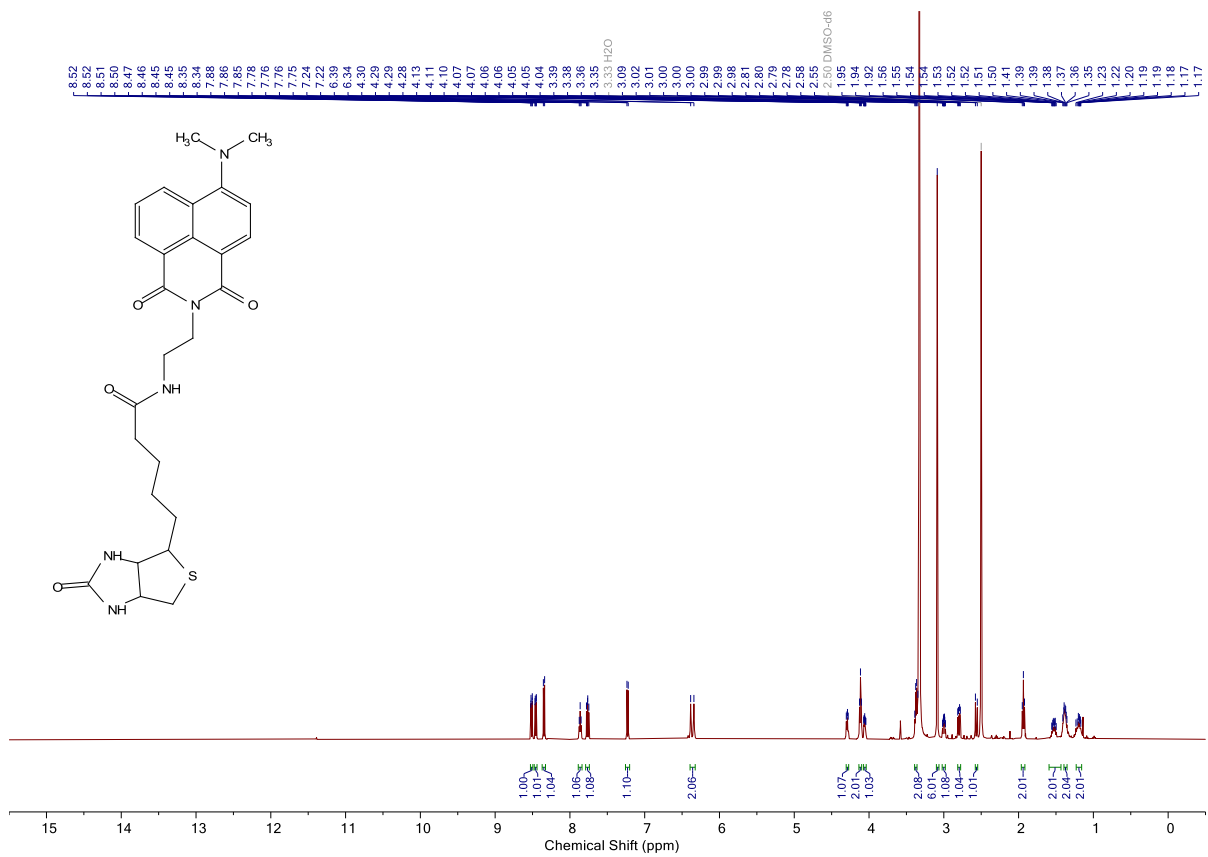
Figure 73 ^{1}H NMR Spectrum of biot-4DMN in DMSO-d$_6$.
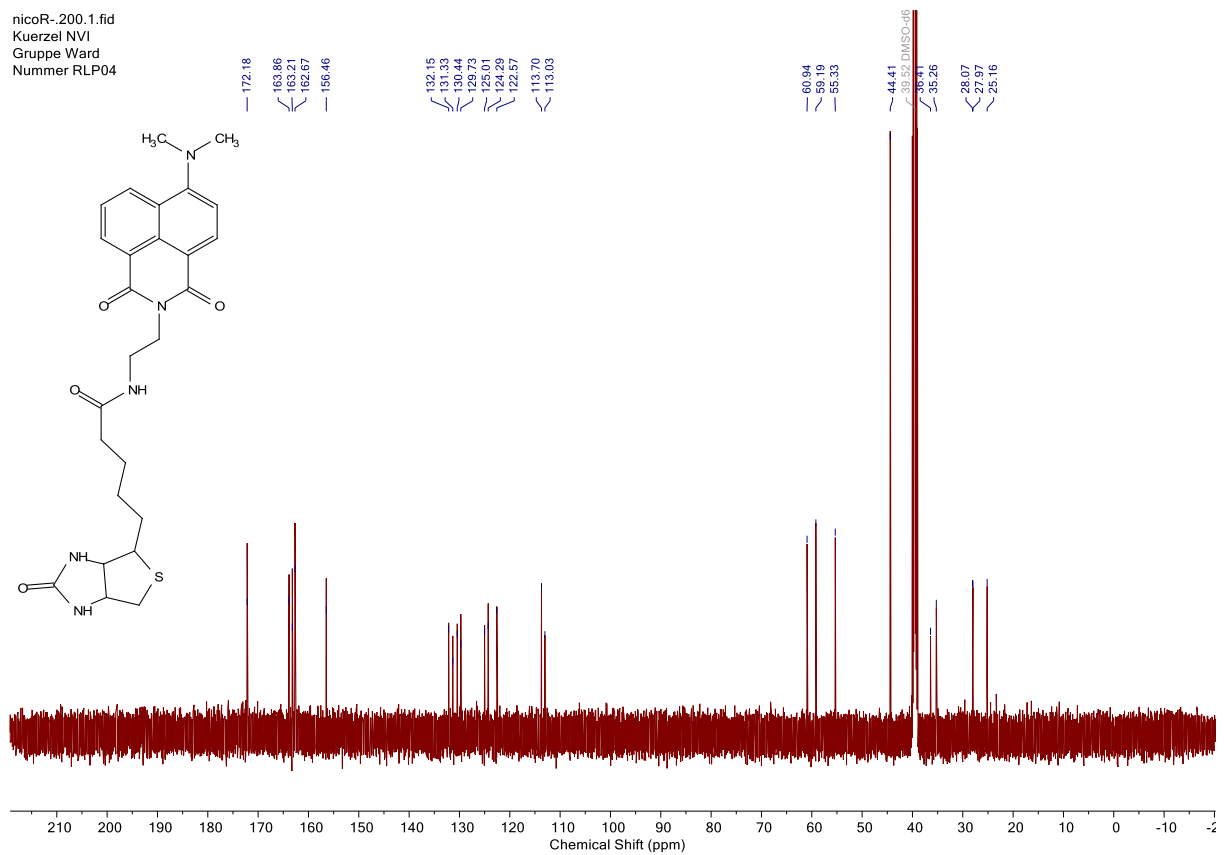
Figure 74 ^{13}C NMR Spectrum of biot-4DMN in DMSO-d$_6$.

3.3 Gold-triggered Drug-Release Systems

3.3.1 Outline and Authors

A modified version of this work was first published by and reproduced with permission from Fadri Christoffel: Christoffel, F. Directed Evolution of Gold-based Artificial Metalloenzymes and Design of Gold-triggered Drug-Release Systems and can be found under the DOI: https://edoc.unibas.ch/84356/

TRW, VS and FC conceived and designed the study. FC and VS developed the gold- catalyzed "close-to-release" cyclization of pro-furan substrates. NVI prepared the Sav isoforms. NVI and BL prepared the LM3 and the TATE peptide on rink amide and trityl resin, respectively. FC and NVI developed the peptide functionalization with gold-NHC complexes. FC and NVI analysed the data.

3.3.2 Introduction

In order to fight diseases, scientists rely also on so-called biorthogonal reactions, which stands for chemical transformations that can take place *in vivo* (*in cellulo*) without interfering with native biochemical processes.[182] One of the strategies used is to decorate a drug with a cleavable protecting group. This strategy minimizes the prodrug's toxicity and releases the drug as a result of the metal-catalyzed uncaging reaction.[183–185] However, until recently, the spectrum of *in-vivo* cleavable protecting groups was mostly limited by the repertoire of natural enzymes.[186] These enzymes, or metalloproteins, contain first-row transition metals such as copper, iron, and zinc and are catalyze a large variety of reactions in the cell. Accordingly, these can also be used to cleave a protecting group from a prodrug. To complement these natural enzymes, abiotic transition metal-catalyzed reactions have been introduced into living cells and organisms. They mainly are based on second row and third row transition metals. The implementation of these transition metal catalyst to living systems comes with a range of challenges. The inherent cytotoxicity of the abiotic catalysts needs to be overcome, the efficient uptake by the cell, the localization, and one needs to circumvent the deactivation of the catalyst by proteins (enzymes), thiols, and other cell components.[99]

Two metals were extensively used over the last decades for these *in vivo* uncaging-reactions, namely ruthenium and palladium. The reactions included the uncaging of allyls[187,188], allyloxycarbonyls (Alloc)[160,189–191], allenes[192], azides[193], propargyls[153,194–196], propargyloxycarbonyls (Proc)[197–200] and napthyls[185] (Figure 75).
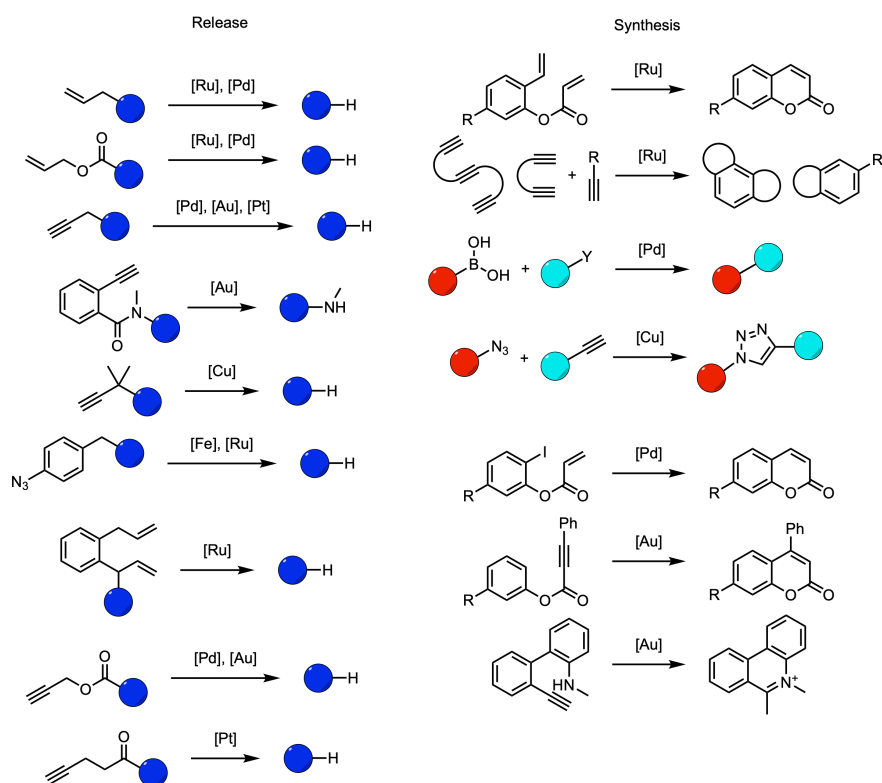
Figure 75 Bioorthogonal reactions applied in vivo. Uncaging reactions releasing the product (blue disk) and the synthesis of bioactive molecules or fluorophores. The figure was adapted from Lozhkin et al.[186] and reproduced with the permission of the authors.

Usually, the reactions were designed to uncage fluorescent molecules such as rhodamine and/or coumarin. Following this approach, the reaction control *in vivo* and *in vitro* is facilitated, as the products can be detected at very low concentrations. The first report of such bioorthogonal uncaging reaction in living systems was published in 2006 by Meggers and co-workers. In their work, they showed that the ruthenium-catalyzed Alloc-cleavage of rhodamine110 can be carried out in HeLa cells.[201] Since then, the reaction was further expanded to other organisms.[189,191,200,202–204] In contrast to these reactions catalyzed by Ru and Pd, the gold catalyzed uncaging reactions are underrepresented. In 2017, the groups of Unciti-Broceta and Tanaka recognized the potential and developed the first gold-catalyzed

propargyl and Poc-uncaging reactions in biological medium or living systems.[96,205] Some of the advantages of the gold catalysts are the remarkable biocompatibility, low toxicity, good reactivity at room temperature, and a propensity to coordinate to alkynes, the most widely used biorthogonal functional group. However, thiols such as glutathione or cysteine, present at significant concentrations in cells, are known to poison the catalysts. Therefore, either by smart catalyst design or by compartmentalization of the catalyst in an oxidazing environment –which leads to the oxidation of glutathione–, the catalyst may be protected and remain active.

In the past 20 years, somatostatin receptors (SSTR) have been used for imaging and treatment purposes of neuroendocrine tumors, as the somatostatin receptors of subtype 2 and 5 are overexpressed on these cells and can be targeted by the cyclic tetradecapeptide somatostatin (SS). However, the short in vivo half-life of somatostatin (1–2 min) precluded clinical applications of radiolabeled agents based on this targetting peptide.[206] Therefore, an intensive effort was invested into identifying a number of new cyclic and acyclic analogs of SS, with favorable *in vivo* properties. The cyclic disulfide containing octapeptide derivative of SS, octreotide, was the first analog to be used in clinical protocols, as it displayed high affinity for $SSTR_2$ and $SSTR_5$, the receptors mainly overexpressed on (human) neuroendocrine tumors.[207] The main feature of this peptide, which was necessary for its properties, was the disulfide tether, stabilizing the type II' s-turn structure centered at the D-Trp^8-Lys^9 sequence, which is considered to be essential for interacting with the cognate SSTRs.[208] In 1996, Bass *et al.* could show that the inversion of chirality at positions 1 and 2 of the octapeptide converted an agonist into a potent antagonist.[209] Simplified, antagonist can be seen as "blockers" of the target-receptor, that upon binding, reduce or inhibit (completely) the activity of the system. Some of the antagonists are widely used as APIs, such as antihistamines and β-blockers. Agonists, on the other hand, are compounds that upon binding to the target receptor, lead to an active response of the cell. Insuline or oxytocin are some examples for agonists.[210] These agonists usually get internalized upon binding, therefore allowing for an attached cargo to be delivered into the cell. Antagonists on the other side, display very low internalization rates and remain primarily on the cell surface. In the field of Theranostics (Therapy & Diagnostics), both agonists and antigonists are used. By attaching a radioactive compound/drug to the targeting peptide, the delivery of the radioactive moiety (in)to the desired target cell allows

for targeted radiotherapy and labelling of the tumors, even at highly metastatic stages.[211] For our studies, we focused on one antagonist (**LM3**) and one agonist (**TATE**), which have been shown to be versatile anchoring motifs for the imaging and therapy of tumor cells (Figure 76).[212–214]
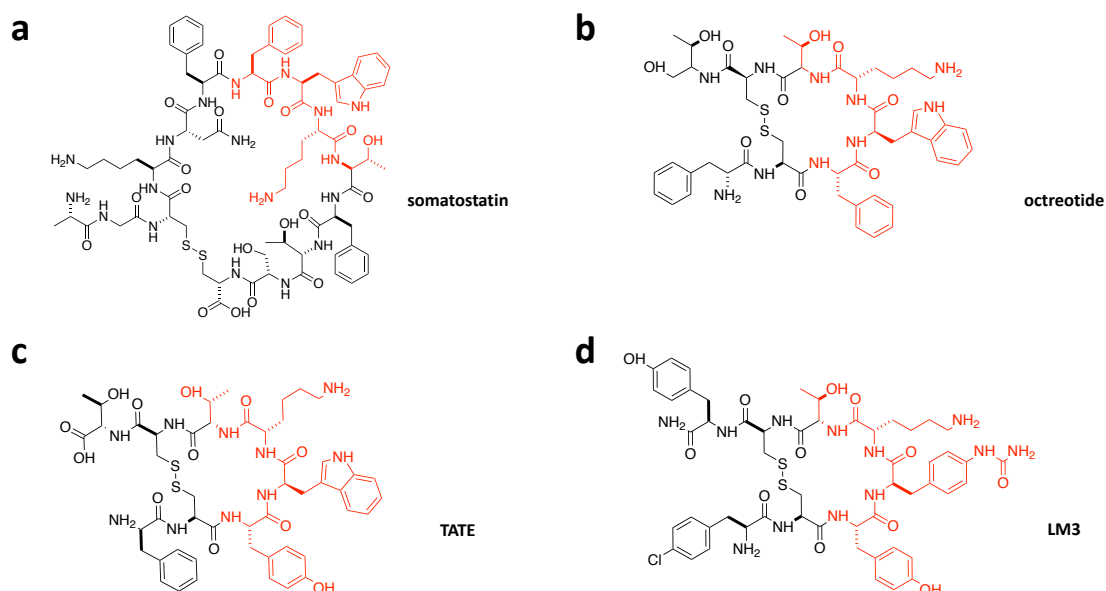
*Figure 76 Somatostatin and their analogues. The recognition site for the receptor is labeled in red for all the peptides. **a** Somatostatin **b** Octreotide **c** TATE **d** LM3*

3.3.3 Results and Discussion

In recent years it has been shown, that the generating aromatic intermediates is an efficient tool to design "close-to-release" probes.[185,215] Driven by our progress in the generation and utilization of highly active gold ArMs, we were interested in adapting them for bioorthogonal uncaging reactions. Nishibayashi et al. showed in 2008 that they could cyclize a substituted 3-butyne-1,2-diols into furans with the elimination of water in the presence of ruthenium catalysts.[216] Inspired by their work, we aimed at designing a novel "close-to-release" protocol based on profuran substrates

Figure 77). The payload of choice was umbelliferone, as coumarin-based fluorophores have a large difference in fluorescence intensity upon release.[217] Furthermore, several active pharmaceutical ingredients (API) contain coumarin moieties.[218–221]
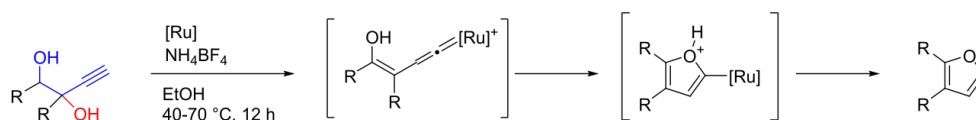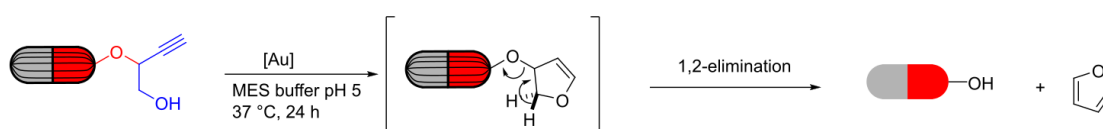
Figure 77 Modification of the ruthenium-catalyzed [Ru] dehydrative intramolecular cyclization to a gold-catalyzed [Au] "close-to-release" reaction. Reproduced with permission of the author from [222].

Capitalizing of the propensity of Au(I) to catalyze intramolecular hydroalkoxylation of alkynes, we investigated the capability of gold-based ArMs[36,161] to catalyze an intramolecular cyclization with subsequent 1,2-elimination under aqueous conditions. To our delight, blue fluorescence could be observed after a couple hours following mixing the substrate with the Au-based ArM. We therefore performed a simple pre-screening for the evaluation of the optimal cofactor protein host system. A black 96 well plate was sampled with different biotinylated gold complexes **biot-Au 1-5** and Sav as well as Sav-SOD mutants in various combinations. A clear preference for **biot-Au 2** as the optimal cofactor and Sav-SOD K121A as the optimal Sav construct was observed (Figure 78). This combination yielded after three hours 68% conversion of caged umbelliferone **4**, using a 5 mol% catalyst loading.
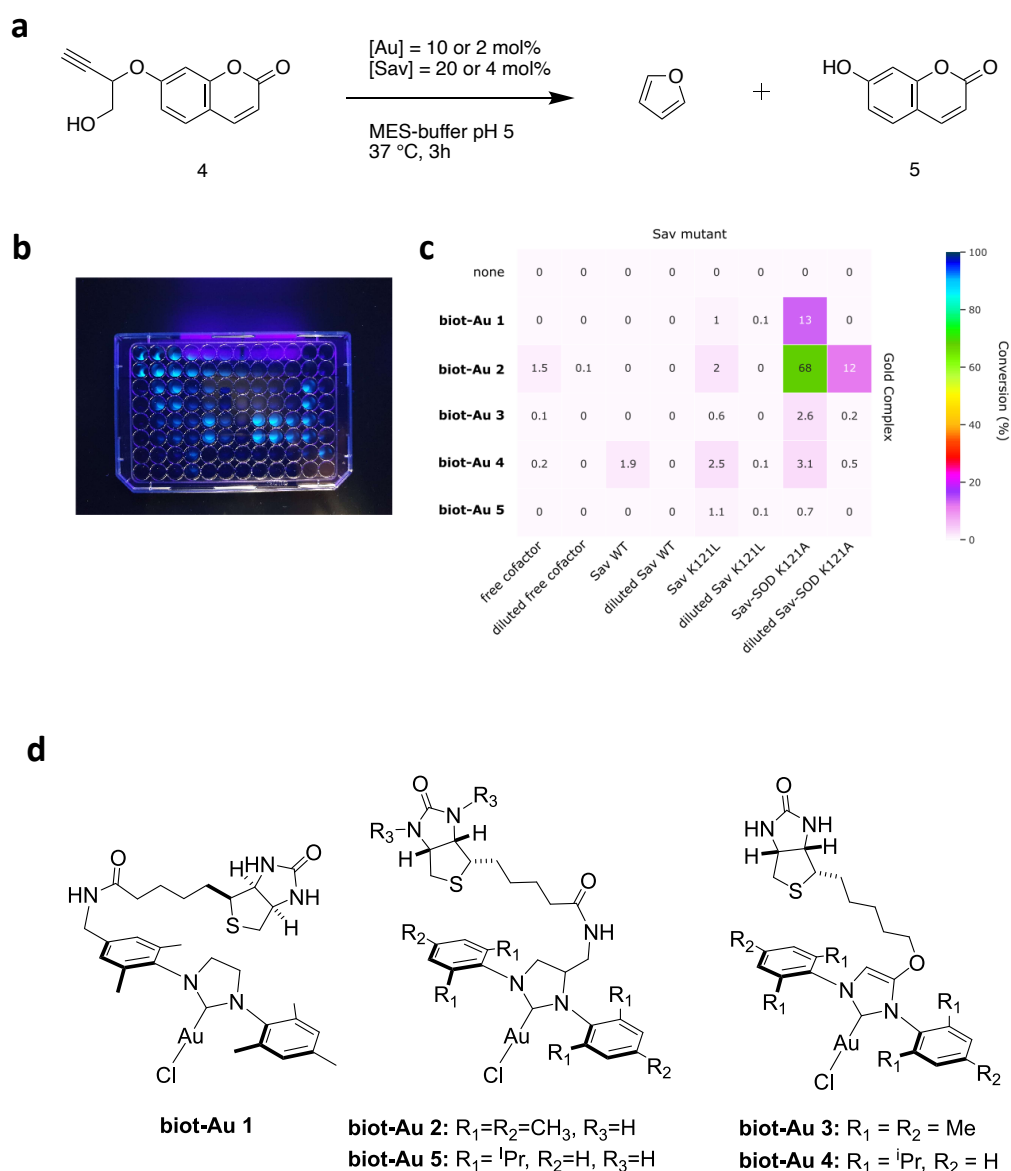
*Figure 78 Fluorescence screening of caged umbelliferone **4**. **a** The "close-to-release" reaction of caged umbelliferone **4** to furan and umbelliferone in MES buffer was screened at a 10 mol% and 5 mol% catalyst loading. **b** Photograph of the dark 96-well plate after 24 h of reaction time under UV illumination. **c** Heatmap displaying the catalytic results of the screening after 3 h of reaction. Reproduced with permission of the author from [222]. **d** Chemical structure of the Au(I) catalyst used.*

Even though the initial results were promising; the turnover remained moderate (12-13TON). Nevertheless, with the aim of using gold catalyzed "close-to-release" reaction *in vivo*, we collaborated with the Fani Group from the University Hospital of Basel. The SSTR targeting cyclic peptides p-Cl-Phe-cyclo(D-Cys-Tyr-D-4-amino-Phe(carbamoyl)-Lys-Thr-Cys)D-Tyr-NH2 (**LM3**) and H-D-Phe-Cys-Phe-D-Trp-Lys-Thr-Cys-Thr-OH (**TATE**) were synthesized. [223,224] As mentioned in the introduction, **LM3** acts as an antagonist, whereas **TATE** acts as an agonist.

Taking advantage of their properties, we planned to functionalize the peptides with the gold catalysts in order to generate cancer-targeting catalysts. This may allow the targeting of the cancer cells, whereas in a second step the caged, and thus less active, chemotherapeutics would be added and converted in or on the cancer cell to the active compound. To test the system, we first relied on the caged umbelliferone for this specific close-to-release reaction. A similar approach as in with the Sav ArMs was pursued. In a first step the NHC ligands were functionalized with gold, yielding a gold complex. In a second step, the protecting group of the amine was removed and the NHC-Au(I) was coupled to the carboxylic acid moiety of a succinic acid modified **LM3** and **TATE** (Figure 80). This carboxylate moiety was installed, by the reaction of the primary amine with succinic anhydride. After washing, the functionalized peptides were cleaved from the resin and precipitated by addition of cold diethyl ether. To monitor the coupling efficiency, we relied on the standard three step Kaiser test[225] which revealed the presence of free amines through the formation of a purple chromophore (Figure 79). To remove the protecting groups from the peptides, both a mild and a harsher cleaving protocol were evalutated. If 5% TFA was used, TATE could be cleaved without removing the amino acid sidechains protecting groups. In contrast, by using 95% TFA we would cleave **TATE-Au** and simultaneously remove all protecting groups on the peptide (Figure 81). For **LM3-Au**, by HRMS analysis the desired compound could not be detected, however a pattern slightly lower in m/z, suggesting the potential decomposition of the urea moiety on the D-4-amino-Phe(carbamoyl) fragment was observed (Figure 86). To exclude the possibility of partial decomposition of the peptide resulting from the ionization mode, the complex was also measured with an MALDI-Tof (HRMS), yielding nearly the same mass peak (ESI m/z = 1637.74 vs. MALDI m/z = 1640.87). Therefore, the exact composition of the peptide complex needs to be investigated in more detail.
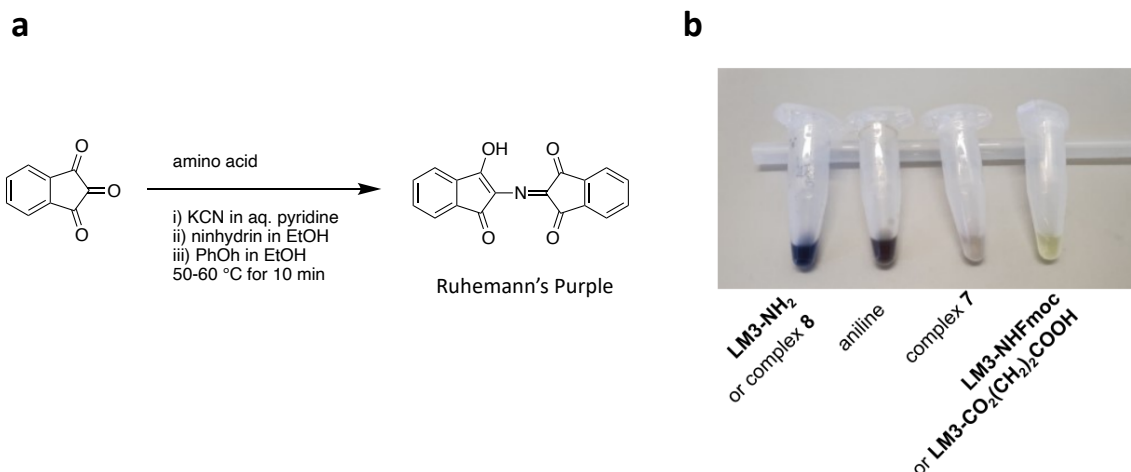
Figure 79 Kaiser Test. **a** *The reaction of ninhydrin with amino acids leads to a complex reaction resulting in the formation of Ruhemann's Purple, a strong purple chromophore.* **b** *Eppendorf tubes displaying common color patterns resulting from samples treated with the Kaiser Test. Reproduced with permission of the author from* [222].
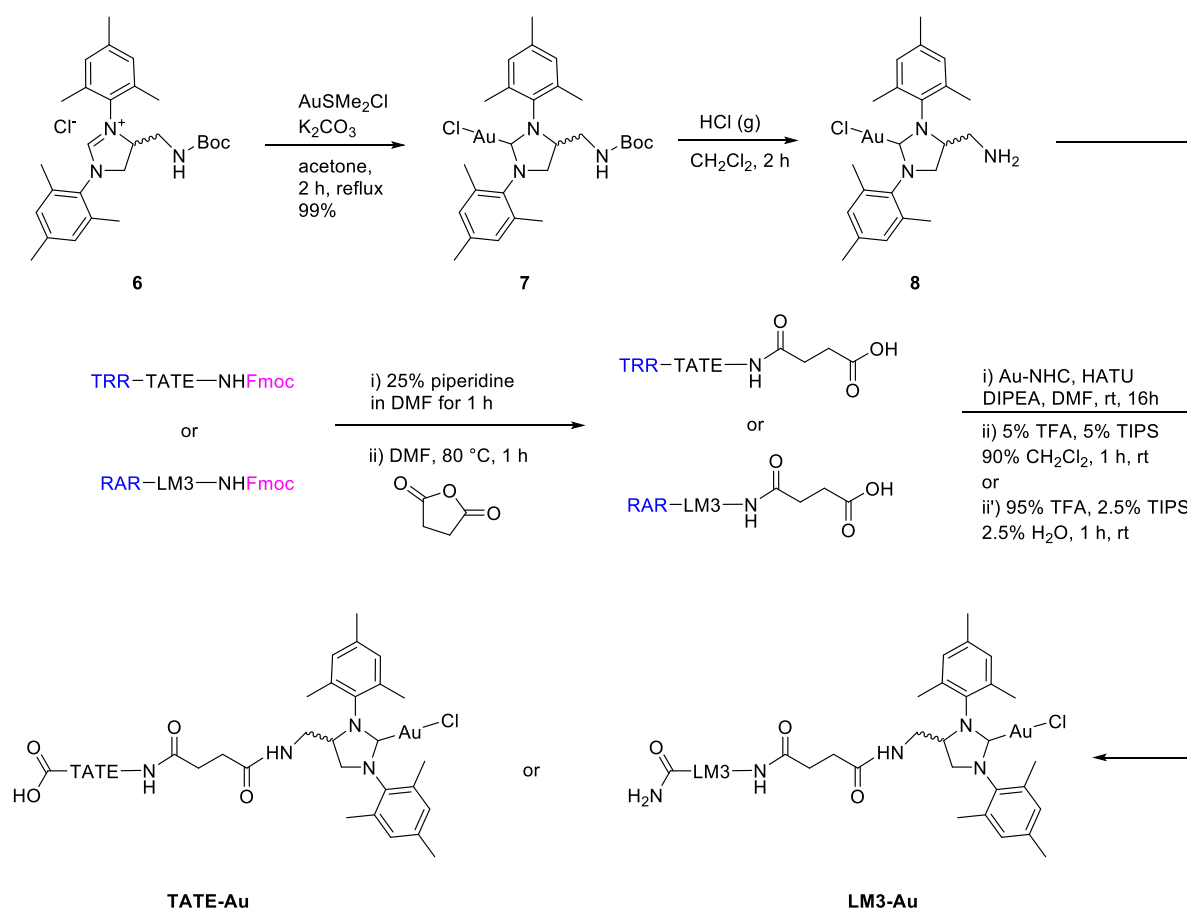
Figure 80 Synthesis of **TATE-Au** *and* **LM3-Au** *on trityl resin and rink amide resin, respectively.* *Reaction monitoring was performed via a Kaiser test and subsequently confirmed via HPLC into HR-MS. As the reaction was carried out on an analytical scale, the exact yields of the gold-functionalized peptides could not be determined precisely (ca. 1 mg crude product for both* **LM3-Au and TATE-Au**). *Reproduced with permission of the author from* [222].

Prior to upscaling, we tested the **TATE-Au** and **LM3-Au** as catalysts for the hydroarylation of pro-coumarin **9** and the close-to-release reaction of pro-furan **4**. The reactivity was also compared to the performance of **biot-Au 2** · Sav-SOD K121A (Figure 82). Initial screenings in MeCN:H_2O (1:1) highlighted the ability of both gold-functionalized peptides to catalyze the hydroarylation of pro-coumarin **9** almost quantitatively, clearly outperforming the ArM under these conditions (Table 30).
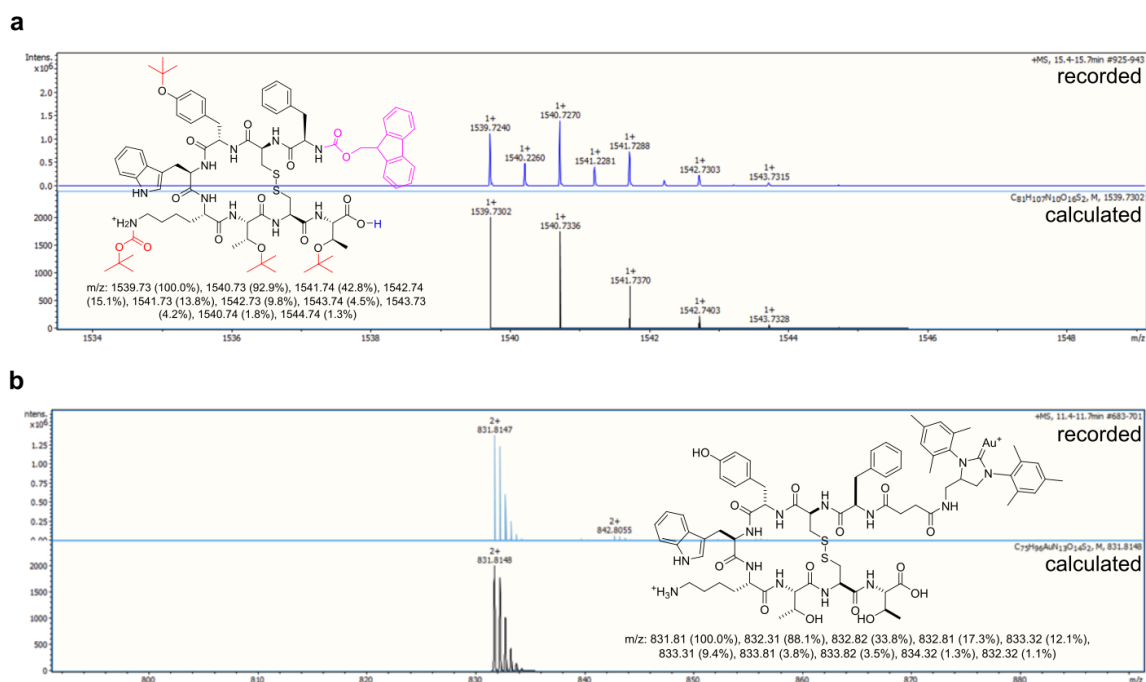
Figure 81 HPLC into HR-MS Spectrum of TATE-COOH[+] and TATE-Au[2+]. Reproduced with permission of the authors from [222].
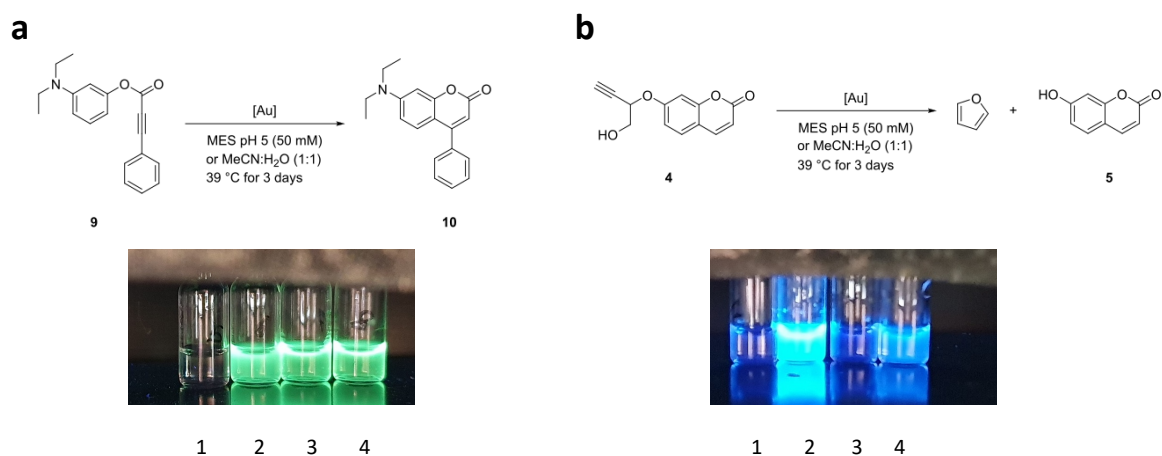
*Figure 82 Screening the gold-functionalized peptides for the intramolecular hydroarylation and the close-to-release reaction. 1 substrate only 2 **biot-Au 2** · Sav-SOD K121A 3 **TATE**-Au (batch 1) 4 **TATE**-Au (batch 2).*

The "close-to-release" reaction did not proceed in aqueous acetonitrile, however, using MES buffer (50 mM, pH 5), the **biot-Au 2 · Sav-SOD K121A** led to quantitative conversion. The functionalized peptides also catalyze the reaction, but with a modest conversion (11% for **LM3-Au** and 7% for **TATE-Au**) at 5% catalyst loading.

*Table 28 Selected Results of **LM3-Au** and **TATE-Au**-catalyzed intramolecular cyclization reactions. Reproduced with permission of the authors from [222].*

Entry	Au-Catalyst	Substrate	Solvent	Conversion[a]
0	-	**4**	MES pH 5	0%
1	-	**4**	MeCN:H_2O (1:1)	0%
2	**biot-Au 2 · Sav-SOD K121A**	**4**	MES pH 5	99%
3	**LM3-Au**	**4**	MeCN:H_2O (1:1)	1% (0%[b])
4	**LM3-Au**	**4**	MES pH 5	11% (0%[b])
5	**TATE-Au**	**4**	MES pH 5	7% (1%[c])
6	-	**9**	MES pH 5	0%
7	-	**9**	MeCN:H_2O (1:1)	0%
8	**biot-Au 2 · Sav-SOD K121A**	**9**	MeCN:H_2O (1:1)	37%
9	**LM3-Au**	**9**	MES pH 5	0%
10	**LM3-Au**	**9**	MeCN:H_2O (1:1)	97%
11	**TATE-Au**	**9**	MeCN:H_2O (1:1)	99% (96%[c])

[a]The conversion was determined via UPLC-MS at 321 nm. Reaction condition V_{tot} 500 µL, [Sub] 1 mM, [**Au**] 20 µM and [Sav] 40 µM (ArM) or ca 0.5 mg/mL (peptide), 39 °C for 3 days. [b] **LM3-Au** synthesized via sequential coupling of protected succinic acid followed by gold complex. [c] **TATE-Au** cleaved under mild conditions (5% TFA, 5% TIPS in CH_2Cl_2)

3.3.4 Conclusion

A new "close-to-release" reaction, resulting from a gold-catalyzed hydroalkoxylation and spontaneous 1,2-elimination, has been developed. The reaction leads to the release of umbelliferone as a payload. It may be amenable to the release of a drug cargo *in vivo*. This reaction is catalyzed by chimeric ArM **biot-Au 2 · Sav-SOD K121A** to afford quantitative conversion (i.e. 50 TONs). With *in vivo* applications in mind, we tested the peptides **LM3-Au** and **TATE-Au** for their catalytic performance on the hydroarylation of pro-coumarin substrate **9**, as well as the releasing umbelliferone from substrate **4**. However, the nature of the solvent significantly affects the conversion, which may complicate complicate *in vivo* applications. To avoid poisoning of the gold catalyst by cellular components, the **LM3-Au** complex may prove

more attractive as it should not be internalized by the somatostatin receptors mechanism upon binding to its target. Herein, we could also establish the first proof-of-concept of gold-catalysis on cancer-targeting peptides by utilizing a newly designed, biorthogonal close-to-release reaction resulting from a gold-catalyzed hydroalkoxylation coupled with a spontaneous 1,2-elimination of a cargo.

3.3.5 Experimental

General information. Chemicals were purchased from Sigma Aldrich, Acros Organics, Alfa Aesar or Fluorochem and used without further purification. High-resolution mass spectrometry (HR-MS) was performed on a Bruker maXis II QTOF ESI mass spectrometer coupled to a Shimadzu LC or a MALDI 8020 (MALDI-Tof) from Shimadzu. Fluorescence assays were recorded on a Tecan fluorimeter Infinite M1000Pro. Samples were prepared on a clear 96-well plate.

Peptide synthesis. The cyclic peptides were prepared with the help of the M. Fani Group at the University Hospital of Basel according to reported procedures by Nico Igareta and Boris Lozhkin.[211,226] The peptide **LM3** was synthesized by parallel standard Fmoc solid-phase synthesis on RinkAmide MBHA resin as previously published.[212] The Fmoc-protected p-Cl-Phe-OH, D-Tyr(tBu)-OH, Cys(Acm)-OH, Thr(tBu)-OH, Lys(Boc)-OH, D-4-amino-Phe-carbamoyl(tBu)-OH, Tyr(tBu)-OH, and D-Cys(Acm)-OH (Acm 5 acetamidomethyl) were used in a 2-equivalent excess. Fmoc removal was achieved with 20% piperidine in N,N-dimethylformamide in two 10-min treatments. Cyclization was performed with 10 eq. of iodine in DMF overnight. Final cleavage of the peptide from the resin and deprotection was achieved by adding a 95 % TFA, 2.5 % triisopropylsilane and 2.5 % H_2O solution to the resin and shaking it for 4h at room temperature. The peptide was precipitate in ice cold diethyl ether and washed several times with ice cold diethyl ether and dried at high vacuum.

Gold complex. The gold complex **7** was prepared according to a reported procedure.[161]
The **LM3** and **TATE** peptides on the resin (rink amide and trityl resin respectively) were transferred into peptide-synthesizer plastic vessel with filter inlet (10 mL) and sealed with a

septum. To ensure a dry and oxygen-free environment, nitrogen was purged from below through the vessel, allowing for continuous mixing and easy access for the addition of further reagents via a syringe. To remove the solvent through the bottom filter, the nitrogen line was swapped from bottom to top. Parafilm was used to seal the system for longer reaction times.

Fmoc-cleaving Procedure: The Peptide-containing resin (ca 25 mg) was soaked in DMF (3 mL) under nitrogen purging (30 min). The solvent was removed and the resin washed with DMF (2 mL). The Fmoc-group was removed by the addition of piperidine in DMF (25%, 3 mL). After purging (30 min), the solvent was removed, the resin was washed with DMF (2 mL) and the deprotection step was repeated one more time. Afterwards, the solvent was removed and the resin washed with DMF (3x2 mL). A spatula tip of resin was transferred to an Eppendorf vial to perform a Kaiser test.[227] In short, three solutions must be prepared for this test: (a) ninhydrin (500 mg) in ethanol (10 mL), (b) phenol (80 mg) in ethanol (10 mL) and (c) aq. KCN (2 mL, 0.001 M) in pyridine (100 mL). From each of these three solutions, ~20 μL are transferred to the Eppendorf vial containing the sample. The vial is heated to reflux (>10 min) upon which the solution turned blue if amines are present in the solution. The presence of a deep blue solution qualitatively confirms the successful deprotection of Fmoc from the peptide.

The resin was removed from the plastic vessel and transferred to a glass vial containing succinic anhydride (15 mg) and DMF (2 mL). The mixture was shaken in a thermoshaker (90 min, 80 °C). After the mixture cooled down (rt), it was transferred with the resin into the plastic vessel. The solvent was removed, and the resin was washed with dry DMF (3x2 mL). A second Kaiser test was performed in which the solution remained clear, thus confirming the absence of free amines.

Deprotection of gold complex **7**: The Boc-protected complex **7** (48 mg, 0.07 mmol) was dissolved in CH_2Cl_2 (4 mL/mmol) and HCl gas was bubbled through the solution in the dark (1 . The HCl gas was generated by the dropwise addition of concentrated sulfuric acid to ammonium chloride in a second flask and transferred via cannula into the reaction mixture (the controlled release of HCl gas from the system must be guaranteed to avoid overpressure). The solution was stirred in the dark (1 h, rt) while the deprotection step was monitored by thin-layer chromatography (MeOH/CH_2Cl_2 1:9) or a Kaiser test. The solvent was evaporated (high vacuum) and the resulting solid was dissolved in dry DMF (2 mL).

Coupling of the gold complex **8** to the peptides: In case the resin was stored temporarily after the addition of succinic anhydride, the resin should be soaked with dry DMF (30 min), purged and then washed with dry DMF (2 x 2 mL) before further proceeding. HATU (27 mg, 0.07 mmol), DIPEA (24 μL, 0.14 mmol) and dry DMF (1.5 mL) were sequentially added to the vessel. The mixture was purged (10 min) and the pH was monitored to ensure basic conditions. If required, DIPEA was added. Finally, the deprotected gold complex **8**, dissolved in dry DMF (2 mL) was added to the mixture. The seal was reinforced with Parafilm and the mixture was purged with N_2 overnight. The solvent was removed and the resin was washed with dry DMF (5 x 2 mL).

Cleaving from the resin:[228] Two cleaving procedures were tested: (1) harsh conditions with TFA:TIPS:H_2O (95.0:2.5:2.5) and (2) milder conditions with CH_2Cl_2:TFA:TIPS (92.5:5.0:2.5). Therein, the corresponding "cleaving mixtures" (2 mL) were added to the resin in the plastic vessel. The mixtures were purged (90 min, rt). Under harsh conditions a color change from yellow to dark grey was observed. The filtrate was collected in a V-shaped microwave vial. The gold-functionalized peptides were precipitated by the addition of ice-cold diethyl ether and isolated via centrifugation. RP-HPLC may be necessary to further purify the desired compounds **LM3-Au2** and **TATE-Au2**.

*Figure 83 **a** Peptide-coupling Setup and Kaiser test samples of **b** Fmoc-cleaved resin and **c** succinic acid coupled resin. Reproduced with permission of the author from [222].*
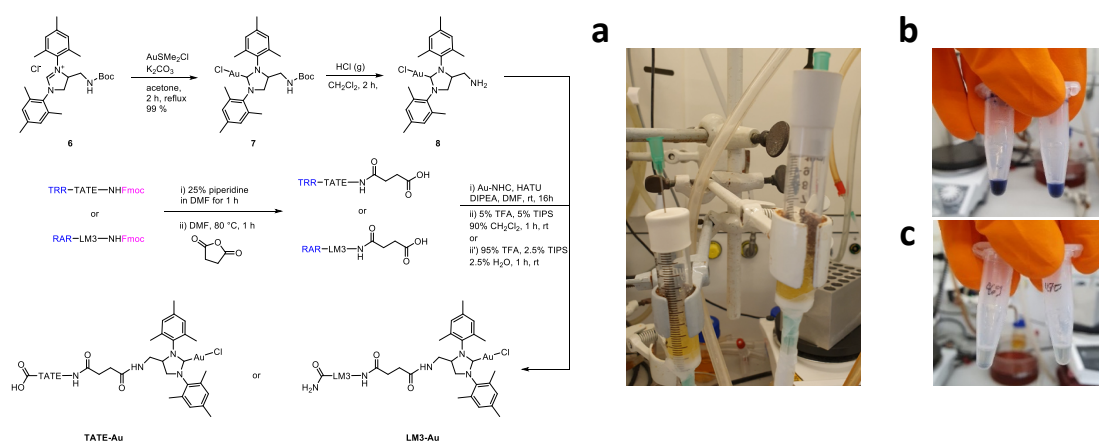
ESI-MS characterization. The peptides were dissolved in Mili-Q water, 0.1% formic acid pH = 2.5 with a final concentration of 0.2 mg/mL and clarified by centrifugation. A HPLC (Shimadzu, equipped with a Jupiter® 5 μm C4 300 Å)-ESI QTof (Bruker maXis II QTOF ESI) system was used to record the data. The ESI-QTof mass spectrometer was calibrated with ESI-Tof TuneMix (Agilent). The Compass DataAnalysis software (Bruker Daltonics) was used for data analysis.

Table 29. MS settings for peptide mass spectral analysis.

Source Type	ESI	Ion Polarity	Positive	Set Nebulizer	2.0 Bar
Focus	Not active	**Set Capillary**	4500 V	**Set Dry Heater**	200 °C
Scan Begin	150 m/z	**Set End Plate Offset**	-500 V	**Set Dry Gas**	8.0 l/min
Scan End	2750 m/z	**Set charging Voltage**	0 V	**Set Divert Valve**	Source
		Set Corona	4000 nA	**Set APCI Heater**	200 °C

MALDI-MS characterization. MALDI-TOF mass spectra were recorded on a Shimadzu MALDI 8020. As a matrix sinapinic acid was used (10 mg/mL). The peptide complex solution (0.5 μl, 1:1000 compared to the matrix, in DCM) was sandwiched within twice 0.5 μl of the matrix solution.
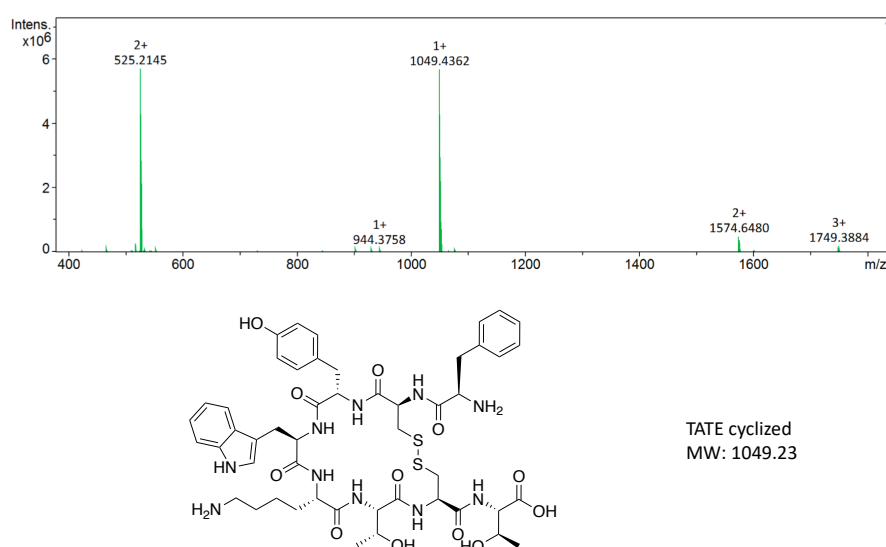
Figure 84 ESI-HRMS spectrum of the cyclized unprotected TATE.
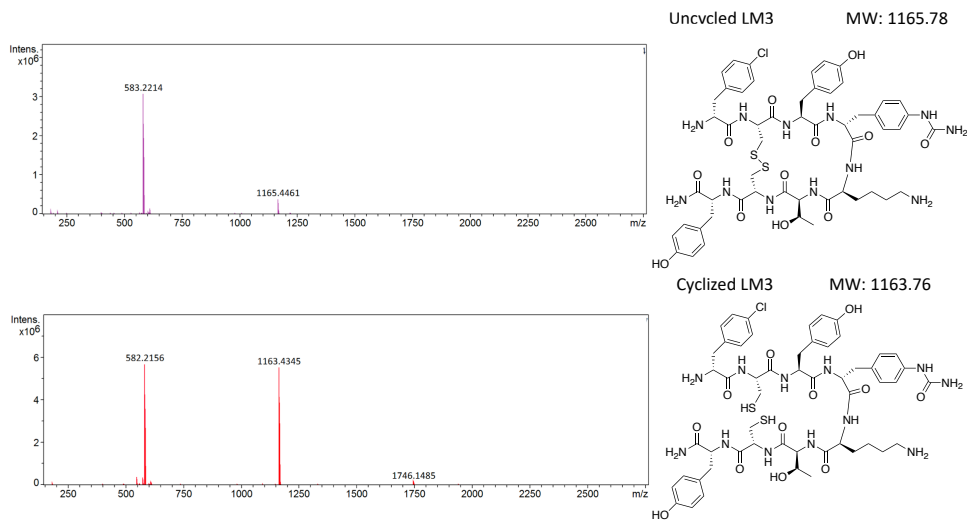
*Figure 85 ESI-HRMS spectra of the un- and cyclized unprotected **LM3**.*
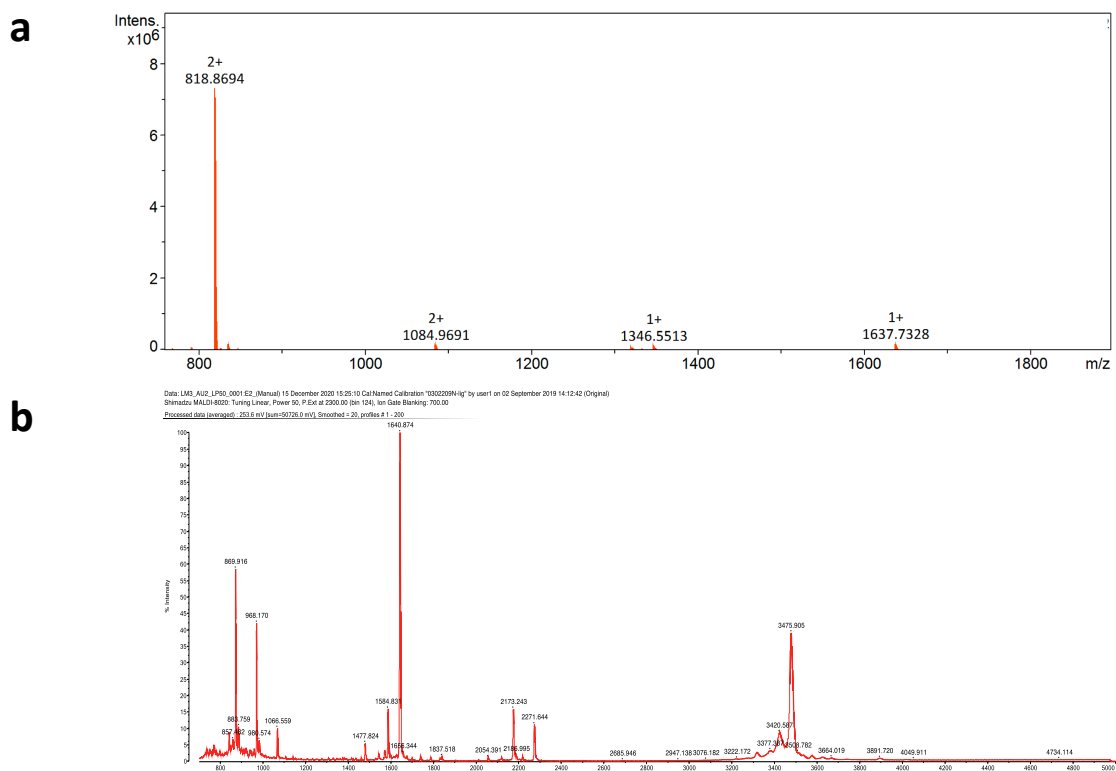
*Figure 86 **a** ESI-MS **b** MALDI MS spectrum of the "**LM3-Au2** complex".*

Chapter 4. Chimeric ATHase

4.1 Introduction

Chiral amines are structural units in nearly 40% of all pharmaceuticals and also widely used in the production of agrochemicals.[68] As such, the synthesis of enantiopure amines has attracted significant research interest in academia and industry. Common synthetic strategies may be broadly categorized as chemical, utilizing homogenous catalysts,[229] or biological,[230] using enzymes such as imine reductases, transaminases, phenylalanine ammonia lyases and amine dehydrogenases.[231–233] More recently, various artificial metalloenzymes have also been used for the synthesis of enantiopure amines.[67,69,94] The Ward group has extensively explored the use of artificial transfer hydrogenases (ATHases) based on the biotin-(strept)avidin technology for the reduction of ketones and imines.[66–69,76] These efforts were inspired by the work of Noyori *et al.*, who showed that imines could be effectively reduced by [(η^6-arene)Ru(*S,S*-TsDPEN)Cl] in the presence of a 5:2 formic acid:triethylamine azeotropic mixture in acetonitrile.[234] For various isoquinolines, an enantiomeric excess (*ee*) of up to 97% was achieved, revealing the critical importance of the chiral metal center in dictating enantioselectivity. As proposed by Noyori *et al.,* and established in subsequent studies, the active species in the imine reduction is a chiral Ir–hydride, which is formed following coordination and decarboxylation of formate.

Initially, the Ward group investigated the diversity of Noyori-type catalyst [(η^n-arene)M(biot-*p*-L)Cl] featuring a biotinylated supporting ligand. The metal and arene (benzene, *p*-cymene, mesitylene, durene, and hexamethyl benzene for M = Ru; pentamethylcyclopentadienyl for M = Rh, Ir) as well as the location of the substituent on the biotin moiety (*ortho*, *meta*, or *para*) were varied (Figure 87a),[40] and the resulting twenty-one racemic metal complexes were used for the reduction of acetophenone, *p*-bromoacetophenone, and *p*-methylacetophenone in the presence of streptavidin. These experiments revealed that the capping arene has a significant influence on the activity and selectivity of the ATHase, and only the *para*-substituted ligand, Biot-*p*-L, afforded significant conversions. Further, it was found that the Ru-based catalysts outperformed the Rh- and

Ir-based systems. From docking studies, using [(*p*-cymene)Ru(biot-*p*-L)H] and wild-type streptavidin, residues S112 and S122 were proposed to lie in the closest proximity to the metal center. Nineteen mutants were generated from saturation mutagenesis at position S112 and used as protein scaffolds for the ATHase, and it was found that aromatic and cationic residues yielded the highest enantioselectivities for (*R*)- and (*S*)-products, respectively.

Metal	η^6-C_nR_n	ee (%)	conv. (%)
Ru	η^6-*p*-cymene	22 (*R*)	97
Ru	η^6-benzene	12 (*R*)	76
Rh	η^6-C_5Me_5	52 (*R*)	94
Ir	η^6-C_5Me_5	57 (*R*)	quant.
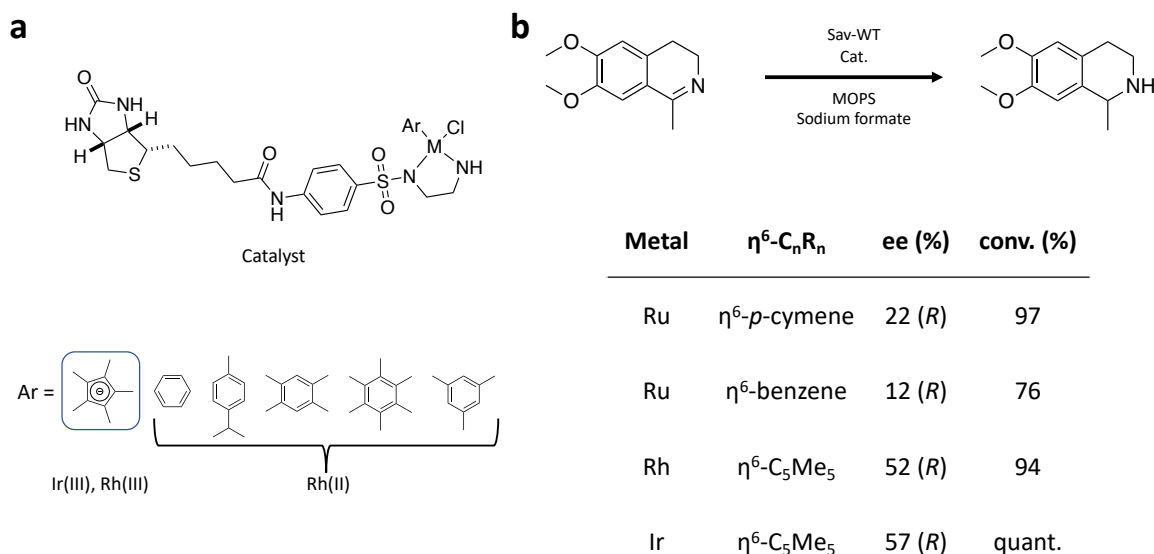
Figure 87: Initial studies of the artificial transfer hydrogenase for cyclic imines based on the biotin-(strept)avidin technology.[67] ***a*** *The Noyori-type catalyst used in the studies with different arene ligands and metals.* ***b*** *Reaction scheme of the reduction of the salsolidine precursor and results thereof.*

In 2008, the Ward laboratory reported the single-crystal X-ray diffraction structure of [(η^6-benzene)Ru(biot-*p*-L)Cl]·Sav-S112K, a promising (*S*)-selective ATHase for the reduction of acetophenone derivatives.[66] This structure helped to identify the K121 residue of an adjacent monomer as another close-lying residue of the metal center. Furthermore, residue L124 was found to be in close contact with the sulfonamide moiety of the ligand. Importantly, the Ru—Cl center was determined to have an absolute (*S*)-configuration, as also observed in the (*S*)-selective homogenous system developed by Noyori *et al.*[66,234] A second round of saturation mutagenesis was aimed at positions K121 and L124, and screening revealed that position K121 is more crucial for optimizing enantioselectivity. It was speculated that this effect might be due to the influence of the side chain 121, both on the piano-stool complex itself and on the trajectory of the incoming prochiral substrate.

In the reduction of imines, the iridium complex [Cp*Ir(biot-*p*-L)Cl] outperformed the Ru- and Rh-based catalysts when incorporated in Sav-WT (Figure 87b).[67] Two different mechanisms have been proposed for the asymmetric transfer hydrogenation of imines using homogeneous catalysts, the first being a concerted outer-sphere mechanism and the second one involving an ionic pathway, where the imine is protonated prior to hydride transfer from the metal center to the carbon of the C=N bond.[8,76,235] As noted above, the absolute configuration of the metal center in the transition state of the reaction is crucial for enantioselectivity. This was shown by Ikariya and Koike, who demonstrated for a homogenous ruthenium system that the chiral ligand dictates the absolute configuration of the metal hydride, which in turn determines the enantioselectivity.[236] In the case of the ATHase, two possible enantioselectivity determining mechanisms may operate, and both may contribute to enantiodiscrimination. Similar to the homogenous system, the second coordination sphere of the host protein may favor the formation of a single [Cp*Ir(biot-*p*-L)H]-diastereomer, which then induces the enantioselectivity (i) or simply one of the two resulting ATHase enantiomers is more reactive towards the substrate, but there is no influence of the second coordination sphere on the selectivity of the hydride formation (ii).

Single-crystal X-ray diffraction analysis of the ATHase [Cp*Ir(biot-*p*-L)Cl]·Sav-S112A revealed that cofactors are able to occupy all four binding sites within the protein without any steric clashes, and the configuration at the Ir—Cl center is (*S*).[67] Thus it was proposed that the protein environment discriminates between two diastereomers of [Cp*Ir(biot-*p*-L)H] and dictates the absolute configuration of the metal. Furthermore, based on the close proximity of the Ir centers of the two adjacent monomers (5.1 Å) (Figure 88a), interactions during the catalytic process seemed very likely. Interestingly, upon increasing the molar ratio of [Cp*Ir(biot-*p*-L)Cl] to Sav-112A from 1:1 to 4:1, gradual erosion of enantioselectivity was observed (Figure 88c).[76] Furthermore, the corresponding catalytic efficiency decreased from k_{cat}/K_M = 0.22 min^{-1}mM^{-1} to 0.011 min^{-1}mM^{-1}. In contrast, the enantioselectivity of [Cp*Ir(biot-*p*-L)H]·Sav-S112K remained nearly constant when the catalyst:tetramer ratio was varied from 1:1 to 4:1, and the corresponding activity decreased only by 1.5 fold. From computational analysis, it was also found that in the absence of Sav, [Cp*Ir(biot-*p*-L)H] is not configurationally stable and instead affords a racemic metal intermediate in the transfer hydrogenation, thus leading to a racemic product mixture. Based on analysis of the catalytic

activity, single-crystal X-ray diffraction data, and computations, it was proposed that the reaction involves a non-concerted mechanism with a transition state wherein there is a CH⋯π interaction between the substrate and the Cp* moiety was proposed (Figure 88c). In the case of the [Cp*Ir(biot-*p*-L)H]·Sav-S112A, the ratio of cofactor to protein is crucial, whereas for [Cp*Ir(biot-*p*-L)H]·Sav-S112K an "induced key-and-lock" mechanism was postulated, where all four biotin-binding sites can accommodate a (R^{Ir})-[Cp*Ir(biot-*p*-L)H] cofactor and therefore a nearly constant rate and enantioselectivity were observed.[40,235]

Sav variant	eq. [Ir]/FBS	ee %
S112A	1.0	93
S112A	2.0	92
S112A	3.0	89
S112A	4.0	45
S112K	1.0	-70
S112K	2.0	-74
S112K	3.0	-76
S112K	4.0	-78
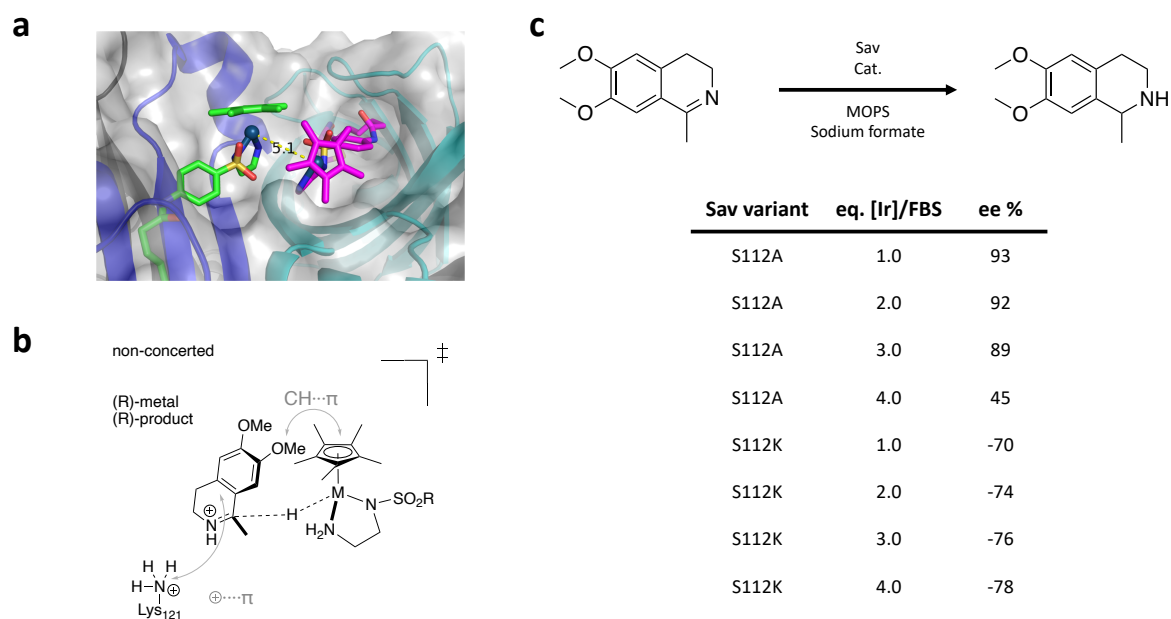
*Figure 88: **a** Crystal structure of [Cp*Ir(Biot-p-L)Cl]·S112A displayed as a tetramer, highlighting the close proximity of the two biotinylated cofactors. **b** Proposed mechanism for the reduction of the salsolidine precursor. The reaction takes place via a non-concerted mechanism as described by Dürrenberger et al.[67] **c** Dependency of the cofactor:biotin binding sites ratio for the reduction of the salsolidne precursor from the studies by Robles et al.[76]*

In 2013, the Ward group reported the optimization of an additional ATHase based on the biotin-streptavidin technology, wherein they linked the biotin anchor to the cyclopentadienyl moiety to generate a widely applicable d^6-piano stool rhodium complex with three coordination sites available for catalysis and/or activation via additional ligands. By introducing a histidine (either K121H or S112H) in close proximity to the rhodium center, a dual anchoring strategy with a supramolecular (biotin) and dative (over His) binding was achieved.[71] This strategy enhanced the enantioselectivity of the ATHase for the reduction of a salsolidine precursor (Figure 88b) compared to the use of Sav-WT. In a follow-up study,

instead of a dual anchoring strategy, a library of bidentate diamines, amino acids and amino amides were screened as ligands to tune the reactivity of the ATHase for the reduction of 1-phenyl-3,4-dihydroisoquinoline to 1-phenyl-1,2,3,4-tetrahydroisoquinoline.[237] ATHases featuring enantiopure α-amino amides as ligands were found to be the most promising, for example [Ir(biot-en-η^5-Cp*)(L-ThrNH$_2$)Cl)]·Sav-WT yielded the (*S*)-product with ee as high as 63% and turnover numbers (TONs) as high as 190. Additionally, the versatility of such ATHases extends to the creation of enzyme cascades.[238–240] In 2018, Hestericová *et al.* carried out a directed evolution campaign on the [Cp*Ir(biot-*p*-L)Cl]·Sav system using a straightforward procedure allowing screening in cell-free extracts.[68] Two mutants with increased reaction rates and enantioselectivities for the reduction of 1-phenyl-3,4-dihydroisoquinoline were found: Sav S112A-N118P-K121A-S122M yielded an *ee* of 95% for the (*R*)-enantiomer and Sav S112R-N118P-K121A-S122M-L124Y afforded 86% *ee* for the (*S*)-enantiomer.

Following the development of highly active and selective ATHases based on the biotin-streptavidin technology, subsequent efforts were directed toward optimizing the nature of the protein host. As mentioned in 2.1 Why engineer (strept)avidin?, Pellizzoni *et al.* introduced different chimeric streptavidins as protein scaffolds and demonstrated that corresponding ATHase activity could be increased by as much as seven-fold compared to Sav-WT.[73] Soon thereafter, Wu *et al.* introduced the so-called single-chain dimer streptavidin to the repertoire [Cp*Ir(biot-*p*-L)Cl]-based ATHases.[69] With this new chimeric streptavidin variant, they addressed the previously mentioned unfavourable interactions between the two adjacent (*cis*) complexes in the biotin-binding vestibule and expanded the genetic optimization potential. Of note, the newly generated ATHase achieved unrivaled levels of activity and selectivity for the reduction of challenging prochiral imines. Building on the previous work summarized above, we endeavored to design a new ATHase featuring chimeric Sav-SOD as the protein scaffold (see Figure 89), and we studied the impact of the protein on the reactivity and selectivity of the ATHase in the reduction of cyclic imines.

4.2 Kinetics and Mechanism

In 1913, Michaelis and Menten introduced a basic kinetic model of enzymes.[241] They postulated that for the case of a simple enzyme-catalyzed reaction involving the conversion

of substrate (S) to product (P), the mechanism consists of two steps. In the first step the enzyme-substrate complex (ES) is formed with an association rate of k_1, followed by the second step, where the chemical transformation takes place with a first-order rate (k_2) before the product (P) is released by the enzyme. Assuming that the second step is rate-limiting ($k_1 \gg k_2$), the enzyme-substrate complex is in thermodynamic equilibrium with the free enzyme and substrate.

$$E + S \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} ES \overset{k_2}{\rightleftharpoons} E + P$$
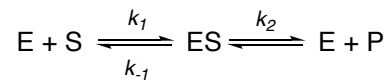
Furthermore, the equilibrium constant for the first step can be described as follows:

$$\frac{k_1}{k_{-1}} = \frac{[E][S]}{[ES]} = K_s$$

Because the concentration of the substrate is typically much higher than that of the enzyme ([S]>>[E]) in biotransformations, the concentration of the enzyme-substrate complex remains constant during the reaction. In 1925 Haldane and Briggs[242] made a more general "steady-state" approximation than the assumption of equilibrium, and based on this expressed the Michaelis-Menten equation as follows:

$$v_0 = \frac{V_{max} + [S]}{K_M + [S]}$$

By plotting the initial rate (v_0) versus substrate concentration [S], a hyperbolic curve, indicating the maximal reaction rate (V_{max}), can be obtained. If the reaction of interest meets the parameters of the basic Michaelis-Menten model, i.e., it involves the formation of only one particular enzyme-substrate complex and if all binding steps are fast, k_2 represents the catalytic constant k_{cat}. This constant is the equivalent of the turnover frequency (TOF) used in the context of homogenous catalysis and can be described as follows:

$$k_{cat} = \frac{V_{max}}{[E]_0}$$

Where $[E]_0$ is the total enzyme concentration ($[E]_0 = [E] + [ES]$). One other important constant for the Michaelis-Menten equation is K_M, which represents the concentration at which the rate v_0 is equal to half V_{max}.

4.3 Shielding the Active Site: A Chimeric Streptavidin Super-Oxide-Dismutase as Host Protein for Asymmetric Transfer Hydrogenation.

4.3.1. Outline of Author Contributions

A modified version of this work was first published by the Royal Society of Chemistry: Igareta, N.V., Tachibana, R., Spiess, D.C., Peterson, R.L., Ward, T.R., Spiers Memorial Lecture: Shielding the active site: a streptavidin superoxide-dismutase chimera as a host protein for asymmetric transfer hydrogenation. *Faraday Discussions,* Advance Article (2023) and can be found under the DOI: https://doi.org/10.1039/D3FD00034F

TRW, RLP, and NVI conceived of and designed the study. NVI, DCS, RLP contributed to mutagenesis, protein expression, protein purification. RLP and NVI performed the catalytic experiments. RT performed the computational calculations. NVI performed the crystallization, X-ray structure determinations and the (native) MS experiments.

4.3.2 Abstract

By anchoring of a metal cofactor within a host protein, so-called artificial metalloenzymes can be generated. Such hybrid catalysts combine the versatility of transition metals to catalyze new-to-nature reactions with the power of genetic-engineering to evolve proteins. With the aim of gaining better control over second coordination-sphere interactions between the streptavidin host-protein (Sav) and a biotinylated cofactor, we engineered a hydrophobic dimerization domain, borrowed from superoxide dismutase C (SOD), on Sav's biotin-binding vestibule. The influence of the SOD dimerization domain (DD) on the performance of an asymmetric transfer hydrogenase (ATHase) resulting from anchoring a biotinylated Cp*Ir-cofactor –[Cp*Ir(biot-p-L)Cl] **1**– with Sav-SOD is reported herein. We show that, depending on the nature of the residue at position Sav-S112, the introduction of the SOD DD on the biotin-binding vestibule leads to an inversion of configuration of the reduction product, as well as a

fivefold increase in catalytic efficiency. The findings are rationalized based QM/MM calculations combined with X-ray crystallography.

4.3.3 Introduction

Artificial metalloenzymes (ArMs) consist of catalytically-competent cofactors anchored within a protein host.[1] These systems uniquely combine the synthetic tunability of homogenous catalysis with the evolutionary malleability of nature's fundamental protein building blocks to yield new chemistries. One of the first ArMs was reported by Wilson and Whitesides in 1978. They demonstrated that anchoring an achiral biotinylated rhodium cofactor in avidin affords a hybrid catalyst for the asymmetric hydrogenation of prochiral alkenes.[2] Since that seminal report, advancements in the fields of homogenous catalysis and molecular biology have resulted in the development of a large variety of ArMs featuring different cofactors and diverse protein scaffolds, including carbonic anhydrase,[3] hemoproteins,[4,5] prolyloligopeptidase,[6] nitrobindin,[7] human serum albumin,[8] and (strept)avidin.[2,9]

Over the last two decades, the Ward laboratory has extensively developed streptavidin (Sav)-biotin technology to create a wide range of ArMs exhibiting unique biochemistry.[10–13] However, due to the inherent structure of the Sav host, the metal cofactors remain relatively solvent-exposed and accessible (Figure 89a and b), whereas in many natural metalloenzymes, the cofactors (e.g., metal ions, heme, pterins, etc.) are often deeply buried. A significant advantage of this feature of biological systems is that so-called secondary coordination sphere interactions can be exploited to alter cofactor reactivity and/or enhance chemoselective substrate modification.

Our group has explored numerous strategies for enclosing the streptavidin biotin-binding vestibule in order to furnish corresponding ArMs with a more defined substrate-binding

pockets. These strategies have included the insertion of (GGS)$_n$ loop extensions, chimeric-Sav genes containing small synthetic and natural peptide domains ranging from 24 to 60 residues in size, and the encapsulation of Sav within proteins or nanoparticle host scaffolds.[14–17] However, these approaches are not amenable to high-throughput methodologies because of limited protein production levels or necessary time-consuming refolding and/or extensive processing workflows. To circumvent the challenges faced with previous Sav designs, we recently introduced a chimeric streptavidin variant (Figure 89c and d) as a scaffold for an artificial hydroaminase (HAMase) based on dual-gold activation of alkynes.[18] In that study, we introduced the dimerization domain (DD) of the superoxide dismutase C (SOD) from *M. tuberculosis* (pdb 1pzs)[19] to the 3,4-loop of streptavidin (Figure 89e). By directed evolution, the second coordination sphere around the abiotic cofactor was further optimized to control the regioselectivity of the hydroamination reaction, leading to two mutants for either the single gold π-activation or dual gold σ,π-activation reaction. Encouraged by these results, we sought to utilize the Sav-SOD chimera as a protein host for an ATHase featuring the "racemic-at-metal" biotinylated d^6-piano stool complex [Cp*Ir(biot-*p*-L)Cl] (**1**) as a cofactor. Herein, we present the results of our investigation for the reduction of cyclic imines.
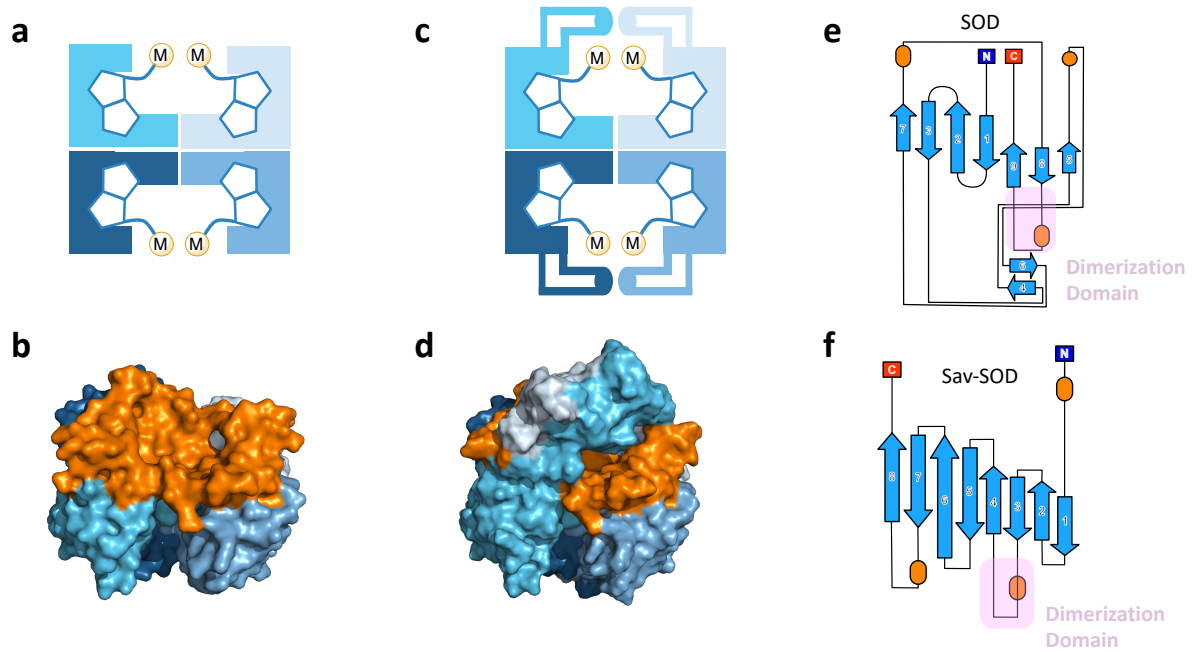
*Figure 89: **a** Schematic representation of tetrameric streptavidin (Sav) and **b** surface representation of Sav (pdb 3pk2) centered on one of the two biotin binding vestibules; each chain is represented by a different color and the surface of the vestibule is highlighted in orange. **c** Schematic representation of Sav-SOD, illustrating the potential shielding of the metal center of the cofactor and **d** surface representation of Sav-SOD centered on the biotin binding vestibule, highlighting the shielding of the vestibule (orange surface) by the DD (turquoise and grey surface). Origami representations of the monomer of **e** SOD including and the DD between the β-sheets 8 and 9 and (pink highlight) and **f** chimeric Sav-SOD including the DD inserted between the β-sheets 3 and 4.[36]*

4.3.4 Results and Discussion

Biochemical properties of the Sav-SOD Chimera. Based on our previous studies[73] and reports of cyclic permuted streptavidin (CPSav),[62,246] we engineered the dimerization domain of SOD into the 3-4, 4-5, 5-6, 7-8 loop of streptavidin, as well as the CPSav, to generate the chimeric Sav-SOD scaffold. Additionally, an N-terminal SUMO-tag was added to decrease the risk of aggregation and increase the solubility of the fused protein. Initial protein production experiments revealed that the chimera with the DD at the 3,4-loop was the only construct that yielded a soluble and biotin-binding tetrameric protein. In order to compare the chimeras more directly with the Sav-WT, and with the goal of streamlining the protein production, a chimeric Sav-SOD variant with the DD at the 3,4-loop without the SUMO-tag was used for protein production in soluble titers approaching that of our native Sav construct. Analysis of the purified isolated chimeric Sav-SOD protein using mass spectrometry confirmed expression fidelity and confirmed that no post-translation processing events occurred during the course

128

of expression and purification (Figure 18). Alterations in the protein sequence in the Sav 3,4-loop have been shown to alter the binding properties of biotin to the Sav host in the so-called "Strep-Tactin", "Traptavidin", and CPSav engineered families of proteins.[247] As we have shown, this is not the case for the Sav-SOD construct.[36] Furthermore, analysis of the binding of [Cp*Ir(biot-p-L)Cl] (**1**) within Sav and Sav-SOD using isothermal titration calorimetry (ITC) revealed very similar binding properties (Figure 98). For example, the systems exhibit similar dissociation constants of 8.7 nM (K_{D_Sav}) and 7.5 nM ($K_{D\_Sav\text{-}SOD}$), which are indicative of high binding affinity (Table 35). With the chimeric Sav in hand, we sought to test the utility of the scaffold for development of an artificial transfer hydrogenase ATHase featuring **1** as a cofactor.

Impact of SOD Domain on ATHase Activity. We selected six structurally-related bicyclic asymmetric imines (**2-7**) to afford the corresponding enantioenriched amine products (**8-13**) (Figure 90a).[11,23] The ATHase based on Sav and cofactor **1** (**1** · Sav hereafter, Figure 90b) is well-studied using sodium formate as a hydride source.[11,17,23–25] For this system, the stepwise reaction involves the formation of the Ir^{III}–H by hydride transfer from formate, followed by hydride transfer to the prochiral imine carbon, yielding the enantioenriched amine (Figure 90c). The C-H bond-forming reaction has been proposed to occur via a non-concerted mechanism involving a CH···π interaction between the substrate and the Cp* ligand of the cofactor,[23] and it has been shown that the Sav host can impact substrate turnover and enantioselectivity.[11,16,17,23–27] Most notably, single alanine substitution targeting residues S112 and K121, located within 5–10 Å from the biotin vestibule, have shown to impact significantly ATHase activity.[11,24,28] Thus, we sought to commence our Sav-SOD chimeric host studies with the single alanine mutant targeting the S112 and K121 positions on the Sav. For simplicity, the amino acid numbering for Sav-SOD was kept the same as for Sav, despite the presence of the SOD 34 amino acid insert in the $L^{3,4}$ loop (Figure 89).
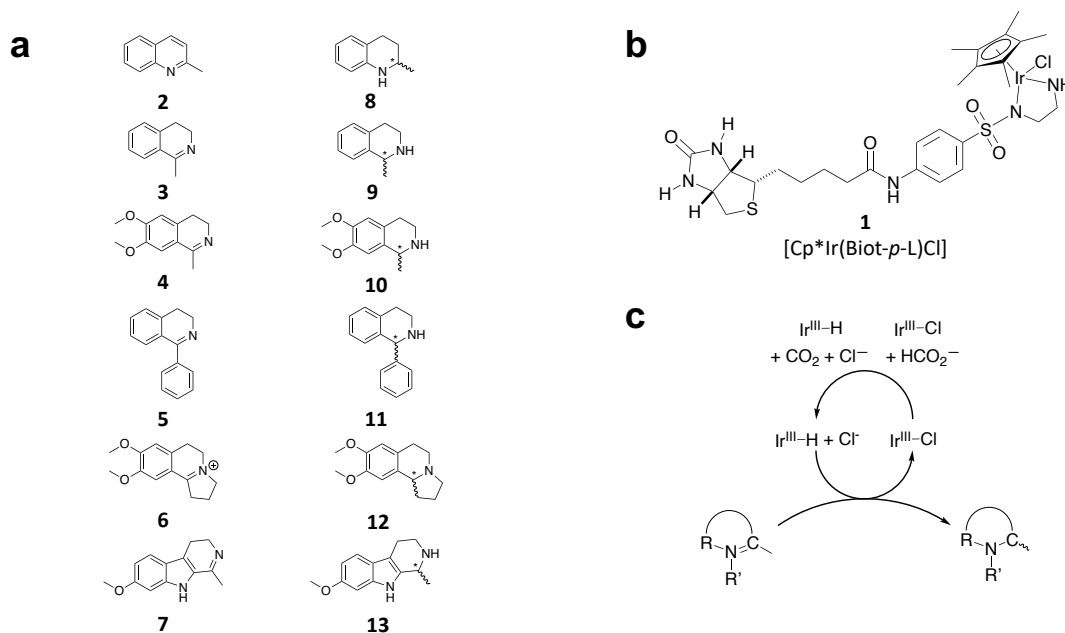
*Figure 90: **a** Structure of the cyclic imine substrates (**2**-**7**) used for the ATHase studies with their corresponding products (**8**-**13**). **b** Illustration of the structure of the cofactor [Cp*Ir(biot-p-L)Cl] (**1**) used in this study. **c** Schematic representation of the transfer hydrogenation mechanism.*
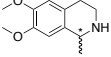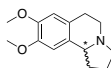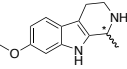
Mass spectrometry was used to confirm the ATHase assembly resulting from the incorporation of **1** within Sav-SOD-S121A and Sav-SOD-K121A, yielding **1**·Sav-SOD-S112A and **1**·Sav-SOD-K121A, respectively. We observed mass ion peaks with ~767 m/z units per equivalent of **1** bound to each tetrameric Sav-SOD host (Figure 93). Both Sav-SOD ATHases effectively reduce prochiral imine substrates in excellent yields in the presence of sodium formate (NaHCO$_2$) as a sacrificial hydride donor, at 37 °C and pH between 6.0 and 7.0. The results from the activity assay using the ATHases, as well as the free cofactor in the absence of a protein host, are collected in .

Table 30. In all cases, the turnover number (TON) achieved with the Sav-SOD ATHases is comparable to or exceeds that achieved with the benchmark Sav ATHases.

For the structurally-related quinoline and dihydroisoquinoline substrates bearing a methyl substituent on the prochiral imine (**2**, **3**, and **4**), the preferred enantiomer of products (**8**, **9**, and **10**) formed by the Sav-SOD ATHases was the same as that formed by the ATHase with the

corresponding Sav scaffold. In contrast, an inversion in the preferred enantiomer was found to occur between the S112A mutants for substrate **5**, which possesses a bulky phenyl moiety at the prochiral carbon. In particular, **1**·Sav-S112A yields primarily the (*R*)-**11** product (64% *ee*) in 78% yield, while **1**·Sav-SOD-S112A generates the (*S*)-**11** product (27% *ee*) in 77% yield. Strikingly, **1**·Sav-K121A and **1**·Sav-SOD-K121A produce the same (*R*)-**11** product with 66% and 58% ee, respectively. Importantly, the conversion after 48 hours achieved by **1**·Sav-SOD-K121A is more than double that achieved by **1**·Sav-K121A (86% vs. 40%).

*Table 30: Selected results for the asymmetric transfer hydrogenation catalyzed by Sav and Sav-SOD based ATHases. The reactions were performed with 10 mM substrate at 37 °C for 48 h (see Supporting Information for details). The conversion is displayed with the enantioselectivity in parentheses. Positive % ee values correspond to the (R)-product and negative % ee values correspond to the (S)-product. For product **12**, the absolute configuration was not determined.*

	Product	No Protein	Sav WT	Sav S112A	Sav K121A	Sav-SOD S112A	Sav-SOD K121A
	8	44 (-4)	49 (-71)	46 (-68)	48 (-46)	47 (-20)	70 (-77)
	9	28 (-5)	63 (41)	74 (20)	93 (28)	84 (62)	100 (59)
	10	82 (-2)	62 (29)	84 (75)	96 (9)	100 (69)	100 (-5)
	11	58 (-1)	80 (58)	78 (64)	40 (66)	77 (-27)	86 (58)
	12	6 (4)	7 (-7)	7 (-5)	7 (2)	21 (-57)	9 (-36)
	13	87 (-2)	87 (-25)	94 (32)	100 (-18)	100 (58)	100 (37)

From previous crystallographic analysis of [Cp*Ir(biot-*p*-L)Cl]·Sav-S112A (pdb 3pk2), combined with QM/MM studies, we suggested that the reduction of **4** proceeds through a non-concerted mechanism involving a CH···π interaction between the substrate and [Cp*Ir(biot-*p*-L)H], with possible involvement of the residue K121 in the protonation step.[23,29,30] However, based on the trends observed for product formation from substrates **2**-**7**, we hypothesize that the DD in the Sav-SOD scaffold renders the ATHase activity and selectivity more sensitive to aminoacid substitutions, thus highlighting the influence of the DD on the positioning of the prochiral substrate. The increased yields obtained with the Sav-

SOD ATHases may result from the more hydrophobic binding vestibule provided by the DD.[18] This is in line with the previous observation that the activity of **1**·Sav was positively affected by the mutation of the cationic residue at position 121 with a hydrophobic residue.[26,31] Distinctions between the Sav and Sav-SOD ATHases are more apparent for the Crispine A precursor (**6**) and harmaline (**7**). The cationic substrate **6** is structurally related to **4** but contains a confined bicyclic iminium moiety. Most notably, **1**·Sav-SOD-S112A produces **12** with 57% *ee*, compared to the modest 5% *ee* observed when using **1**·Sav-S112A. Furthermore, the chimeric streptavidin converted 3 times more substrate than the cofactor alone (6 % conv. vs. 21 % conv.). With the ring-expanded harmaline (**7**), we observe a moderate level of enantioselectivity for the dihydroharmaline product **13** with **1**·Sav-S112A ((*R*)-**13** with 32% *ee*) and **1**·Sav-K121A ((*S*)-**13** with 18% *ee*). However, both Sav-SOD S112A and Sav-SOD K121A based ATHases yielded the same (*R*)-product **13** with 58% and 37% *ee*, respectively. These results support the hypothesis that the DD shielding of the biotin binding vestibule overides the enantioselectivity-preference enforced by the mutations at position S112 or K121.

To support the latter hypothesis, we collected X-ray diffraction data for single crystals of **1**·Sav-SOD-S112A, prepared by co-crystallizing Sav-SOD-S112A with [Cp*Ir(biot-*p*-L)Cl] **1** over the course of 35 days (see Supporting Information for details). The structure was solved with a resolution of 1.8 Å. Residual electron density in the biotin-binding vestibule was apparent from the $2F_o$–F_c map and could be modeled with the [Cp*Ir(biot-*p*-L)] cofactor. Further, the location of iridium was well localized from anomalous dispersion (Figure 91b). Overlaying the structure of the iridium complex in **1**·Sav-S112A (pdb 3pk2)[23] with the structure of the complex in **1**·Sav-SOD-S112A (pdb 7b74) enables a direct comparison of the position of the cofactor and metal center within the protein scaffolds (Figure 91b). In both cases, the ATHase bears an Ir$_{(S)}$-configuration metal center, and the location of the both cofactors are similar (RMSD = 0.260 Å). It should be noted that the dimerization domain in **1**·Sav-SOD-S112A could not be fully resolved due to its high flexibility. Given the similar location of the cofactors in the Sav and Sav-SOD ATHases, we propose that the distinct enantioselectivity of **1**·Sav-SOD-S112A is by-and-large influenced by secondary coordination sphere interactions introduced by the dimerization domain. In order to investigate this possibility further, we used quantum mechanics/molecular mechanics (QM/MM) to identify the most likely transition state involved in the reduction of imine **5** by **1**·Sav-SOD-S112A.
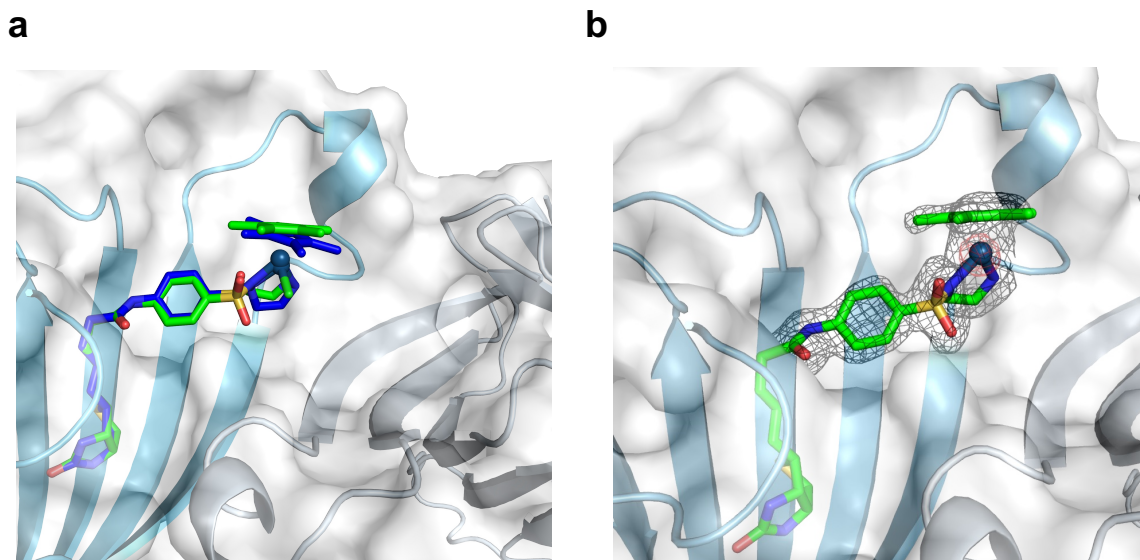
*Figure 91: Superposition of crystal structures of ATHse isoforms **1**·Sav-S112A (pdb 3pk2) and **1**·Sav-SOD-S112A (pdb 7b74). The [Cp*Ir(biot-p-L)Cl] of 3pk2 is represented as blue stick model (atoms are color-coded; nitrogen = blue, oxygen = red, carbon = blue and sulfur = yellow) with the Ir as a sphere. The [Cp*Ir(biot-p-L)Cl] of 7b74 is represented as green stick model (atoms are color-coded; nitrogen = blue, oxygen = red, carbon = green and sulfur = yellow) with the Ir as a sphere. The Sav-SOD-S112A protein is represented as a surface and cartoon model. The monomers are color coded in cyan and gray. The residues 50-83 could not be resolved in the crystal structure, probably due to disorder. **a** Close-up view of the biotin binding vestibule, showing the well-fitting superpostion of the [Cp*Ir(biot-p-L)Cl] . **b** Close-up view of the the biotin binding vestibule of 7b74 with the modeled [Cp*Ir(biot-p-L)Cl] in the residual electron density obtained from the diffraction measurements. The 2Fo-Fc difference map is displayed as dark grey mesh (1σ) and the anomalous electron density is displayed as red mesh (8σ). The occupancy of the Iridium was set to 60%.*

Quantum mechanics/molecular mechanics calculations. As noted above, the SOD lid could not be fully resolved in the crystal structure of **1**·Sav-SOD-S112A. In order to generate a structure for computational analysis, the crystal structure obtained in this study was supplemented with a sampled structure obtained from molecular dynamics MD simulations in our previous study of the Sav-SOD based HAMAse.[18] Only the Ir$_{(S)}$ configuration of [Cp*Ir(biot-*p*-L)H] –corresponding Ir$_{(R)}$- [Cp*Ir(biot-*p*-L)Cl] observed in the X-ray structure– was considered in Sav-SOD-S112A. Indeed, with Ir$_{(R)}$-[Cp*Ir(biot-*p*-L)H], the hydride cannot be delivered to the prochiral imine as it points towards the bottom of the biotin-binding vestibule. Several initial poses were examined for binding of imine **5** in the ATHase, and all of

these converged to four possible transition states (see Figure 100 and Table 31). Of these possible transition states, TS1 and TS3 lead to the (*S*)-product **11**, whereas TS2 and TS4 lead to the (*R*)-product **11**. The calculations revealed that the TS3 conformation is the most stable for **1**·Sav-SOD-S112A (Figure 92a, b). We also performed analogous calculations for **1**·Sav-S112A, and in this case, the cofactor is more solvent-exposed, so that both Ir-configurations of the [Cp*Ir(biot-*p*-L)H] were included in the computation. This leads to four additional possible transition states in addition to TS1-TS4 (see Figure 101 and Table 31). The most favorable transition state for the conversion of **5** by the Sav-S112A ATHase is TS4, which preferentially yields the (*R*)-product **11** (Figure 92c, d). Interestingly, in absence of the SOD lid, substrate **5** can approach so that the bulky phenyl group is oriented towards the solvent, thereby reducing steric hindrance and making this particular conformation more stable than those found in the other transition states. In contrast, the presence of the SOD lid in **1**·Sav-SOD-S112A cause steric hindrance and destabilization of the substrate in TS4. For both **1**·Sav-SOD-S112A and **1**·Sav-S112A, the transition state is positively charged, and the presence of the lysine at position 121 may therefore be destabilizing. On the other hand, in the case of the Sav-SOD ATHase, the SOD lid may stabilize the transition state, given that it contains many negatively-charged amino acid residues. Therefore, the negatively tilted electric field around the catalyst may lower the transition state energy, which may have increased the reaction rate.
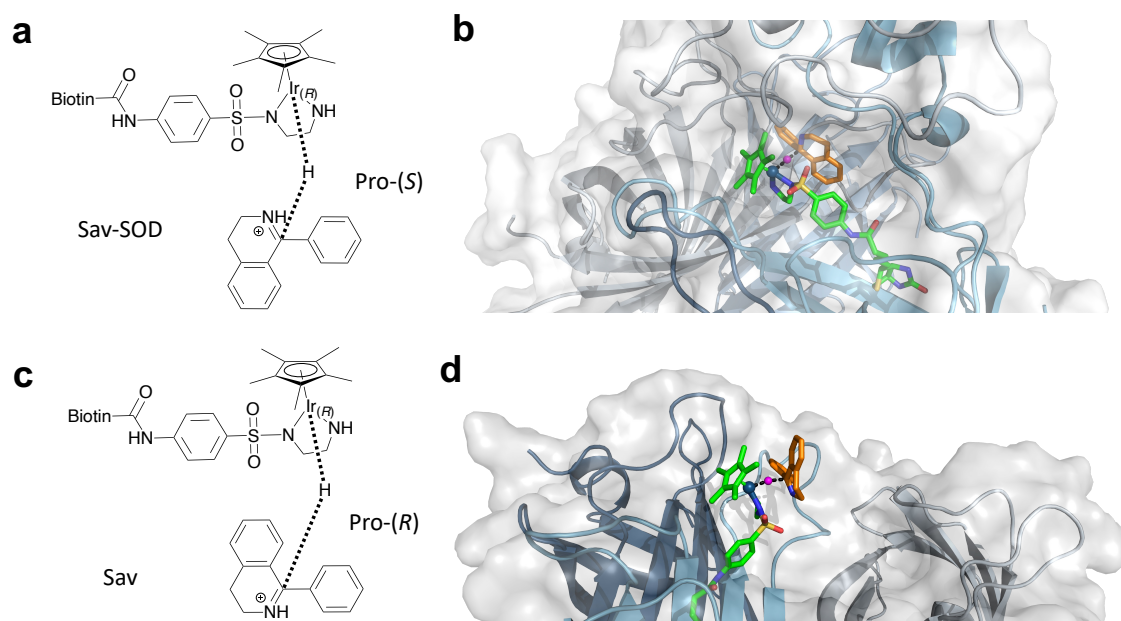
Figure 92: ***a*** *Representation of favored transition state TS3 for the reduction of* ***5*** *by* ***1****·Sav-SOD-S112A.* ***b*** *Modeled transition state TS3 in* ***1****·Sav-SOD-S112A.* ***c*** *Major product formed from TS4.* ***d*** *Proposed favored transition state TS4 for the reduction of* ***5*** *by* ***1****·Sav-S112A. In* ***b*** *and* ***d****, the protein is represented as cartoon with different colors for each Sav monomer. The solvent-accessible surface of the protein is represented in transparent white. Cofactor* ***1*** *(green) and substrate* ***5*** *(orange) are represented as a stick model (atoms are color coded; N = blue, O = red, C = green and S = yellow). The hydrogen atom is displayed as a purple sphere.*

Table 31 Calculated major products and energies differences of the predicted possible transition states for reduction of substrate ***5*** *by* ***1****·Sav-SOD-S112A or* ***1****·Sav-S112A to the lowest energy transition state.*

Transition state	Major product	Energy of transition state (kcal/mol)	
		Sav-SOD	Sav
TS1	(S)	27.13	13.16
TS2	(R)	32.47	25.21
TS3	(S)	0.00	17.29
TS4	(R)	4.65	0.00
TSi	(S)	-	12.17
TSii	(R)	-	1.46
TSiii	(R)	-	21.16
TSiv	(S)	-	14.82

Kinetics of Imine reduction using PDQ. The marked differences between the Sav and Sav-SOD ATHase products for substrate **5** warranted further kinetic investigation. Thus, we performed a Michaelis-Menten kinetic study in 300 mM MOPS buffer (pH = 6.0) with fixed concentrations of sodium formate (2 M). As determined previously for Sav-based ATHases,[11,24,32] we observed imine saturation kinetic profiles. A summary of these data is presented in Table 32 and Figure 95-96.

Substrate-dependent ATHase activity profiles were obtained at 25 and 37 °C. At 37 °C, the Sav-SOD ATHase systems exhibited approximately a doubled turnover number (k_{cat}) those of their Sav-based counterparts, and the Michaelis constant (K_M) for the Sav-SOD ATHases were approximately 2.5 times those of the parent Sav ATHases. At 25 °C, the catalytic activity for **1**·Sav-SOD-K121A and **1**·Sav-K121A were equivalent, with a k_{cat} values of 3.92 min^{-1} and 3.61 min^{-1}, respectively. However, at 25 °C, **1**·Sav-SOD-S121A maintains a high activity of 11.71 min^{-1} that is approximately five times higher than that of **1**·Sav-S121A.

*Table 32: Results from the saturation kinetic experiments with the substrate **5**. The PDQ-dependent kinetic profiles were determined at 25 and 37 °C with a total reaction volume of 400 μL and a fixed concentration for streptavidin (Sav) or Sav-SOD (30 μM), cofactor (15 μM), MOPS (300 mM) and sodium formate (2 M). The reaction was initiated upon addition of the substrate stock solution yielding a final substrate concentration ranging from 5–50 mM. Aliquots (50 μL) of the reaction were collected at 10, 20, 30, 45, 60, 90 min time intervals, quenched and analyzed by GC-MS. Full details are collected in the Supporting Information.*

	ATHase	k_{cat} (min^{-1})	K_M (mM)	k_{cat}/K_M (min^{-1} mM^{-1})
	Sav S112A	2.13	6.02	0.35
	Sav K121A	3.93	9.51	0.41
25 °C	Sav-SOD S112A	11.71	14.19	0.83
	Sav-SOD K112A	3.61	4.10	0.88
	Sav S112A	6.44	15.15	0.43
	Sav K121A	7.88	10.64	0.74
37°C	Sav-SOD S112A	13.83	38.96	0.35
	Sav-SOD K112A	16.00	29.20	0.55

We next sought to investigate whether the rate acceleration observed for the Sav-SOD ATHases results from greater efficiency in Ir–H complex formation or in sequential imine reduction. We therefore collected additional ATHase activity profiles with varying concentrations of the sacrificial hydride donor sodium formate from 0.25 to 4.0 M. We performed these experiments at 25 °C with an initial concentration of 50 mM for the imine substrate **5** to ensure all ATHases were operating above their respective K_M values. The kinetic results of this study are presented in Figure 97. We found that the K_M of sodium formate is significantly lower for **1**·Sav-SOD-S112A than **1**·Sav-S112A (approximately 0.78 M vs. >2.74 M), which might be a result of interactions with the DD, facilitating the formation of the active reductant [Cp*Ir(biot-*p*-L)H]. Furthermore, in the case of **1**·Sav-S112A, we found that the activity decreases with increasing formate concentration, although this inhibition is not observed for **1**·Sav-SOD-S112A. The trends in k_{cat} are less clear. In the case of **1**·Sav-S112A, a substrate inhibition seems to occur. Due to the limited solubility of sodium formate in aqueous solution, it was not possible to reach the saturation points for either Sav or Sav-SOD. However, the catalytic efficiency (k_{cat}/K_M) was found to be higher for the Sav-SOD ATHases (17.71 vs 1.87 $M^{-1}min^{-1}$ and 6.70 vs. 2.83 $M^{-1}min^{-1}$ for Sav-SOD-S112A vs. Sav-S112A and Sav-SOD-K121A vs. Sav-K121A respectively). The iridium–hydride forming reaction was found to be 5.5-fold more efficient for **1**·Sav-SOD-S112A than **1**·Sav-S112A (11.71 vs. 2.13 $M^{-1}min^{-1}$) and about the same (3.61 vs. 3.93 $M^{-1}min^{-1}$) for **1**·Sav-SOD-K121A and **1**·Sav-K121A. Therefore, we propose that the increased yields observed for the Sav-SOD constructs may be in part due to more efficient formation of the hydride catalyst.

Conclusion. In summary, we have tested the chimeric streptavidin Sav-SOD as a scaffold for an ATHase incorporating [Cp*Ir(biot-*p*-L)Cl] **1** as a cofactor. The dimerization domain presents a secondary coordination environment that is inherently hydrophobic and capable of shielding tethered biotinylated cofactors from the aqueous milieu, akin to natural enzymatic systems. Our investigation of Sav-SOD based ATHases has shown that, depending on the nature of the substrate, the dimerization domain can alter the catalytic efficiency and selectivity of iridium-catalyzed imine transfer hydrogenation reactions. In particular, in the case of the substrate 1-phenyl-3,4-dihydroisoquinoline (**5**), **1**·Sav-SOD and **1**·Sav ATHases exhibit different enantioselectivities. Single-crystal X-ray diffraction data revealed that the

137

configuration of the cofactor in both systems, **1**·Sav-S112A and **1**·Sav-SOD-S112A, is the same, although both ATHases afford opposite enantiomers of amine **11.** QM/MM calculations suggest that the transition states formed during reduction differ for the two ATHAses, **1**·Sav-S112A and **1**·Sav-SOD-S112A, providing a rationale for their differing enantioselectivity. Further catalytic experiments will need to be performed in order to fully characterize the reactivity of the Sav-SOD ATHases. However, this study further highlights the versatility of the Sav-SOD scaffold and the power of directed evolution for tuning and optimizing the selectivity and reactivity of ArMs.[33]

4.3.4 Experimental

General information. Chemicals were purchased from Sigma Aldrich, Acros Organics, Alfa Aesar or Fluorochem and used without further purification. Water used for molecular biology and in the catalytic reactions was purified by Milli-Q Advantage system. All catalytic reactions were carried out with non-degassed solvents under air. Temperature was maintained using Thermowatch-controlled heating blocks. Gas chromatography–mass spectrometry (GC-MS) analysis was run on a Shimadzu GCMS-QP2020. A normal phase HPLC instrument from Agilent with a Chiralpak IB column (5 μm, 4.6 mm × 250 mm, Daicel) and a UPC2 system (Waters) was used to analyze the samples. High-resolution mass spectrometry (HR-MS) was performed on a Bruker maXis II QTOF ESI mass spectrometer coupled to a Shimadzu LC. The transfer hydrogenation co-factor [Cp*Ir(biot-p-L)Cl] was synthesized as previously reported.[40] Imine reduction yields were determined using an internal standard or substrate to production response factor. Chiral amine retention times and enantiomeric excess (*ee*) yields were determined using authentic and racemic standards. Molecular biology reagents were purchased from New England Biolabs (NEB), Integrated DNA Technologies (IDT), and Macherey-Nagel using accompanying protocols. Protein concentrations were determined by the calculated A_{280} molar absorptivity coefficients using the ProtParam tool (ExPASy) of 41940 M^{-1} cm^{-1} and 43430 M^{-1} cm^{-1} for wild type Sav and Sav-SOD respectively.

Cloning and expression of Sav and Sav-SOD mutants. The Sav-SOD chimera gene was synthesized and cloned into the NcoI and EagII restriction sites of the pET28 vector by Gene

Universal Inc. (Newark, DE). Site directed mutagenesis was achieved using the primers listed in Table 33, followed by BsaI/DpnI digestion and ligated using T4 ligase. Mutations were verified by Sanger sequencing performed by Microsynth (Balgach, Switzerland). The expression and purification of Sav and Sav-SOD proteins were achieved using auto induction media (ZYP-5052) followed by cell lysis and purification using iminobiotin agarose affinity chromatography as previously described.[105]

Table 33 Primers used for mutagenesis.

Oligo	Sequence	
oRP222A	CATAT*ggtctc*TTTGCTTCAGTAGTGCC<u>GGC</u>GGTCAGCAGCCACTGGGTATT	S112A
oRP225S	CATAT*ggtctc*TTTGCTTCAGTAGTGCC<u>GCT</u>GGTCAGCAGCCACTGGGTATT	S112S (WT)
oRP226K	TATAA*ggtctc*AGCAAATGCCTGG<u>AAA</u>AGCACCCTGGT	K121K (WT)
oRP231A	TATAA*ggtctc*AGCAAATGCCTGG<u>GCA</u>AGCACCCTGGT	K121A

ATHase Kinetic Measurements. The 1-phenyl-3,4-dihydroisoquinoline (PDQ)-dependent kinetic profiles were determined at 25 and 37 °C, with a total reaction volume of 400 μL. The final concentrations of the reactions were as follows: streptavidin (Sav) or Sav-SOD (30 μM), cofactor (15 μM), MOPS buffer (300 mM) and sodium formate (2 M). The *in situ* assembly of the ATHase was achieved by adding a solution of [Cp*Ir(biot-*p*-L)Cl] in *N,N*-dimethylformamide (DMF) (4 mM) to a solution of the scaffold protein for 10 min at room temperature. The volume was adjusted to 160 μL with Mili-Q water and 200 μL of a of MOPS (300 mM) and sodium formate (2 M) buffer solution was added. The reaction was initiated upon addition of a 10× substrate stock solution (40 μL) yielding a final substrate concentration ranging from 5–50 mM. An aliquot (50 μL) of the reaction was collected at 10, 20, 30, 45, 60, 90 min time intervals and quenched by the addition of 40 μL freshly prepared 0.2 M reduced glutathione. Reaction aliquots were diluted with Mili-Q water (200 μL) and made basic with the addition of 50 μL of 20% wt/vol aqueous NaOH. The organic products were extracted with dichloromethane (600 μl), dried with anhydrous sodium sulfate and analyzed by GC-MS. Formate dependent reaction kinetics were performed using the same procedure using 200 μL of MOPS (300 mM) buffer stock solutions (pH = 6.0) with varying sodium formate concentrations. The reaction was initiated upon the addition of 40 μL 0.5 M PDQ.

Protein ESI-MS determination. Proteins were dissolved in Mili-Q water, 1% formic acid (pH = 2.5) with a final concentration of 0.2 mg/mL and clarified by centrifugation. A HPLC (Shimadzu, equipped with a Jupiter® 5 μm C4 300 Å)-ESI QTof (Bruker Daltonics, ESI MaxisII QTof MS) system was used to record the data. The ESI-QTof mass spectrometer was calibrated with ESI-Tof TuneMix (Agilent). The charge envelope from 2800–4500 m/z for the Sav-SOD and from 2500–4000 m/z for the Sav, was deconvoluted using the Compass DataAnalysis software (Bruker Daltonics) with the Maximum Entropy setup. For samples with [Cp*Ir(biot-*p*-L)Cl], the proteins were dissolved in Mili-Q water with a concentration of 0.5 mg/mL and 4 equiv. of the cofactor. After 15 min, the samples were diluted with Mili-Q water, 1% of formic acid (pH = 2.5) to a final protein concentration of 0.2 mg/mL and clarified by centrifugation. The charge envelope from 2800–4500 m/z was deconvoluted using the Compass DataAnalysis software (Bruker Daltonics) with the Maximum Entropy setup.
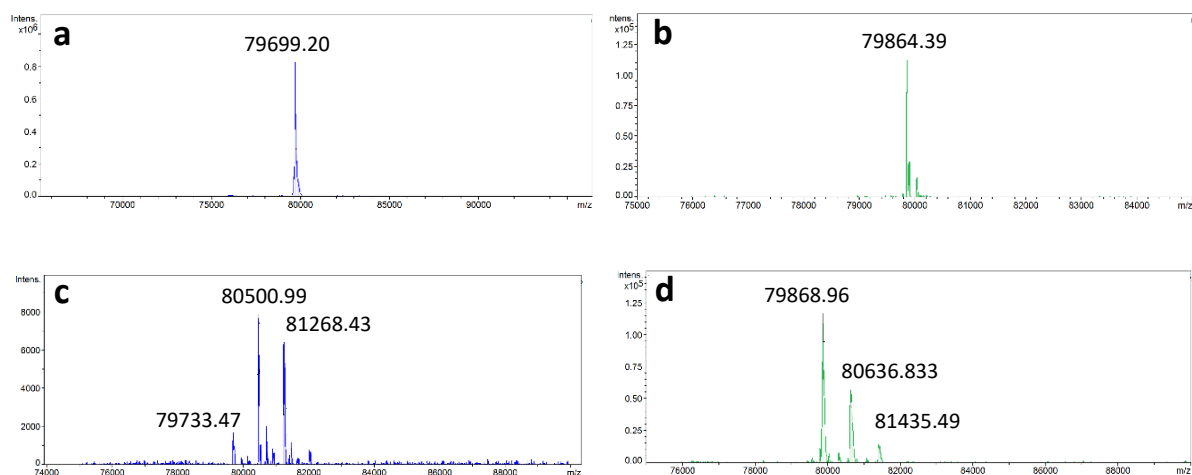
Figure 93: Exact mass of Sav-SOD chimeras and ArMs. **a** *Deconvoluted mass spectra of Sav-SOD-K121A chimera.* **b** *Deconvoluted mass spectra of **1**·Sav-SOD-K121A.* **c** *deconvoluted mass spectra of Sav-SOD-S112A chimera.* **d** *Deconvoluted mass spectra of **1**·Sav-SOD-S112A.*

Single-crystal X-ray diffraction analysis of [Cp*Ir(biot-*p*-L)Cl]·Sav-SOD-S112A. Protein crystals were obtained using a sitting drop vapor diffusion experiment. The lyophilized protein mutant (Sav-SOD-S112A) was dissolved in a 20 mM Tris-HCl solution (pH = 7.0) to a concentration of 2.5 mg/mL. A 10 mM solution of [Cp*Ir(biot-*p*-L)Cl] in dimethylsulfoxide (DMSO) (5 μL) was added to the solution and the mixture was incubated at room temperature overnight. After concentration to 25 mg/mL by ultracentrifugation (Amicon® ultra centrifugal

filters, Merck; cut-off of 10 kDa), crystals were grown by sitting drop vapor diffusion using 20 %w/v PEG 3350, 0.2 M KF as crystallization buffer. Crystals grew after 30 days and were flash frozen in liquid nitrogen without further cryoprotection. Protein crystal diffraction data was collected at 100 K at the Swiss Light Source beam line PXI with a wavelength of 1.0 Å and analyzed with CCP4i2 Suite.[110] Crystal indexing, integration and scaling were carried out with XDS and reflections were merged with AIMLESS. PHASER MR was used to solve the structure by molecular replacement with pdb 2bc3 as input model. For structure modeling, water picking and electron density visualization the software COOT was used. Amino acid residues 1-14, 51-85 and 172-196 are not resolved in the electron density, presumably due to disorder. Residual electron density was observed the biotin binding pocket and in the biotin vestibule from the F_o–F_c map. Furthermore, anomalous dispersion was observed in the vestibule. Modeling of cofactor [Cp*Ir(biot-*p*-L)Cl] into the electron density projected the iridium in the position of the anomalous density peak (Figure 94).

a **b**
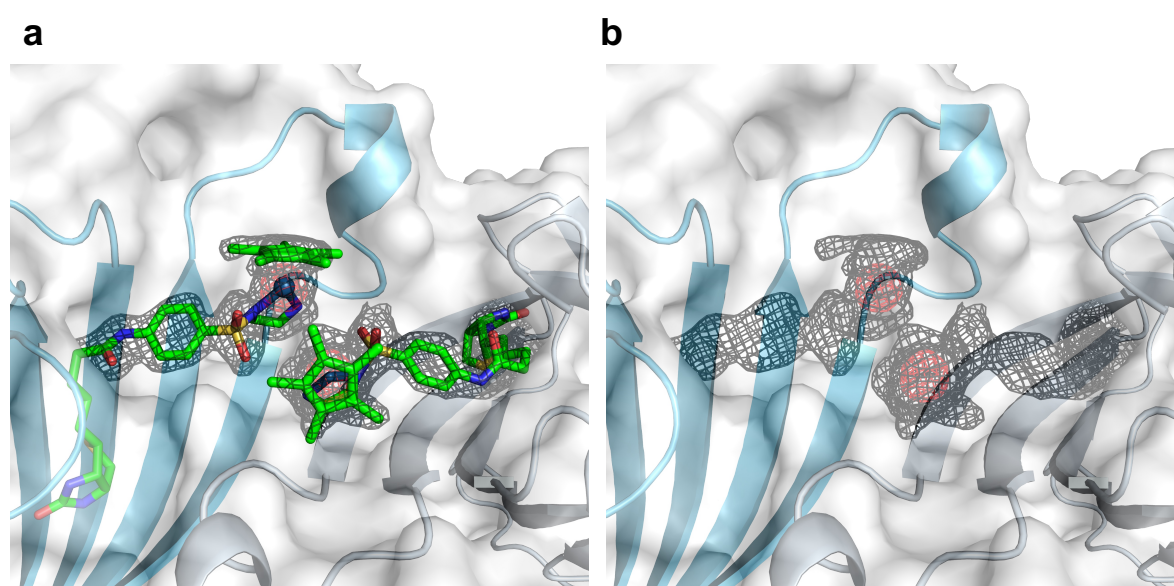
*Figure 94 Single-crystal X-ray diffraction structure of [Cp*Ir(biot-p-L)Cl]·Sav-SOD-S112A highlighting the biotin binding site of two facing monomers of Sav-SOD-S112A with **a** and without **b** [Cp*Ir(biot-p-L)Cl] modeled as a stick model, with the iridium represented as a sphere. The 2F$_o$–F$_c$ map is shown as a grey mesh with 1.0 σ. The anomalous density obtained from the measurement is shown in red (5.0 σ).*

Table 34 Data processing and crystal structure refinement statistics.

Data collection	[Cp*Ir(biot-p-L)Cl]·Sav-SOD S112A
Space group	P 1 2$_1$ 1
Cell dimension	
a, b, c (Å)	57.28, 57.31, 88.03
α, β, γ (°)	90, 94.62, 90
Resolution (Å)	47.98–1.0
R_{merge}	15.9 (127.0)
I/σ(I)	9.2 (1.56)
CC$_{1/2}$	1.00 (0.55)
Completeness (%)	98.3 (98.0)
Refinement	
Resolution (Å)	47.98–1.85
No. reflections	15314
R_{work}/R_{free}	20.2 / 22.6

Transfer hydrogenation reactions for the determination of the product yield. To test the reduction of the prochiral imines with the Sav-SOD-based ATHases, reactions were run for 24 hs at 37 °C. The following stock solutions were prepared: (i) Substrates (400–500 mM in DMSO, ethyl acetate or H$_2$O); (ii) [Cp*Ir(biot-p-L)Cl] (4 mM in DMF); (iii) protein scaffold (400 μM in H$_2$O) and (iv) MOPS-formate buffer (600 mM MOPS with 4 M sodium formate, pH = 6.0). These solutions were added in the following order and concentrations to a HPLC glass vial (total reaction volume 200 μL): 1) [Cp*Ir(biot-p-L)Cl] 50 μM, 2) Sav 100 μM, 3) buffer 300 mM, 2 M and 4) substrate 10 mM. Solutions 1)-3) were added sequentially, and incubated for 10 min at room temperature for assembly of the ATHase. The reaction was initiated upon the addition of 4) and run in an Eppendorf thermocycler (37 °C, 1000 rpm) for 24 h. The reactions were quenched by the addition of 20 % w/v NaOH (50 μL), at which point the internal standard was added to the reaction vessel and extracted with 600 μL ethyl acetate. The organic phase was collected and analyzed.

Chiral amine analysis. The conversions and enantiomeric excesses were carried out either on an Agilent HPLC or an ACQUITY UPC2 Waters system using chiral columns and running conditions as detailed below.

Analysis of **2**, (*S*)-**8** and (*R*)-**8** was performed with a Chiralpak IC column (5 μm, 4.6 mm × 250 mm, Daicel) using a constant flow rate (2.5 m/min). Biphenyl was used as internal standard. The following program was set up for analysis: 98% CO_2 and 2% IPA (0.1% DEA) from 0–4 min; linear increase to 80% CO_2 and 20% IPA (0.1% DEA) from 4–6 min; linear decrease to 98% CO_2 and 2% IPA (0.1% DEA) from 6–8 min; keep at 98% CO_2 and 2% IPA (0.1% DEA) from 8–10 min. Retention times: 2.3 min for biphenyl, 3.7 min for (*R*)-**8**, 3.9 min for (*S*)-**8**, 7.3 min for **2**. The compounds were quantified based on the relative substrate areas and internal standard area (absorbance at 254 nm) using a response factor of 0.5023.

Analysis of **3**, (*S*)-**9** and (*R*)-**9** was performed with a Chiralpak IC column (5 μm, 4.6 mm × 250 mm, Daicel) using constant 90% CO_2 and 10% IPA (0.1% DEA) at a constant flow rate (2.5 mL/min) on a Waters UPC2. 1-Methoxynaphthalene was used as internal standard. The retention times were 2.1 min for 1-methoxynaphthalene, 5.6 min for **3**, 9.3 min for (*R*)-**9**, 10.2 min for (*S*)-**9**. The compounds were quantified based on the relative substrate peak areas and the internal standard area (absorbance at 210 nm) using a response factor of 0.1213.

Analysis of **4**, (*S*)-**10** and (*R*)-**10** was performed using a Chiralpak ID column (5 μm, 4.6 mm × 250 mm, Daicel) at a constant flow rate (2.5 mL/min) UPC2. Biphenyl was used as an internal standard. The program was set up as follows: 85% CO_2 and 15% IPA (with 0.3% diethylamine, DEA) from 0–1 min; then linear increase to 75% CO_2 and 25% IPA (0.3% DEA) from 1–10 min; then linear decrease to 85% CO_2 and 15% IPA (0.3% DEA) from 10-12 min; keep at 85% CO_2 and 15% IPA (0.3% DEA) from 12–14 min. Retention times were 1.5 min biphenyl, 6.2 min for **4**, 8.6 min for (*R*)-**10**, 9.5 min for (*S*)-**10**. The compounds were quantified based on the relative substrate and product peak areas (absorbance at 280 nm) using a response factor of 2.9108.

Analysis of **5**, (*S*)-**11** and (*R*)-**11** was performed with a Chiralpak IA column (5 μm, 4.6 mm × 250 mm, Daicel) by UPC2. Biphenyl was used as internal standard. The program was set up with 80% CO_2 and 20% IPA (0.1% DEA) at a constant flow rate (2.5 mL/min). The retention times were 1.7 min for biphenyl, 2.4 min for **5**, 5.0 min for (*R*)-**11**, 8.1 min for (*S*)-**11**. The compounds were quantified based on the relative substrate peak areas and the internal standard area (absorbance at 210 nm) using a response factor of 1.2368.

Analysis of **6**, (*S*)-**12** and (*R*)-**12** was performed with a a Chiralpak ID column (5 μm, 4.6 mm × 250 mm, Daicel) by UPC2. 2-methylindoline was used as internal standard. The program was set up as follows: 80% CO_2 and 20% IPA (with 0.3% diethylamine, DEA) at a constant flow rate (2.5 mL/min). Retention times: 1.77 min for 2-methylindoline, 7.9 min for (*1*)-**12** and 8.9 for (*2*)-**12**. The compounds were quantified based on the relative peak areas and the internal standard area (absorance at 286 nm) using a response factor of 0.7281.

Analysis of **7**, (*S*)-**13** and (*R*)-**13** was performed with a Chiralpak ID column (5 μm, 4.6 mm × 250 mm, Daicel) by UPC2. 2-methylindoline was used as internal standard. The program was set up as follows: 85% CO_2 and 15% IPA (with 0.3% diethylamine, DEA) from 0–1.5 min; then linear increase to 63% CO_2 and 37% IPA (0.3% DEA) from 1.5–2.3 min; this ratio was kept to 6 min; then the ratio was linear decrease to 85% CO_2 and 15% IPA (0.3% DEA) from 6–6.2 min; keep at 85% CO_2 and 15% IPA (0.3% DEA) from 6.2–6.5 min. Retention times: 2.0 min for 2-methylindoline, 4.6 min for **7**, 5.4 min for (*R*)-**13** and 5.8 for (*S*)-**13**. The compounds were quantified based on the relative substrate peak areas and the internal standard area (absorbance at 293 nm) using a response factor of 6.1247.

Kinetic measurements analysis. GC-MS measurements for reaction kinetics on substrate **5** (PDQ) were carried out on a Shimadzu, GCMS-QP2020 equipped with an Agilent HP1-1MS column (length: 30 m; Diameter: 0.25 mm; Film: 0.25 μM). A single acquisition mode was used with m/z values of 175–180. The product was quantified based on the relative product peak area and the substrate area using a response factor of 1.19.
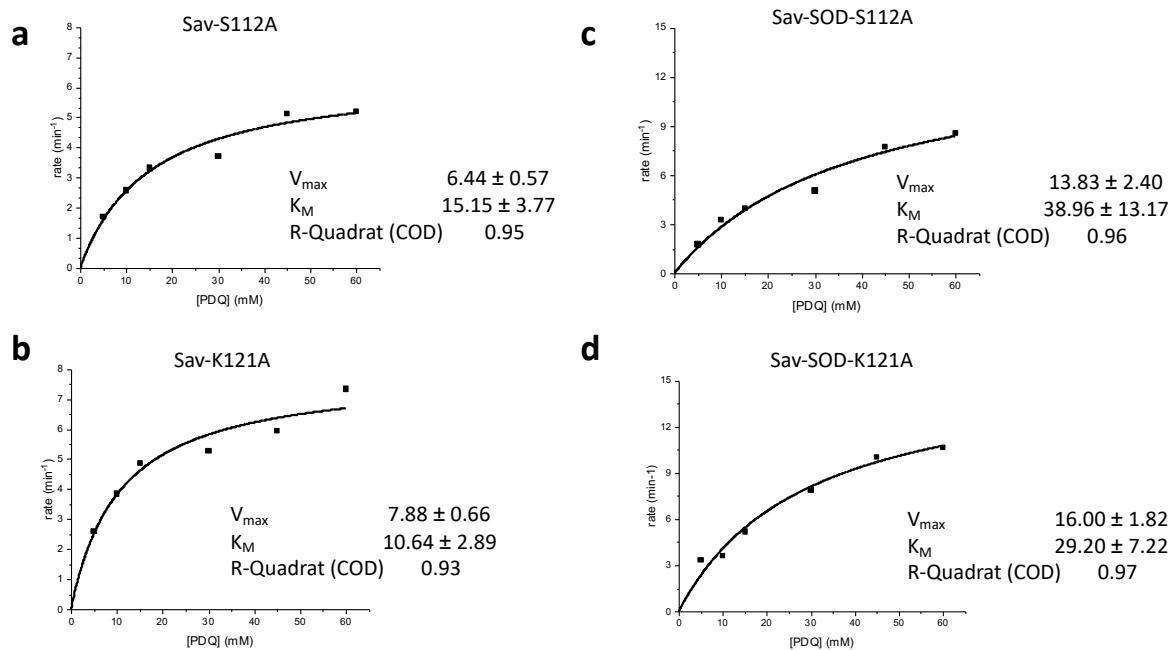
Figure 95 Michaelis-Menten kinetic profiles for the reduction of 5 (PDQ) in 2 M sodium formate buffer pH at 37 °C.
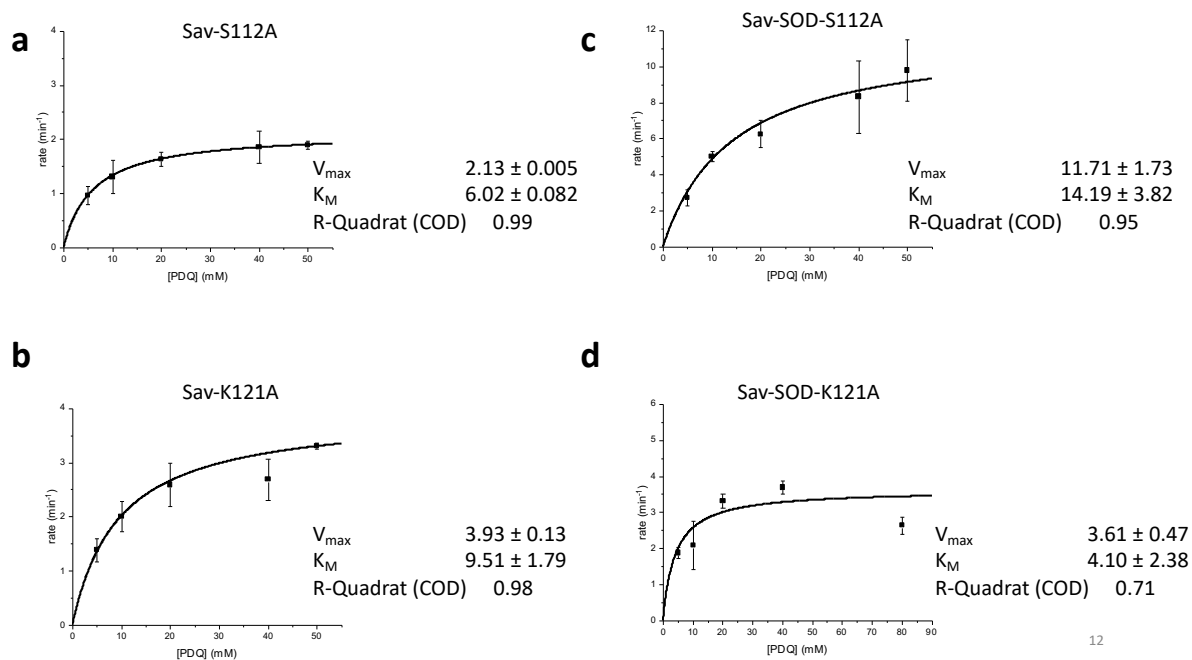
Figure 96 Michaelis-Menten kinetic profiles for the reduction of 5 (PDQ) in 2 M sodium formate buffer pH at 25 °C

145
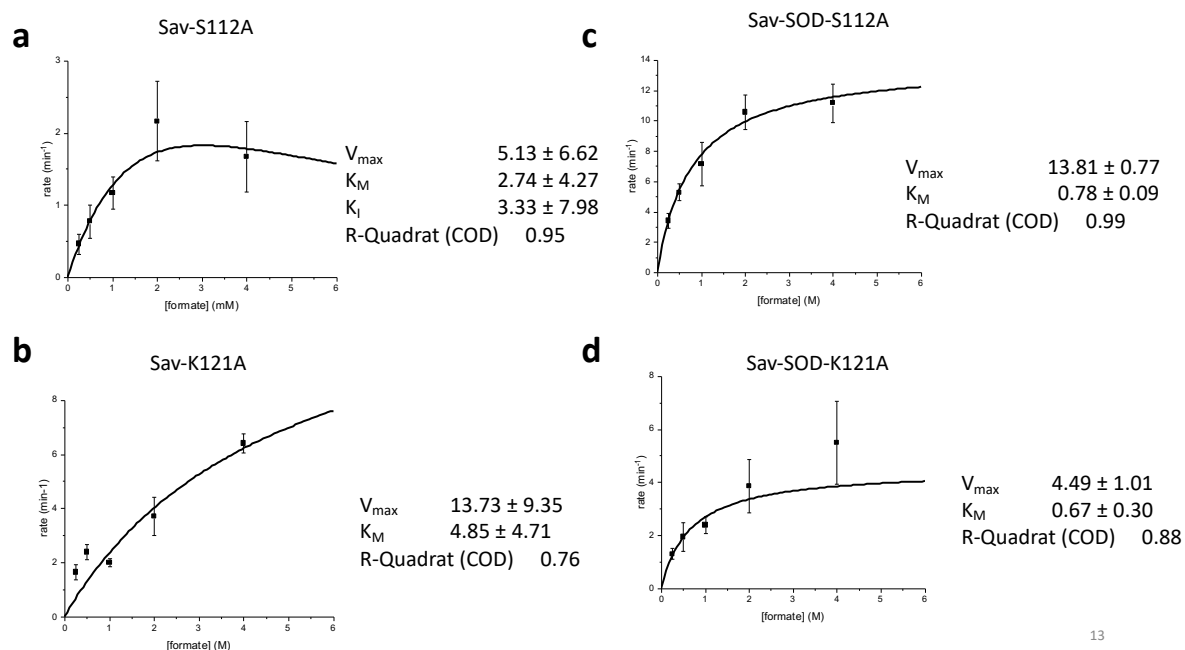
Figure 97 Michaelis-Menten kinetic profiles for the reduction of 5 (PDQ) (50 mM) with varying concentrations of sodium formate at 25 °C.

ITC analysis. Biotin binding affinities were measured using a Microcal VP-ITC as previously described by Stayton and Coworkers.[181] A 10 μM solution of [Cp*Ir(biot-*p*-L)Cl] in 50 mM sodium phosphate buffer containing 100 mM sodium chloride at pH = 7.75 with 2.5 % DMSO was titrated with 5 μL injections of Sav-S112A or Sav-SOD-S112A solutions (100 μM, 50 mM sodium phosphate, 100 mM sodium chloride, 2.5% DMSO, pH = 7.75). The reference cell contained the same buffer as the protein and cofactor solution. Measurements were performed at 25 °C. The [Cp*Ir(biot-*p*-L)Cl] binding constant K_a, enthalpy h, and number of binding sites at each temperature were calculated using ITC data analysis origin software (MicroCal).

*Table 35 Thermodynamic parameters for [Cp*Ir(biot-p-L)Cl] binding to Sav and Sav-SOD mutant proteins determined in this work. All data were collected at 25 °C. Standard deviations are given in parentheses.*

Protein	K_a (M^{-1})	ΔG° (kcal/mol)	ΔH° (Kcal/mol)	TΔS° (Kcal/mol)
Sav-S112A	$1.15(0.20) \times 10^8$	−14.8	−14.5(0.6)	−0.3
Sav-SOD-S112A	$1.27(0.31) \times 10^8$	−13.8	−13.6(0.9)	−0.2
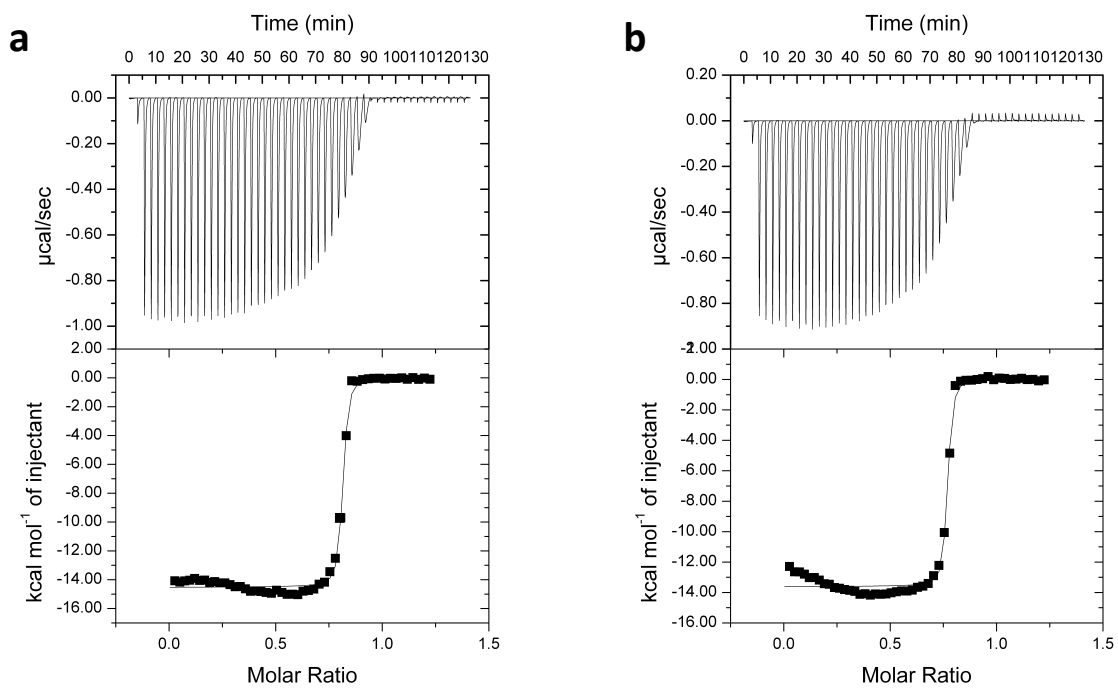
*Figure 98: ITC analysis of **a** [Cp*Ir(biot-p-L)Cl]·Sav-S112A and **b** [Cp*Ir(biot-p-L)Cl]·Sav-SOD-S112A .*

4.3.5 Appendix

In the case of the ATHase, the Michaelis Menten saturation kinetics is a bit more challenging, as two substrates (hydride (A) and imine (B)) react to give two products (CO_2 (P_1) and amine (P_2)). Therefore, this reaction might not be applicable *per se* to the Michaelis-Menten mechanism. Nevertheless, the reaction proceeds most probably via a so-called "ping pong" mechanism, where the first substrate (formate (A)) binds to the enzyme and is then released as the first product (CO_2 (P_1)), leaving a hydride bound to the enzyme. In a second step, the modified enzyme (E_A) reacts with the second substrate (imine (B)) to yield the second product (amine (P_2)). The mechanism can therefore be illustrated as follows:

$$E + A \underset{}{\overset{k_{1A}}{\rightleftharpoons}} EA \overset{k_2}{\longrightarrow} E_A + P_1 + B \underset{}{\overset{k_{1B}}{\rightleftharpoons}} E_AB \overset{k_3}{\longrightarrow} E + P_2$$
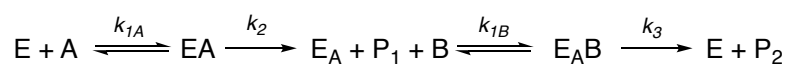
The rate equilibrium for this mechanism is:

$$v = \frac{V_{max}\,[A][B]}{K_A\,[B] + K_B[A] + [A][B]}$$

where V_{max} is the maximum velocity and K_a and K_b are the limiting Michaelis constants for A and B that are seen at saturation with B or A, respectively. If the reaction indeed follows the "ping pong" mechanism, the Lineveawer-Burk plot would look like that displayed in Figure 99. In a double-reciprocal plot, the reciprocal initial velocity is plotted against the reciprocal substrate concentration (in this case, the PDQ concentration). A linear regression analysis can then be carried out in order to obtain the negative reciprocal value for the K_M as the *x*-axis intercept, the reciprocal value for the k_{cat} as the *y*-axis intercept and the catalytic activity (k_{cat}/K_M) as the slope. However, more importantly, by varying the concentration of the second substrate (in our case, formate) and measuring the initial velocity at different concentrations of the first substrate (PDQ), we would obtain the same slope but different K_M and k_{cat} values. Therefore, by using the Lineveawer-Burk plot and the right experimental conditions, we can gain more information about the reaction mechanism of the ATHase based on Sav and Sav-SOD.

148
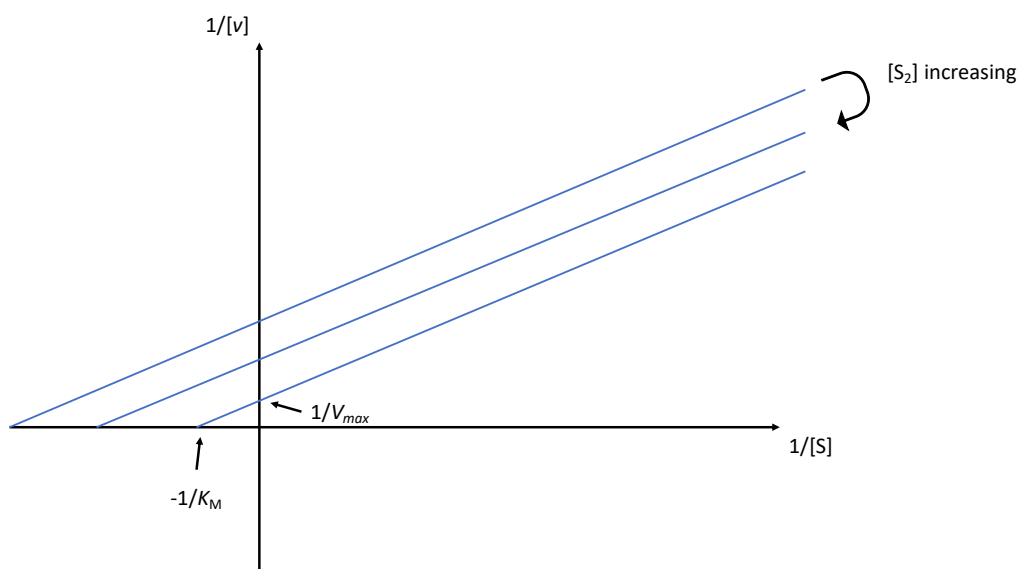
Figure 99: The characteristic parallel-lined reciprocal plots (Lineweaver-Burk) of "ping-pong" kinetics. As the concentration of the second substrate increases, V_{max} increases, as well as does the K_M for the first substrate.

One further aspect which was inspected previously by the group is the influence of the ratio of [Cp*Ir(biot-*p*-L)Cl] and the biotin binding sites. This ratio was fixed at 1:2 in order to minimize the full occupation of the binding sites, with the goal of minimizing possible steric repulsion between cofactors that might lead to a decreased enantioselectivity. In order to further optimize the reaction conditions, a screening of the catalyst:protein ratio would be beneficial.
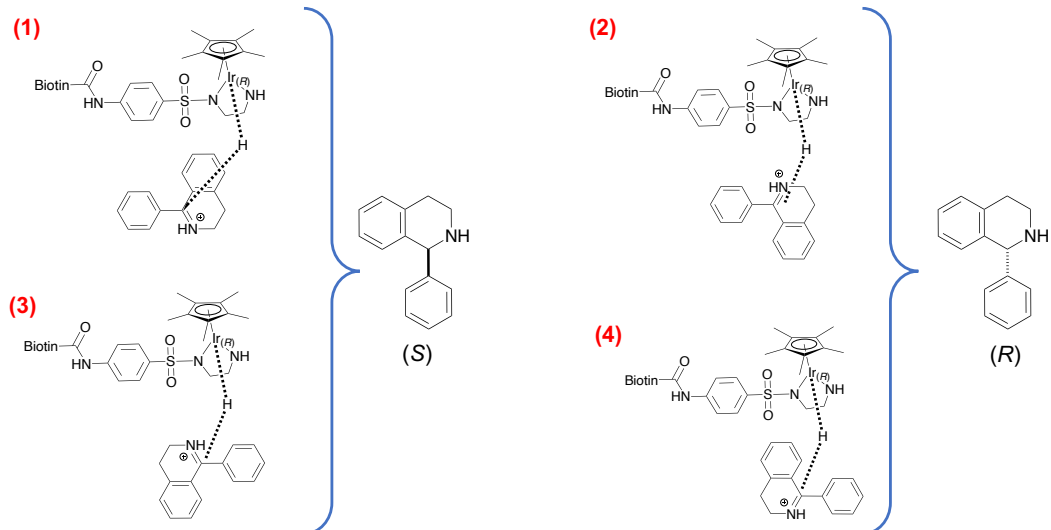
Figure 100 Possible transition states for the ATHase studied.
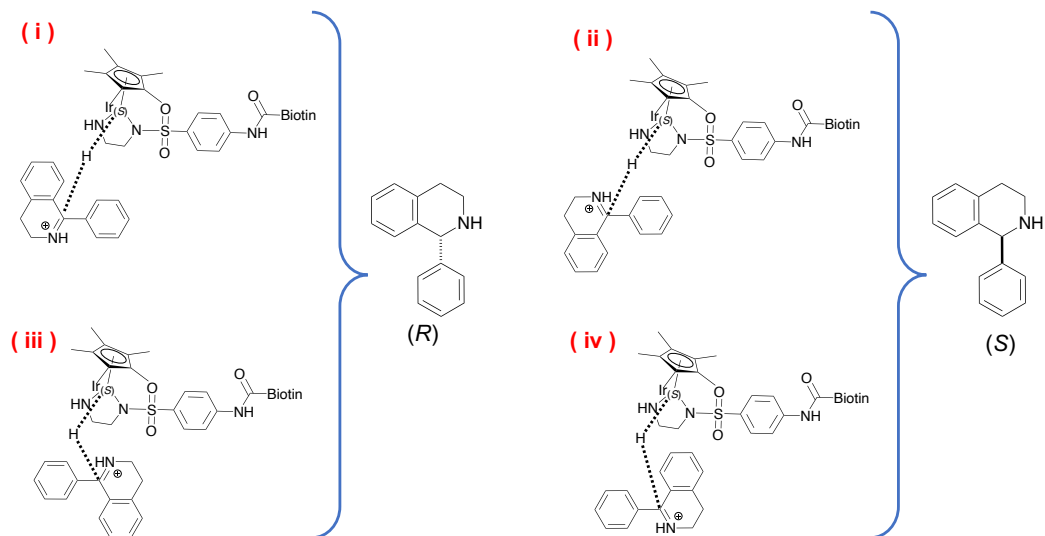
*Figure 101 Additional transition states for the Sav based ATHase studied with substrate **5**.*

150
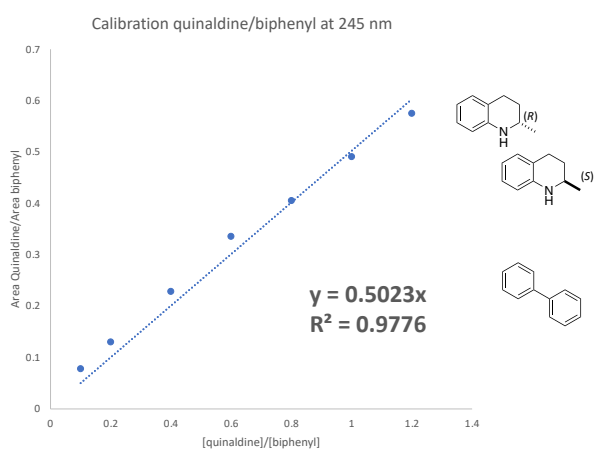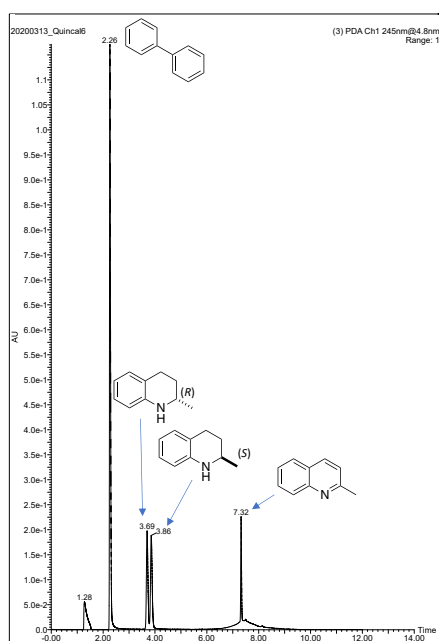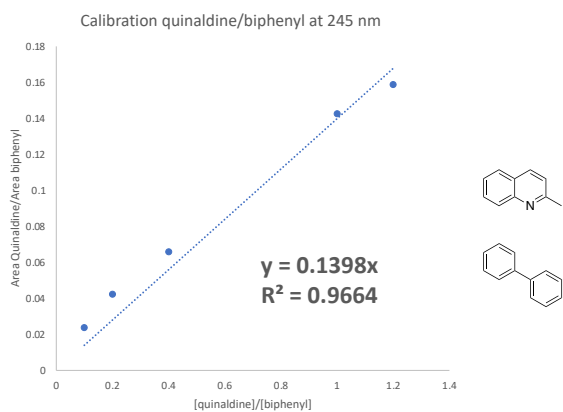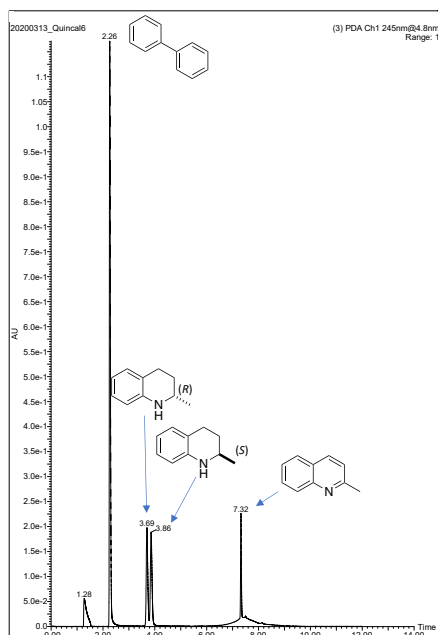
*Figure 102 Chromatogram and calibration curve for **2**. Left: UPC2 trace for the imine **2**, amine (R)- and (S)-**8** and the internal standard biphenyl in 10 mM concentrations. Right: calibration curve for (R)- and (S)-**8** versus internal standard. The calibration was made in a range from 0.5 mM to 5 mM.*

151
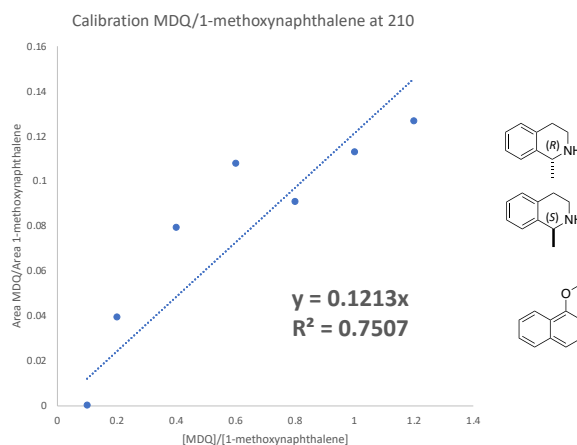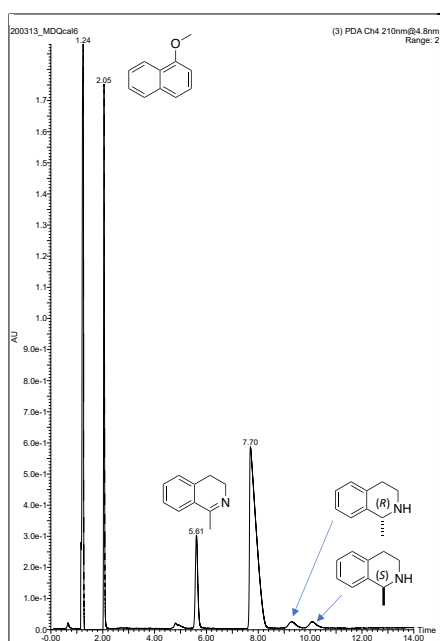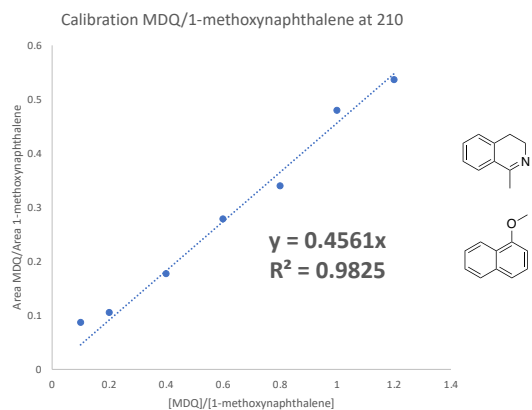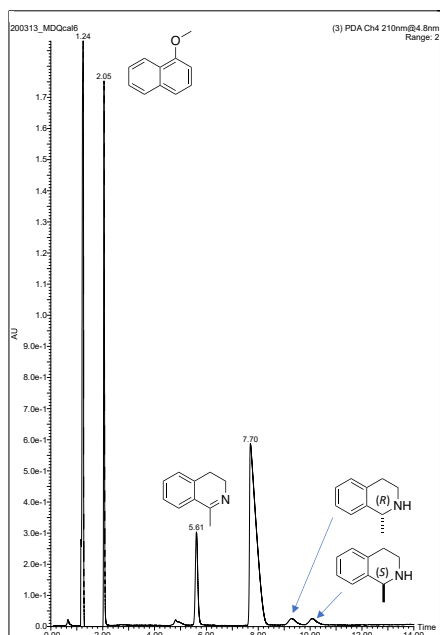
*Figure 103 Chromatogram and calibration curve for **3**. Left side UPC2 trace for the imine **3**, amine (R)- and (S)-**9** and the internal standard 1-methoxynaphthalene at 10 mM concentrations. On the right side the calibration curve for (R)- and (S)-**9** versus internal standard using the UPC2 method. The calibration was made in a range from 0.5 mM to 5 mM.*
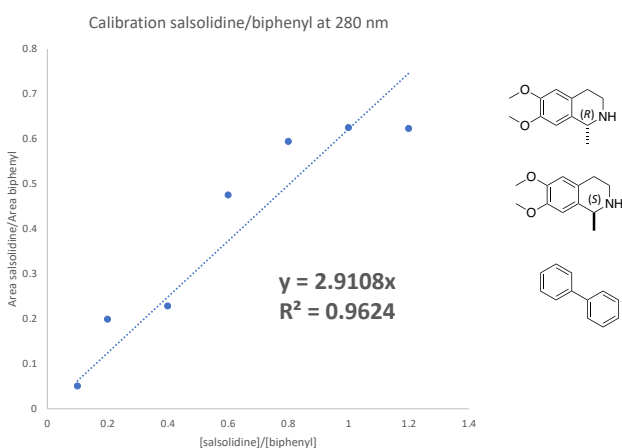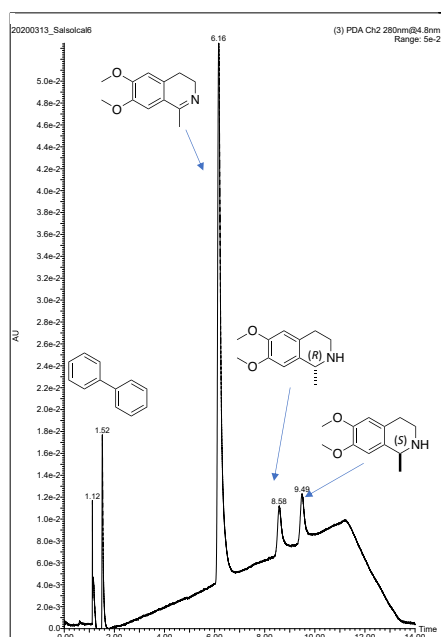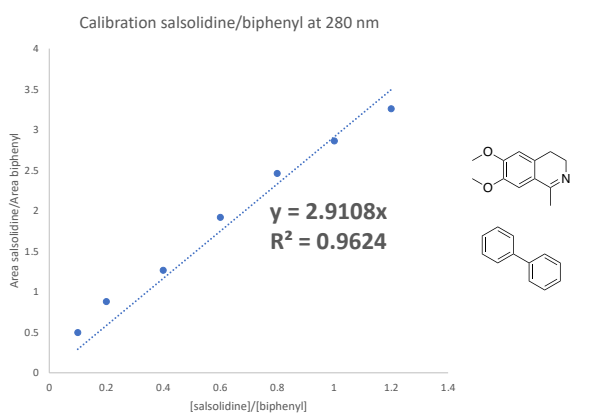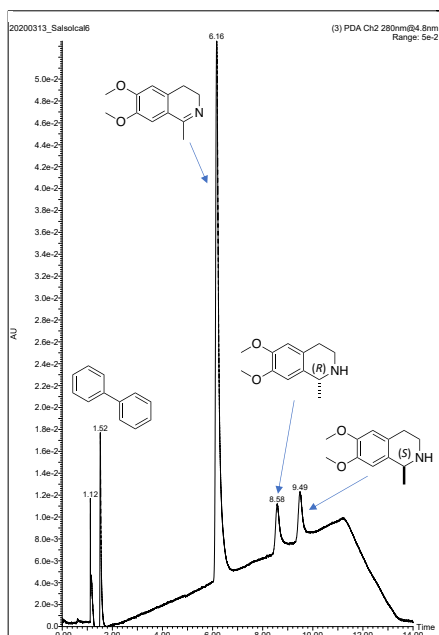
*Figure 104 Chromatogram and calibration curve for **4**. Left side UPC2 trace for the imine **4**, amine (R)- and (S)-**10** and the internal standard biphenyl at 10 mM concentrations. On the right side the calibration curve for (R)- and (S)-**10** versus internal standard using the UPC2 method. The calibration was made in a range from 0.5 mM to 5 mM.*
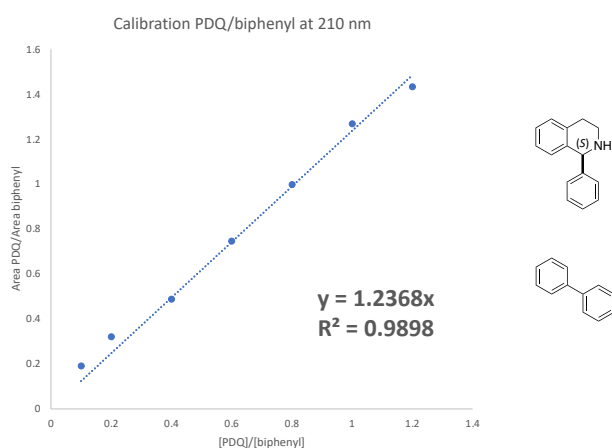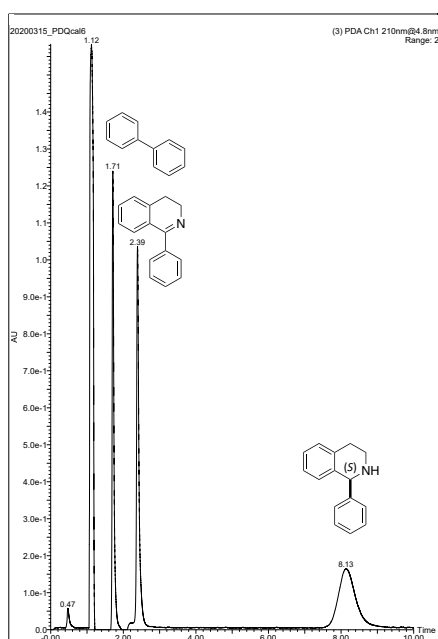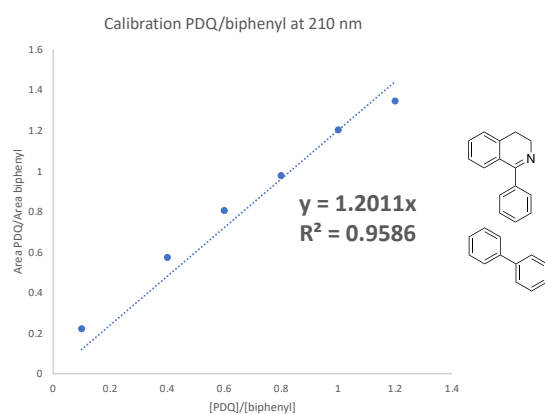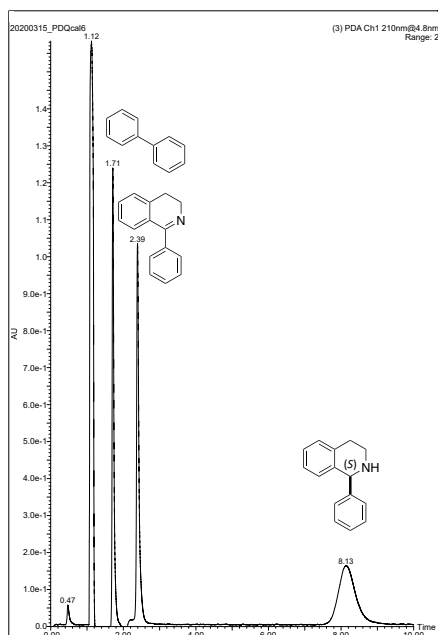
*Figure 105 Chromatogram and calibration curve for **5**. Left side UPC2 trace for the imine **5**, amine (S)-**11** and the internal standard biphenyl in 10 mM concentrations. On the right side the calibration curve for (S)-**11** versus internal standard using the UPC2 method. The calibration was made in a range from 0.5 mM to 5 mM.*
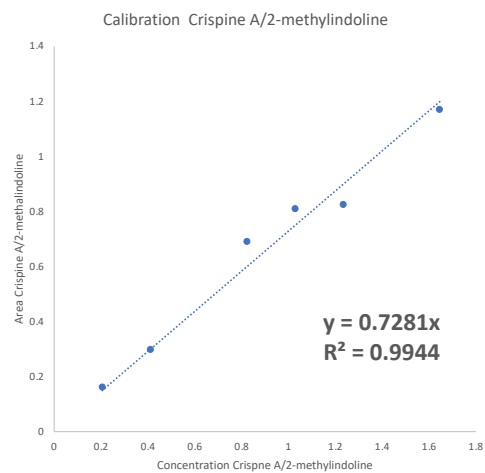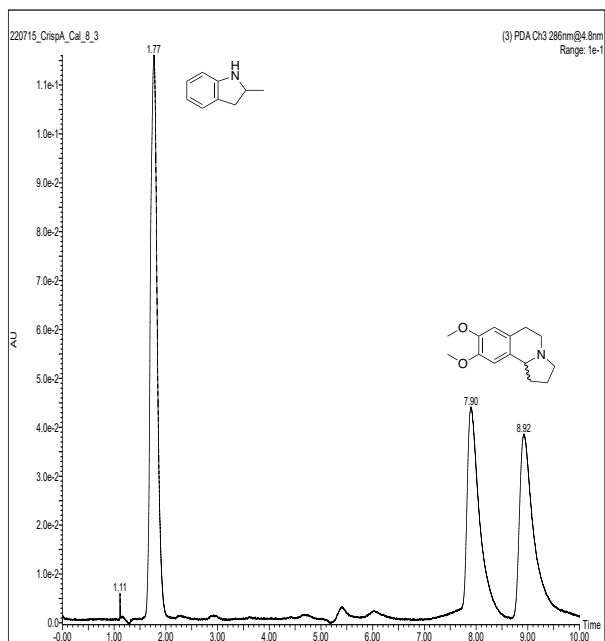
*Figure 106 Chromatogram and calibration curve for **12**. Left side UPC2 trace for the amines **12** and the internal standard 2-methylindoline at 10 mM concentration. On the right side the calibration curve for **12** versus internal standard using the UPC2 method. The calibration was made in a range from 0.5 mM to 5 mM.*

155
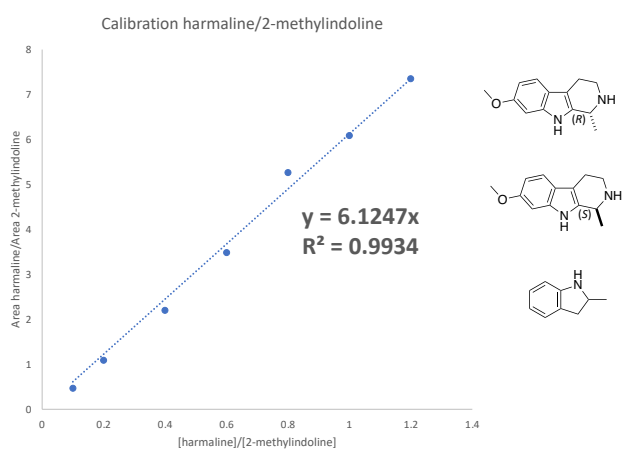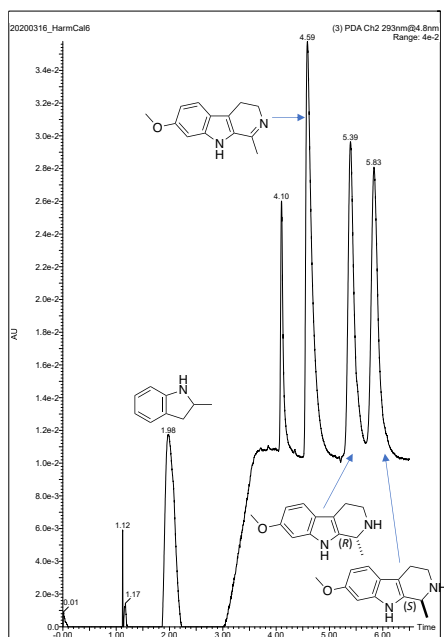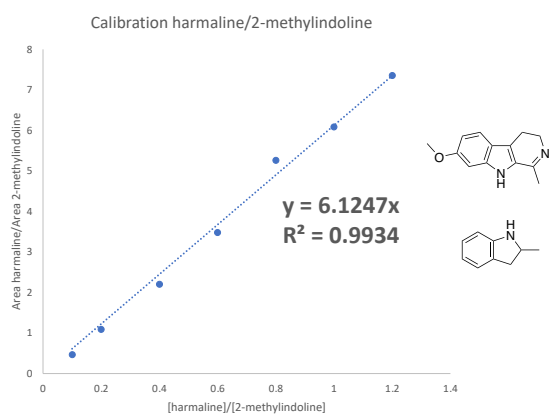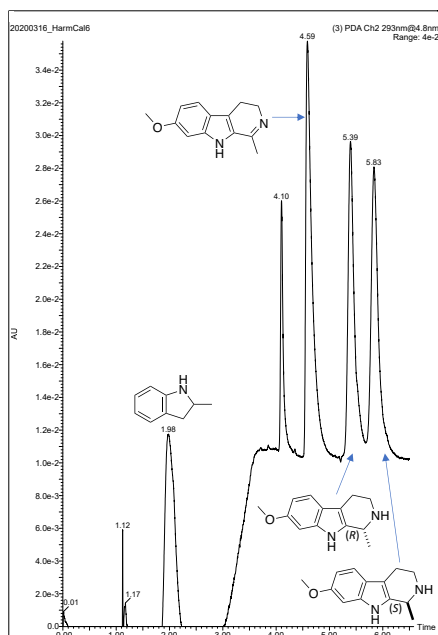
*Figure 107 Chromatogram and calibration curve for **7**. Left side UPC2 trace for the imine **7**, amine (R)-**13** and (S)-**13** and the internal standard 2-methylindoline in 10 mM concentrations. On the right side, the calibration curve for (R)-**13** and (S)-**13** versus internal standard using the UPC2 method. The calibration was made in a range from 0.5 mM to 5 mM.*

Chapter 5. Characterization of ArMs

5.1 Structural Characterization of ArMs

5.1.2 Native Mass Spectroscopy and X-Ray Crystallography

Depending on the information needed, different approaches are can be routinely used to analyze the ArMs. Mass spectral analysis (tryptic digestion, e.g.) may be used to characterize the protein's primary structure. However, if high-resolution structural information for the protein is needed, X-ray crystallography may be the way to go.[251,252] This high-resolution approach allows identifying ligand binding sites and conformational changes in a protein structure after binding a ligand, which may offer the possibility to identify close-lying positions that may be targeted by site-saturation mutagenesis to improve the ArMs' performance. A complementary method would be to use NMR analysis of an isotopically-labeled protein.

Furthermore, to gain information about the composition, stoichiometry, and subunit assemblies of proteins, native mass spectroscopy (native MS) can be used. An advantage of native MS, compared to the other methods, is the possibility of identifying the ligand with the highest affinity by inspecting a mixture of the protein target and different ligands simultaneously. Furthermore, the high sensitivity, speed, selectivity, and the requirement of a low sample amount (concentration pmol-fmol, V < 100 μl) render this method attractive for the investigation of ArMs.[253,254] In this chapter, I summarize my efforts to characterize ArMs by X-ray crystallography and (native) mass spectral analysis.

5.2 Native Mass Spectroscopy for Streptavidin based ArMs

 At the end of the last century, the introduction of matrix-assisted laser desorption/ionization (MALDI) by Tanaka *et al.*[255] and electrospray ionization (ESI) by Fenn *et al.*[256] as ionization methods for mass spectral analysis opened the door for the inspection of biomolecular compounds by mass spectroscopy (MS). The great potential of the methods was noticed and

honored with the Nobel Prize in chemistry in 2002. Shortly after introducing the ionization methods, non-covalent assemblies were shown to be preserved in the gas phase and analyzed by mass spectrometry.[257–259] Even though the experiments showed that it was possible to transfer intact protein complexes into the gas phase and analyze them by MS, the term native MS was introduced only in 2004.[260]

The ionization mode of the samples is different in both methods. In MALDI, a rapid photo-volatilization of biomolecules embedded in an ultraviolet-absorbing matrix followed by mass analysis is used, usually yielding a singly charged species. In ESI, a high electrical potential is applied to a capillary containing the solution, causing an accumulation of ions at the tip of the capillary, which in turn promotes the evolution of charged droplets carrying the precursors of the charged gas-phase ions. By electrostatic repulsion, finally, droplet fission is achieved, leading to the gas-phase ions (multi-charged).

After the introduction of soft ionization methods, significant developments in mass analyzers also had to be achieved to increase the applications of native mass spectrometry for the analysis of biomolecules. Nowadays, usually a quadrupole time-of-flight (Q-ToF) mass spectrometer, composed of a quadrupole mass filter, a hexapole collision-gas cell, and a ToF mass analyzer[261] is used for native MS experiments. The m/z values of the compounds are detected by ToF mass analyzers by the principle, that the time taken by ions to travel a known field-free distance under a high vacuum is directly related to their m/z ratios. By the addition of a reflector at the end of the flight tube, furthermore the difference in the kinetic energy distribution of the same m/z ions is corrected, and peak broadening and decreased resolution is avoided.[262] While ToF detectors have no theoretical upper mass limit, conventional quadrupoles allow an acquisition up to only m/z 2000 - 4000.[261] By reducing the frequency, however, in native MS, the m/z-selection range can be increased up to m/z 32'000.[263]

Most of the native MS studies up to date rely on ESI, ionizing the protein complexes from an aqueous ammonium acetate buffer. However, other buffers like ammonium bicarbonate or triethylammonium bicarbonate (TEAB) at a near-neutral pH can be used. Nevertheless, the transfer into the gas phase can lead to a pH change, as shown for ammonium acetate, therefore the term buffer for ammonium acetate in native MS studies might be misleading.[264] Usually, concentrations between 10 mM and 500 mM are used, nevertheless, if needed the concentration can be raised up to 1 M. Furthermore, different disrupting agents like DMSO,

methanol and others can be used, in order to study the stability of the protein complexes. When the protein complex gets transferred to the gas phase, ion signals at different m/z values, forming a Gaussian distribution, can be observed. These peaks are also called the charge-state envelope or charge-state distribution (CSD). The charges thereof result from the different amounts of the protons at the surface of the protein complex. Each charge state observed in the spectrum differs from its adjacent charge states by one proton. This useful characteristic is exploited to calculate the mass of the protein complex.[254] Furthermore, it is accepted that the folded protein structure can accommodate fewer charges as an denatured species, and therefore, the charge state distribution can be used as a reference for protein intactness.[265] Nevertheless, in native MS, the experimentally determined mass of a protein or protein complex is usually higher than its theoretical mass. The shift arises from incomplete desolvation and salt adducts, and remaining buffer. However, by activation of the ions, using carefully balanced energy, the difference can be minimized. Furthermore, using a method of activation, which is based on collisions with gas molecules and is called collision-induced dissociation (CID), investigations of the stability of the complexes can be run.

As mentioned in the chapter's introduction, information about the composition, stoichiometry and the subunit assemblies of proteins can be gained by native MS of protein complexes. Furthermore, it has been shown that dissociation constants of noncovalent complexes can be obtained from native MS experiments.[266–268] Therefore, we investigated different ArMs for their composition by native MS. This method can be of particular interest when no crystal structure of the ArM can be obtained (yet).

5.2.1 Outline of the Authors Contribution

TRW and NVI conceived and designed the study. NVI performed the (native) MS experiments. MP trained NVI on the instrument and gave advices.

5.2.2 Investigation of streptavidin ArMs by native MS

Already in 1995 Schwartz *et al.* inspected the binding properties of biotin by streptavidin by native MS.[269] Inspired by their work we aimed at introducing the method to our lab and adapt

it to the Bruker maXis II QTOF ESI system used within our department. As mentioned in the introduction, next to the use of a volatile buffer, soft ionization settings are needed, in order to transfer the protein complexes to the gas phase without disrupting their quaternary structure. Therefore, we adapted the settings of the mass spectrometer accordingly. When using denaturing conditions (0.2 mg/mL protein in H_2O, 0.1 % formic acid pH 1) we observed as expected a wider charge state distribution from +13 to +21 (Figure 108), in contrast to the charge states observed (+16 to +19) during native MS conditions (2 mg/mL, 50 mM ammonium acetate pH 6.8). This difference arises, as in the folded state, some of the amino acid residues which might be charged, are not accessible for the buffer. If the protein was incubated prior to the measurement with 10 eq. of biotin, the charge states for the streptavidin·biotin complex had a lower charge of exactly 4 each (Figure 110), as expected by the incorporation of the, at neutral pH, singly negatively charged biotins. To our delight, the deconvoluted spectrum was in accordance to the calculated mass (Table 36). As not only the absolute mass, giving information about the stoichiometry of the ligand protein complex, but also the charge of the ligand can be evaluated by native MS, this method could be used to complement the identification methods already used in our lab for the ArMs. Therefore, we tried to use native MS analysis to identify some ArMs, for which the crystal structure was yet to be obtained. One of the examples is the complex of Sav-S112A-K1211A with a bis-biotinylated iron porphyrin (**biot$_2$-Fe-pp)** (Figure 111, Figure 112). To our surprise, two different species could be observed, one being the tetrameric protein complex with two bis-biotinylated ligands and the other being the tetrameric protein complex containing three bis-biotinylated ligands. We hypothesized that it might be possible for streptavidin to bind more than two bis-biotinylated ligands, if the second biotin point outwards of the binding pocket. Nevertheless, the crystal structure, which will be discussed in the next section, revealed that in the soaked crystals, exclusively contained two bis-biotinylated ligands per tetramer. Furthermore, by increasing the CID energy to the maximum of 200 eV we were able to both disrupt the tetrameric complex, and obtain the apo-streptavidin monomer (Figure 113). This property could be used in future to determine the dissociation constant.[266] Another bis-biotinylated ligand used in our lab is the iron sulfur cluster (**biot$_2$-Fe$_4$S$_4$**, Figure 111). As the crystal structure of the ArM consisting of the cluster ligand and streptavidin was only obtained recently, we first aimed at determining the incorporation of the ligand into streptavidin by

native MS. To our delight, even though the cluster is oxygen sensitive, we were able to obtain the mass spectrum of the ArM in its tetrameric state. As can be appreciated in (Figure 114) the deconvoluted spectrum correlates well with the expected mass for the tetrameric Sav-WT di-bis-biotinylated iron sulfur cluster. Furthermore, as explained for the case of the streptavidin·biotin complex, the charge states of the cluster complex have a difference of -3 compared to the apo-protein. This difference arises from the charged bis-biotinylated iron sulfur cluster. However, the expected charge would be only -2, therefore the results should be taken with a grain of salt. In order to displace the cluster during the measurement and detect it, equally to the experiment for the iron porphyrin (**biot$_2$-Fe-pp**), the CID energy was increased and the isotope pattern at the expected m/z region for the Fe$_4$S$_4$-cluster (**biot$_2$-Fe$_4$S$_4$**) was inspected and compared to the simulated one (Figure 115). As mentioned, the cluster has a charge of -2, therefore to see in the positive mode of the MS, the species needs to be 3 times positively charged. The chemical formula is then $C_{40}H_{55}Fe_4N_8O_6S_{10}$ instead of $C_{40}H_{52}Fe_4N_8O_6S_{10}$. Both spectra match and the isotope pattern is identical. This supports the hypothesis that the cluster remains intact incorporation and subsequent displacement from Sav. We could thus show that the incorporation of the cluster into streptavidin is possible and could be displaced by the isCID.
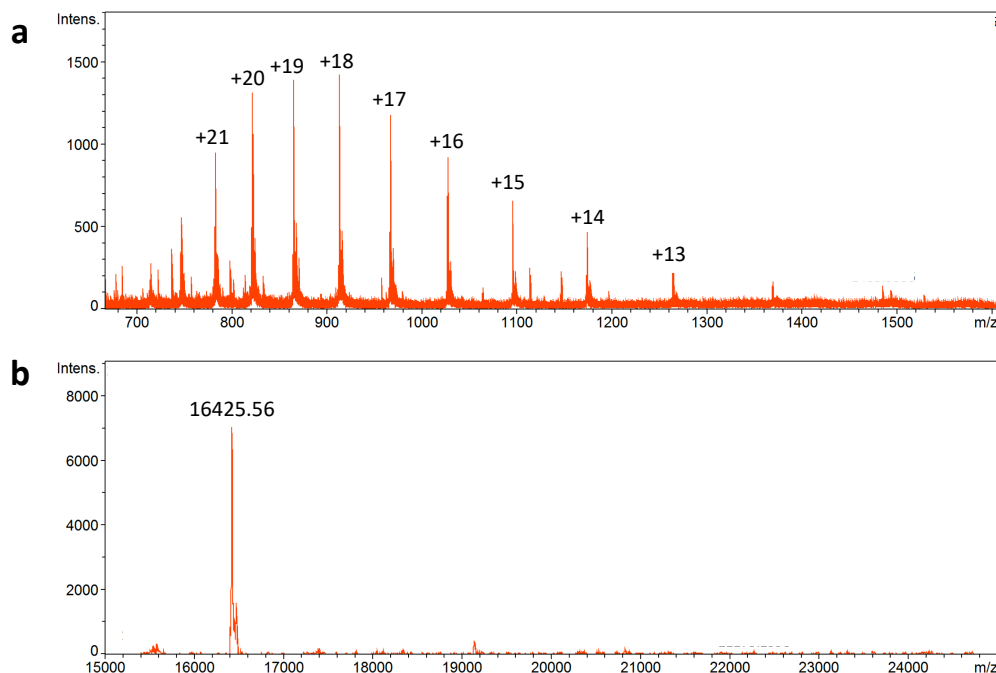
*Figure 108 Denaturing MS of Sav-WT. **a** Charge state distribution of the unfolded, monomeric protein, with various charge states of higher charge (13-21 per monomer). **b** Deconvoluted spectrum.*
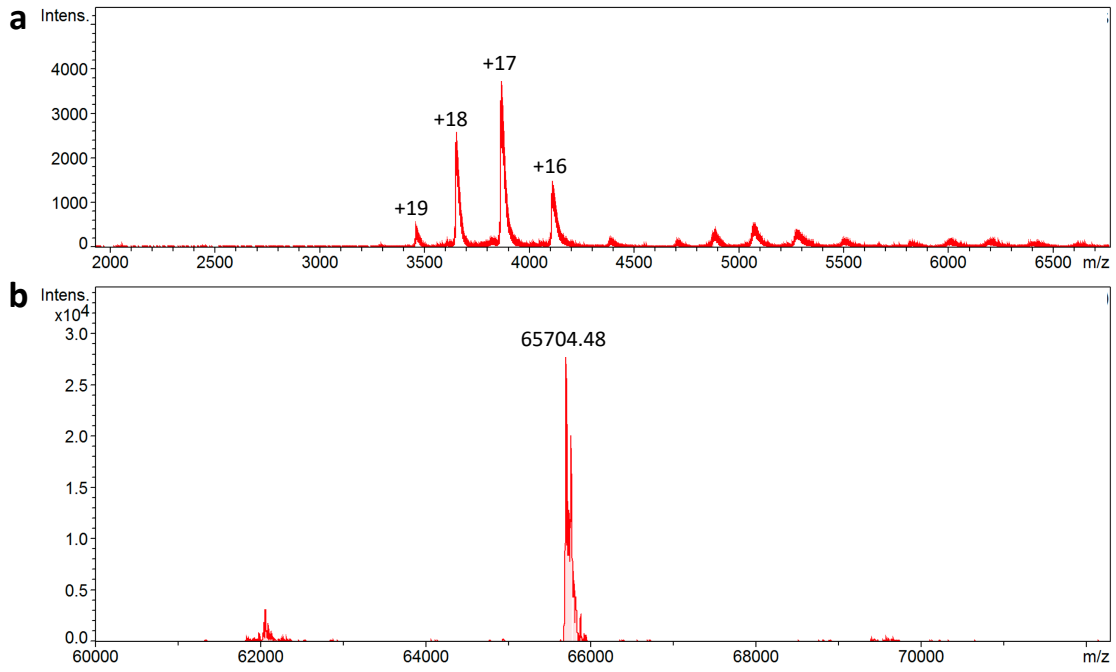
*Figure 109 Native mass spectrum of Sav-WT. **a** Charge state distribution of the folded tetrameric protein, with lower charge states (16-19, per tetramer). **b** Deconvoluted spectrum.*
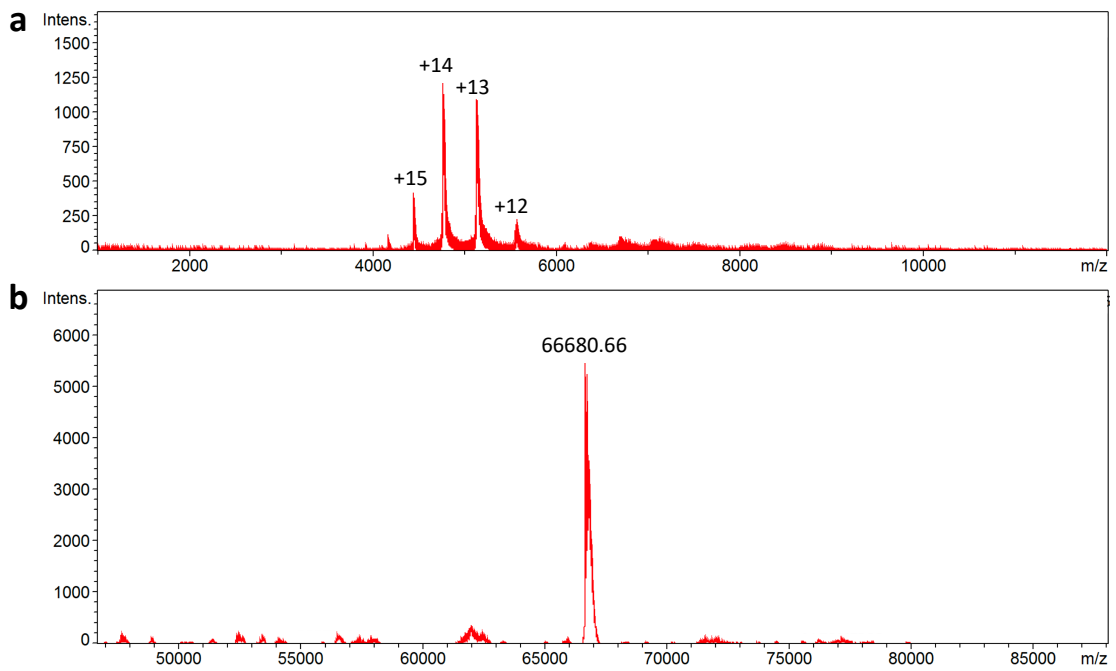
*Figure 110 Native mass spectrum of Biotin·Sav-WT. **a** Charge state distribution of the folded tetrameric protein with 4 bound biotins, with lower charge states (12-15, per tetramer). **b** Deconvoluted spectrum.*

a

Chemical Formula: $C_{48}H_{56}FeN_{12}O_{12}S_4$
Molecular Weight: 1177.13

b

Chemical Formula: $C_{40}H_{52}Fe_4N_8O_6S_{10}$
Molecular Weight: 1284.89
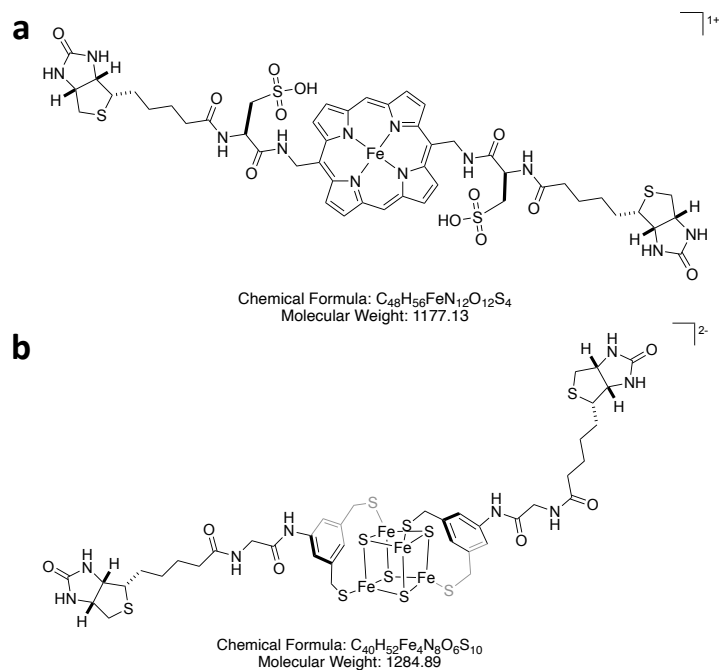
*Figure 111 **a** Iron porphyrin cofactor (**biot₂-Fe-pp**) used for the native MS studies. **b biot₂-Fe₄S₄** cofactor used for the native MS studies.*
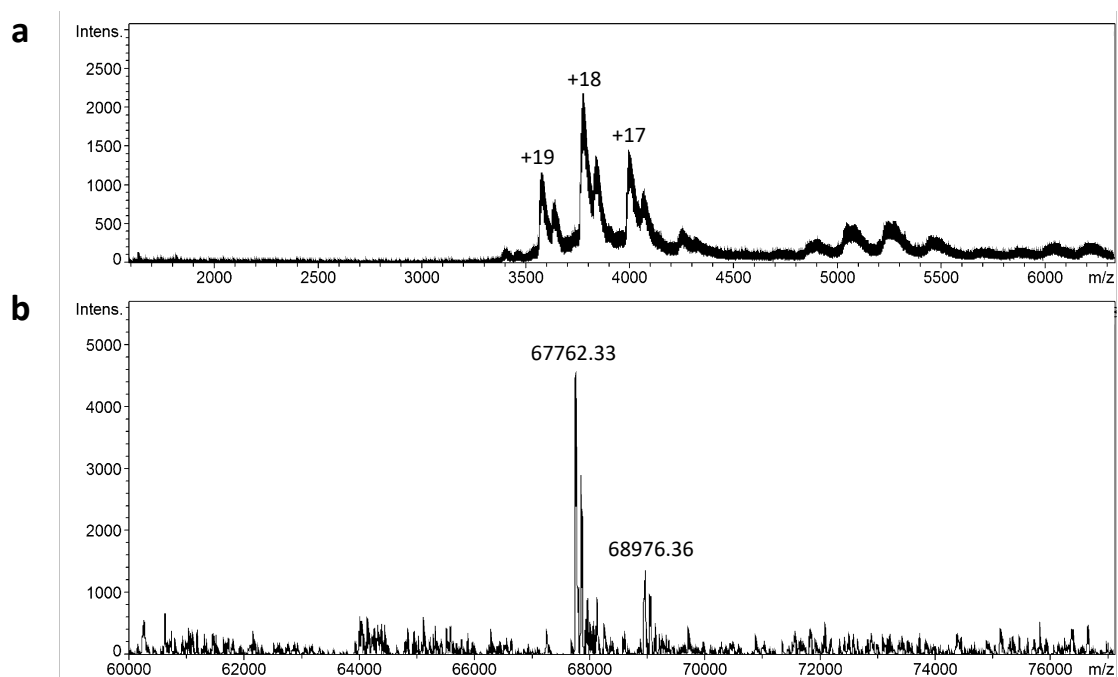
*Figure 112 Native mass spectrum of **biot₂-Fe-pp**·Sav-S112AK121A. **a** Charge state distribution of the folded tetrameric protein, with lower charge states (17-19, per tetramer). **b** Deconvoluted spectrum.*

163
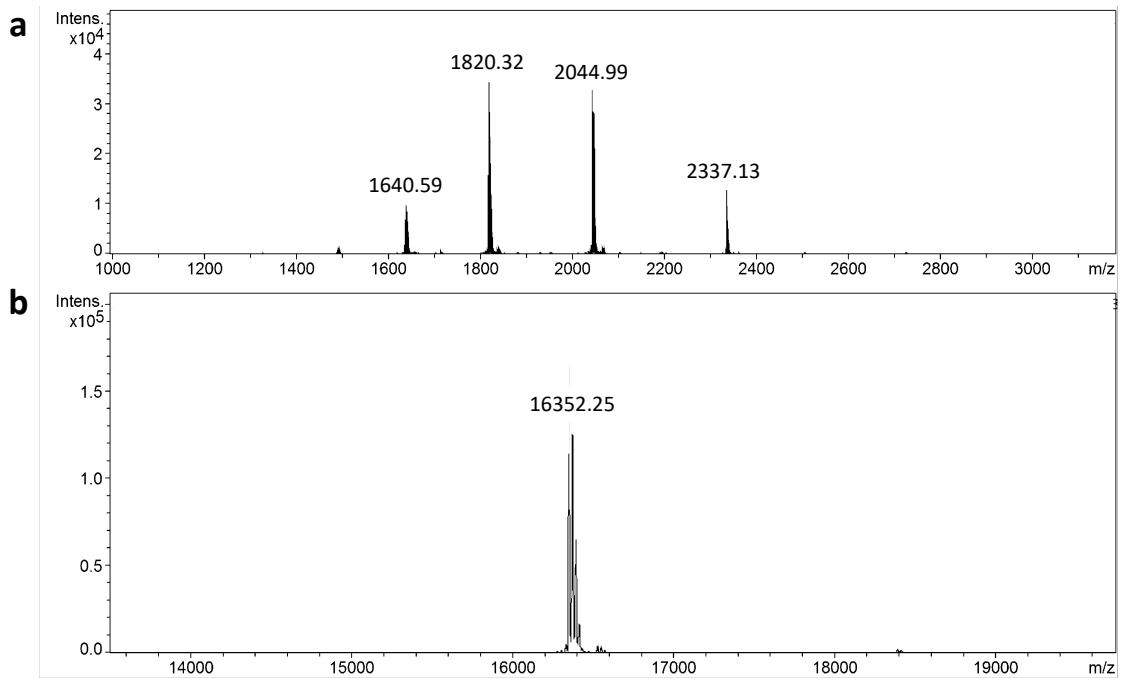
*Figure 113 Native mass spectrum of **biot$_2$-Fe-pp**·Sav-S112AK121A with isCID set to 200 eV. **a** Charge state distribution of the folded monomeric protein **b** Deconvoluted spectrum.*
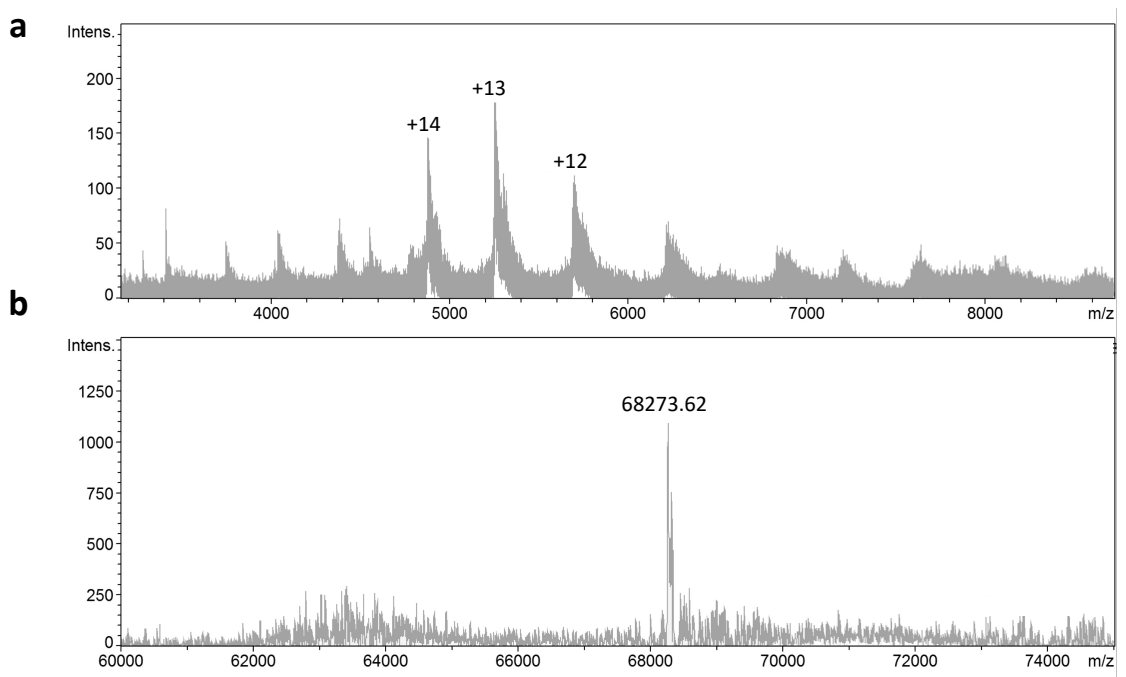
*Figure 114 Native mass spectrum of **biot$_2$-Fe$_4$S$_4$**·Sav-WT. **a** Charge state distribution of the folded monomeric protein **b** Deconvoluted spectrum.*
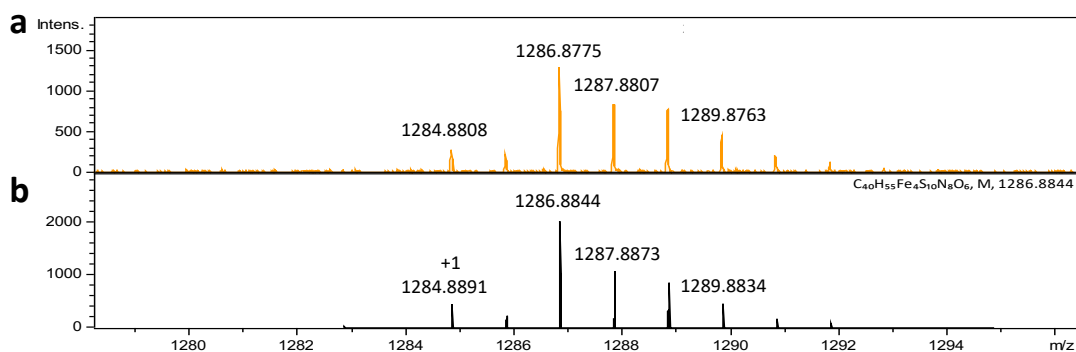
*Figure 115 Native mass spectrum of **biot₂-Fe₄S₄**-Sav-WT with isCID set to 200 eV focusing on the m/z region for the Fe₄S₄-cluster. **a** measured spectrum. **b** simulated spectrum of the singly positively charged (3 times protonated) Fe₄S₄-cluster.*

Table 36 Calculated and observed masses for the different streptavidin variants measured by MS.

Protein	Calculated mass (Da)	Observed mass (Da)
Streptavidin WT (monomer)	16424.93	16425.56
Streptavidin WT (tetramer)	65699.72	65704.48
Streptavidin WT + 4 biotin	66676.96	66680.66
Sav S112A K121A + 2 **biot₂-Fe-pp**	67765.50	67762.33
Streptavidin WT + 2 **biot₂-Fe₄S₄**	68273.38	68273.62

As discussed previously, native MS analysis is not restricted to streptavidin and can be adapted to any protein, which is ionizable by ESI. Therefore, this method should be considered, in case that there is no crystal structure available (yet), especially, as next to the stoichiometric information, also the dissociation constant could be calculated from the observed spectrum. In the next two sections native and "denaturing" mass spectral analysis were used to investigate chimeric streptavidin proteins, as well as human carbonic anhydrase II ArMs. Especially, in the case of the hCAII based ArMs we were able to show that native MS is a good complementary method for analysis of ArMs and should be implemented to the labs working with ArMs.

5.2.3 Experimental

Streptavidin (0.2 mg/mL) was dissolved in water: formic acid (1000:1). The protein concentration was determined by a NanoDrop microvolume spectrophotometer (Thermo

Fisher Scientific, USA) and adjusted to 0.2 mg/mL (MQ-H_2O). A HPLC (Shimadzu, equipped with a Jupiter® 5 µm C4 300 Å)-ESI QTof (Bruker Daltonics, ESI Maxis II QTof MS) system was used to record the data. As eluent water: formic acid (A, 1000 : 1) and acetonitrile : formic acid (B, 1000 : 1) were used. The time program was as follows; 2 min 10% B, 7 min 10-55%, 2 min 55%, 2 min 55-65%, 2 min 65%, 2 min 65-10%, 8 min 10%.

Table 37. MS settings for protein mass spectral analysis.

Source Type	ESI	Ion Polarity	Positive	Set Nebulizer	2.0 Bar
Focus	Active	Set Capillary	4500 V	Set Dry Heater	220 °C
Scan Begin	150 m/z	Set End Plate Offset	-500 V	Set Dry Gas	8.0 l/min
Scan End	2750 m/z	Set charging Voltage	0 V	Set Divert Valve	Source
		Set Corona	4000 nA	Set APCI Heater	200 °C

To assemble the ArMs, the protein was dissolved in ammonium acetate (50 mM , pH 6.8) or ammonium carbonate (50 mM, pH 8.5) to a concentration of 2 mg/mL, supplemented with 0.1 mM cofactor and incubated while shaking (18 h, 25°C, 600 rpm). Electrospray ionization (ESI) low concentration tuning mix from Agilent was used as calibrant. The samples were directly injected into the high-resolution MS (HRMS) with a syringe pump using a flow rate of 5 µL/min. The ESI source parameters were optimized for the streptavidin analysis. The charge-state distribution from 3000-6000 m/z was deconvoluted using the Compass Data Analysis software (Bruker Daltonics) with the maximum entropy setup.

Table 38 MS settings for native MS analysis of the tetrameric streptavidin (complex).

Source Type	ESI	Ion Polarity	Positive	Set Nebulizer	1.8 Bar
Focus	Active	Set Capillary	4000 V	Set Dry Heater	220 °C
Scan Begin	1000 m/z	Set End Plate Offset	-500 V	Set Dry Gas	3.0 l/min
Scan End	12000 m/z	Set charging Voltage	2000 V	Set Divert Valve	Source
		Set Corona	0 nA	Set APCI Heater	0 °C

One problem we faced with the bis-biotinylated iron sulfur cluster is its stability in the presence of traces of oxygen. Therefore, the samples were prepared in degassed solutions

and the incubation was set up in the glove box. After 18h of incubation time, the samples were transferred in the glove box to a 50 μL Hamilton® syringe and directly injected to the MS. To verify, if the stability of the complex might be increased by using a different buffer, ammonium carbonate (50 mM $(NH_4)_2CO_3$, pH 8.5) was used instead of ammonium acetate.

5.3 Mass spectral analysis of scdSav

The results were first published by and reproduced with permission from ACS:
Wu, S., Zhou, Y., Rebelein, J. G., *et al.* Breaking Symmetry: Engineering Single-Chain Dimeric Streptavidin as Host for Artificial Metalloenzymes. *J. Am. Chem. Soc.* **141**, 40, 15869–15878 (2019) and can be found under the DOI: https://doi.org/10.1021/jacs.9b06923

5.3.1 Mass spectral analysis

In order to detect the dimer of scdSav method of the mass spectrometer was optimized as collected in Table 39. The experiments were run as follows: A scdSav solution (0.2 mg/mL) in a water-methanol (1:1) mixture supplemented with 0.1 % formic acid and was clarified by centrifugation. The supernatant was analyzed by direct injection on an electrospray ionization time-of-flight mass spectrometer (Bruker maXis II QTOF ESI), calibrated to a streptavidin standard. To calculate the mass of the (scdSav)2 the charge envelope from 2600 to 48000 m/z was used (Figure 116). The results of the scdSav mass spectrometry are provided in Table 40.

Table 39 MS settings for protein mass spectral analysis.

Source Type	ESI	Ion Polarity	Positive	Set Nebulizer	1.4 Bar
Focus	Active	Set Capillary	4500 V	Set Dry Heater	220 °C
Scan Begin	500 m/z	Set End Plate Offset	-500 V	Set Dry Gas	9.0 l/min
Scan End	6000 m/z	Set charging Voltage	0 V	Set Divert Valve	Source
		Set Corona	0 nA	Set APCI Heater	0 °C

*Figure 116 Mass spectra of the dimers of scdSav. **a** Charge state envelope of AASK. **B** Deconvoluted mass spectra of AASK. **c** Charge state envelope of SKSA. **d** Deconvoluted mass spectra of SKSA. In Table 40 the masses of the rest of proteins used in the study are shown and compared to the theoretical mass.*

*Table 40 Results from the mass spectral analysis of the study with the scdSav. * The scdSav mutants KKKK and KAKK could not be ionized.*

scdSav () *	Theoretical mass of the dimer (Da)	Measured mass (Da)
SKSK	69701.66	69701.11
AKSK	69669.66	69670.83
SKAK	69669.66	69670.21
AKAK	69637.66	69639.8
RKSK	69839.88	69840.32
SKRK	69839.88	69840.03
RKRK	69978.1	69980.32
KKSK	69783.86	69786.42
SKKK	69783.86	69786.02
SASK	69587.48	69587.3
AASK	69555.48	69556.8
SAAK	69555.48	69556.36

SAKK	69523.48	69521.89
RASK	69725.7	69725.09
SARK	69725.7	69724.05
RARK	69863.92	69863.29
KASK	69669.66	69670.32
SAKK	69669.66	69670.26
SKSA	69587.48	69587.05
AKSA	69555.48	69554.49
SKAA	69555.48	69557.76
AKAA	69523.48	69523.09
RKSA	69725.7	69726.36
SKRA	69725.7	69725.78
RKRA	69863.92	69863.46
SASA	69473.28	69472.77
AASA	69441.28	69439.79
SAAA	69441.28	69440.23
AAAA	69409.28	69406.73
RASA	69611.5	69612.33
RARA	69749.72	69751.39
scd(SARK)mv1	69695.66	69693.94
scd(SARK)mv2	69607.64	69606.81

5.4 A Dual Anchoring Strategy for the Directed Evolution of Improved Artificial Transfer Hydrogenases Based on Carbonic Anhydrase

5.4.1 Outline of the Authors Contribution

T.R.W., J.G.R. and C.Y. conceived and designed the study. C.Y. and D.C. contributed to the synthesis of the substrates, products and complexes. J.G.R. and A.S. contributed to mutagenesis, protein expression and protein purification. J.G.R and A.S. performed the crystallization, X-ray structure determinations. N.V.I. performed the LC-MS and native MS experiments. Y.C., J.G.R., D.C., A.S. performed the catalytic experiments, designed the screening protocol and recorded the data. T.R.W., J.G.R., Y.C., A.S., D.C. and N.V.I. analyzed the data. T.R.W., J.G.R., Y.C., A.S., D.C. and N.V.I. wrote the manuscript. All authors have given approval to the final version of the manuscript.

5.4.2 Abstract

Artificial metalloenzymes result from anchoring a metal cofactor within a host protein. Such hybrid catalysts combine the selectivity and specificity of enzymes with the versatility of (abiotic) transition metals to catalyze new-to-Nature reactions in an evolvable scaffold. With the aim of improving the localization of an arylsulfonamide-bearing iridium-pianostool catalyst within human carbonic anhydrase II (hCAII) for the enantioselective reduction of prochiral imines, we introduced a covalent linkage between the host and the guest. Herein we show that a judiciously positioned cysteine residue reacts with a *p*-nitropicolinamide ligand bound to iridium to afford an additional sulfonamide covalent-linkage. Three rounds of directed evolution, performed on the dually-anchored cofactor, led to improved activity

and selectivity for the enantioselective reduction of harmaline (up to 97 % *ee* (*R*) and > 350 turnovers on a preparative scale). To evaluate the substrate scope, the best hits of each generation were tested with eight substrates. X-ray analysis, carried out at various stages of the evolutionary trajectory, was used to scrutinize i) the nature of the covalent linkage between the cofactor and the host as well as ii) the remodeling of the substrate-binding pocket.

5.4.3 Introduction

Artificial metalloenzymes (ArMs) expand the repertoire of biocatalysed reactions.[93,98,148,185,270–275] ArMs combine attractive features of organometallic catalysts –such as broad substrate scope and large reaction repertoire– with characteristics of enzymes, including high turnover numbers, unrivalled stereoselectivity and biocompatibility. Such hybrid catalysts result from the incorporation of an abiotic metal cofactor within a protein.[276–281] Four distinct anchoring strategies have been pursued to date: covalent,[282–284] supramolecular,[26,285,286] dative[287–289] and metal substitution.[18,290,291] As the "repurposed active site" seldom matches snugly the structure of the cofactor · substrate complex, the cofactor is poorly localized, potentially limiting the effect of mutations on the catalytic performance of ArMs.[67] With the aim of firmly localizing an abiotic cofactor, several groups have engineered additional interactions between the cofactor and the host protein, ultimately leading to improved activities and selectivities for ArMs based on myoglobin,[292,293] LmrR,[150,151,294,295] streptavidin, etc.[71]

To complement our efforts centered around streptavidin-based ArMs, our group has reported the use of human carbonic anhydrase II (hCAII) as host protein for the assembly and optimization of an artificial transfer hydrogenase[102,154,296] (ATHase hereafter) and artificial metathase.[297] hCAII provides an attractive scaffold due to its monomeric globular structure (30 kDa), high affinity for aromatic-sulfonamides and its large funnel-shaped hydrophobic access to the active site.[298,299] Thanks to its monomeric nature, hCAII can readily be displayed on *E. coli*'s outer-membrane, significantly simplifying the ArMs assembly and the whole-cell screening procedure. In a previous study, we identified cofactor **1** (Figure 117a) as a promising cofactor for the transfer hydrogenation using wild-type hCAII (hCAIIWT) as a protein

scaffold.[102] Enantiopure amines represent attractive targets both as pharmaceutical ingredients and agrochemicals.[300–303] Following a seminal report in 2011,[304] imine reductases have firmly highlighted their versatility for the production of enantiopure amines.[305–308] These complement homogeneous catalysts for the reduction of prochiral enamines and imines.[309–311] Capitalizing on our previous efforts in ATHases, we set out to introduce a covalent anchor between the IrCp* cofactor and hCAII. We hypothesized that the resulting firm localization of the cofactor may positively affect both the TONs and the stereoselectivity of the resulting ATHase.

Building on the abundant literature covering ligation of cysteines with electrophiles[312] including maleimides,[313] nitro pyridines,[314] halogenoalkanes,[315] alpha-halocarbonyls,[316] Mukaiyama regents[317] and sulfones,[318] we envisioned engineering a suitably-positioned cysteine residue that would undergo a nucleophilic addition on the cofactor bearing an electrophilic moiety. Herein, we report on our efforts to design and evolve an ATHase equipped with a dual-anchoring system.

5.4.4 Results

Design of a dual anchoring strategy for carbonic anhydrase. Building on our previous hCAII-based ATHase,[102] the requirements for the cofactor design were set as follows: i) an arylsulfonamide anchor for anchoring to hCAII, ii) an ethyl spacer between the anchor and the IrCp* moiety, iii) a picolinamide chelate for increased transfer hydrogenation activity[319] and iv) an electrophilic substituent on the pyridine, susceptible of nucleophilic attack by a suitably-positioned cysteine (Figure 117). We envisioned that cofactor **2** would efficiently bind to hCAII thanks to the arylsulfonamide linker and would be further anchored via a nucleophilic aromatic substitution (S_NAr) with a suitably-positioned cysteine residue present in the protein vestibule (Figure 117c). Moreover, we hypothesized that the S_NAr may further increase the activity of the IrCp* picolinamide complex by increasing the electron-donating ability of the ligand.[320] Cofactor **2** was synthesized in two steps from commercially-available starting materials (Scheme 1). Next, hCAII^WT was crystallized by sitting drop vapor diffusion and the crystals were crosslinked with glutaraldehyde prior to soaking with cofactor **2**. The resulting X-ray diffraction data was solved at 1.41 Å resolution. Inspection of the X-ray structure **2** · hCAII^WT reveals that both glutamate 69 (E69) and isoleucine 91 (I91) residues lie

closest to the electrophilic nitro group of the picolinamide. The shortest contacts to one oxygen of the nitro group to cofactor **2** for these two residues are C_γ of E69 and $C_{\gamma 1}$ of I91 (3.4 Å and 3.6 Å, respectively, Figure 117b). Accordingly, we expressed and purified the single mutants hCAII^{E69C} and hCAII^{I91C} and evaluated their catalytic performance towards the transfer hydrogenation of harmaline **3** and dihydroisoquinolinium **4** precursor to crispine A **4**-H$_2$. ATHases **1** · hCAII^{WT}, **1** · hCAII^{E69C} and **1** · hCAII^{I91C} perform poorly for the reduction of substrate **3** (Table 47). In contrast, ATHases **2** · hCAII^{WT}, **2** · hCAII^{E69C} and **2** · hCAII^{I91C} catalyze the enantioselective reduction of both substrates. Strikingly, the position of the cysteine residue affects the resulting enantioselectivity: for both substrates, opposite enantiomers are preferentially produced in the presence of **2** · hCAII^{E69C} and **2** · hCAII^{I91C}, respectively (Table 41). To assemble the ArMs, cofactor **2** (0.1 mM) and hCAII$^{WT,\ E69C\ or\ I91C}$ (0.1 mM) were incubated (6 h) at room temperature using carbonate buffer (50 mM, pH 9.4). Transfer hydrogenation was performed using substrate (2 mM) and ArMs (10 µM) in MOPS buffer containing NaHCO$_2$ (1 M) as the hydride source. Reduction of harmaline **3** using **2** · hCAII^{WT} gave (*R*)-**3**-H$_2$ with a good yield (68%, after 16 h) and a moderate enantioselectivity (26% *ee*). In contrast, **2** · hCAII^{E69C} gave (*R*)-**3**-H$_2$ with a comparable yield (66%) but with an increased *ee* = 50%. Interestingly, **2** · hCAII^{I91C} also gave a good yield (66%) and a reverse *ee* = -38% for (*S*)-**3**-H$_2$ (Table 41, entry 1-3). The difference in enantioselectivity suggests different binding poses for harmaline in the larger binding pocket of **2** · hCAII^{I91C} compared to **2** · hCAII^{E69C} (Figure 136 and Figure 138). More strikingly, the reduction of iminium **4** by **2** · hCAII^{WT} gave a low conversion (6%) and hardly any enantioselectivity (Table 41, entry 4). In contrast, **2** · hCAII^{E69C} led to a slightly higher yield (13 %) and *ee* (-50%) (Table 41, entry 5). Mutant **2** · hCAII^{I91C} gave higher yield (30%) but low enantioselectivity (15%, Table 41, entry 6). Taken together, these results highlight the power of the dual anchoring strategy, giving rise to an increased enantioselectivity for substrates **3** and **4** compared to the parent hCAII^{WT}. To further improve the catalytic performance of the dually-anchored ATHases, we relied on directed evolution.[20,321,322]
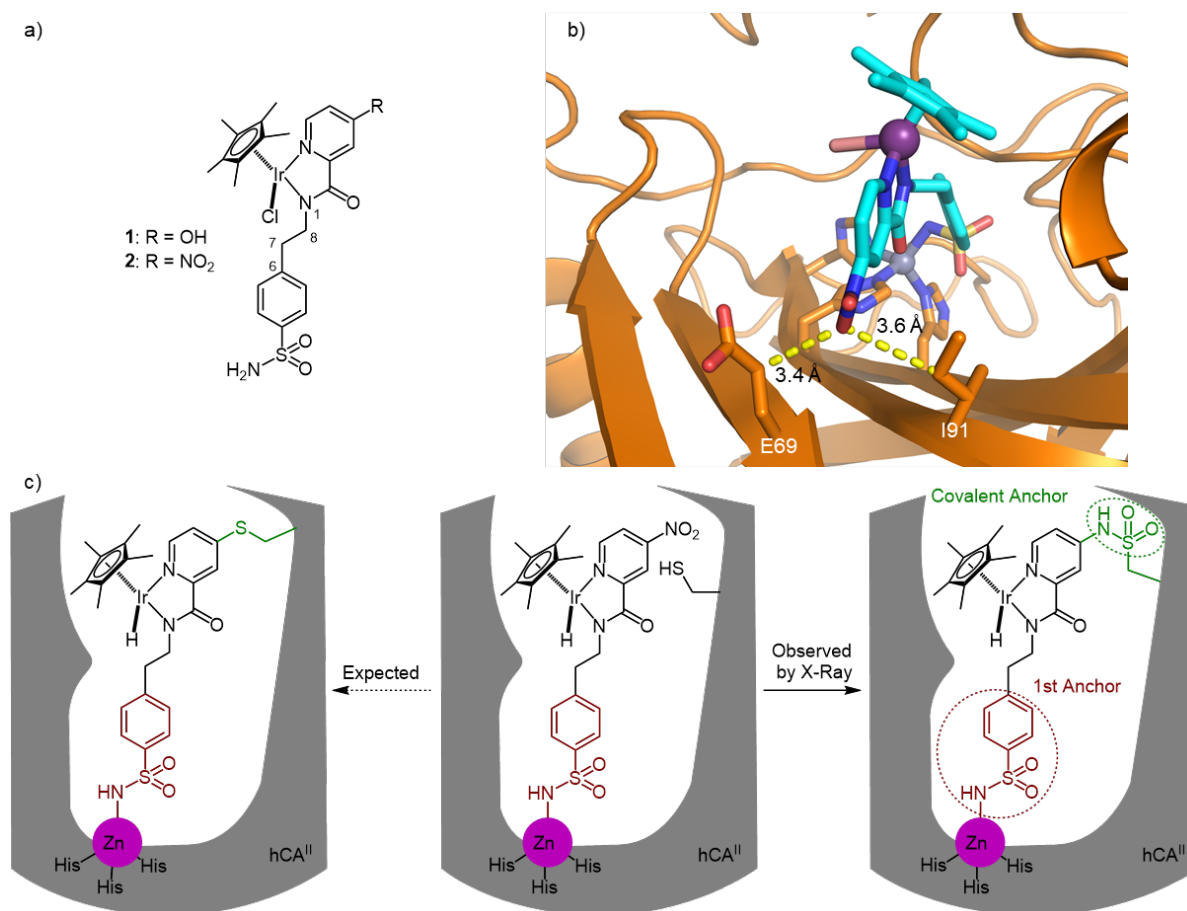
*Figure 117 Strategy to assemble dually anchored ArMs. (**a**) Structure of the cofactors **1–2**, (**b**) X-ray structure of cofactor **2** (depicted as sticks, Ir and Zn as color-coded spheres) bound to hCA^{WT} (depicted as cartoon representation). Distances between the O$_{nitro}$ and the residue glutamate E69 and isoleucine I91 (depicted as sticks) are highlighted, (c) Schematic representation of the expected and observed dual anchoring within engineered hCAII^{E69C} and hCAII^{I91C}.*

*Table 41 Catalytic Performance of ATHases Constructed by Anchoring Cofactor **2** to hCAII^{WT}, hCAII^{E69C}, and hCAII^{I91C} ^{a}Reactions were performed for 16 h using hCAII (10 µM), cofactor 2 (10 µM), and substrate (2 mM) in MOPS/NaHCO2 buffer (pH 7.4, 0.34 M, 0.85 M).*

	Protein	Substrate	Yield (%)	TON	*ee* (%)
1	WT		68	136	26 (R)
2	E69C		68	132	50 (R)
3	I91C	**3**	66	132	38 (S)
4	WT		6	12	7
5	E69C		13	26	-50
6	I91C	**4**	30	60	15

To confirm the dual anchoring of the cofactor **2** within hCAII[E69C] and hCAII[I91C], the two variants were incubated with cofactor **2** (16 h, pH 9.4, 50 mM $(NH_4)_2CO_3$). The unbound cofactor **2** was removed and the buffer was exchanged (25 mM Tris-HCl, pH 7.4) by ultrafiltration. The resulting ArMs were crystallized by sitting drop vapor diffusion (see SI). To our delight, the crystals diffracted to a good resolution for **2** · hCAII[E69C] (1.51 Å) and for **2** · hCAII[I91C] (1.04 Å). The electron density highlights the cofactor's localization within hCAII[E69C] (Figure 134) and hCAII[I91C] (Figure 119b-c). As observed for **2** · hCAII[WT], the arylsulfonamide anchor of cofactor **2** is bound to the zinc ion of hCAII. An additional bond was identified between the picolinamide moiety and the engineered cysteine residues. Unexpectedly, the well-resolved electron density revealed the presence of a pyridine-sulfonamide linkage (Figure 119b-c, Figure 134 and Figure 135b), rather than the anticipated thioether moiety (Figure 117). For **2** · hCAII[I91C], the pyridine-sulfonamide bond is fully formed, resulting in a complete occupancy (100 %) for the IrCp*(picolinamide) moiety (Figure 119c, Figure 135b and Figure 135e). For **2** · hCAII[E69C], the IrCp*(picolinamide) moiety with the pyridine-sulfonamide linkage has a lower occupancy (~70%) (Figure 134). Two additional cofactor positions with a free cysteine E69C were refined (Figure 134c). The incomplete formation of the pyridine-sulfonamide bond for hCAII[E69C] may be caused by the somewhat unfavorable geometry that cofactor **2** has to adopt to form the pyridine-sulfonamide linkage (Figure 136). The dihedral angle between C6-C7-C8-N1 serves as an indicator for the strain on cofactor **2**. Within hCAII[E69C] the dually-anchored cofactor **2** has an acute dihedral angle (102.2°), which is significantly smaller than the dihedral angle observed for the non-covalently bound cofactor **2** in hCAII[WT] (158.6°). The dihedral angle for **2** in hCAII[I91C] (134.0°) is closer to the angle of **2** in hCAII[WT] (158.6°) suggesting a more relaxed geometry of **2** in hCAII[I91C]. The observed difference of the dihedral angles of **2** in hCAII[E69C] and hCAII[I91C] might explain the difference between the incomplete and quantitative formation of the pyridine-sulfonamide bond for these two hCAII mutants, respectively.

Compared to **2** · hCAII[WT], the presence of the pyridine-sulfonamide linkage leads to the relocation of the Ir-picolinamide complex in hCAII[I91C], providing slightly more space for substrate binding (Figure 136 and Figure 138). The Ir-ion of ATHase **2** · hCAII[I91C] is displaced (1.0 Å) as well as the N2 of the nitro precursor is shifted (3.4 Å) in **2** · hCAII[I91C] compared to **2** · hCAII[WT] (Figure 137). Cofactor **2** is stabilized via H-bonds with glutamine 92 (Q92) and threonine 198 (T198) and by π-stacking with phenylalanine (F130) in hCAII[I91C] (Figure 138). In

all three X-ray structures of **2** · hCAII[WT, E69C and I91C], the Ir in **2** has an (*S*)-configuration. In contrast to previous reports on ATHases, the absolute configuration of the metal does not correlate with the configuration of the products.[76] Indeed, both **2** · hCAII[E69C] and **2** · hCAII[I91C] contain cofactor **2** with an (*S*)-configuration at the Ir-Cl, but convert substrate **3** and **4** preferentially to opposite enantiomers. H aving established a powerful method to dually anchor the cofactor 2 to hCAII scaffolds bearing a judiciously-positioned cysteine residue, we set out to optimize the ATHase activity by directed evolution.[20,321,322]

Directed evolution. To improve the activity of ATHases based on the dual anchoring strategy, we selected hCAII[I91C] as a scaffold for directed evolution. The selection of hCAII[I91C] over hCAII[E69C] is based on the following considerations: i*)* the bioconjugation is quantitative, thus affording a single and well-defined localization of the cofactor **2** within the "active site", and ii) more space is available for the prochiral substrate to approach the Ir–H moiety. To streamline directed evolution, we selected an *E. coli* surface display for hCAII.[102] This strategy was favored over the periplasmic compartmentalization in light of the following factors: i) the cofactor concentrations are roughly similar –as determined by ICP-MS– and ii) the accessibility for the substrate are coupled with a defined reaction environment (i.e. extracellular vs. peptidoglycan environment).[102] For *E. coli* surface display, hCAII was fused at its *N*-terminus with a truncated *E. coli* lipoprotein Lpp (residues 1–9), followed by the first five *β*-sheets of OmpA (residues 46–159, Figure 118). Initial experiments were performed with hCAII[WT] and hCAII[I91C] evaluating the effect of varying concentrations of cofactor **2** in the presence of either substrate **3** or **4**. In line with the reactions performed with purified hCAII samples (Table 41), the conversion of substrate **3** was higher than that of substrate **4**. Accordingly, harmaline **3** was selected as the substrate for the directed evolution campaign. To minimize the background reaction resulting from unspecifically-bound cofactor **2** (e.g., in the presence of *E. coli* bearing an empty plasmid), its concentration was set to 5 μM (Figure 139).
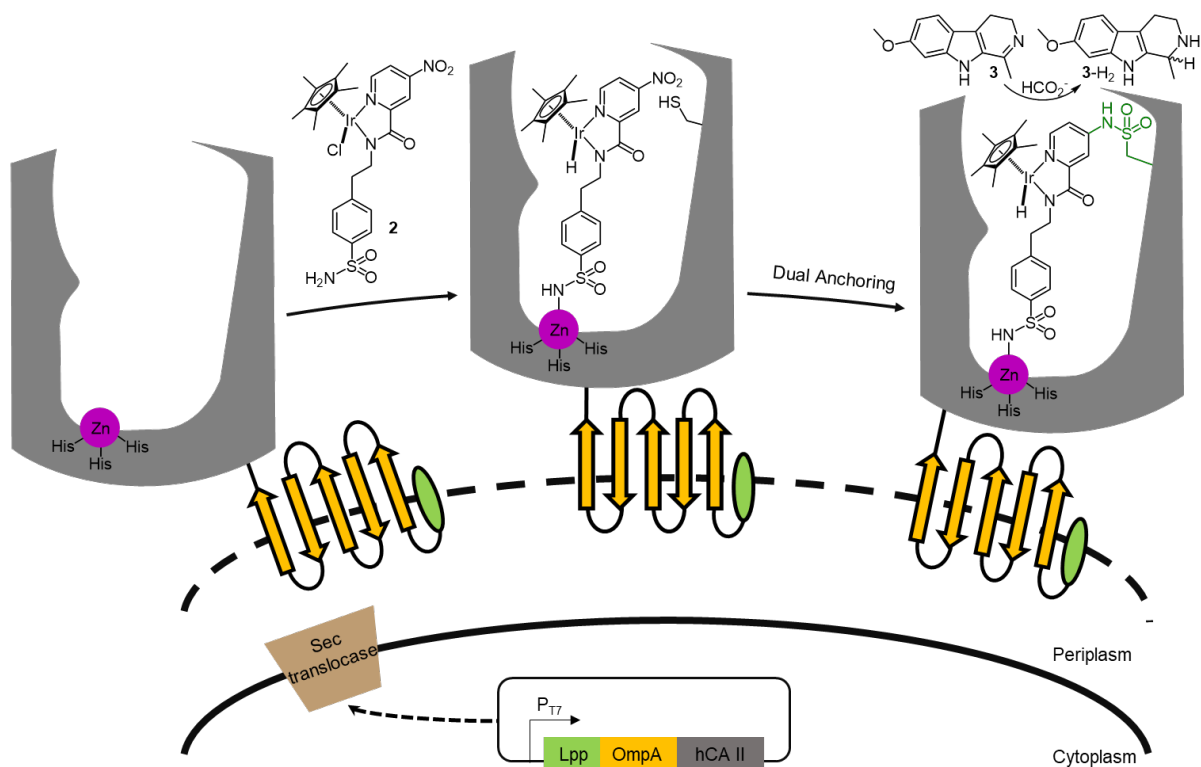
Figure 118 In cellulo assembly of artificial transfer hydrogenases (ATHases). hCAII is displayed on the cell surface of E. coli by fusion to a truncated lipoprotein and the outer membrane protein A (Lpp-OmpA). Following incubation with the sulfonamide-bearing IrCp piano-stool complex **2**, the resulting ArM catalyzes the enantioselective reduction of harmaline **3** in the presence of NaHCO$_2$ as hydride source.*

Next, we designed site saturation mutagenesis libraries for hCAII^{I91C}, focusing on the five closest residues to the chloride ligand of **2** as determined in the X-ray structure of **2** · hCAII^{I91C} (Figure 119b). The closest lying positions are: L60, N62, N67, E69 and Q92. To minimize the screening effort, we relied on a "small-intelligent" focused library approach by combining NDT, VMA and TGG codons.[323] We omitted the ATG codon to exclude methionine, thus affording a non-redundant library encoding 19 amino acids. We developed a workflow for the growth, expression and catalysis by *E. coli* displaying hCAII mutants in a 96-well format (Figure 119a). After induction of hCAII, followed by overnight expression (30 °C), the cells were harvested by centrifugation and washed with a [Cu(gly)$_2$] buffer to oxidize traces of thiols.[324,325] Next, the cells were resuspended in a buffer (pH 9.4, 50 mM (NH$_4$)$_2$CO$_3$) containing the cofactor **2** and incubated (1 h at 30 °C) to afford the dually-anchored ArMs on the cell surface. The supernatant was removed by centrifugation, the cells were resuspended in the reaction buffer (pH 7.4, 0.4 M MOPS) containing substrate **3** (50 µM) and NaHCO$_2$ (1

M). The 96-well plate was then shaken (16 h at 30°C, 300 rpm). Cells were sedimented by centrifugation, and the supernatant was transferred to a reaction tube (1.5 mL) for work-up. Finally, each sample was analyzed by supercritical fluid chromatography (SFC, using 2-methyl indoline as internal standard) to determine the product concentration and enantioselectivity. This procedure enables the screening of hundreds of mutants per week, substantially increasing the throughput compared to the conventional workflow that relies on purified mutants. The first round of directed evolution was performed by constructing a combinatorial library, mutating simultaneously asparagine 67 (N67) and glutamate 69 (E69), which are located 5.6 Å and 6.5 Å from the chloride ligand of cofactor **2**, respectively. Upon screening (> 650 colonies), we identified two mutants displaying an increased cell-specific activity compared to **2** · hCAII^{WT} and **2** · hCAII^{I91C}. The ATHase **2** · hCAII$^{N67G-E69R-I91C}$ had a higher cell specific activity (1.8 ± 0.2 (average ± s.d.) fold) compared to **2** · hCAII^{I91C} and an enhanced enantioselectivity for (*R*)-**3**-H$_2$ (80.0 ± 2.9% *ee*). To explore the influence of the pyridine-sulfonamide anchor, we expressed hCAII$^{N67G-E69R}$ lacking the I91C mutation. The corresponding **2** · hCAII$^{N67G-E69R}$, afforded a markedly reduced *ee* (10.0 ± 1.2 %) and a reduced activity compared to **2** · hCAII$^{N67G-E69R-I91C}$ for the reduction of harmaline **3**. Interestingly, the second identified mutant, **2** · hCAII$^{N67L-E69Y-I91C}$ afforded (*S*)-**3**-H$_2$ with a slightly higher cell specific activity compared to **2** · hCAII^{I91C} (1.2 ± 0.1-fold) and moderate *ee* (-48.0 ± 10.7%). Two additional rounds of directed evolution at positions L60 and N62 afforded a quintuple mutant hCAII$^{L60F-N62Y-N67G-E69R-I91C}$, which displayed a higher cell specific activity (2.8 ± 0.3-fold) and a similar *ee* (79 ± 0.4%) for enantiomer (*R*)-**3**-H$_2$ compared to **2** · hCAII^{I91C}. Starting from the (*S*)-selective ATHase **2** · hCAII$^{N67L-E69Y-I91C}$, the directed evolution campaign afforded the quadruple mutant hCAII$^{L60W-N67L-E69Y-I91C}$ which displayed the highest cell-specific activity (4.3 ± 0.4 fold) compared to hCAII^{I91C} and an improved *ee* (-59 ± 1.8%) for (*S*)-**3**-H$_2$ (Figure 119e). The purified mutants hCAII$^{N62D-N67L-E69Y-I91C}$ and hCAII$^{L60W-N62D-N67L-E69Y-I91C}$ displayed higher activities than the parent hCAII$^{N67L-E69Y-I91C}$ for several substrates as summarized below (Table 42). All mutants resulting from the site saturation mutagenesis library at position Q92 displayed a decreased activity. We hypothesize that the H–bond interaction between Q92 and the cofactor **2** plays an important role in maintaining the cofactor's position within the active site of the ATHase to favor the pyridine-sulfonamide linkage (Figure 119b and Figure 120).

178
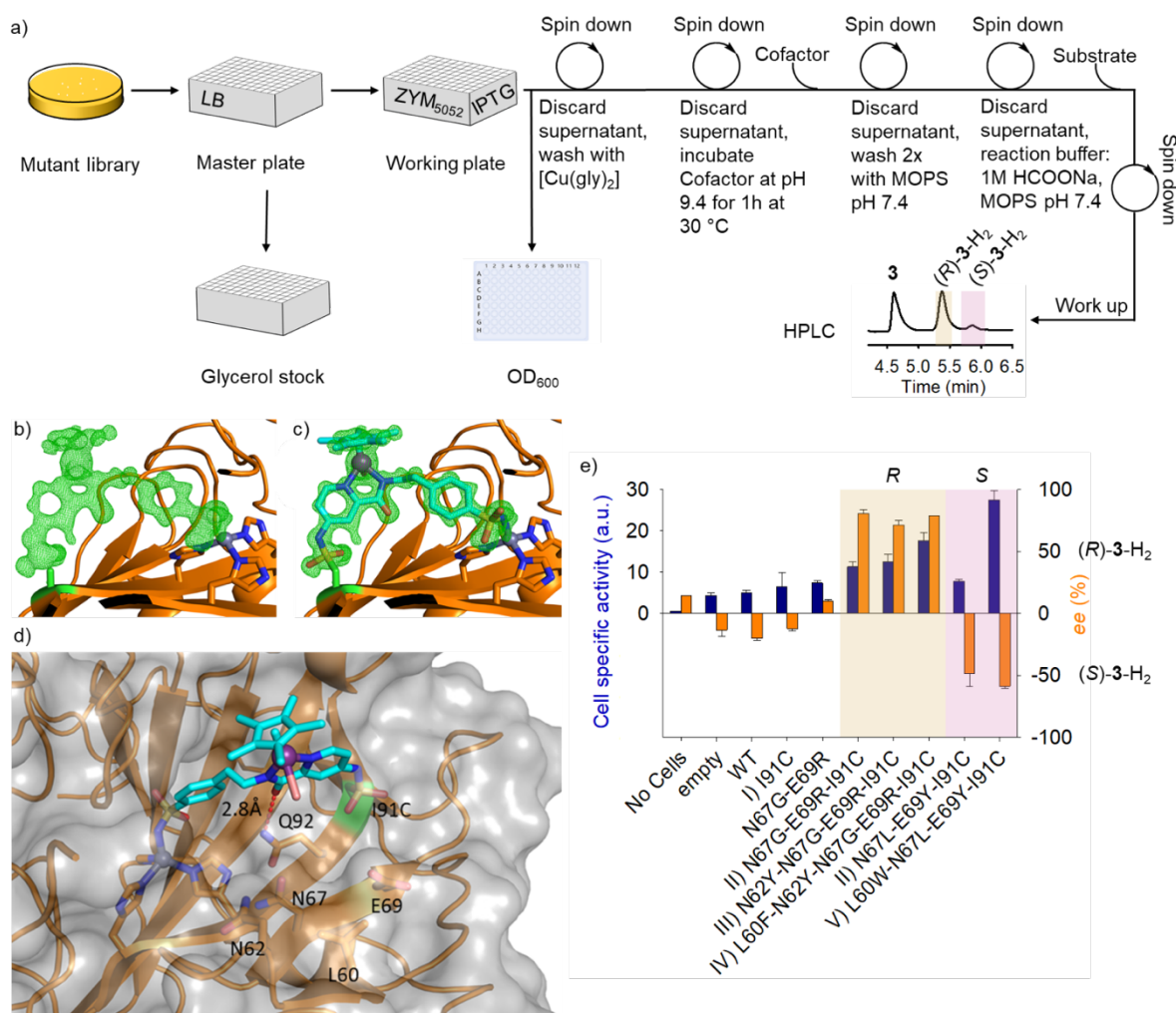
*Figure 119 Directed evolution of ArMs resulting from the dual anchoring of cofactor 2. **a**) A streamlined screening workflow was developed that enables expression of hCAII mutants and subsequent evaluation of their activity and selectivity in a 96 well-plate format. b-d) X-ray characterization of the ATHase **2** · hCAII^{I91C}. **b**) Fo-Fc difference map displayed as green mesh at 3σ without and **c**) with cofactor **2**. **d**) hCAIII91C is depicted as transparent grey surface and orange cartoon. Amino acids in the proximity of the cofactor are highlighted as sticks and labelled, the H-bond between Q92 and cofactor **2** is depicted as red dashes. The five residues (L60, N62, N67, E69 and Q92) closest to the catalytic iridium in hCAIII91C were randomized by site saturation mutagenesis. **e**) Summary of the results of the directed evolution for the production of tetrahydroharmine **3**-H₂. For each round of site saturation mutagenesis, the cell specific activity (product concentration/O.D.) and the enantiomeric excess (ee) of selected mutants are displayed as bar chart (TON, dark blue; ee, orange; enantiomer: R, positive, highlighted in gold; S, negative, highlighted in pink). Data are presented as the average of three independent replicate cultures and include the standard deviation. For the site-saturation library of Q92 no further improved mutant was identified.*

The hCAII mutant proteins were expressed in 2 L cultures and purified (up to 100 mg lyophilized protein). Transfer hydrogenation was then performed on harmaline **3** with purified proteins and cofactor **2** to validate the results obtained from the whole cell experiments. From the *in vitro* validation, we can draw the following conclusions: i) TON and

stereoselectivity of the evolved ATHases are higher than the parent cofactor **2** and **2** · hCAII^{WT} (Table 42, entry 2-3), confirming the efficiency of our strategy, ii) concerning (*R*)-selectivity, **2** · hCAII$^{N67G-E69R-I91C}$ is the best mutant, affording enantiomer (*R*)-**3**-H$_2$ (96% *ee*, 451 TON; Table 42, entry 5), iii) the best ATHase for (*S*)-**3**-H$_2$, only moderate *ee* are observed, with **2** · hCAII$^{N67L-E69Y-I91C}$ (-62% *ee*) and **2** · hCAII$^{L60W-N67L-E69Y-I91C}$ (-49% *ee*; Table 42, entry 8-9), iv) compared to **2** · hCAII$^{N67G-E69R-I91C}$, a reduced enantioselectivity was observed for the third and fourth generation ArMs. For the (*S*)- but also for the (*R*)-selective ArMs, the *in vitro* activity deviates from the observed whole cell transfer hydrogenation. *In vitro,* we observed that **2** · hCAII$^{N67L-E69Y-I91C}$ (458 TON) is more active than **2** · hCAII$^{L60W-N67L-E69Y-I91C}$ (221 TON) this activity profile contrasts with the whole cell assays. These observations suggest that the expression levels of hCAII$^{N67L-E69Y-I91C}$ and hCAII$^{L60W-N67L-E69Y-I91C}$ may differ from each other on the surface of *E. coli*. Next, we performed the transfer hydrogenation on a preparative scale in the presence of harmaline **3** (0.5 mmol) with **2** · hCAII$^{N67G-E69R-I91C}$ and **2** · hCAII$^{L60W-N67L-E69Y-I91C}$ (0.2 mol%). The ATHase **2** · hCAII$^{N67G-E69R-I91C}$ performs very well (97% *ee*, 75% isolated yield, 84 h) for (*R*)-**3**-H$_2$ and ArM **2** · hCAII$^{L60W-N67L-E69Y-I91C}$ afforded (*S*)-**3**-H$_2$ with good yield and moderate *ee* (86% yield, -43% *ee*, 84 h). Relying on hCAII^{I91C} as a starting point, the directed evolution of a dually-anchored ATHase led to the identification of significantly improved artificial metalloenzymes, both in terms of activity (up to 451 TONs) and selectivity (up to 97% ee for (*R*)-**3**-H$_2$ and -62 % ee for (*S*)- **3**- H$_2$).

Table 42 Selected Catalytic Results Obtained with Purified 2·hCAIImutants for the Transfer Hydrogenation of Harmaline 3a. [a]Reactions were performed over 16 h using hCAII (10 μM), cofactor 2 (10 μM), and substrate (5 mM) in MOPS/NaHCO2 buffer (pH 7.4, 0.34 M, 0.85 M). Values are the mean values of duplicate experiments. [b]Experiments were performed on a preparative scale with 0.5 mmol of harmaline 3.

	Mutants	TON	*ee* (%)
1	None	311	0
2	hCAII	149	31 (R)
3	hCAII[E69C]	278	48 (R)
4	hCAII[I91C]	265	39 (S)
5	hCAII[N67G-E69R-I91C]	451	96 (R)
6	hCAII[N62Y-N67G-E69R-I91C]	460	73 (R)
7	hCAII[L60F-N62Y-N67G-E69R-I91C]	415	74 (R)
8	hCAII[N67L-E69Y-I91C]	458	62 (S)
9	hCAII[N60W-N67L-E69Y-I91C]	221	9 (S)
10	hCAII[N67G-E69R-I91C]*	375	97 (R)
11	hCAII[N60W-N67L-E69Y-I91C]*	430	43 (S)

We attempted to crystallize the complexes of cofactor **2** bound to the best mutant of each directed evolution round. The most (*R*)-selective ATHase **2** · hCAII[N67G-E69R-I91C] afforded crystals suitable for X-ray analysis. The overall structures of **2** · hCAII[I91C] and **2** · hCAII[N67G-E69R-I91C] are virtually identical, reflected by a C$_\alpha$ RMSD (0.340 Å). Moreover, the position of cofactor **2**, including the pyridine-sulfonamide linkage resulting from the dual anchoring, the interactions between hCAII[N67G-E69R-I91C] and the protein as well as the (*S*)-configuration of the Ir–Cl, is the same in these two mutants (Figure 137c and Figure 138d). These observations confirm that the dual anchoring locks the cofactor **2** in a defined configuration and position. On the one hand, the mutation of asparagine N67 to glycine generates more space (63.7 Å^3)[326] for the substrate to approach the Ir–H moiety. On the other hand, the exchange of glutamate E69 by a cationic arginine might lead to cation-π interactions with the electron-rich substrate **3**. We thus hypothesize that mutations N67G and E69R may improve the binding of the substrate, thereby giving rise to increased TONs and enantioselectivity (Table 42 and Table 43). Due to the dual anchoring of cofactor **2**, N67 and E69 are not needed for cofactor binding, as previously described for the computational improvement of the hCAII binding pocket.
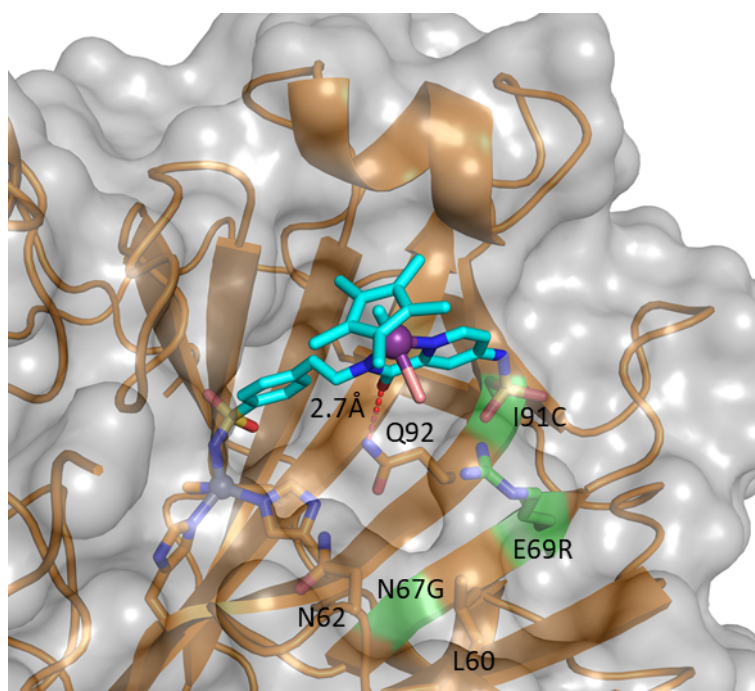
Figure 120 Crystal structure of the evolved $2 \cdot hCAII^{N67G-E69R-I91C}$. The protein is depicted as an orange cartoon and as a transparent gray surface. The cofactor 2 is displayed in color-coded sticks, and Ir-atoms and Zn-atoms are depicted as purple and gray spheres, respectively. Amino acids in the proximity of the cofactor are highlighted as green sticks, labeled, mutated amino acids are highlighted in green, and the H-bond between Q92 and cofactor 2 is depicted as red dashes.

Covalently anchoring. As pointed out by an insightful referee, Flemming and coworkers reported that a sulfonamide linkage was formed as a product resulting from the photoreduction of an aromatic nitro group which was trapped by a thiol, followed by redox chemistry at both the nitrogen and sulfur atoms.[327]

In order to confirm that the bioconjugation step is not the result of photoreduction caused by the X-ray beam, the ATHases $2 \cdot hCAII^{E69C}$, $2 \cdot hCAII^{I91C}$ and $2 \cdot hCAII^{N67G-E69R-I91C}$ as well as the control ATHase $2 \cdot hCAII^{WT}$ were subjected to MS analysis.

i) Under native MS conditions, $2 \cdot hCAII^{WT}$ $2 \cdot hCAII^{E69C}$, $2 \cdot hCAII^{I91C}$ and $2 \cdot hCAII^{N67G-E69R-I91C}$ afford, after deconvolution, a main mass signal at 29836.9 Da, 29812.1 Da, 29827.0 Da and 29797.7 Da, respectively (Table 43 and Figure 121-123). These masses correspond to the ATHases (hCAII plus cofactor 2). To further investigate the formation of the sulfonamide linkage, cofactor 2 and hCAII$^{N67G-E69R-I91C}$ were incubated in the dark and under ambient light. Native MS revealed almost identical masses, either with or without light: 29797.7 Da and

29797.8 Da, respectively (Figure 124 and Figure 125). We thus conclude that the formation of the covalent sulfonamide linkage does not require light and is not the result of X-ray photo-damage.

ii) Under denaturing MS conditions, $2 \cdot hCAII^{E69C}$, $2 \cdot hCAII^{I91C}$ $2 \cdot hCAII^{N67G-E69R-I91C}$ afford signals at 29749.0 Da, 29763.9 Da, and 29733.5 Da, respectively (Table 43 and Figure 126-128). These correspond to the mass of the bioconjugates, minus the Zn ion. For the above ArMs, a mass of ~675 Da was observed for cofactor 2 under native and denaturing conditions. This corresponds to the cofactor 2 without the Cl ion. This confirms that cofactor 2 remains tightly bound to $hCAII^{E69C}$, $hCAII^{I91C}$ and $hCAII^{N67G-E69R-I91C}$, despite the loss of the Zn-sulfonamide dative bond under denaturing conditions. For the ArM lacking the covalent sulfonamide linkage (i.e., $2 \cdot hCAII^{WT}$), only the mass of apo $hCAII^{WT}$ (29098.2 Da, minus the Zn ion and no cofactor bound) is detected (Table 43). iii) Tryptic digest of $2 \cdot hCAII^{E69C}$, $2 \cdot hCAII^{I91C}$ and $2 \cdot hCAII^{N67G-E69R-I91C}$ revealed that the fragments containing the cysteine had a modification of ~675 Da (Table 48, Figure 130-132). The modification of 675 Da corresponds to the cofactor lacking Cl – as observed for both native and denaturing MS. Furthermore, the modification of ~675 Da was detected either for the fragment containing C69 (for $2 \cdot hCAII^{E69C}$) or C91 (for $2 \cdot hCAII^{I91C}$ and $2 \cdot hCAII^{N67G-E69R-I91C}$). Again here, the modification was observed even in the absence of UV light. In contrast, analysis of $2 \cdot hCAII^{WT}$ digested by trypsin did not reveal any peptide containing a modification of ~675 Da.

From this data, we conclude that the sulfonamide bioconjugation results from the proximity of the aromatic nitro group and the thiol and does not require UV irradiation. It should be noted however that, based on X-ray and MS data, we cannot exclude the formation of a sulfonate rather than sulfonamide covalent linkage.

To confirm the positive influence of the covalent bioconjugation on catalytic performance, the enantioselective reduction of harmaline 3 was monitored in the presence of $2 \cdot hCAII^{WT}$, $2 \cdot hCAII^{E69C}$ and $2 \cdot hCAII^{N67G-E69R-I91C}$ (Figure 140). From these data, the following trends emerge:

i) The catalytic performance (both yield and *ee*) of $2 \cdot hCAII^{WT}$ and $2 \cdot hCAII^{E69C}$ is little affected by incubation time following the mixing of the cofactor 2 and the host hCAII (Figure 140a-b). As revealed by X-ray crystallography (Figure 134c) and MS analysis (Figure 122, Figure 127),

prolonged incubation only leads to partial formation of the sulfonamide covalent linkage in **2** · hCAII[E69C].

ii) In contrast, for **2** · hCAII[N67G-E69R-I91C], the catalytic performance is affected by the incubation time, prior to the addition of substrate **3** and formate. The catalytic performance of the covalently-anchored ArM **2** · hCAII[N67G-E69R-I91C] (resulting from 6 hours incubation) clearly outperforms the non-preincubated **2** · hCAII[N67G-E69R-I91C] (Figure 140c).

iii) As the pH requirements for the bioconjugation step and catalysis differ (pH = 9.4 and 7.4, respectively), the covalent anchor is not formed in significant amounts under catalytic conditions.

We thus conclude that both cofactor localization and catalytic performance of **2** · hCAII[N67G-E69R-I91C] are positively affected by the presence of an additional covalent linkage between the host protein and the catalytically active piano stool cofactor **2**.

Table 43 Mass spectrometry analysis of selected ATHases [a]Protein mass plus Zn (65.4 Da) and one H2O (18.0 Da). [b]Protein mass plus Zn (65.4 Da) and cofactor 2 without Cl (675.1 Da). [c]Protein mass without Zn. [d]Protein mass plus cofactor 2 without Cl (675.1 Da). [e]Similar masses are highlighted in light green.

Native MS

Protein	$M_{obs.}$ main peak (Da)	$M_{calc.}$ apoprotein[a] (Da)	$M_{calc.}$ ATHase[b] (Da)
2 · hCAII[WT]	29836.9	29181.2	29838.3
2 · hCAII[E69C]	29812.1	29155.3	29812.4
2 · hCAII[I91C]	29827.0	29171.2	29828.3
2 · hCAII[N67G-E69R-I91C]	29797.7	29141.3	29798.4
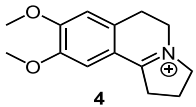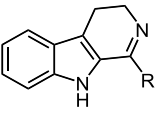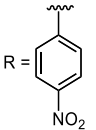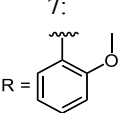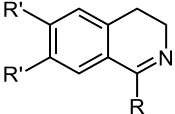
Denaturing MS
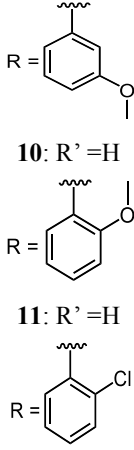
Protein	$M_{obs.}$ main peak (Da)	$M_{calc.}$ apoprotein[a] (Da)	$M_{calc.}$ ATHase[b] (Da)
2 · hCAII[WT]	29098.2	29097.8	29772.9
2 · hCAII[E69C]	29749.0	29071.9	29747.0
2 · hCAII[I91C]	29763.9	29087.8	29762.9
2 · hCAII[N67G-E69R-I91C]	29733.5	29057.9	29733.0

Substrate screening. To evaluate the substrate scope, the purified hCAII mutants were screened for the transfer hydrogenation of various substrates **4**-**11**. The substrates **5**-**7** are derived from β-carbolines and substrates **8**-**11** from dihydroisoquinoline. Selected results are displayed in Table 44 and the full screening data can be found in the SI (Table 49-52). The same conditions were used for all substrates. Mutants hCAII[N62D-N67L-E69Y-I91C] and hCAII[L60W-N62D-N67L-E69Y-I91C] obtained during the directed evolution were also tested for catalysis. From the screening, the following trends are apparent: All substrates were converted with good TON (> 200 for at least one ArM) and with moderate to good enantioselectivity. Substrates with electron-rich substituents give higher TON compared to electron-poor substrates. Finally, lower TON give lower enantioselectivity. As model substrate we investigated the enantioselective reduction of substrate **4** to crispine A. Mutations from the parent hCAII[I91C] are deleterious for **2** · hCAII[N67G-E69R-I91C] but further mutations at positions N62 and L60 improve the TON (219) and the *ee* (62%) for **2** · hCAII[L60F-N62Y-N67G-E69R-I91C] (Table 44, Entry 1). Mutations N67L and E69Y have a positive effect on enantioselectivity (76% *ee*, SI) and the last generation improved both *ee* (86%) and TON (355) for **2** ·hCAII[L60W-N62D-N67L-E69Y-I91C] (Table 44, Entry 2). Next, we tested harmalane **5**, whose structure is similar to harmaline **3** but lacks the electron donating methoxy group in position 7. The ArM **2** · hCAII[N62Y-N67G-E69R-I91C] gives the best results with complete conversion and good *ee* (500 TON, 87% *ee* (*R*)) (Table 44, entry 3) highlighting the tolerance towards substitutions on the indole moiety. For the electron-deficient harmalane-analog **6**, the TONs decrease compared to substrate **5** (Table 52): **2** · hCAII[N67G-E69R-I91C] yields the best *ee* (80% (*S*)) and moderate TON (191). Instead, **2** · hCAII[L60W-N62D-N67L-E69Y-I91C] produces the opposite enantiomer (72% *ee* (*R*), 249 TON) (Entry 5-6). I nterestingly, **2** · hCAII[N67G-E69R-I91C] is the most selective mutant for harmaline **3** as well as substrate **6.** The reduced activity of the ATHases for this substrate may be traced back to steric hindrance and mismatched electronics caused by the *p*-nitrophenyl group. In contrast, the electron-rich substrate **7** affords higher TONs at the cost of a lower *ee* (Table 53). The best ArM **2** · hCAII[I91C] gave (*S*)-**7**-H$_2$ (296 TON, 68% *ee*, Table 44, Entry 7). Next, we turned our attention to dihydroisoquinolines. The less-hindered substrate **8** gives high TON but moderate enantioselectivity, with the most enantioselective ArM being **2** · hCAII[N67L-E69Y-I91C] (500 TON, 59% *ee* (*S*), Table 44, Entry 9). To evaluate the tolerance of transfer hydrogenation towards substrates presenting steric hindrance next to the imine moiety, we tested substrates **9**-**11**.[328]

Substrate **9** is reduced with good activities but low stereoselectivity (Table 55). The most promising ArM is **2** · hCAII^L60F-N62Y-N67G-E69R-I91C (415 TON, -55% *ee*, Table 44, Entry 12). Next, the steric hindrance was increased by introducing an *o*-methoxyphenyl substituent in substrate **10**. The results (Table 56) are rather different from the previous ones, the mutants derived from **2** · hCAII^N67G-E69R-I91C give lower turnovers compared to the mutants derived from **2** · hCAII^N67L-E69Y-I91C. The best ArM for substrate **10** is **2** · hCAII^N62D-N67L-E69Y-I91C (338 TON, 75% *ee*) (Table 44, Entry 13). It should be noted that **2** · hCAII^E69C gives the opposite enantiomer (-71% *ee*, 242 TON, Table 44, Entry 14). Finally, substrate **11** incorporating an *o*-chloro substituent was evaluated. In this case, both TON and enantioselectivities are low compared to the other substrates (Table 57) with the most promising catalyst being **2** · hCAII^N62D-N67L-E69Y-I91C (Table 44, Entry 15).

Table 44 Selected Results of the Substrate Screening in the presence of Evolved ATHases^a. aThe reactions were performed using cofactor 2 (10 µM), hCAII (10 µM), and substrate (5 mM) in NaHCO 2 /MOPS buffer (850 mM/ 340 mM, 25 °C, 16 h). The reported TON and ee are the mean values of independent duplicates. Complete screening data are included in the Supporting Information (Table 49-41).

	Substrate		Mutants	TON	*ee* (%)
1	6,7-dimethoxy-... **4**		hCAII^L60F-N62Y-N67G-E69R-I91C	219	62
2			hCAII^L60W-N62D-N67L-E69Y-I91C	355	86
3		**5**: R = Me	hCAII^N62Y-N67G-E69R-I91C	500	87 (*R*)
4		**6**: R = (4-NO2-phenyl)	hCAII^L60W-N67L-E69Y-I91C	189	72 (*S*)
5			hCAII^N67G-E69R-I91C	191	80 (*S*)
6			hCAII^L60W-N62D-N67L-E69Y-I91C	249	72 (*R*)
7		**7**: R = (2-methoxyphenyl)	hCAII^I91C	296	68 (*S*)
8			hCAII^L60W-N62D-N67L-E69Y-I91C	372	43 (*R*)
9	R' ... N, R	**8**: R' = OMe R = Me	hCAII^N67L-E69Y-I91C	500	59 (*S*)
10			hCAII^N62Y-N67G-E69R-I91C	422	58 (*R*)
11		**9**: R' = H	hCAII^E69C	250	-56
12			hCAII^L60F-N62Y-N67G-E69R-I91C	415	-55

13		R = (3-methoxyphenyl)	hCAII$^{N62D-N67L-E69Y-I91C}$	338	75
14		**10**: R' =H	hCAII^{E69C}	242	-71
15		R = (2-methoxyphenyl) **11**: R' =H	hCAII$^{N62D-N67L-E69Y-I91C}$	215	-51
16		R = (2-chlorophenyl)	hCAII$^{N62Y-N67G-E69R-I91C}$	47	47

5.4.5 Conclusion

With the aim of firmly anchoring the Ir-pianostool cofactor **2** within hCAII, we introduced a nucleophilic residue hCAII^{I91C} in the proximity of the p-nitropicolinamide ligand bound to the catalytically-competent Ir-ion. Rather than affording the anticipated SNAr reaction, a pyridine-sulfonamide linkage was formed, thus resulting in a dually-anchored cofactor. Starting from **2** · hCAII^{I91C}, three rounds of directed evolution were performed using E. coli cell surface-display. This led to the identification of the (R)-selective ATHase **2** · hCAII$^{N67G-E69R-I91C}$ (up to 96% ee and 451 TONs) for tetrahydro-harmine (R)-3-H$_2$. The opposite enantiomer (S)-3-H$_2$ was obtained using **2** · hCAI$^{IL67L-E69Y-I91C}$ (-63% ee, 458 TON). We hypothesize that mutation N67G generates more space for the substrate to access the cofactor whereas the mutation E69R stabilizes the substrate by cation-π interactions.

A substrate scope, carried out with the best mutants, revealed a marked substrate-dependency, thus requiring independent directed evolution efforts for each substrate. This suggests that, starting from **2** · hCAII^{I91C}, the evolutionary trajectory has crafted a close fit between the active site and the Ir–H····harmaline transition state, which is less suitable for structurally-unrelated substrates. The presence of a single and well-defined cofactor conformation within hCAII resulting from the dual anchor, leads to increased TONs and an improved enantioselectivity. Locking the cofactor in place via an additional covalent bond allows randomizing neighboring amino acids to improve catalytic performance without affecting the position of the cofactor. Accordingly, an ArM with a dual anchor is a propitious starting point for a directed evolution campaign. Finally, the unanticipated reaction between

an aromatic nitro group and a cysteine is a welcome addition to the well-established cysteine bioconjugation tools.

5.4.6 Experimental

Electron-Spray Ionization Mass Spectra (ESI-MS) were recorded on a Bruker FTMS 4.7T bioAPEX II. High-resolution mass spectra (HRMS) were measured on a Bruker maXis II QTOF ESI mass spectrometer.

Native MS analysis. The experiments were adapted from the inhibitor binding studies to hCAI examined by Zoppi *et al*.[329] To assemble the ArMs, the protein was dissolved in ammonium carbonate buffer (50 mM $(NH_4)_2CO_3$, pH 9.4) to a concentration of 2 mg/mL, supplemented with 0.1 mM cofactor **2** (10 mM stock in DMSO) and incubated while shaking (18 h, 25°C, 600 rpm). For hCAII$^{N67G-E69R-I91C}$ two sets of samples were prepared. The first set of samples were prepared in amber vials and wrapped with aluminum foil to avoid irradiation. For the second set, transparent vails were used. After incubation, the buffer was exchanged to remove the unbound cofactor by ultracentrifugation (Amicon® ultra centrifugal filters, Merck; cut-off of 10 kDa; 14'000 g, 5 min RT; dilution buffer: 20 mM $NH_4C_2H_3O_2$, pH 8.5). The concentration and dilution steps were repeated three times. The protein concentration was determined with a NanoDrop microvolume spectrophotometer (Thermo Fisher Scientific, USA) and adjusted to 2 mg/mL (20 mM $NH_4C_2H_3O_2$, pH 8.5). Electrospray ionization (ESI) low concentration tuning mix from Agilent was used as calibrant. The samples were directly injected into the high-resolution MS (HRMS) with a syringe pump using a flow rate of 5 µL/min. The ESI source parameters were optimized for the hCAII analysis. The charge-state distribution from 2200-3000 m/z was deconvoluted using the Compass Data Analysis software (Bruker Daltonics) with the maximum entropy setup.

Table 45 MS settings for native MS analysis of the hCAII complexes.

Source Type	ESI	Ion Polarity	Positive	Set Nebulizer	1.8 Bar
Focus	Active	**Set Capillary**	4000 V	**Set Dry Heater**	220 °C
Scan Begin	1000 m/z	**Set End Plate Offset**	-500 V	**Set Dry Gas**	3.0 l/min
Scan End	8000 m/z	**Set charging Voltage**	2000 V	**Set Divert Valve**	Source

Denaturing MS analysis. To assemble the ArMs, the protein was dissolved in ammonium carbonate buffer (50 mM $(NH_4)_2CO_3$, pH 9.4) to a concentration of 2 mg/mL, supplemented with 0.1 mM cofactor **2** (10 mM stock in DMSO) and incubated while shaking (2 h, 25°C, 600 rpm). The buffer was exchanged to remove the unbound cofactor by ultracentrifugation (Amicon® ultra centrifugal filters, Merck; cut-off of 10 kDa; 14'000 g, 5 min RT; dilution buffer: ($NH_4C_2H_3O_2$, pH 8.5)). The concentration and dilution steps were repeated three times. The protein concentration was determined by a NanoDrop microvolume spectrophotometer (Thermo Fisher Scientific, USA) and adjusted to 0.2 mg/mL ($MQ-H_2O$). A HPLC (Shimadzu, equipped with a Jupiter® 5 μm C4 300 Å)-ESI QTof (Bruker Daltonics, ESI Maxis II QTof MS) system was used to record the data. As eluent water: formic acid (A, 1000 : 1) and acetonitrile : formic acid (B, 1000 : 1) were used. The time program was as follows; 2 min 10% B, 7 min 10-55%, 2 min 55%, 2 min 55-65%, 2 min 65%, 2 min 65-10%, 8 min 10%. The formic acid in the mobile phase leads to a pH drop in the sample, which denatures the protein leading to the release of the Zn ion and the cofactor **2** if it is not covalently bound (e.g. hCAII[WT]). An internal calibration with Na-formate was performed for each sample. The charge-state distribution from 650-1300 m/z was deconvoluted using the Compass Data Analysis software (Bruker Daltonics) with the maximum entropy setup.

Table 46. MS settings for protein mass spectral analysis.

Source Type	ESI	**Ion Polarity**	Positive	**Set Nebulizer**	2.0 Bar
Focus	Active	**Set Capillary**	4500 V	**Set Dry Heater**	220 °C
Scan Begin	150 m/z	**Set End Plate Offset**	-500 V	**Set Dry Gas**	8.0 l/min
Scan End	2750 m/z	**Set charging Voltage**	0 V	**Set Divert Valve**	Source
		Set Corona	4000 nA	**Set APCI Heater**	200 °C

Figure 121 Native MS of 2 · hCAIIWT a) Charge-state distribution of 2 · hCAII^WT. b) Deconvoluted spectrum of a). The peak at 29836.9 Da corresponds to the mass of 2 · hCAII^WT with Zn (Mcalc = 29838.3 Da) and the adduct with one water molecule 29855.3 Da (Mcalc = 29856.3 Da).
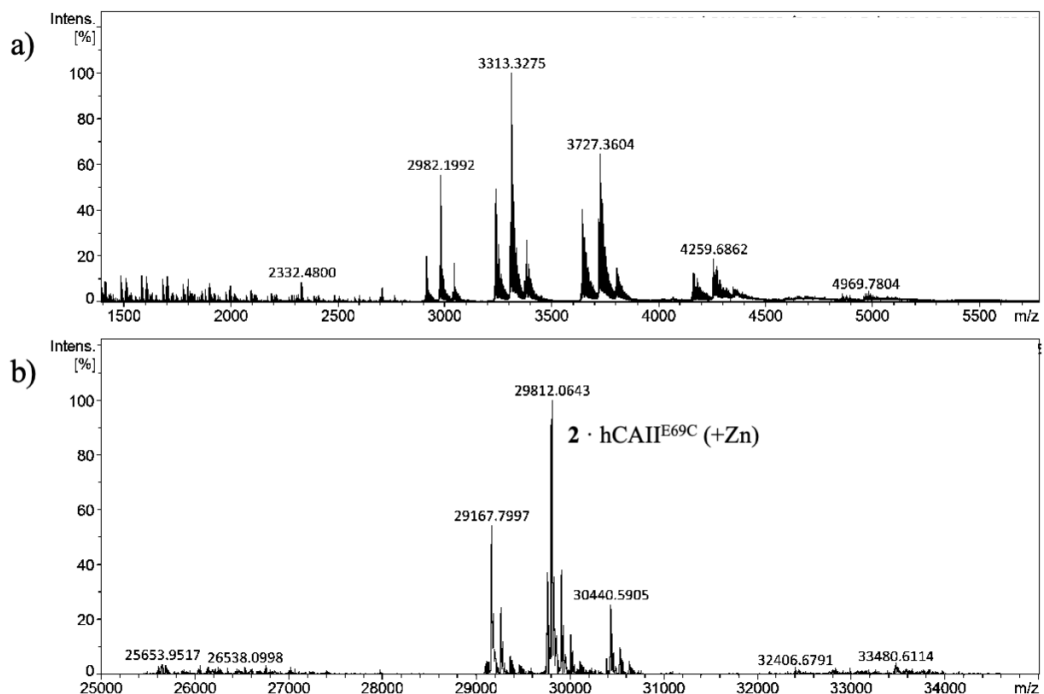
*Figure 122 Native MS of 2 · hCAII^E69C a) Charge-state distribution of **2** · hCAII^E69C. b) Deconvoluted spectrum of a). The main peak at 29812.1 Da corresponds to the mass of **2** · hCAII^E69C with Zn (M_{calc} = 29812.4 Da).*
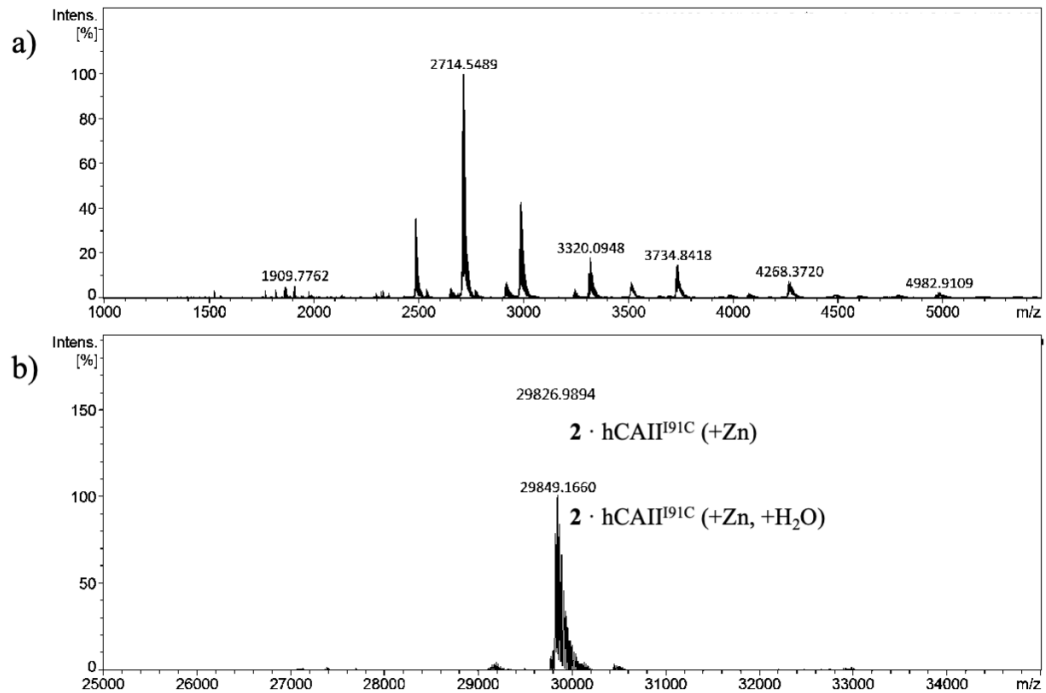
*Figure 123 Native MS of 2 · hCAII^{I91C} a) Charge-state distribution of **2** · hCAII^{I91C}. b) Deconvoluted spectrum of a). The peak at 29849.2 Da corresponds to the mass for **2** · hCAII^{I91C} with Zn (M_{calc} = 29828.3 Da) and the adduct with one water molecule 29849.2 Da (M_{calc} = 29846.3 Da).*
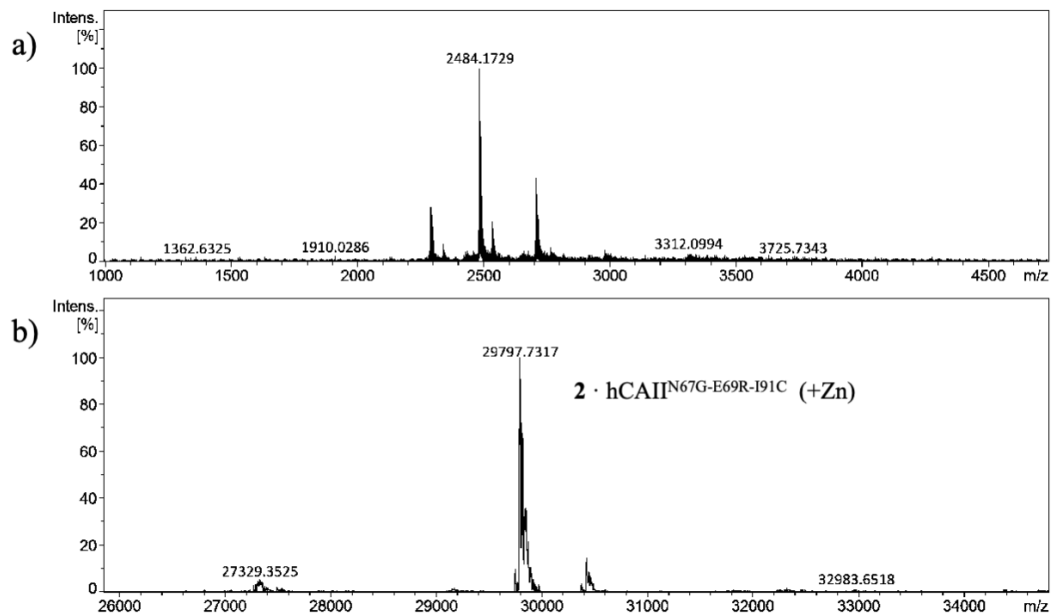
*Figure 124 Native MS of 2 · hCAII^{N67G-E69R-I91C}. a) Charge-state distribution of **2** · hCAII^{N67G-E69R-I91C}. b) Deconvoluted spectrum of a). The main peak at 29797.7 Da corresponds to the mass of **2** · hCAII^{N67G-E69R-I91C} with Zn (M_{calc} = 29797.7 Da). The sample was incubated under laboratory light.*
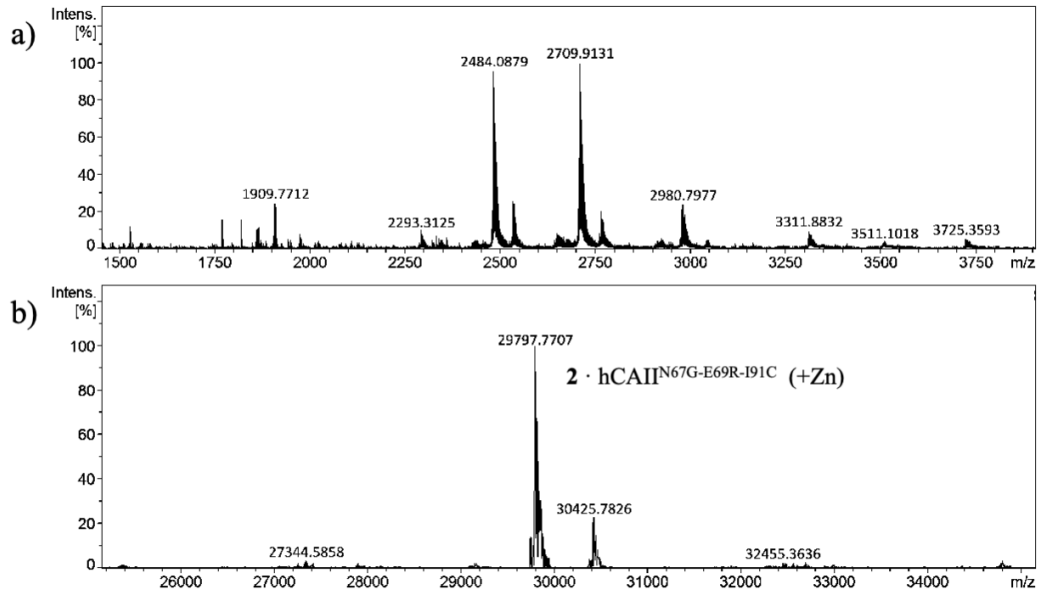
*Figure 125 Native MS of 2 · hCAII^{N67G-E69R-I91C} incubated and measured in the dark a) Charge-state distribution of **2** · hCAII^{N67G-E69R-I91C}. b) Deconvoluted spectrum of a). The main peak at 29797.8 Da corresponds to the mass of **2** · hCAII^{N67G-E69R-I91C} with Zn (M_{calc} = 29797.7 Da).*
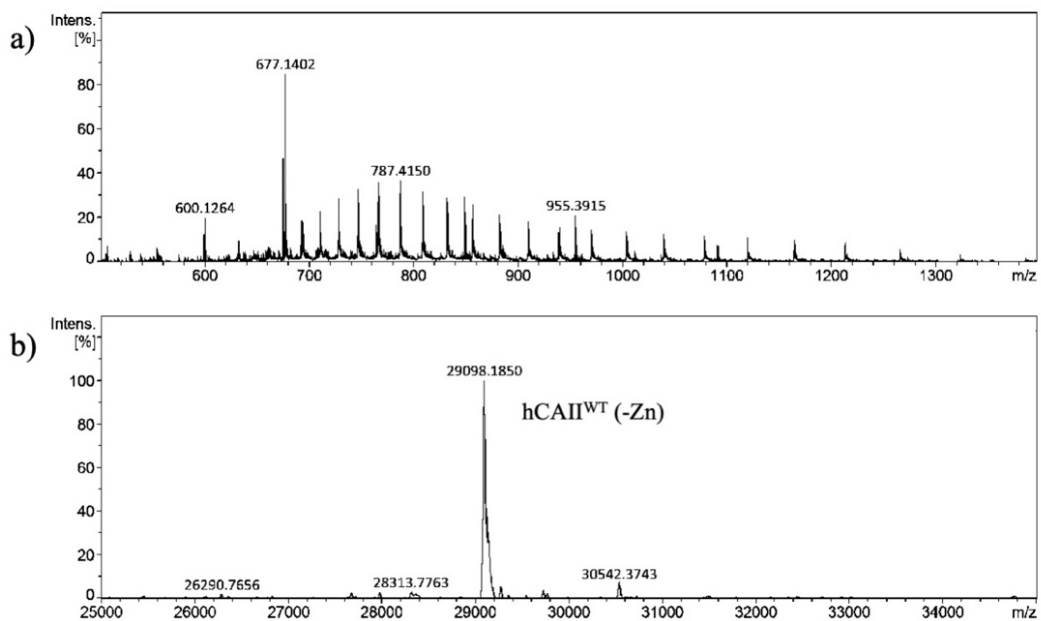
*Figure 126 Denaturing MS of 2 ·hCAII^{WT}. a) Charge-state distribution of **2** ·hCAII^{WT}. The charge state at 677.14 m/z corresponds to the free cofactor, which is released. b) Deconvoluted spectrum of a). The main peak at 29098.2 Da corresponds to the mass of hCAII^{WT}, without Zn (M_{calc} = 29098.2 Da). Due to the pH drop caused by the addition of 0.1% formic acid from the mobile phase, the tertiary structure is disrupted and the Zinc, as well as the cofactor **2**, are displaced from the ArM.*
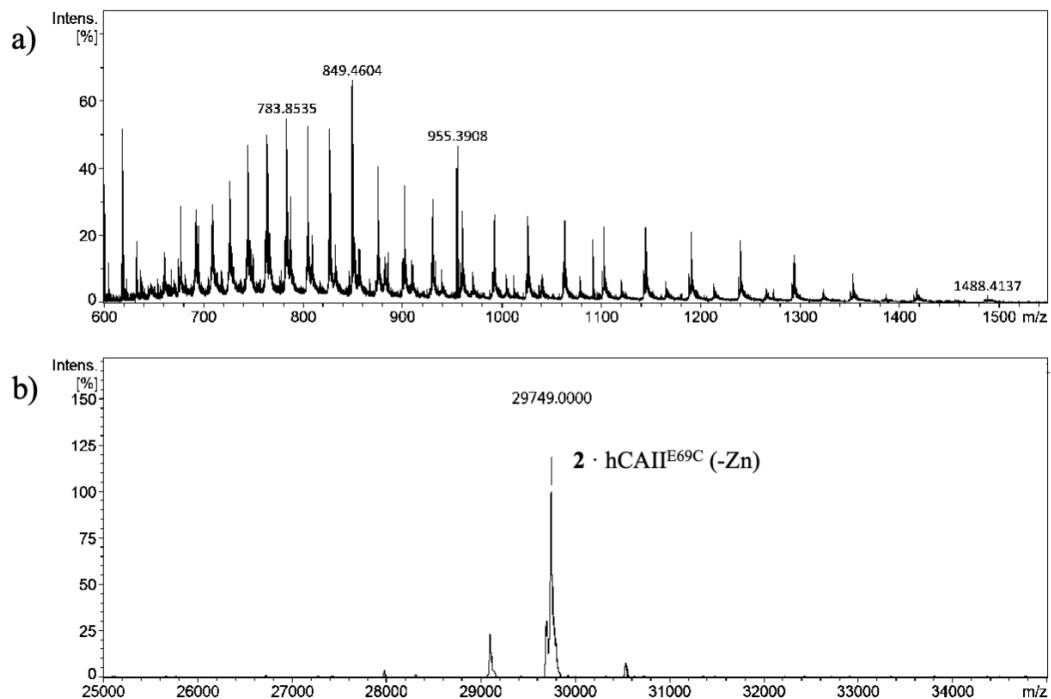
Figure 127 Denaturing MS of 2 ·hCAII^{E69C}. a) Charge-state distribution of 2 ·hCAII^{E69C}. b) Deconvoluted spectrum of a). The main peak at 29749.0 Da corresponds to the mass of 2 · hCAII^{E69C}, without Zn (M_{calc} = 29747.0 Da). Due to the pH drop caused by the addition of 0.1% formic acid from the mobile phase, the tertiary structure is disrupted and the Zinc is displaced from the ArM.
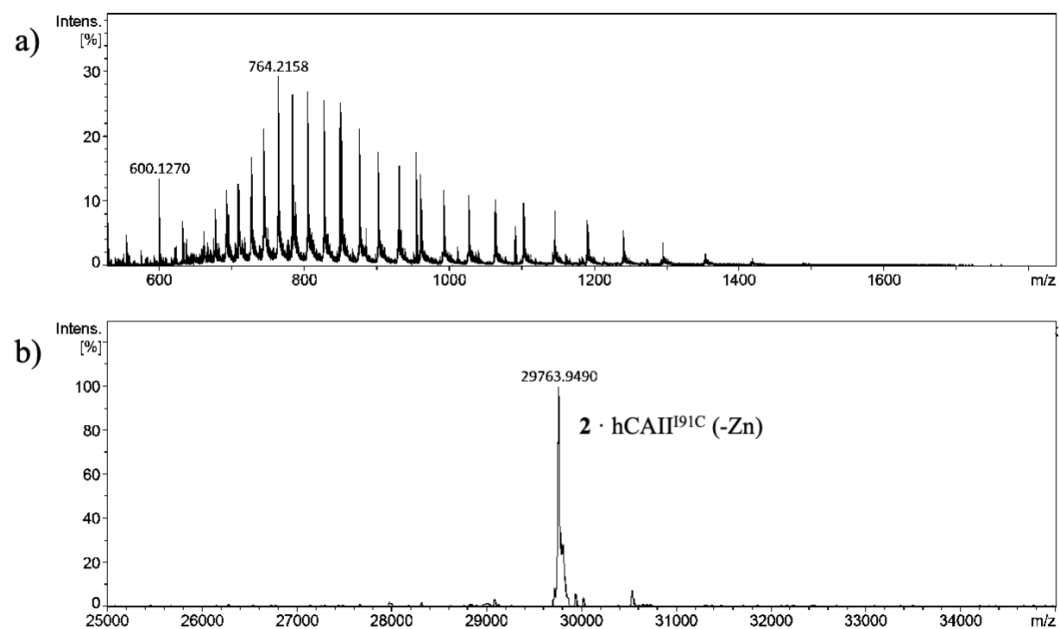
Figure 128 Denaturing MS of 2 ·hCAII^{I91C}. a) Charge-state distribution of 2 ·hCAII^{I91C}. b) Deconvoluted spectrum of a). The main peak at 29764.0 Da corresponds to the mass for 2 · hCAII^{I91C} without Zn (M_{calc} = 29762.9 Da). Due to the pH drop caused by the addition of 0.1% formic acid, the tertiary structure is disrupted and the Zinc is displaced from the ArM.

193
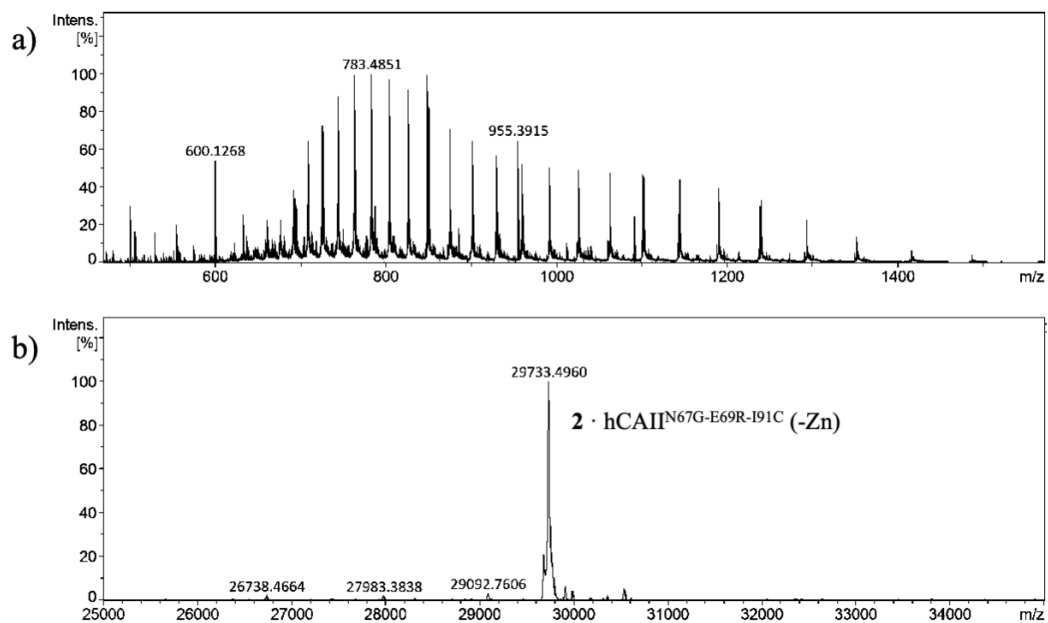
Figure 129 Denaturing MS of 2 ·hCAII^{N67G-E69R-I91C}. a) Charge-state distribution of 2 ·hCAII^{N67G-E69R-I91C}. b) Deconvoluted spectrum of a). The main peak at 29733.5 Da corresponds to the mass of 2 · hCAII^{N67G-E69R-I91C} without Zn (M_{calc} = 29733.0 Da). Due to the pH drop caused by the addition of 0.1% formic acid, the tertiary structure is disrupted and the Zinc is displaced from the ArM.
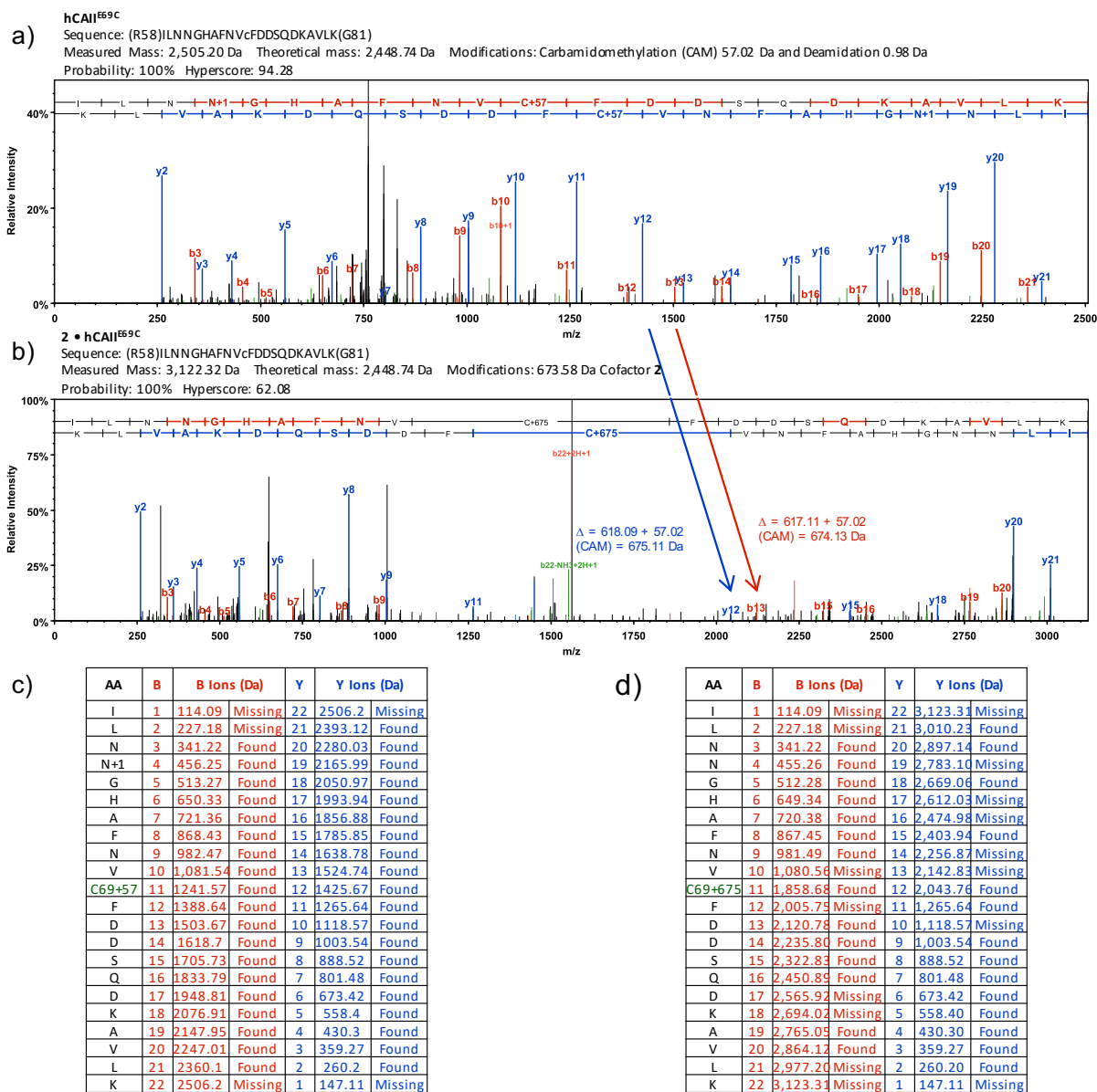
Figure 130 **Tryptic digest of hCAII^{E69C} and 2 · hCAII^{E69C} analyzed by LC-MS/MS.** *MS/MS spectrum of hCAII^{E69C} (a) and* **2 ·** *hCAII^{E69C} (b). The analyzed peptide sequence, measured and theoretical mass, modification and quality indicators are shown. The list of analyzed ions from (a) and (b) are given in (c) and (d), respectively. Carbamidomethylation is abbreviated with CAM.*

a) hCAII^{I91C}

Sequence: (R89)LcQ.FHFHWGSLDGQGSEHTVDK(K112)
Measured Mass: 2,584.16 Da   Theoretical mass: 2,528.74 Da   Modifications: Carbamidomethylation (CAM) 57.02 Da
Probability: 100% Hyperscore: 97.00

b) 2 • hCAII^{I91C}

Sequence: (R89)LcQ.FHFHWGSLDGQGSEHTVDKK(K113)
Measured Mass: 3,330.36 Da   Theoretical mass: 2,656.94 Da   Modifications: 673.42 Da Cofactor **2**
Probability: 100% Hyperscore: 67.80

Δ = 618.09 + 57.02 (CAM) = 675.09 Da

c)

AA	B	B Ions (Da)		Y	Y Ions (Da)	
L	1	114.09	Missing	22	2,585.16	Missing
C+57	2	274.12	Found	21	2,472.08	Found
Q	3	402.18	Found	20	2,312.05	Found
F	4	549.25	Found	19	2,183.99	Found
H	5	686.31	Found	18	2,036.92	Found
F	6	833.38	Found	17	1,899.86	Found
H	7	970.44	Found	16	1,752.79	Found
W	8	1,156.51	Found	15	1,615.73	Found
G	9	1,213.54	Found	14	1,429.66	Found
S	10	1,300.57	Found	13	1,372.63	Found
L	11	1,413.65	Found	12	1,285.60	Found
D	12	1,528.68	Found	11	1,172.52	Found
G	13	1,585.70	Found	10	1,057.49	Found
Q	14	1,713.76	Found	9	1,000.47	Found
G	15	1,770.78	Found	8	872.41	Found
S	16	1,857.81	Found	7	815.39	Found
E	17	1,986.86	Found	6	728.36	Found
H	18	2,123.91	Found	5	599.31	Found
T	19	2,224.96	Found	4	462.26	Found
V	20	2,324.03	Found	3	361.21	Found
D	21	2,439.06	Found	2	262.14	Found
K	22	2,585.16	Missing	1	147.11	Missing

d)

AA	B	B Ions (Da)		Y	Y Ions (Da)	
L	1	114.09	Missing	23	3,331.35	Missing
C+675	2	892.21	Found	22	3,218.26	Missing
Q	3	1,020.27	Missing	21	2,440.14	Found
F	4	1,167.34	Found	20	2,312.08	Missing
H	5	1,304.40	Found	19	2,165.02	Found
F	6	1,451.46	Found	18	2,027.96	Found
H	7	1,588.52	Found	17	1,880.89	Found
W	8	1,774.60	Found	16	1,743.83	Found
G	9	1,831.62	Found	15	1,557.75	Found
S	10	1,918.66	Found	14	1,500.73	Found
L	11	2,031.74	Found	13	1,413.70	Found
D	12	2,146.77	Found	12	1,300.61	Found
G	13	2,203.79	Missing	11	1,185.59	Found
Q	14	2,331.85	Found	10	1,128.56	Found
G	15	2,388.87	Found	9	1,000.51	Found
S	16	2,475.90	Missing	8	943.48	Found
E	17	2,604.94	Missing	7	856.45	Found
H	18	2,742.00	Found	6	727.41	Found
T	19	2,843.05	Missing	5	590.35	Found
V	20	2,942.12	Found	4	489.3	Found
D	21	3,057.15	Found	3	390.23	Found
K	22	3,185.24	Found	2	275.21	Found
K	23	3,331.35	Missing	1	147.11	Missing

*Figure 131 **Tryptic digest of hCAII^{I91C} and 2 · hCAII^{I91C} analyzed by LC-MS/MS.** MS/MS spectrum of hCAI^{I91C} (a) and **2 · hCAII^{I91C}** (b). The analyzed peptide sequence, measured and theoretical mass, modification and quality indicators are shown. The list of analyzed ions from (a) and (b) are given in (c) and (d), respectively. Carbamidomethylation is abbreviated with CAM.*

a) **hCAII^N67G-E69R-I91C**
Sequence: (K80)GGPLDGTYRLcQFHFHWGSLDGQGSEHTVDK(K112)
Measured Mass: 3,500.60 Da   Theoretical mass: 3,445.73 Da   Modifications: Carbamidomethylation (CAM) 57.02 Da
Probability: 100%   Hyperscore: 76.63

b) **2 · hCAII^N67G-E69R-I91C**
Sequence: (K80)GGPLDGTYRLcQFHFHWGSLDGQGSEHTVDKK(K113)
Measured Mass: 4,246.81 Da   Theoretical mass: 3,573.90 Da   Modifications: 672.91 Da Cofactor **2**
Probability: 100%   Hyperscore: 58.59

Δ = 618.09 + 57.02 (CAM) = 675.11 Da

Δ = 746.19 −128.17 (K) + 57.02 (CAM) = 675.01 Da

c)

AA	B	B Ions (Da)		Y	Y Ions (Da)	
G	1	58.03	Missing	31	3,501.60	Missing
G	2	115.05	Missing	30	3,444.58	Missing
P	3	212.1	Missing	29	3,387.56	Found
L	4	325.19	Found	28	3,290.51	Found
D	5	440.21	Found	27	3,177.42	Found
G	6	497.24	Missing	26	3,062.40	Found
T	7	598.28	Missing	25	3,005.37	Found
Y	8	761.35	Found	24	2,904.33	Missing
R	9	917.45	Missing	23	2,741.26	Missing
L	10	1,030.53	Missing	22	2,585.16	Missing
C+57	11	1,190.56	Missing	21	2,472.08	Missing
Q	12	1,318.62	Missing	20	2,312.05	Found
F	13	1,465.69	Missing	19	2,183.99	Found
H	14	1,602.75	Found	18	2,036.92	Missing
F	15	1,749.82	Found	17	1,899.86	Found
H	16	1,886.88	Found	16	1,752.79	Missing
W	17	2,072.96	Found	15	1,615.73	Found
G	18	2,129.98	Found	14	1,429.66	Found
S	19	2,217.01	Found	13	1,372.63	Found
L	20	2,330.09	Found	12	1,285.60	Found
D	21	2,445.12	Found	11	1,172.52	Found
G	22	2,502.14	Found	10	1,057.49	Found
Q	23	2,630.20	Found	9	1,000.47	Found
G	24	2,687.22	Found	8	872.41	Found
S	25	2,774.25	Missing	7	815.39	Found
E	26	2,903.30	Found	6	728.36	Found
H	27	3,040.35	Found	5	599.31	Found
T	28	3,141.40	Found	4	462.26	Found
V	29	3,240.47	Found	3	361.21	Found
D	30	3,355.50	Found	2	262.14	Found
K	31	3,501.60	Missing	1	147.11	Missing

d)

AA	B	B Ions (Da)		Y	Y Ions (Da)	
G	1	58.03	Missing	32	4,247.79	Missing
G	2	115.05	Missing	31	4,190.77	Missing
P	3	212.1	Missing	30	4,133.75	Found
L	4	325.19	Found	29	4,036.69	Found
D	5	440.21	Found	28	3,923.61	Found
G	6	497.24	Missing	27	3,808.58	Found
T	7	598.28	Missing	26	3,751.56	Missing
Y	8	761.35	Missing	25	3,650.51	Missing
R	9	917.45	Missing	24	3,487.45	Missing
L	10	1,030.53	Missing	23	3,331.35	Missing
C+675	11	1,808.65	Missing	22	3,218.26	Missing
Q	12	1,936.71	Missing	21	2,440.14	Missing
F	13	2,083.78	Missing	20	2,312.08	Missing
H	14	2,220.84	Missing	19	2,165.02	Missing
F	15	2,367.91	Missing	18	2,027.96	Missing
H	16	2,504.97	Missing	17	1,880.89	Found
W	17	2,691.05	Missing	16	1,743.83	Found
G	18	2,748.07	Missing	15	1,557.75	Found
S	19	2,835.10	Missing	14	1,500.73	Found
L	20	2,948.18	Found	13	1,413.70	Found
D	21	3,063.21	Missing	12	1,300.61	Found
G	22	3,120.23	Missing	11	1,185.59	Found
Q	23	3,248.29	Missing	10	1,128.56	Found
G	24	3,305.31	Missing	9	1,000.51	Found
S	25	3,392.34	Missing	8	943.48	Found
E	26	3,521.39	Missing	7	856.45	Found
H	27	3,658.45	Missing	6	727.41	Found
T	28	3,759.49	Missing	5	590.35	Found
V	29	3,858.56	Found	4	489.3	Found
D	30	3,973.59	Found	3	390.23	Found
K	31	4,101.68	Missing	2	275.21	Found
K	32	4,247.79	Missing	1	147.11	Missing

*Figure 132 **Tryptic digest of hCAII^N67G-E69R-I91C and 2 · hCAII^N67G-E69R-I91C analyzed by LC-MS/MS.** MS/MS spectrum of hCAII^N67G-E69R-I91C (a) and **2** · hCAII^N67G-E69R-I91C (b). The analyzed peptide sequence, measured and theoretical mass, modification and quality indicators are shown. The list of analyzed ions from (a) and (b) are given in (c) and (d), respectively. Carbamidomethylation is abbreviated with CAM.*

2 · hCAII^(WT)
Sequence: (R58)ILNNGHAFNVEFDDSQDK(A77)
Measured Mass: 2,061.94 Da   Theoretical mass: 2,063.17 Da   Modifications: None
Probability: 100%   Hyperscore: 64.86

2 · hCAII^(WT)
Sequence: (R89)LIQFHFHWGSLDGQGSEHTVDK(K112)
Measured Mass: 2,537.21 Da   Theoretical mass: 2,538.76 Da   Modifications: None
Probability: 100%   Hyperscore: 64.21

c)

AA	B	B Ions (Da)		Y	Y Ions (Da)	
I	1	114.09	Missing	18	2,062.95	Missing
L	2	227.18	Missing	17	1,949.86	Missing
N	3	341.22	Found	16	1,836.78	Found
N	4	455.26	Found	15	1,722.74	Found
G	5	512.28	Found	14	1,608.69	Found
H	6	649.34	Found	13	1,551.67	Found
A	7	720.38	Found	12	1,414.61	Found
F	8	867.45	Found	11	1,343.58	Found
N	9	981.49	Found	10	1,196.51	Found
V	10	1,080.56	Found	9	1,082.46	Missing
E69	11	1,209.60	Found	8	983.4	Found
F	12	1,356.67	Found	7	854.35	Found
D	13	1,471.70	Found	6	707.28	Found
D	14	1,586.72	Found	5	592.26	Found
S	15	1,673.76	Found	4	477.23	Found
Q	16	1,801.81	Found	3	390.2	Found
D	17	1,916.84	Found	2	262.14	Found
K	18	2,062.95	Missing	1	147.11	Missing

d)

AA	B	B Ions (Da)		Y	Y Ions (Da)	
L	1	114.09	Missing	22	2,538.22	Missing
I91	2	227.18	Missing	21	2,425.13	Found
Q	3	355.23	Found	20	2,312.05	Found
F	4	502.3	Missing	19	2,183.99	Found
H	5	639.36	Found	18	2,036.92	Found
F	6	786.43	Found	17	1,899.86	Found
H	7	923.49	Found	16	1,752.79	Found
W	8	1,109.57	Missing	15	1,615.73	Missing
G	9	1,166.59	Found	14	1,429.66	Found
S	10	1,253.62	Found	13	1,372.63	Found
L	11	1,366.71	Missing	12	1,285.60	Found
D	12	1,481.73	Found	11	1,172.52	Found
G	13	1,538.75	Found	10	1,057.49	Found
Q	14	1,666.81	Found	9	1,000.47	Found
G	15	1,723.83	Missing	8	872.41	Found
S	16	1,810.87	Missing	7	815.39	Found
E	17	1,939.91	Missing	6	728.36	Found
H	18	2,076.97	Missing	5	599.31	Found
T	19	2,178.02	Missing	4	462.26	Found
V	20	2,277.08	Missing	3	361.21	Found
D	21	2,392.11	Found	2	262.14	Found
K	22	2,538.22	Missing	1	147.11	Missing

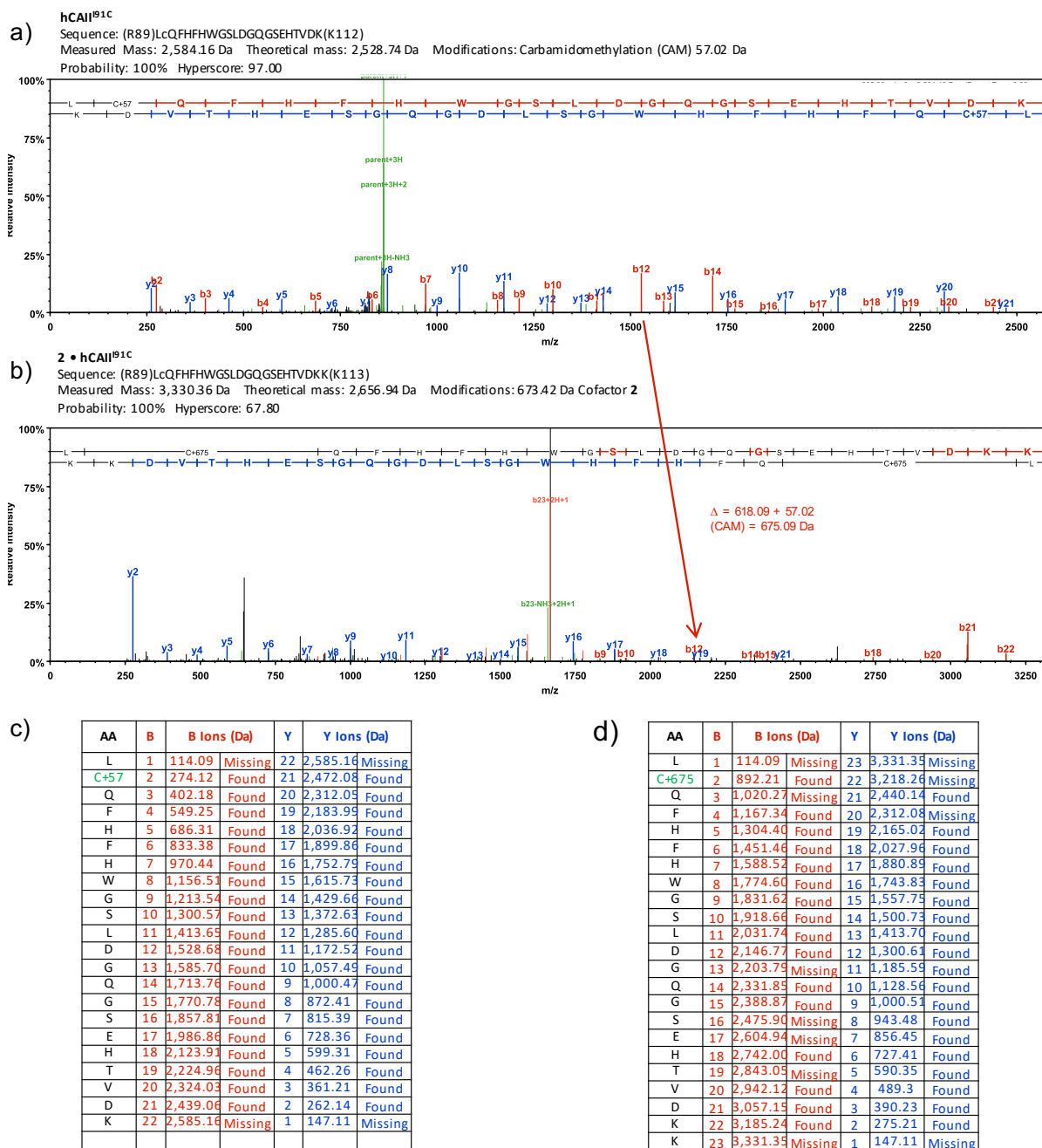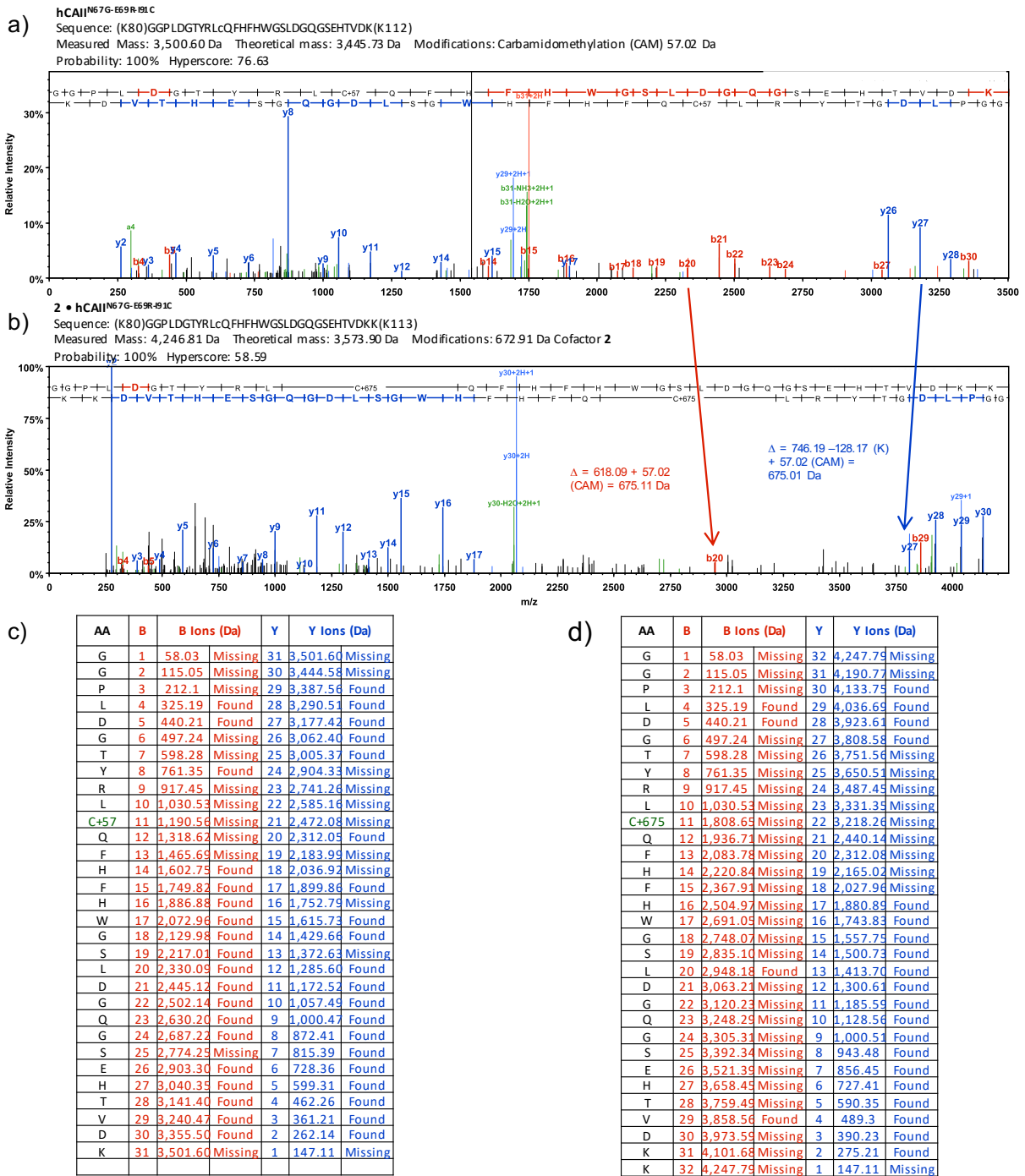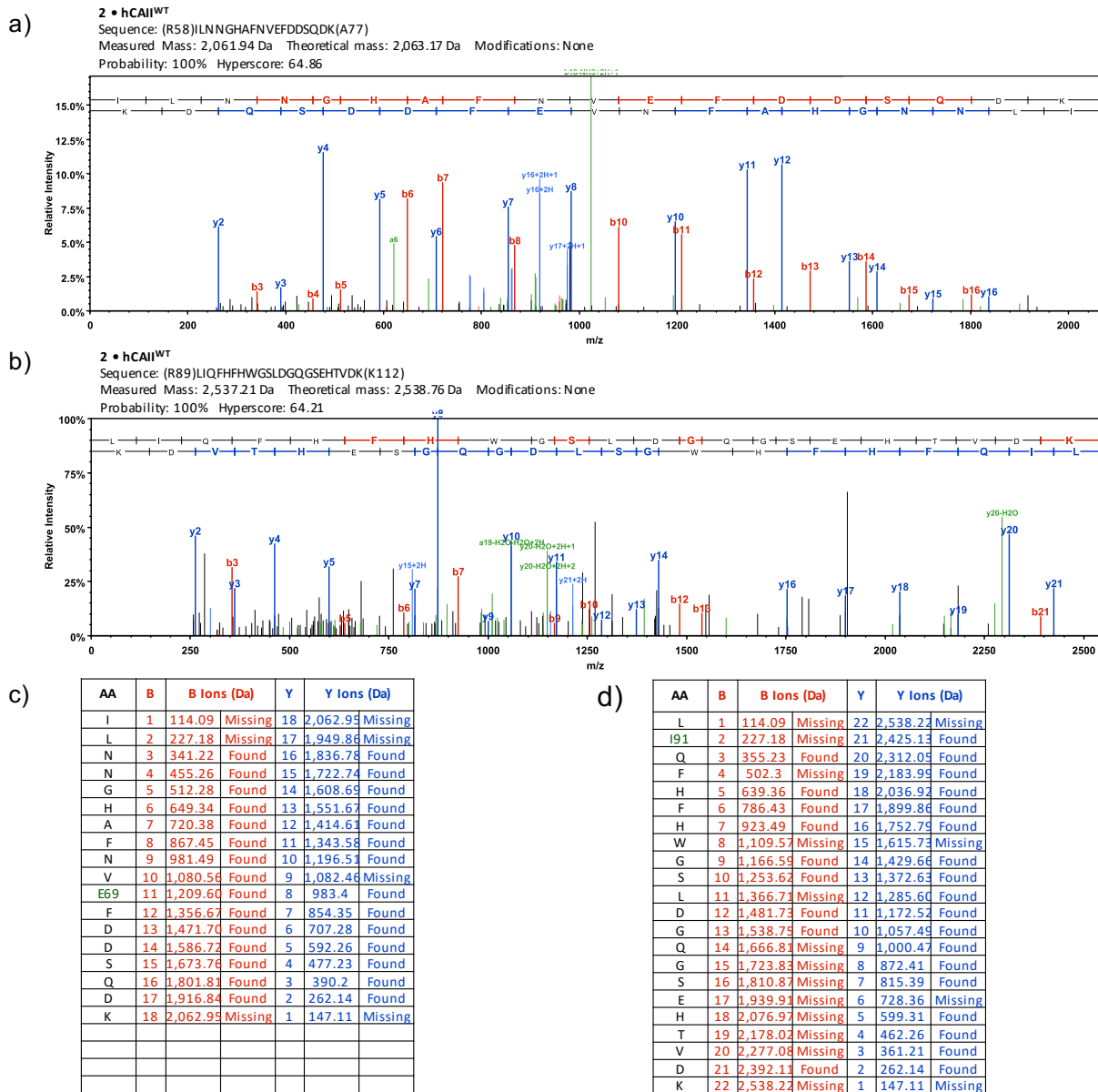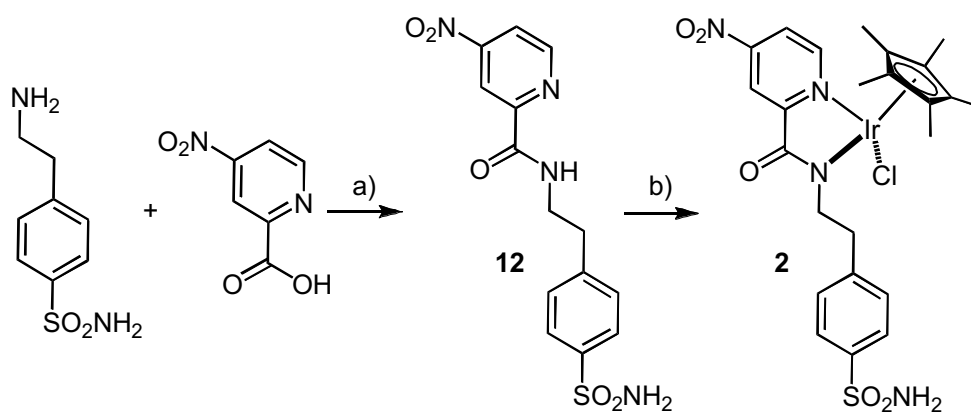*Figure 133 **Tryptic digest of 2 · hCAII^(WT) analyzed by LC-MS/MS.** MS/MS spectrum of **2** · hCAII^(WT) from the peptide with amino acid E69 (a) and I91 (b). The analyzed peptide sequence, measured and theoretical mass, and quality indicators are shown. The list of analyzed ions from (a) and (b) are given in (c) and (d), respectively.*

Scheme 1: *a)* HOBt, EDCI·HCl, DIPEA, DMF, 70 °C, 16 h, 32%. b) [IrCp*Cl₂]₂, EtOH, 80 °C, 2 h, 64%.
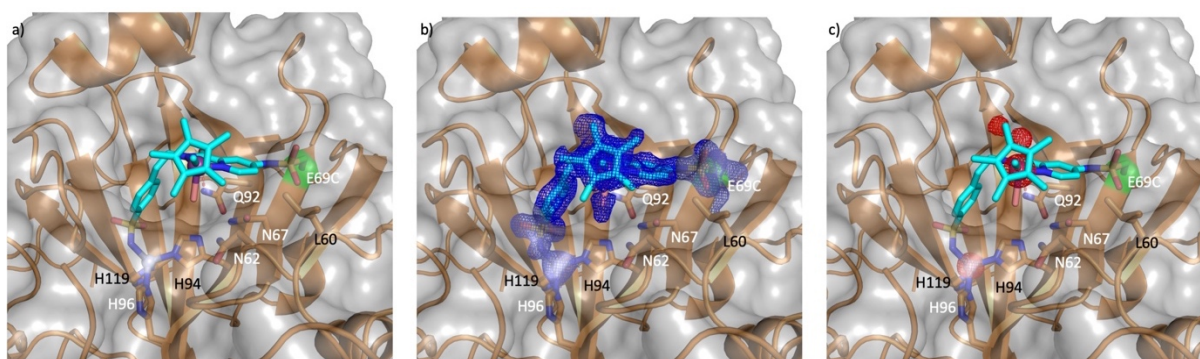
Figure 134 Dual anchoring of cofactor 2 to hCAII[E69C]. (a, b, c) Dually-anchored cofactor 2 bound to hCAII[E69C] (pdb: 7ONQ); (b) the 2Fo-Fc difference map is displayed as blue mesh (1s); (c) the anomalous electron density of the Zn and Ir-atom is highlighted as red mesh (5s). hCAII is depicted as orange cartoon and as transparent grey surface. The cofactor 2 is shown in cyan sticks, the Ir-atoms and Zn-atoms are depicted as purple and grey spheres respectively. Amino acids in the proximity of the cofactor are highlighted as sticks and labelled, mutated amino acids are highlighted in green. Chloride, pink; nitrogen, blue; oxygen, red; sulfur, yellow.
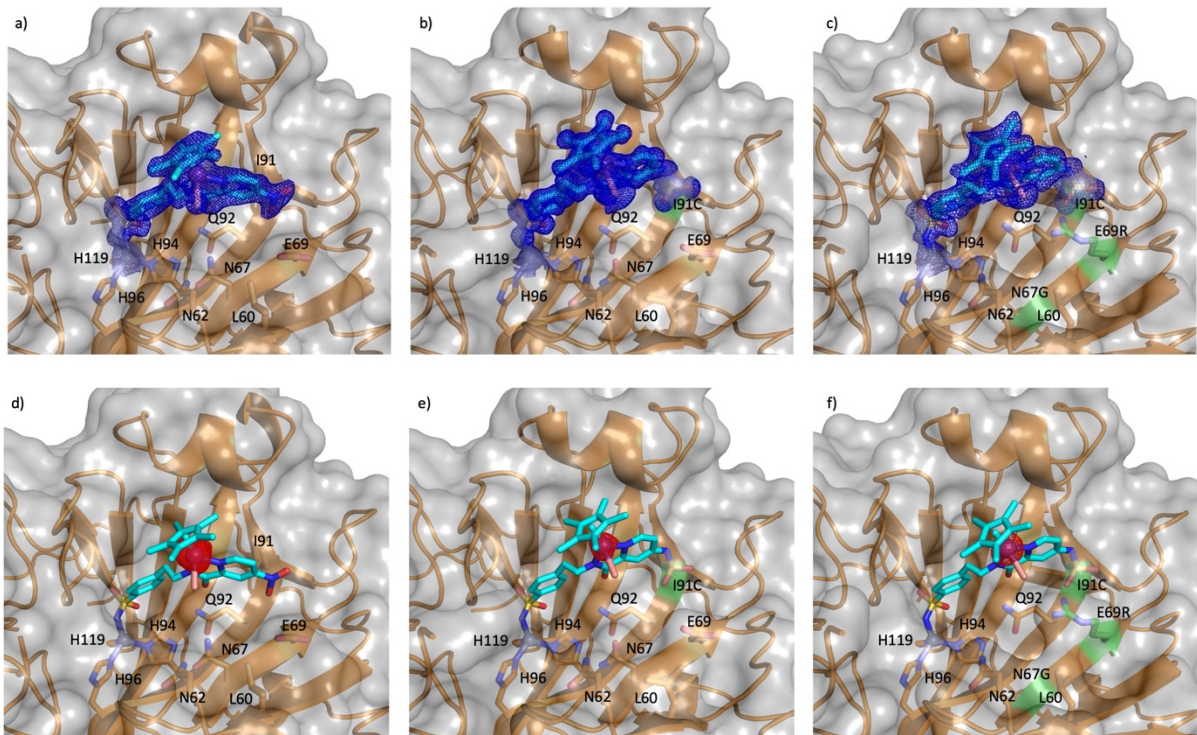
Figure 135 Crystal structures of cofactor 2 bound to hCAII^WT, hCAII^I91C, hCAII^I91C-N67G-E69R. (a, d) Wild type hCAII with bound cofactor 2 (PDB: 7ONP); (b, e) dually-anchored cofactor 2 bound to hCAII^I91C through a sulfonamide moiety (pdb: 7ONV); (c, f) evolved hCAII^I91C-N67G-E69R with the dually- anchored cofactor 2 (pdb: 7ONM). hCAII is depicted as orange cartoon and as transparent grey surface. Cofactor 2 is shown in cyan sticks, the Ir-atoms and Zn-atoms are depicted as purple and grey spheres respectively. The 2Fo-Fc difference map is shown as blue mesh (1s) in (a-c). The anomalous electron density of the Ir-atom is shown as red mesh (5s) in (d-f). Amino acids in the proximity of the cofactor are highlighted as sticks and labelled, mutated amino acids are highlighted in green. Chloride, pink; nitrogen, blue; oxygen, red; sulfur, yellow. The S-O distance in (b, c, e, f) ranges from 1.43 to 1.57 Å.
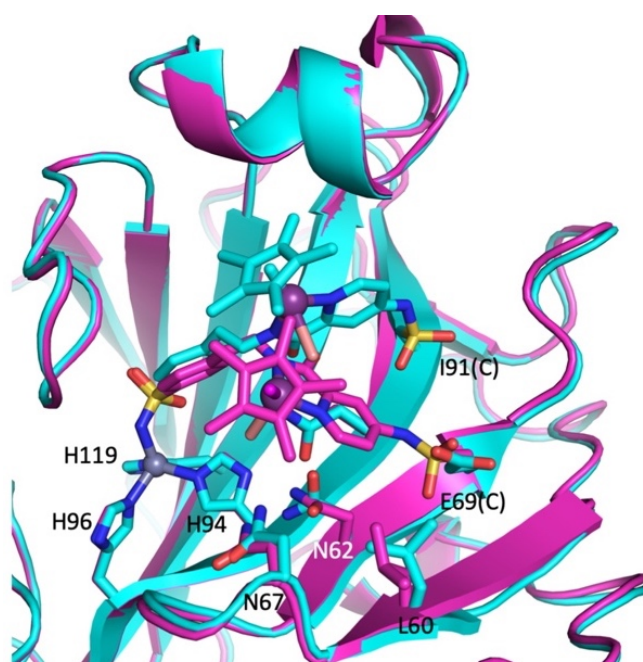
Figure 136 Comparison of dually-anchored cofactor 2 to hCAII^{E69C} and hCAII^{I91C}. The ATHase 2 · hCAII^{I91C} is displayed in cyan and 2 · hCAII^{E69C} in magenta. The Ir-atoms and Zn-atoms are depicted as purple and grey spheres, respectively. Amino acids in the proximity of the cofactor are highlighted as sticks and labelled; chloride, pink; nitrogen, blue; oxygen, red; sulfur, yellow.
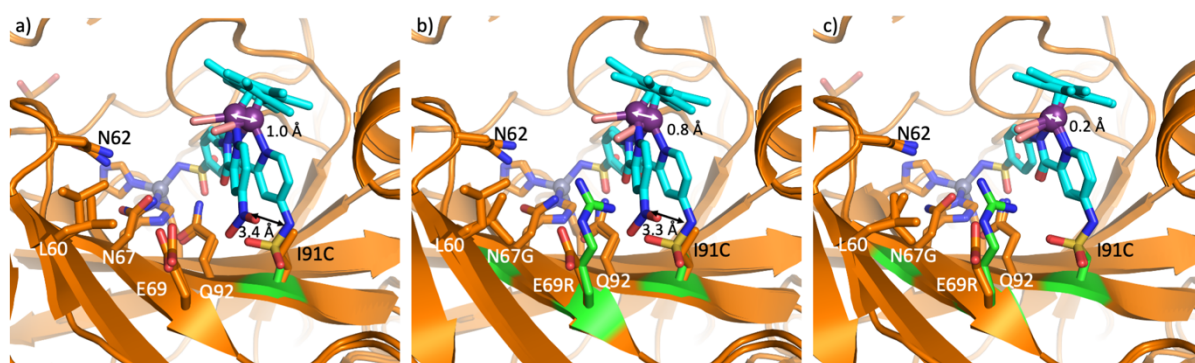
Figure 137 Comparison of the cofactor's position bound to hCAI^{WT}, hCAII^{I91C} and hCAII$^{I91C-N67G-E69R}$. (a) Overlay of hCAII^{WT} and hCAII^{I91C} with bound cofactor 2; (b) overlay of hCAII^{WT} and hCAII$^{I91C-N67G-E69R}$ with bound cofactor 2; (c) overlay of hCAII^{I91C} and hCAII$^{I91C-N67G-E69R}$ with bound cofactor 2. hCAII is depicted as orange cartoon. The cofactor 2 is displayed in cyan sticks, the Ir-atoms and Zn-atoms are depicted as purple and grey spheres, respectively. The movement of the Ir and N2 atoms between hCAII^{WT} and the mutants is highlighted. Amino acids in the proximity of the cofactor are highlighted as sticks and labelled, mutated amino acids are highlighted in green. Chloride, pink; nitrogen, blue; oxygen, red; sulfur, yellow.
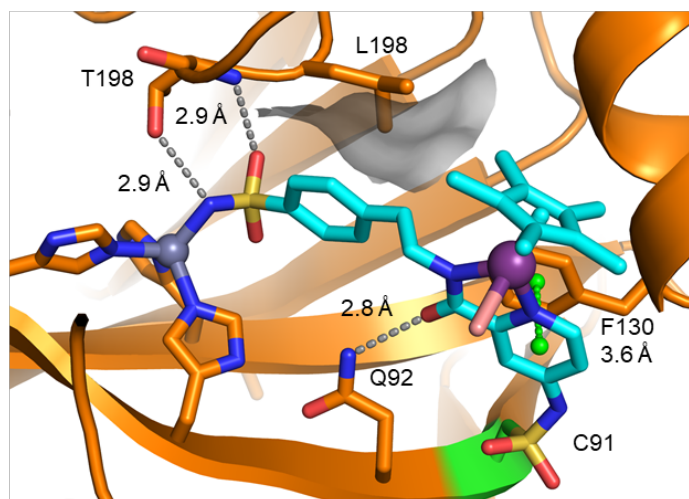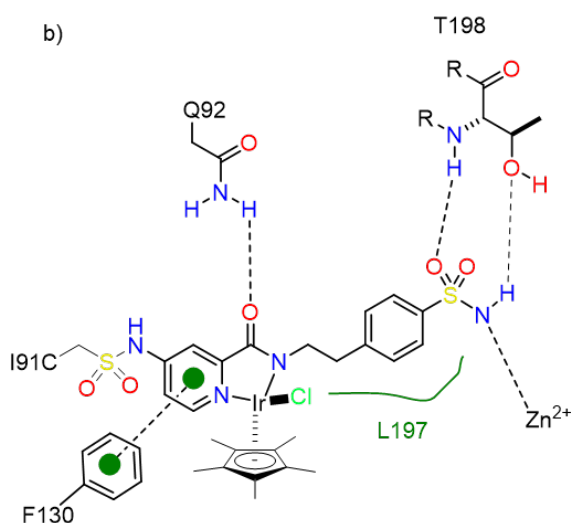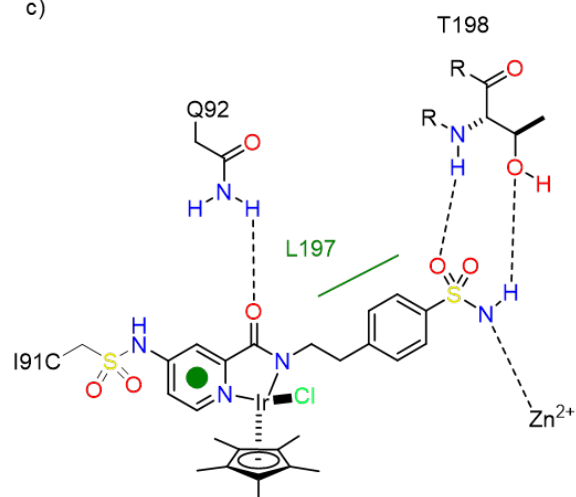
Figure 138 Visualizing the binding of cofactor 2 to hCAII^{I91C}, hCAII^{E69C} and hCAII$^{I91C-N67G-E69R}$. (a) Visualization of the interactions between hCAII^{I91C} and cofactor **2** identified by PoseView[330]. The protein is depicted. as orange cartoon, amino acids in the proximity of the cofactor are highlighted as sticks and labelled, mutated amino acids are highlighted in green. Chloride, pink; nitrogen, blue; oxygen, red; sulfur, yellow. Schematic representation of the interaction between hCAII^{I91C}/hCAII$^{I91C-N67G-E69R}$ (b), hCAII^{E69C} (c) and cofactor **2**. The interactions of hCAII^{I91C} and hCAII$^{I91C-N67G-E69R}$ with cofactor **2** are identical.
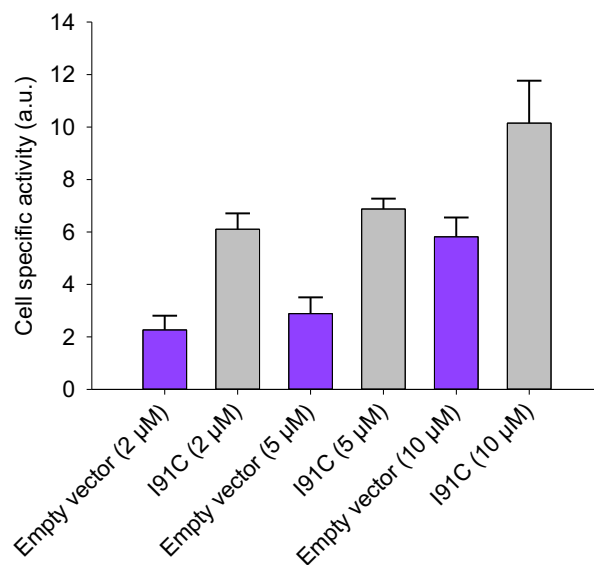
Figure 139 Comparison between background activity (purple) and cell activity (grey) at different cofactor 2 concentrations (µM).
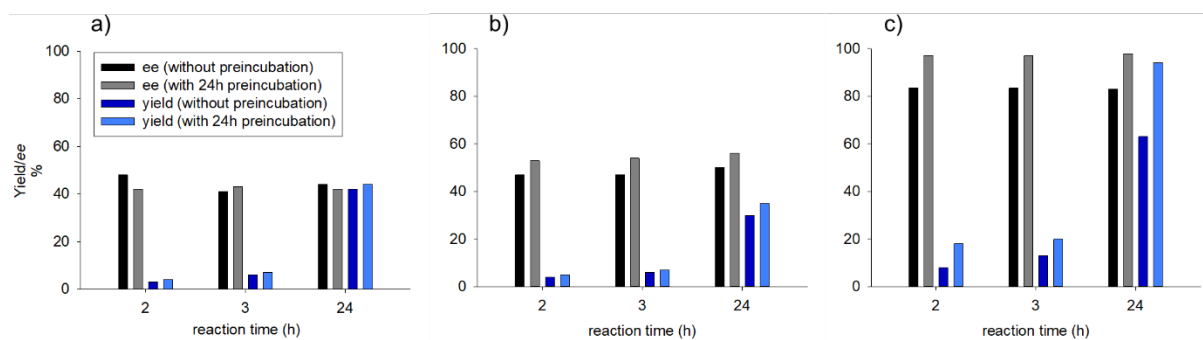
*Figure 140 Effect of cofactor 2 incubation time on the catalytic activity of ArMs 2 · hCAII. All reactions were performed using harmaline (5 mM), ArM (10 µM), formate/MOPS buffer (850 mM/340 mM, pH 7.4). a) ArM **2** · hCAII[WT], b) ArM **2** · hCAII[E69C], c) ArM **2** · hCAII[N67G-E69R-I91C]. The ee and yield are identical after 6 and 24 preincubation for ArM **2** · hCAII[N67G-E69R-I91C].*

*Table 47 Catalysis results obtained from **1** ·hCAII ArMs for the asymmetric transfer hydrogenation of harmaline **3**.*
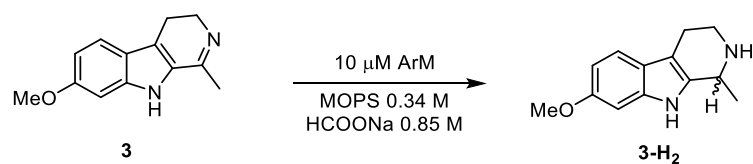
Catalyst	Yield (%)	TON	*ee* (%)
Cofactor **1**	15 (± 2)	30	2 (± 0) (*R*)
1 ·hCAII^{WT}	9 (± 2)	18	19 (± 3) (*R*)
1 ·hCAII^{I91C}	13 (± 1)	26	26 (± 3) (*R*)
1 ·hCAII^{E69C}	11 (± 1)	22	14 (± 5) (*S*)
1 ·hCAII$^{N67G-E69R-I91C}$	8 (± 2)	16	54 (± 3) (*S*)

Reaction conditions: substrate (2 mM), ArM (10 μM), formate/MOPS buffer (850 mM/340 mM, pH 7.4), 25 °C, 16 h. The reactions were performed in separate duplicates, standard deviation in parentheses.

Table 48 LC-MS/MS analysis of trypically digested proteins and ATHases

Protein/ATHase	Counts	
	Total spectrum	Cofactor **2** modification (675.1 Da)
hCAII^{WT}	4652	0
2 · hCAII^{WT}	2923	0
hCAII^{E69C}	2564	0
2 · hCAII^{E69C}	3108	74
hCAII^{I91C}	5184	0
2 · hCAII^{I91C}	4485	232
hCAII$^{N67G-E69R-I91C}$	4909	0
2 · hCAII$^{N67G-E69R-I91C}$	2693	143

Table 49 Screening of purified hCAII mutant proteins for the asymmetric transfer hydrogenation of harmaline 3.
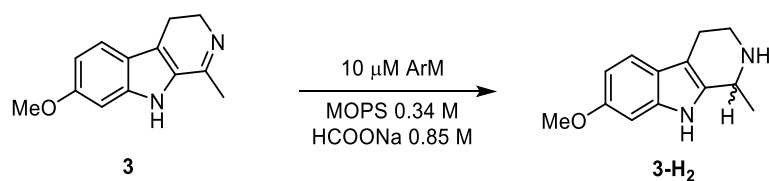
Catalyst	yield (%)	TON	*ee* (%)
Cofactor **2**	62 (± 4)	311	1 (± 2) (*S*)
2 ·hCAII^{WT}	49 (± 1)	247	37 (± 1) (*R*)
2 ·hCAII^{I91C}	76 (± 2)	381	35 (± 0) (*S*)
2 ·hCAII^{E69C}	56 (± 1)	278	48 (± 3) (*R*)
2 ·hCAII$^{N67G-E69R-I91C}$	90 (± 4)	451	96 (± 1) (*R*)
2 ·hCAII$^{N62Y-N67G-E69R-I91C}$	92 (± 1)	460	73 (± 0) (*R*)
2 ·hCAII$^{L60F-N62Y-N67G-E69R-I91C}$	83 (± 7)	415	74 (± 1) (*R*)
2 ·hCAII$^{N67L-E69Y-I91C}$	92 (± 1)	458	63 (± 1) (*S*)
2 ·hCAII$^{N62D-N67L-E69Y-I91C}$	92 (± 1)	460	46 (± 1) (*S*)
2 ·hCAII$^{L60W-N62D-N67L-E69Y-I91C}$	88 (± 1)	440	23 (± 1) (*S*)
2 ·hCAII$^{L60W-N67L-E69Y-I91C}$	44 (± 2)	221	49 (± 0) (*S*)
[a]**2** ·hCAII$^{N67G-E69R-I91C}$	93 (± 2)	467	98 (± 1) (*R*)
[a]**2** ·hCAII$^{N62Y-N67G-E69R-I91C}$	97 (± 7)	484	75 (± 0) (*R*)
[a]**2** ·hCAII$^{L60F-N62Y-N67G-E69R-I91C}$	60 (± 4)	301	74 (± 4) (*R*)

Reaction condition: substrate (5 mM), ArM solution (10 µM), formate/MOPS buffer (850 mM/340 mM, pH 7.4), 25 °C, 16 h. The reactions were performed in separate duplicates, standard deviation in parentheses. [a]The reaction was performed under the degassing condition for 16 h.

Table 50 Screening of purified hCAII mutants for the asymmetric transfer hydrogenation of substrate 4.
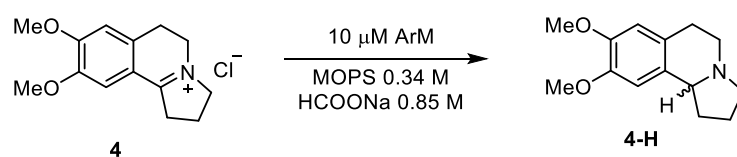
Catalyst	Yield (%)	TON	*ee* (%)
Cofactor **2**	4 (± 0)	22	19 (± 1)
2 ·hCAII^{WT}	5 (± 0)	24	19 (± 17)
2 ·hCAII^{I91C}	21 (± 1)	104	12 (± 0)
2 ·hCAII^{E69C}	16 (± 0)	81	-57 (± 2)
2 ·hCAII$^{N67G-E69R-I91C}$	13 (± 1)	67	16 (± 5)
2 ·hCAII$^{N62Y-N67G-E69R-I91C}$	26 (± 3)	130	55 (± 4)
2 ·hCAII$^{L60F-N62Y-N67G-E69R-I91C}$	44 (± 4)	219	62 (± 4)
2 ·hCAII$^{N67L-E69Y-I91C}$	24 (± 2)	119	75 (± 1)
2 ·hCAII$^{N62D-N67L-E69Y-I91C}$	49 (± 1)	243	82 (± 2)
2 ·hCAII$^{L60W-N62D-N67L-E69Y-I91C}$	71 (± 7)	355	86 (± 3)
2 ·hCAII$^{L60W-N67L-E69Y-I91C}$	8 (± 0)	42	78 (± 1)

Reaction condition: substrate (5 mM), ArM solution (10 µM), formate/MOPS buffer (850 mM/340 mM, pH 7.4), 25 °C, 16 h.

The reactions were performed in separate duplicates, standard deviation in parentheses.

Table 51 Screening of purified hCAII mutants for the asymmetric transfer hydrogenation of substrate 5.
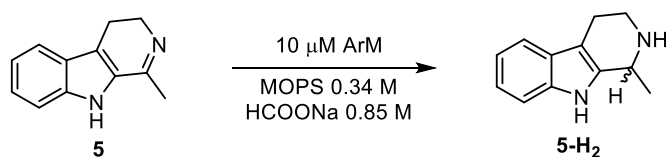
Catalyst	Yield (%)	TON	*ee* (%)
Cofactor **2**	21 (± 2)	107	11 (± 3) (*S*)
2 ·hCAII^{WT}	14 (± 1)	70	30 (± 3) (*R*)
2 ·hCAII^{I91C}	36 (± 2)	182	54 (± 1) (*S*)
2 ·hCAII^{E69C}	20 (± 5)	101	2 (± 8) (*R*)
2 ·hCAII$^{N67G-E69R-I91C}$	42 (± 0)	211	80 (± 5) (*R*)
2 ·hCAII$^{N62Y-N67G-E69R-I91C}$	100 (± 9)	500	87 (± 1) (*R*)
2 ·hCAII$^{L60F-N62Y-N67G-E69R-I91C}$	71 (± 1)	355	72 (± 6) (*R*)
2 ·hCAII$^{N67L-E69Y-I91C}$	50 (± 8)	250	67 (± 5) (*S*)
2 ·hCAII$^{N62D-N67L-E69Y-I91C}$	28 (± 7)	141	52 (± 3) (*S*)
2 ·hCAII$^{L60W-N62D-N67L-E69Y-I91C}$	50 (± 2)	252	14 (± 3) (*R*)
2 ·hCAII$^{L60W-N67L-E69Y-I91C}$	38 (± 21)	189	72 (± 0) (*S*)

Reaction condition: substrate (5 mM), ArM solution (10 μM), formate/MOPS buffer (850 mM/340 mM, pH 7.4), 25 °C, 16 h.

The reactions were performed in separate duplicates, standard deviation in parentheses.

Table 52 Screening of purified hCAII mutants for the asymmetric transfer hydrogenation of substrate 6.
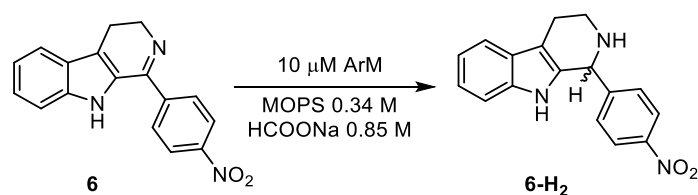
Catalyst	Yield (%)	TON	*ee* (%)
Cofactor **2**	17 (± 9)	87	4 (± 12) (*R*)
2 ·hCAII^{WT}	11 (± 2)	57	54 (± 3) (*R*)
2 ·hCAII^{I91C}	16 (± 3)	80	18 (± 5) (*S*)
2 ·hCAII^{E69C}	20 (± 6)	99	25 (± 2) (*R*)
2 ·hCAII$^{N67G-E69R-I91C}$	38 (± 2)	191	80 (± 5) (*S*)
2 ·hCAII$^{N62Y-N67G-E69R-I91C}$	12 (± 1)	60	42 (± 4) (*S*)
2 ·hCAII$^{L60F-N62Y-N67G-E69R-I91C}$	8 (± 2)	40	66 (± 4) (*S*)
2 ·hCAII$^{N67L-E69Y-I91C}$	10 (± 3)	49	25 (± 2) (*R*)
2 ·hCAII$^{N62D-N67L-E69Y-I91C}$	35 (± 1)	175	2 (± 3) (*S*)
2 ·hCAII$^{L60W-N62D-N67L-E69Y-I91C}$	50 (± 9)	249	73 (± 1) (*R*)
2 ·hCAII$^{L60W-N67L-E69Y-I91C}$	19 (± 1)	95	50 (± 3) (*R*)

Reaction condition: substrate (5 mM), ArM solution (10 μM), formate/MOPS buffer (850 mM/340 mM, pH 7.4), 25 °C, 16 h.

The reactions were performed in separate duplicates, standard deviation in parentheses.

*Table 53 Screening of purified hCAII mutants for the asymmetric transfer hydrogenation of substrate **7**.*

Catalyst	Yield (%)	TON	*ee* (%)
Cofactor **2**	100 (± 4)	500	2 (± 2) (*R*)
2 ·hCAII^{WT}	45 (± 3)	224	35 (± 1) (*S*)
2 ·hCAII^{I91C}	59 (± 7)	296	68 (± 0) (*S*)
2 ·hCAII^{E69C}	43 (± 8)	214	39 (± 0) (*S*)
2 ·hCAII$^{N67G-E69R-I91C}$	62 (± 9)	310	26 (± 0) (*S*)
2 ·hCAII$^{N62Y-N67G-E69R-I91C}$	36 (± 8)	181	15 (± 2) (*S*)
2 ·hCAII$^{L60F-N62Y-N67G-E69R-I91C}$	47 (± 5)	236	6 (± 0) (*R*)
2 ·hCAII$^{N67L-E69Y-I91C}$	66 (± 2)	329	0 (± 0)
2 ·hCAII$^{N62D-N67L-E69Y-I91C}$	74 (± 2)	372	43 (± 0) (*R*)
2 ·hCAII$^{L60W-N62D-N67L-E69Y-I91C}$	78 (± 2)	389	33 (± 1) (*S*)
2 ·hCAII$^{L60W-N67L-E69Y-I91C}$	52 (± 1)	260	56 (± 0) (*S*)

Reaction condition: substrate (5 mM), ArM solution (10 μM), formate/MOPS buffer (850 mM/340 mM, pH 7.4), 25 °C, 16 h. The reactions were performed in separate duplicates, standard deviation in parentheses.

Table 54 Screening of purified hCAII mutants for the asymmetric transfer hydrogenation of salsolidine 8.

Catalyst	Yield (%)	TON	*ee* (%)
Cofactor **2**	58 (± 3)	289	3 (± 2) (*R*)
2 ·hCAII^{WT}	44 (± 3)	218	7 (± 2) (*S*)
2 ·hCAII^{I91C}	65 (± 4)	323	4 (± 0) (*S*)
2 ·hCAII^{E69C}	57 (± 3)	284	15 (± 3) (*R*)
2 ·hCAII$^{N67G-E69R-I91C}$	74 (± 8)	370	34 (± 0) (*R*)
2 ·hCAII$^{N62Y-N67G-E69R-I91C}$	84 (± 0)	422	58 (± 0) (*R*)
2 ·hCAII$^{L60F-N62Y-N67G-E69R-I91C}$	80 (± 6)	398	36 (± 3) (*R*)
2 ·hCAII$^{N67L-E69Y-I91C}$	91 (± 2)	454	58 (± 3) (*S*)
2 ·hCAII$^{N62D-N67L-E69Y-I91C}$	88 (± 2)	440	33 (± 2) (*S*)
2 ·hCAII$^{L60W-N62D-N67L-E69Y-I91C}$	92 (± 5)	461	8 (± 3) (*S*)
2 ·hCAII$^{L60W-N67L-E69Y-I91C}$	69 (± 0)	343	25 (± 2) (*S*)

Reaction condition: substrate (5 mM), ArM solution (10 μM), formate/MOPS buffer (850 mM/340 mM, pH 7.4), 25 °C, 16 h.

The reactions were performed in separate duplicates, standard deviation in parentheses.

*Table 55 Screening of purified hCAII mutants for the asymmetric transfer hydrogenation of substrate **9**.*
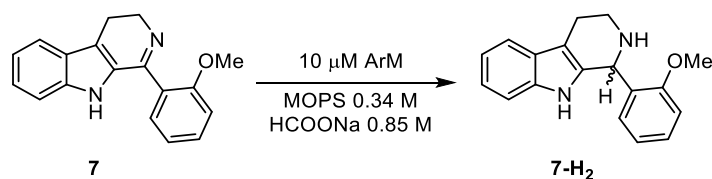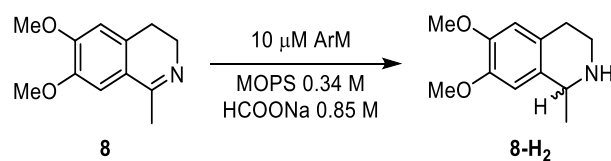

Catalyst	Yield (%)	TON	*ee* (%)
Cofactor **2**	76 (± 4)	379	1 (± 1)
2 ·hCAII^{WT}	56 (± 5)	278	-42 (± 3)
2 ·hCAII^{I91C}	66 (± 10)	328	-53 (± 0)
2 ·hCAII^{E69C}	50 (± 8)	250	-56 (± 2)
2 ·hCAII$^{N67G-E69R-I91C}$	71 (± 10)	355	-25 (± 3)
2 ·hCAII$^{N62Y-N67G-E69R-I91C}$	63 (± 15)	315	-38 (± 3)
2 ·hCAII$^{L60F-N62Y-N67G-E69R-I91C}$	83 (± 13)	416	-55 (± 0)
2 ·hCAII$^{N67L-E69Y-I91C}$	82 (± 14)	411	-31 (± 1)
2 ·hCAII$^{N62D-N67L-E69Y-I91C}$	75 (± 21)	374	24 (± 3)
2 ·hCAII$^{L60W-N62D-N67L-E69Y-I91C}$	89 (± 20)	443	5 (± 4)
2 ·hCAII$^{L60W-N67L-E69Y-I91C}$	31 (± 11)	156	-56 (± 1)

Reaction condition: substrate (5 mM), ArM solution (10 µM), formate/MOPS buffer (850 mM/340 mM, pH 7.4), 25 °C, 16 h.

The reactions were performed in separate duplicates, standard deviation in parentheses.

Table 56 Screening of purified hCAII mutants for the asymmetric transfer hydrogenation of substrate 10.
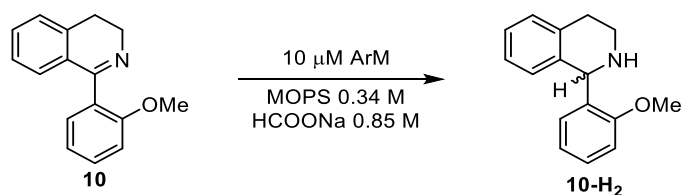
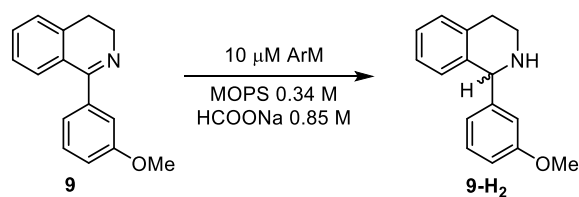

Catalysts	Yield (%)	TON	*ee* (%)
Cofactor **2**	75 (± 12)	373	2 (± 2)
2 ·hCAII^{WT}	57 (± 1)	287	-44 (± 1)
2 ·hCAII^{I91C}	55 (± 3)	274	-56 (± 1)
2 ·hCAII^{E69C}	48 (± 4)	242	-71 (± 7)
2 ·hCAII$^{N67G-E69R-I91C}$	44 (± 5)	219	-14 (± 1)
2 ·hCAII$^{N62Y-N67G-E69R-I91C}$	33 (± 2)	167	24 (± 1)
2 ·hCAII$^{L60F-N62Y-N67G-E69R-I91C}$	42 (± 8)	210	42 (± 1)
2 ·hCAII$^{N67L-E69Y-I91C}$	70 (± 16)	348	-18 (± 4)
2 ·hCAII$^{N62D-N67L-E69Y-I91C}$	68 (± 8)	338	75 (± 8)
2 ·hCAII$^{L60W-N62D-N67L-E69Y-I91C}$	57 (± 4)	286	72 (± 0)
2 ·hCAII$^{L60W-N67L-E69Y-I91C}$	27 (± 4)	137	-33 (± 0)

Reaction condition: substrate (5 mM), ArM solution (10 μM), formate/MOPS buffer (850 mM/340 mM, pH 7.4), 25 °C, 16 h.

The reactions were performed in separate duplicates, standard deviation in parentheses.

*Table 57 Screening of purified hCAII mutants for the asymmetric transfer hydrogenation of substrate **11**.*
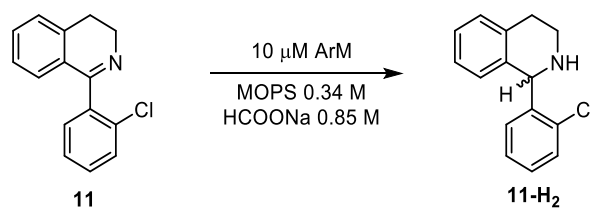
Catalysts	Yield (%)	TON	*ee* (%)
Cofactor **2**	36 (± 4)	179	0 (± 2)
2 ·hCAII^{WT}	16 (± 1)	82	-25 (± 0)
2 ·hCAII^{I91C}	24 (± 1)	118	-55 (± 4)
2 ·hCAII^{E69C}	15 (± 1)	77	-42 (± 4)
2 ·hCAII$^{N67G-E69R-I91C}$	21 (± 2)	107	5 (± 1)
2 ·hCAII$^{N62Y-N67G-E69R-I91C}$	9 (± 2)	47	47 (± 10)
2 ·hCAII$^{L60F-N62Y-N67G-E69R-I91C}$	7 (± 1)	37	31 (± 3)
2 ·hCAII$^{N67L-E69Y-I91C}$	22 (± 1)	108	-63 (± 1)
2 ·hCAII$^{N62D-N67L-E69Y-I91C}$	43 (± 0)	215	-51 (± 3)
2 ·hCAII$^{L60W-N62D-N67L-E69Y-I91C}$	39 (± 4)	193	-57 (± 2)
2 ·hCAII$^{L60W-N67L-E69Y-I91C}$	12 (± 0)	60	-35 (± 2)

Reaction conditions: substrate (5 mM), ArM solution (10 μM), formate/MOPS buffer (850 mM/340 mM, pH 7.4), 25 °C, 16 h.

The reactions were performed in separate duplicates, standard deviation in parentheses.

5.5 X-ray Crystallography for Streptavidin based ArMs and more

5.5.1 Outline of the Authors Contribution

T.R.W and N.V.I conceived and designed the study. N.V.I, V.W., M.M, D.C., K.Y. contributed to the synthesis of the substrates, probe and complexes. S.K. did the UV-Vis stability experiments. F.P.S., R.P.J. and F.L. contributed with helpful advices concerning the crystallization. R.T. performed the computational modelling.

5.5.2 Introduction

Alongside mass spectral analysis, protein crystallography is the most widely used method for structural characterization of ArMs. However, compared to other methods, X-ray crystallography is arguably best-suited to determine the exact atomic coordinates. Protein crystallization is based on the phenomenon that, once the solubility limit of the protein is reached, a new state or phase will appear, which might finally lead to protein crystal(s). The crystallization of proteins depends on creating a solution that is super-saturated in the macromolecule but exhibits conditions that do not significantly perturb its natural state.[251] Protein crystallization is a thermodynamic process and, once it begins, will continue under kinetic control until super-saturation is lifted.[331] In our lab, the crystallization conditions for the two scaffolds mainly used, Sav and hCAII, were optimized over the last 15 years.[65] In the initial studies of Stenkamp *et al.*, it was shown that the C-terminus of the T7-tagged full-length streptavidin variant used in our lab occupies the biotin binding site in the crystal structure. This property influences the affinity to biotin, as the C-terminus needs to be displaced by biotin. Importantly, there are two different space groups in which streptavidin crystallizes: $I4_122$ and C121. They differ in the composition of the asymmetric unit. Whereas, in the $I4_122$ space group, only one monomer is found and based on this structure the tetramer is generated by translation and rotation operations. In the C121 space group the whole tetramer is found in the asymmetric unit.

Our group focuses on combining chemical and genetic tools to optimize the catalytic performance of the corresponding ArMs. In this process, protein crystallography provides

precious insight in the cofactor's localization, in order to get chemo genetic optimization of the system, protein crystallography yield information about the cofactor's precise location to guide the optimization process. There are two methods to form the cofactor protein complex in crystals: soaking a (fully) grown crystal in a solution containing the cofactor or co-crystallization, where a protein solution is incubated with a cofactor solution prior to crystallization. The advantage of soaking is that we can rely on well-established conditions for crystallization. In contrast, new screens need to be set up with the incubated sample to identify conditions where the ArMs crystallize, once assembled. Depending on the pH and buffer compatibility of the cofactor, the conditions from which the protein crystals grow can be adapted slightly. For example, the pH of the standard conditions (pH 4.0) can be increased to pH 8.0 in two steps prior to soaking if required by the cofactor's stability.

Nevertheless, finding the right conditions under which the crystals grow and the metal catalyst is stable might be challenging. Therefore, to obtain structural information of some ArMs, both methods were used, starting with the soaking approach under standard conditions. The following part presents the crystallization approach for different streptavidin-cofactor assemblies.

5.5.3 Hydrophobic probe (**biot-4DMN**) and fluorescent substrate **1**

We recently introduced a "hydrophobic" probe to highlight that the SOD loop increases the hydrophobicity in the biotin binding vestibule of the Sav-based ArMs (Figure 141).[36] To verify the binding of the fluorescent moiety within the biotin binding vestibule, Sav-WT crystals were grown under the following conditions: 20 mg/mL protein, 0.1 M sodium acetate, 25 % w/v PEG 1500 pH 4.0, 1:1 protein to buffer solution, were soaked for 5 h and overnight (20 h).

The resulting X-ray diffraction data was solved at 1.45 Å (pdb 8aqd) and 1.85 Å (pdb 8aqj) resolution. Inspection of the X-ray structures revealed that in both cases, residual density in the biotin biding vestibule and the fluorescent probe **biot-4DMN** could be modeled into the density. In the case of short soaking (5h), two different conformations (**1** and **2**, 50% occupancy each) were found (Figure 142a). However, after overnight soaking, only one conformation (100% occupancy) was found (Figure 142b). Nevertheless, by superposing the

two structures, it could be seen that the conformation of the extended soaking structure is identical to conformation **1** of the shorter soaking structure (Figure 143).

biot-4DMN   substrate **1**
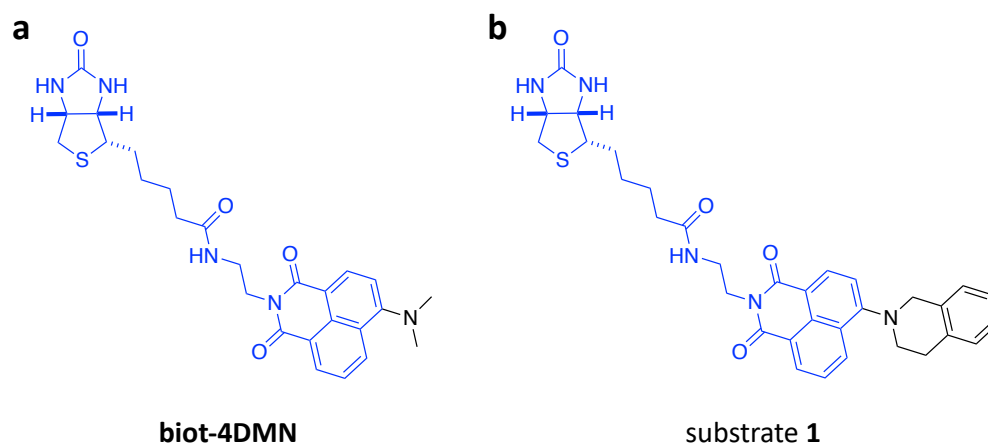
*Figure 141 Chemical structure of **biot-4DMN a** and the fluorescent substrate **1 b**. The structural basis on which the substrate **1** is based on, is highlighted in blue.*

Furthermore, the serine at position 112 of Sav forms a hydrogen bond to the nitrogen in the aromatic system and one of the flanking oxygens in conformation **2** for the shorter soaking time (pdb 8aqd). Taking this into account, one would guess that conformation **2** is the more stable one, therefore being the one observed for structures with prolonged soaking times. However, the conformation found was not the conformation **1**. Nevertheless, the structures indicate the probe's ability to adapt a conformation pointing into the biotin binding vestibule. Therefore, it can be used to detect increased hydrophobicity within the vestibule.

*Figure 142 **a** Crystal structures of **biot-4DMN**·Sav 5h (pdb 8aqd) and **b biot-4DMN**·Sav 20h (pdb 8aqj). The **biot-4DMN** is represented as stick model in green for **a** and in cyan for **b** (atoms are color coded; nitrogen = blue, oxygen = red). The 2Fo-Fc difference maps are displayed as black mesh (1σ). The protein is represented as a surface model and the biotin binding vestibule is labeled in orange color in order to visualize it. The conformations of 8aqd are labeled with **1** and **2**.*

*Figure 143 Superposition of crystal structures of **biot-4DMN·Sav 5h** (pdb 8aqd) and **biot-4DMN·Sav 20h** (pdb 8aqj). The **biot-4DMN** is represented as stick model in green for **biot-4DMN·Sav 5h** and in cyan for **biot-4DMN·Sav 20h** (atoms are color cod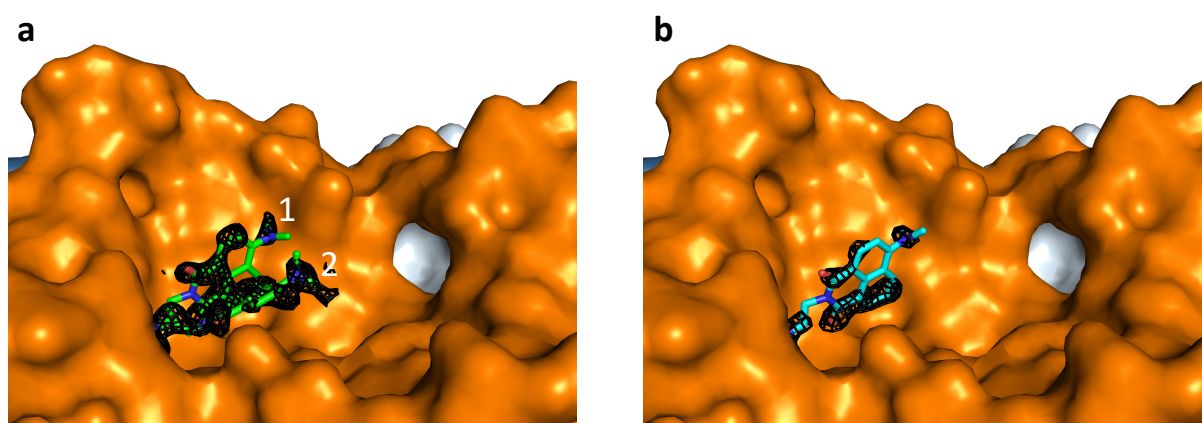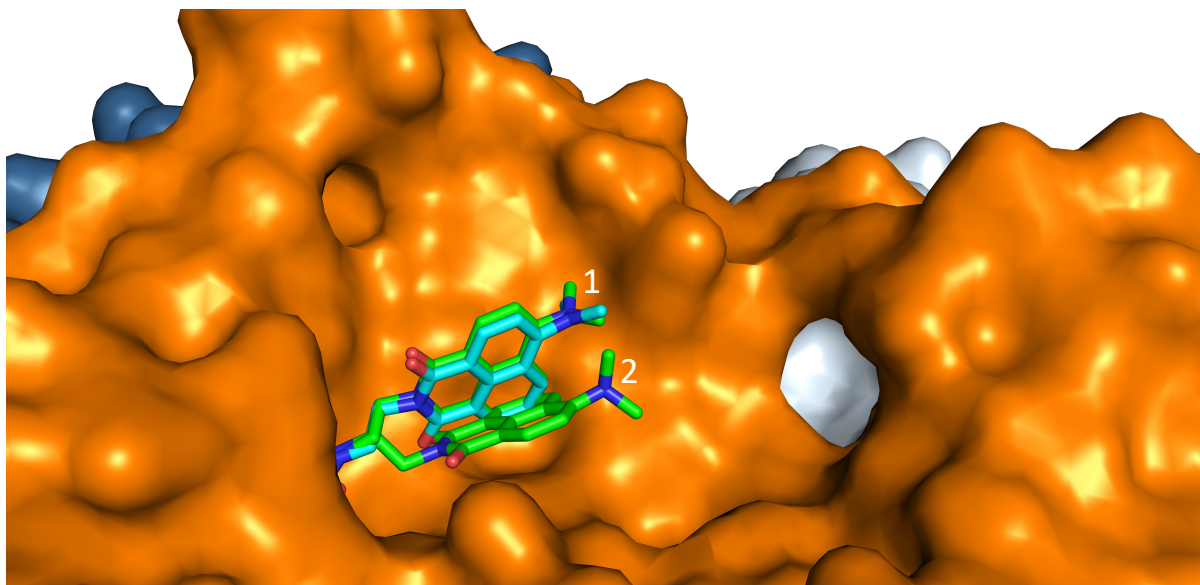ed; nitrogen = blue, oxygen = red). The protein is represented as a surface model with different shades of blue for each monomer. The biotin binding vestibule is labeled in orange color in order to visualize it. Conformations of 8aqd are labeled with **1** and **2**.*

In another study, we modified the scaffold of **biot-4DMN** to create a substrate **1** (Figure 141) susceptible to an alkylation reaction on the isoquinoline moiety. The ultimate goal was to use an iridium photocatalyst sitting in the facing biotin binding site to catalyze the (enantioselective) alkylation of substrate **1** at the α-position of the nitrogen. In order to highlight that substrate **1** binds in the desired way, crystals were grown under standard conditions (20 mg/mL protein, 2 M ammonium sulfate, 0.1 M sodium acetate pH 4.0, 1:1 protein to buffer solution) and soaked overnight (20 h). The resulting X-ray diffraction data was solved at 1.55 Å resolution (pdb 7zx9). Inspection of the X-ray structures revealed residual density in the biotin biding vestibule, and the fluorescent substrate **1** could be modeled into the density (Figure 144b). It could be seen that there are two conformations (50% occupancy each) for the substrate. However, one of them, the one sitting closer to the protein surface, was the preferred one as it was better resolved. When superposing this conformation with one of the crystal structures with the **biot-4DMN** (pdb 8aqd), it could be shown that they sit in a very similar way in the biotin binding pocket (Figure 144c). PoseView, an online tool that calculates the possible interactions in a crystal structure, was used to investigate the possible interactions. Not surprisingly, the main interactions are still the biotin binding and the

hydrophobic interactions with L124 (Figure 144a). With the three structures, we could show that the fluorescent moiety of the **biot-4DMN** and the substrate **1** are binding within the biotin binding vestibule and interact with the protein. Therefore, **biot-4DMN** could be used to detect the hydrophobicity of the Sav biotin binding vestibule, whereas substrate **1** for the change in fluorescence upon alkylation. Accordingly, it could be used for high-throughput screening by fluorescence measurements relying on a plate reader.
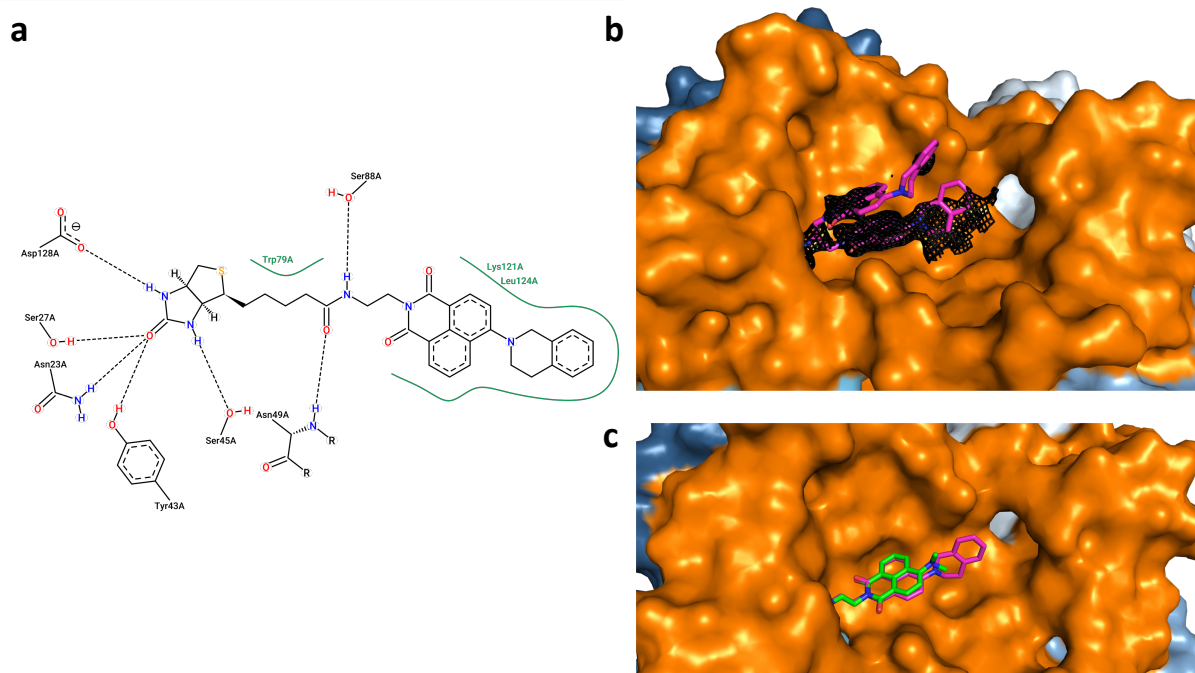
*Figure 144 **a** 2D representation of the interactions of Sav with substrate **1** generated by PoseView (https://proteins.plus/). The hydrogen bonds are represented as dotted lines, the hydrophobic interactions are represented as a green line. **b** Close-up view of the X-ray structure of substrate **1**·Sav (pdb 7zx9). The substrate **1** is represented as stick model in purple (atoms are color coded; nitrogen = blue, oxygen = red). The 2Fo-Fc difference map is displayed as black mesh (1σ). The protein is represented as a surface model with different shades of blue for each monomer. The biotin binding vestibule is highlighted in orange. **c** superposition of the crystal structure 7zx9 and 8aqd with one (matching) conformation only.*

5.5.4 Iron-porphyrin (**biot₂-Fe-pp**)

In 5.2.2 Investigation of streptavidin ArMs by native MS, we investigated the streptavidin complex with a bis-biotinylated iron-porphyrin (Figure 111a). As the information provided by this tool was ambiguous concerning the stoichiometry and the corresponding geometry of the metalloprotein, crystallization of the assembly was attempted. Various soaking experiments under aerobic conditions and with different Sav variants (WT, S112A, K121A,

S112A K121A) were unsuccessful. Therefore, cocrystallization and soaking experiments were run under anaerobic conditions with the Sav S112A K121A mutant. Finally, the crystal structure was obtained by soaking the anaerobically grown Sav S112A K121A crystals under the following conditions; 0.1 M MIB 7 pH 22.5 %v/v PEG 1500 and soaking for 48h. The resulting X-ray diffraction data of the dark red crystals (Figure 146a) was solved at 1.74 Å (pdb 7zof). Inspection of the X-ray structures revealed residual density in the biotin biding vestibule, and the iron porphyrin could be modeled into the density (Figure 145a). Modeling of the iron porphyrin into this electron density projected the iron (80 % occupancy) in the position of the anomalous density peak (Figure 145b).

As expected, the bis-biotinylated iron porphyrin **biot$_2$-Fe-pp** binds to both biotin binding sites flanking the vestibule. Nevertheless, residual electron density was found on both axial positions of iron, which could be modeled with imidazole present in the buffer. Furthermore, by an X-ray fluorescence measurement prior to subjecting the crystal to X-ray crystallography, the typical K-edge energy (7112 eV) for iron could be detected (Figure 146b), indicating that iron is indeed present in the crystal. Inspection of second coordination sphere interactions between cofactor **biot$_2$-Fe-pp** and close lying residues reveal a hydrogen bond between the sulfonic acid oxygen and the hydrogen from the residue N48 of Sav. With this structure, we could localize and verify the binding of the iron porphyrin into the two facing biotin binding pockets. However, the imidazole binding to the axial position of the iron porphyrin enforces a tilted orientation of the porphyrin ring compared to the protein surface. Accordingly, identifying amino acid residues that might be act as axial ligands in the absence of imidazole, is challenging based on this structure alone. Nevertheless, this structure unambiguously validates the design principle of a bis-biotinylated porphyrin complex that occupies the biotin-binding vestibule.
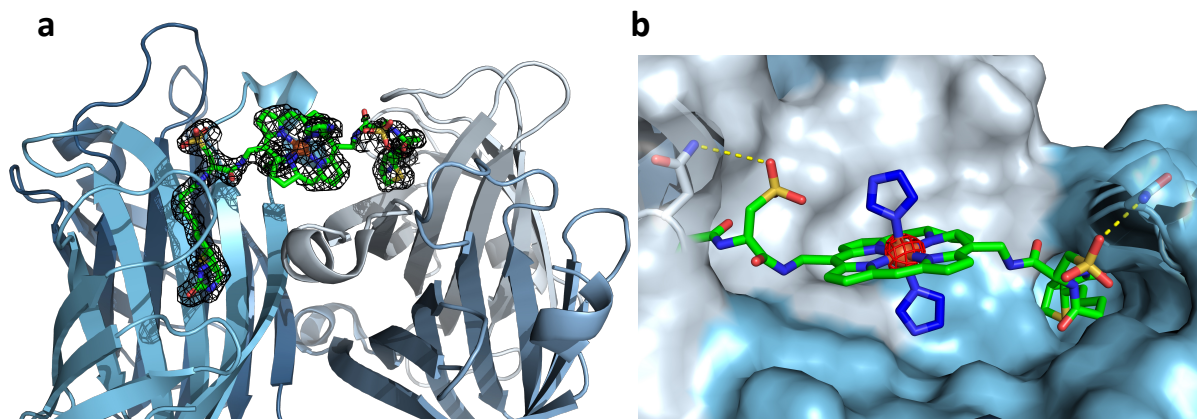
*Figure 145 **a** X-ray Crystal structure of the iron porphyrin **biot₂-Fe-pp**·SavS112AK121A (pdb 7zof). The iron porphyrin is represented as stick model in green (atoms are color coded; nitrogen = blue, oxygen = red, iron = orange and sulfur = yellow). The 2Fo-Fc difference map is displayed as black mesh (1σ). The protein is represented as a cartoon model with different shades of blue for each monomer. **b** Close-up view of the biotin binding vestibule revealing the exact location of the iron by the anomalous electron density highlighted as red mesh (6σ). The iron porphyrin is represented as stick model in green (atoms are color coded; nitrogen = blue, oxygen = red). The protein is represented as a surface model with different shades of blue for each monomer. The hydrogen bonds to residues N49 are shown in yellow. Residues N49 are represented as sticks.*
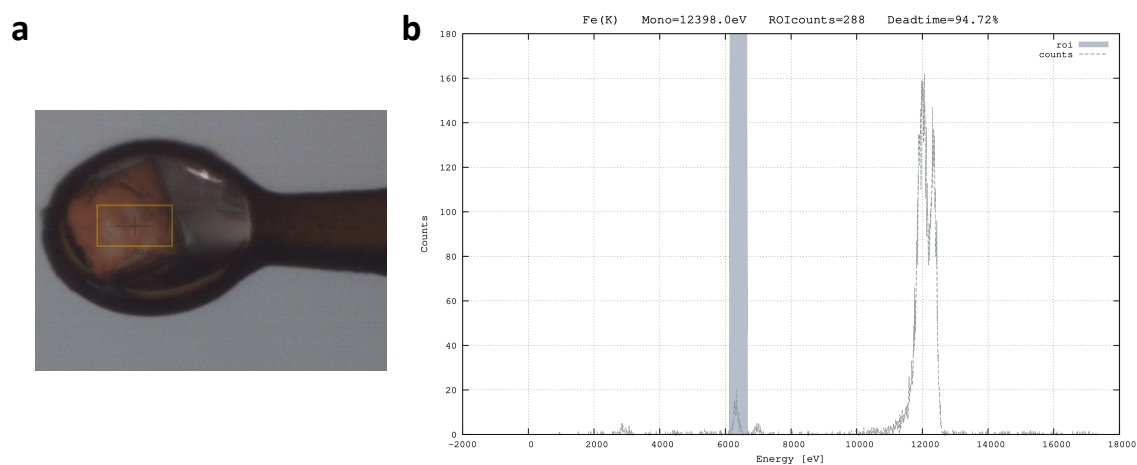
*Figure 146 **a** Picture of the soaked crystal revealing the dark red color. **b** X-ray fluorescence recorded at the SLS X06DA – PXIII beamline. The typical K-edge energy (7112 eV) for iron is highlighted with a grey bar.*

5.5.5 Fe$_4$S$_4$-cluster (**biot$_2$-Fe$_4$S$_4$**)

In 5.2.2 Investigation of streptavidin ArMs by native MS, we investigated the streptavidin complex with a bis-biotinylated Fe$_4$S$_4$-cluster **biot$_2$-Fe$_4$S$_4$** (Figure 114 and Figure 115). The results from the experiment can be used as proof of concept that the complex is indeed formed and the Fe$_4$S$_4$-cluster (**biot$_2$-Fe$_4$S$_4$**) remains intact after complexation and ionization under native MS conditions. In order to gain some information about the exact location of the cluster and possible interactions with the streptavidin scaffold, crystallization trials were set-up. Because of the fragile nature of the **biot$_2$-Fe$_4$S$_4$** under aerobic conditions, the crystallization trials were run in an inert gas vinyl tent (Coy Laboratory Products Inc.) containing a 95/5 % mixture of N$_2$/H$_2$. Different commercial screens were used (Index, Crystal Screen, PEG/Ion screen, JCSG Plus, BCS Screen, PACT premier, SG1 screen), and various conditions yielded streptavidin crystals within 21 days. However, soaking approaches were initially unsuccessful. Therefore, co-crystallizations with a reduced set of screens were evaluated (PEG/Ion screen, JCSG Plus, PACT premier, and Crystal Screen). No crystal was observed even after 80 days.

Going one step back, soaking experiments with different ligands were performed. The ligands contained different spacers (glycine (Gly), aspartate (Asp) and proline (Pro)) between the valeric acid side chain of biotin, and amine of the di-thiol (Figure 147). Crystals were obtained using 0.1 M MIB (malonic acid, imidazole and boric acid (2:3:3 Molar Ratio)) pH 7.0, 25 % w/v PEG 1500 after 7 days and could be soaked in the same solution with ligand **2** (Figure 149b). The resulting X-ray diffraction data of the soaked crystals was solved at 1.45 Å (pdb 7zxz) with a space group of C121, affording one Sav tetramer per asymmetric unit. Inspection of the X-ray structure revealed residual density in the biotin biding vestibule, and ligand **2** could be modeled into the density (Figure 150b). However, cyclization of the aspartic acid was observed (Figure 148a). This cyclization is known to happen in solution[332], nevertheless, usually, this cyclization is only an intermediate step, and the protein is stabilizing the succinimide (Figure 149a).
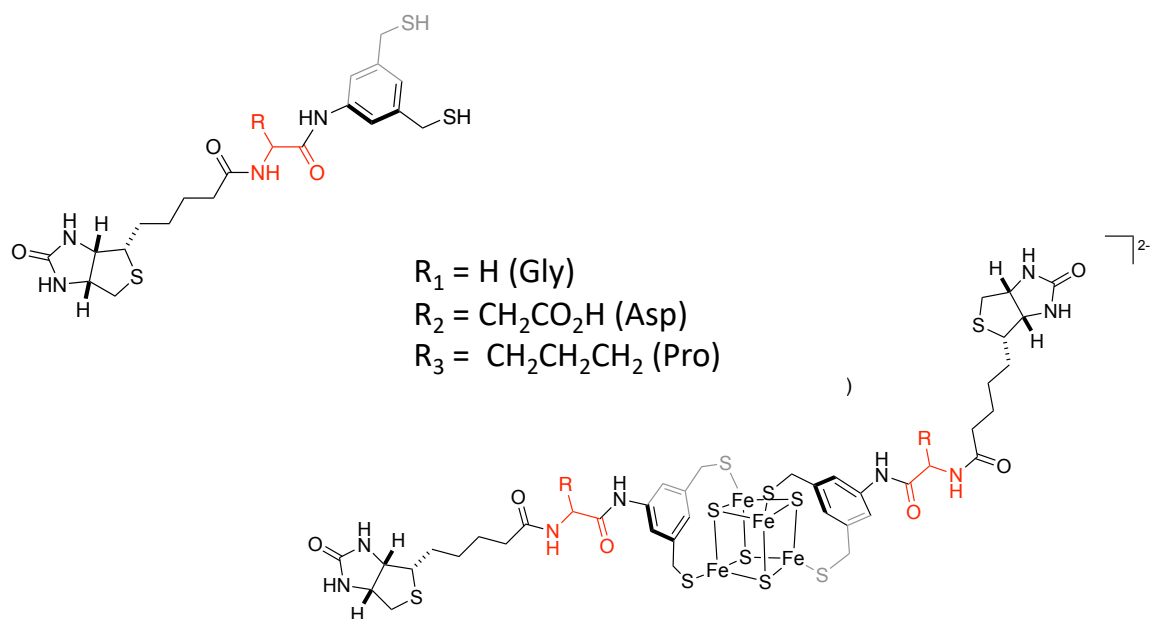
Figure 147 Chemical structure of the ligands and the Fe₄S₄ cluster thereof, evaluated in the crystallography studies. The linker region is highlighted in red.

$R_1 = H$ (Gly)
$R_2 = CH_2CO_2H$ (Asp)
$R_3 = CH_2CH_2CH_2$ (Pro)

a

b
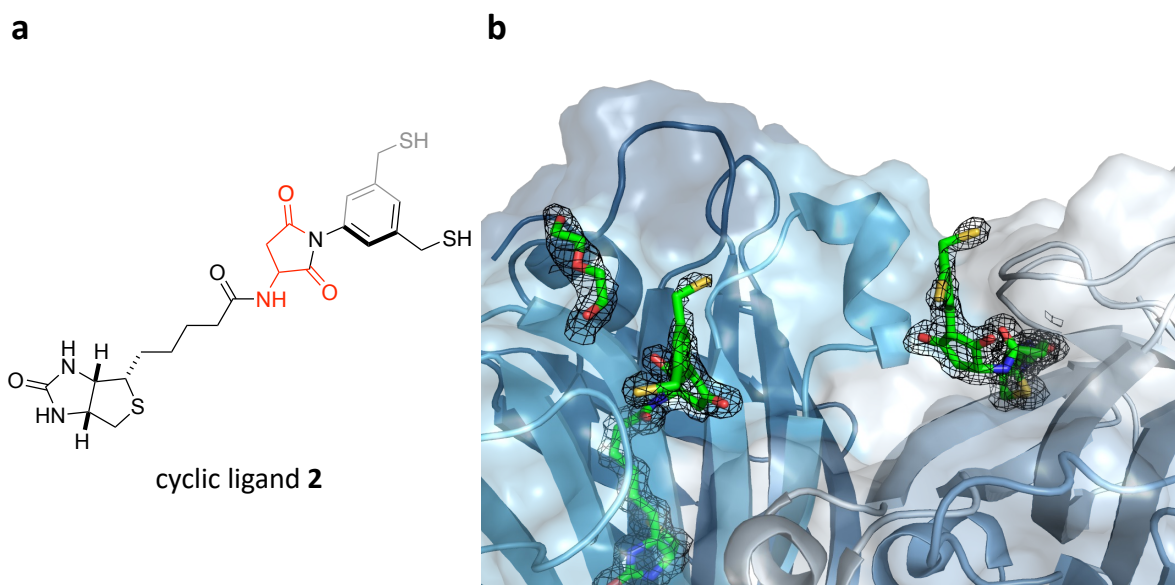
cyclic ligand **2**

*Figure 148 **a** Chemical structure of the cylized ligand **2**. **b** X-ray Crystal structure of the Fe₄S₄ cluster biot₂-Fe₄S₄·SavS112AK121A (pdb 7zxz). The Fe₄S₄ cluster is represented as stick model in green (atoms are color coded; nitrogen = blue, oxygen = red, carbon = green and sulfur = yellow). The 2Fo-Fc difference map is displayed as black mesh (1σ). The protein is represented as a cartoon and surface model with different shades of blue for each monomer.*
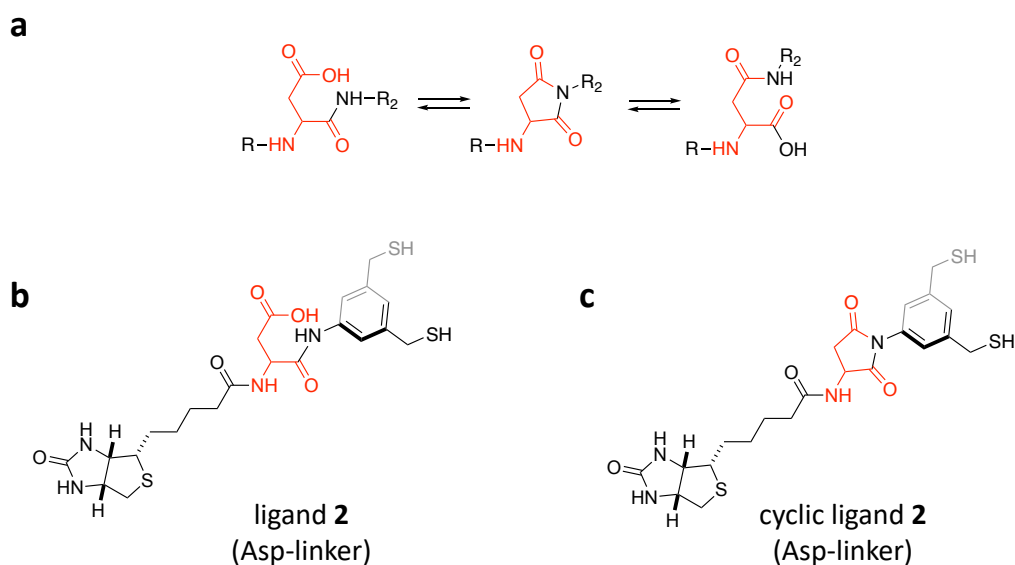
*Figure 149 **a** Equilibrium reaction leading to the cyclisation of aspartate. **b** Chemical structure of the ligand **2** used for the crystallographic study. **c** Chemical structure of the cyclized ligand **2**.*

Based on the small molecule crystal structure (ccdc 925636) a model structure of the Fe_4S_4-cluster (**biot$_2$-Fe$_4$S$_4$**) was computationally modeled (Figure 150c). From the model, it is apparent that a suitable spacer would be glycine and the distance between the two facing thiol groups binding to the cluster would be 6.5 Å. For comparison, the aspartate spacer (ligand **2**) yields a distance of 11.7 Å between the two facing thiol groups. Therefore, the focus was set on the crystallization of the glycine spacer Fe_4S_4-cluster (**biot$_2$-Fe$_4$S$_4$**). Next to the use of an oxygen scavenging system[333] for co-crystallization under the newly identified conditions, also the *in-situ* formation of the Fe_4S_4-cluster after soaking with the ligands was evaluated. However, both approaches were unsuccessful, therefore a new screen dividing each buffer component and combining them in different ratios and pH ranges was set up. Streptavidin crystals could be obtained from 0.1 M borate pH 7-8.5 with 25 % w/v PEG 1500.
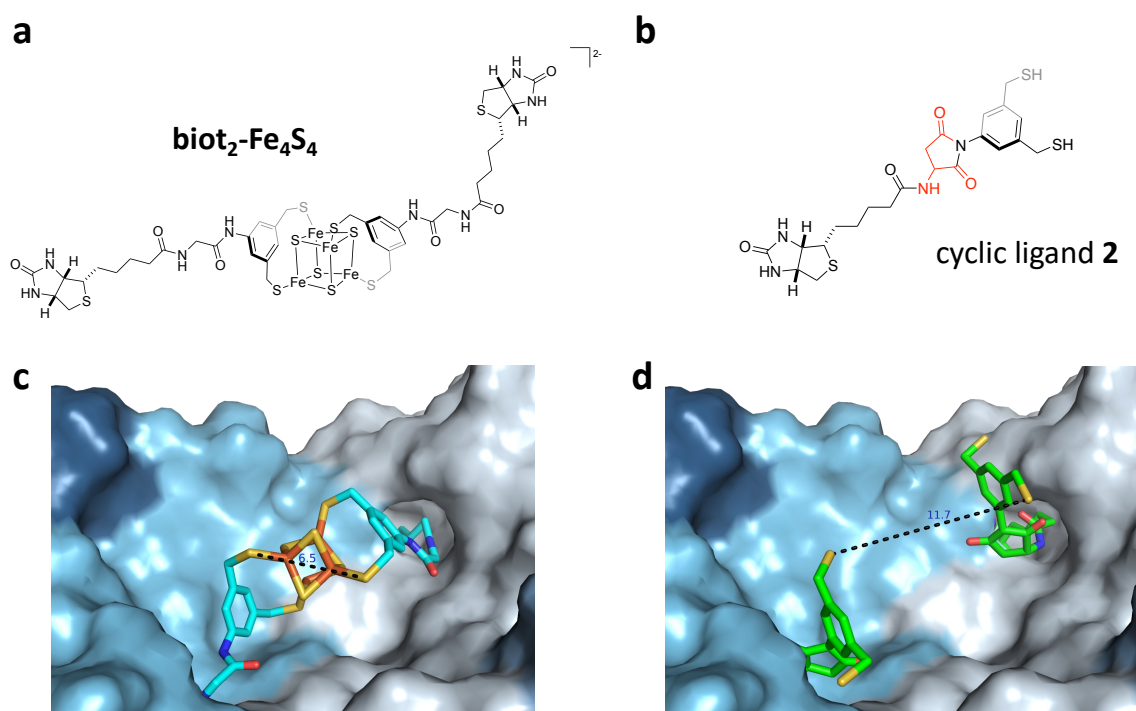
*Figure 150 **a** Chemical structure of **biot₂-Fe₄S₄**. **b** Chemical structure of the cyclic ligand **2**. **c** Close-up view of the biotin binding vestibule from the computationally modeled structure with the marked distance between the thiol groups of the facing edges of the cluster (**biot₂-Fe₄S₄**). **d** Close-up view of the biotin binding vestibule of the crystal structure of ligand **2**·SavS112AK121A (pdb 7zxz). Distance between the thiol groups of facing ligands is labeled with black dotted lines. The protein is represented as a surface model (**c** and **d**) with different shades of blue for each monomer. The Fe₄S₄-cluster is represented as stick model in cyan and the ligand **2** is represented as stick model in green (atoms are color coded; nitrogen = blue, oxygen = red, iron = orange and sulfur = yellow). The 2Fo-Fc difference map is displayed as black mesh (1σ).*

In order to confirm that the Fe₄S₄-cluster (**biot₂-Fe₄S₄**) remains intact under the crystallization conditions, the UV/Vis spectra of the complex in aqueous solution spiked with different DMSO concentrations (Figure 151) and in the buffer (Figure 152) at different timepoints were recorded. As can be appreciated from the characteristic absorption band at around 430 nm, **biot₂-Fe₄S₄** remained intact in 99% water after 17h under anaerobic conditions. This was a good starting point, as Sav crystals typically tend to dissolve in high concentrations of organic solvents. Furthermore, as presented in Figure 152, **biot₂-Fe₄S₄** remained intact in the crystallization buffer, even for 4 days at a pH of 7.5 and 8.5 under strict anaerobic conditions.
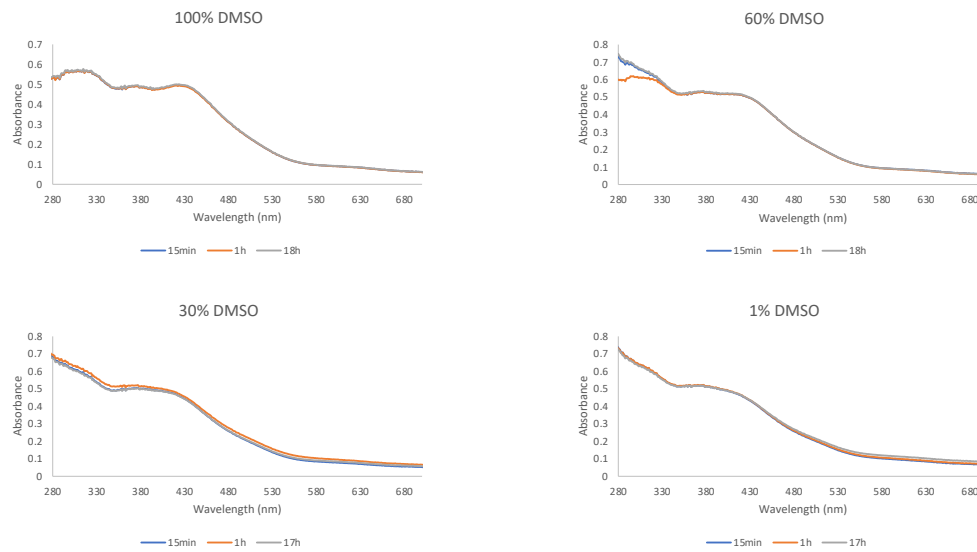
*Figure 151 UV/Vis spectra of the **biot₂-Fe₄S₄** in DMSO/water mixtures under anaerobic conditions recorded after 15 min, 1h and 17h.*
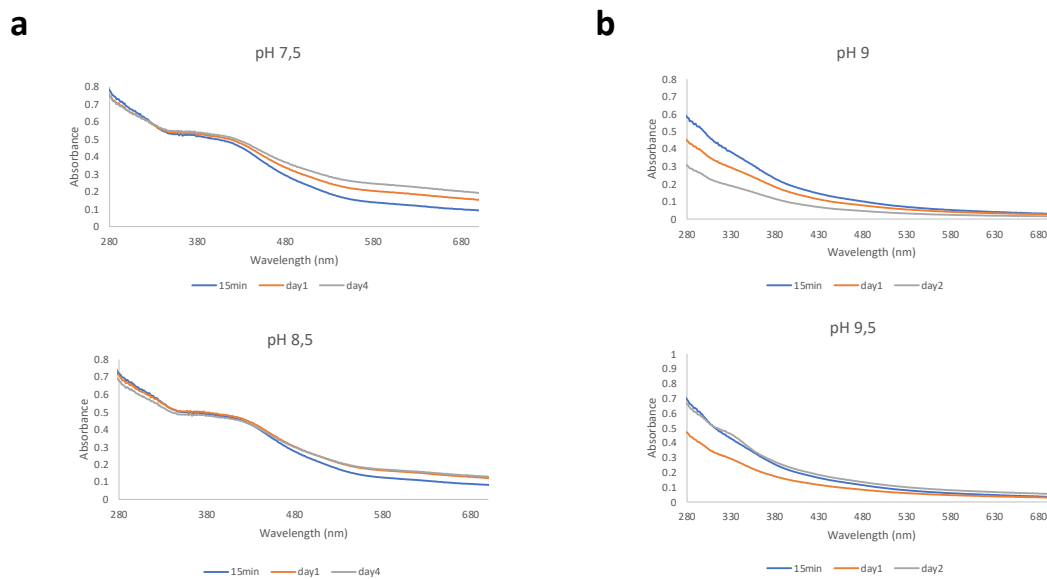
*Figure 152 UV/Vis spectra of the **biot₂-Fe₄S₄** in buffers under anaerobic conditions recorded after 15 min, 24h and 48h. **a** 50 mM borate, 25 % w/v PEG 1500 **b** 50 mM MIB (malate, imidazole, boric acid) 25 % w/v PEG 1500.*

Having identified conditions in which the protein crystallizes and the cluster (**biot₂-Fe₄S₄**) remains intact, soaking experiments were run (1, 30, 60 min, 3h, 6h, 18h, and 24h of soaking prior to freezing the resulting crystal). After a soaking time of 24h, dark brown crystals were

obtained (Figure 153c). In order to have proof that the color comes from iron, an X-ray fluorescence spectrum of the crystal was recorded prior to data diffraction measurement. A signal for the K-edge of iron was unambiguously identified (Figure 153b).
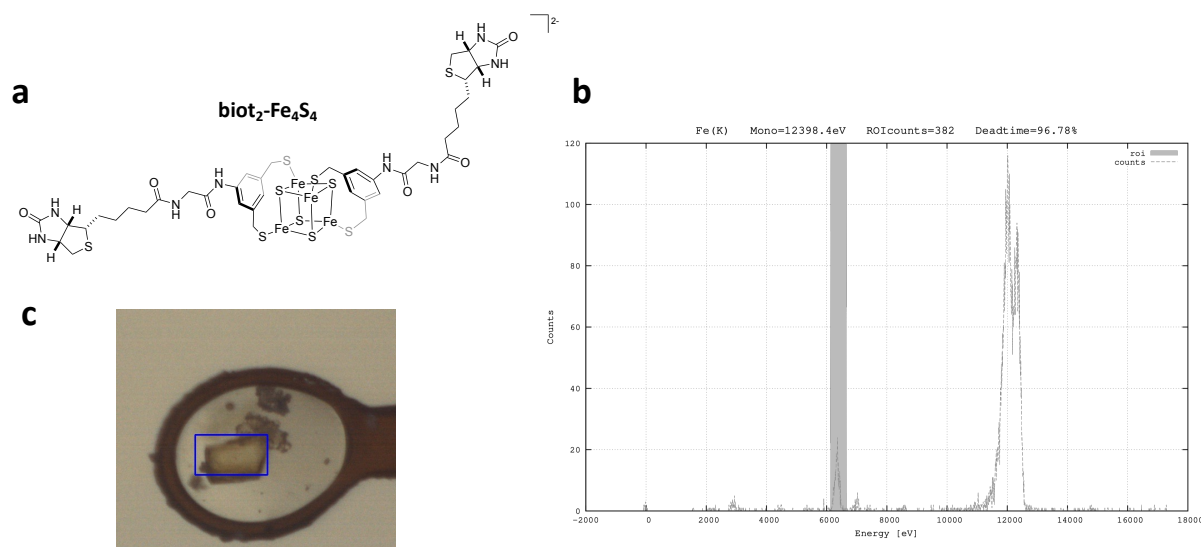
*Figure 153 **a** Fe4S4-cluster **1 b** X-ray fluorescence recorded at the SLS X06DA – PXIII beamline. The typical K-edge energy (7112 eV) for iron is marked with a grey bar. **c** Picture of the soaked crystal showing the dark brown color typical for iron complexes.*

The resulting X-ray diffraction of the dark brown crystals was collected and the structure was solved at 1.85 Å (pdb 8aq0) in the C121 space group. One tetramer was obtained per asymmetric unit. Inspection of the X-ray structure revealed residual density in the biotin biding vestibule, and the **biot$_2$-Fe$_4$S$_4$** could be modeled into the density (Figure 154a and c). Modeling of the **biot$_2$-Fe$_4$S$_4$** into this electron density projected the **biot$_2$-Fe$_4$S$_4$** in the position of the anomalous density peak (Figure 154b and d). However, the anomalous density was not well refined, and the electron density for the aromatic di-thiol part was not well defined either. This might be caused by the flexibility of the cluster or partial decomposition during the measurement or sample handling.
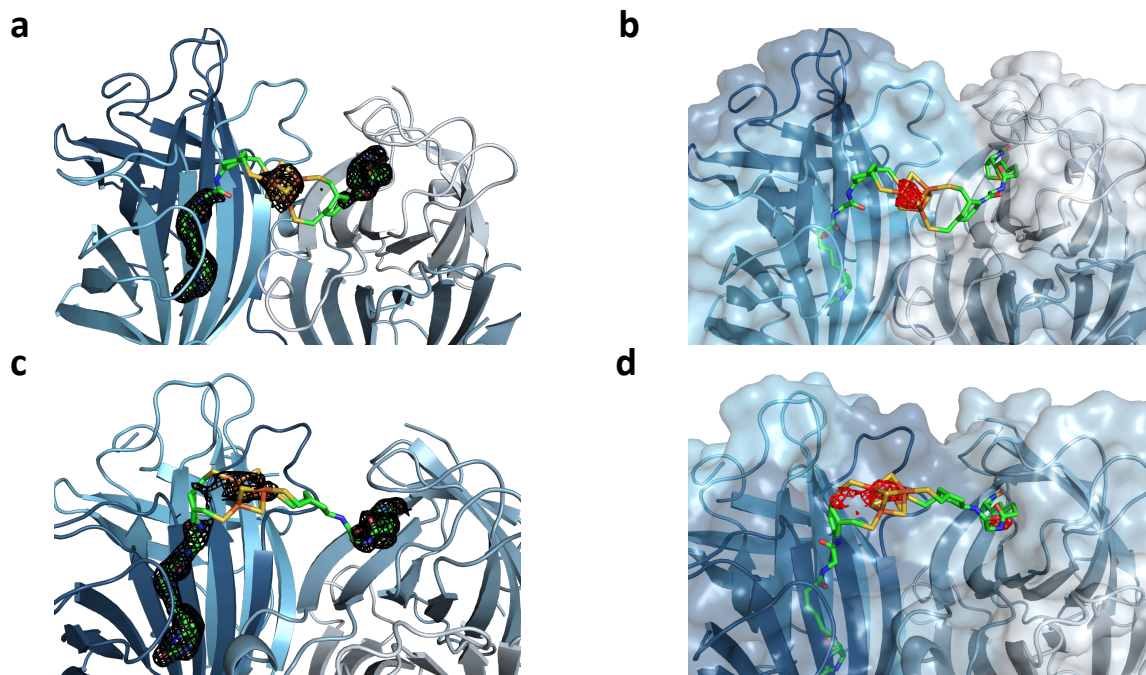
*Figure 154 Crystal structure of **biot₂-Fe₄S₄**·SavS112AK121A (pdb 8aq0). The protein is represented as a cartoon model (**a** and **c**) and surface model (**b** and **d**) with different shades of blue for each monomer. The Fe_4S_4-cluste (**biot₂-Fe₄S₄**) is represented as stick model in green (atoms are color coded; nitrogen = blue, oxygen = red, iron = orange and sulfur = yellow). The 2Fo-Fc difference map is displayed as black mesh (1σ). The anomalous electron density highlighted as red mesh (4σ). **a** and **b** Close-up view of the biotin binding vestibule A. **c** and **d** Close-up view of the biotin binding vestibule B.*

When superimposing the QM/MM-calculated structure with the crystal structure (Figure 155), differences are apparent. For example, the location of the cluster in one of the biotin binding vestibules is different. This is most probably caused by the input model for the calculations, where the position of the biotin-spacer part is not as tightly fixed as it should be.
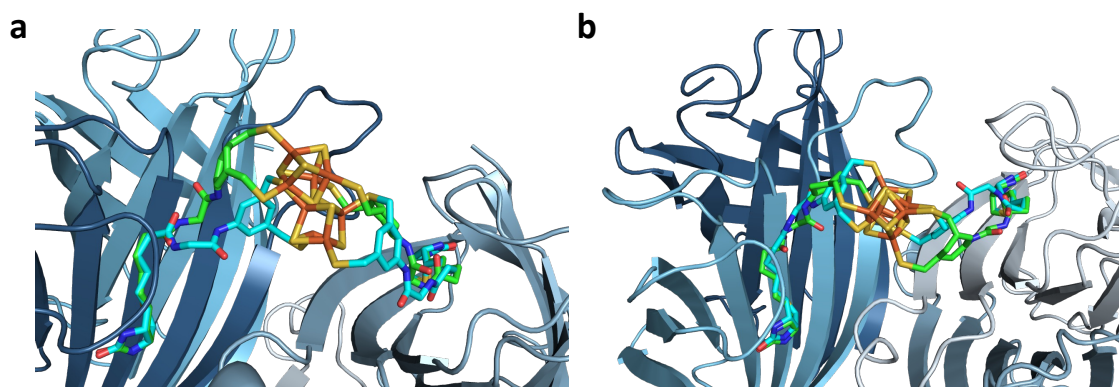
*Figure 155 Overlay of the crystal structure of **biot₂-Fe₄S₄**·SavS112AK121A (pdb 8aq0) and the computationally-modeled structure. The protein is represented as a cartoon with different shades of blue for each monomer. The Fe₄S₄-cluster (**biot₂-Fe₄S₄**) is represented as stick model in green (atoms are color coded; nitrogen = blue, oxygen = red, iron = orange and sulfur = yellow) for the crystal structure and in cyan for the computationally modeled structure. **a** Close-up view of the biotin binding vestibule A. **b** Close-up view of the biotin binding vestibule B.*

5.5.6 [Cp*Ir(biot-AQ)Cl] for C-H activation

One of the newly studied reactions in our group is the C-H activation by the [Cp*Ir(biot-AQ)Cl] (Figure 156a) cofactor incorporated into streptavidin. In order to identify residues to target for directed evolution of the ArM, the structure of two variants [Cp*Ir(biot-AQ)Cl]·SavS112A and [Cp*Ir(biot-AQ)Cl]·SavS112I were determined by protein crystallography. Those two variants were selected, as the mutation on S112 had a significant influence on the enantioselectivity of the reaction. In order to obtain crystal structures with well localized ligands, we relied on co-crystallization as this approach often affords a better localization of the ligand.

Crystals of [Cp*Ir(biot-AQ)Cl] ·SavS112A (pdb 8aqy) were obtained using 0.1 M HEPES 7.5 pH, 10 % w/v PEG 8K and 8 % v/v EG as crystallization buffer within 25 days. The resulting X-ray diffraction data of the yellow crystals was solved at 1.65 Å (pdb 8aqy) in a C121 space group. Inspection of the X-ray structures revealed residual density in the biotin biding vestibules, and [Cp*Ir(biot-AQ)Cl] could be modeled into the density (Figure 156b). Modeling of the [Cp*Ir(biot-AQ)Cl] into this electron density projected the iridium in the position of the anomalous density peak (Figure 156b). No further anomalous density was found for iridium. The occupancy of the metal and Cp* ligand was set to 60% in order to fit better to the

observed data. This partial occupancy may arise from distortion or decomposition during the crystallization, sample handling or measurement.

Crystals of [Cp*Ir(biot-AQ)Cl] ·SavS112I (pdb 8aqx) were obtained using 0.2 M NaCl, 0.1 M BIS-TRIS 6.5 pH and 25 % w/v PEG 3350 as crystallization buffer within 21 days. The resulting X-ray diffraction data of the yellow crystals was solved at 1.85 Å (pdb 8aqx) in a C121 space group. Inspection of the X-ray structures revealed residual density in the biotin biding vestibules, and [Cp*Ir(biot-AQ)Cl] could be modeled into the density (Figure 156c). Modeling of the [Cp*Ir(biot-AQ)Cl] into this electron density projected the iridium in the position of the anomalous density peak (Figure 156c). No additional anomalous density was found for residual iridium. The occupancy of the metal and Cp* ligand was set to 80% in order to fit better to the observed data. This partial occupancy can arise from decomposition during the crystallization, sample handling or measurement.
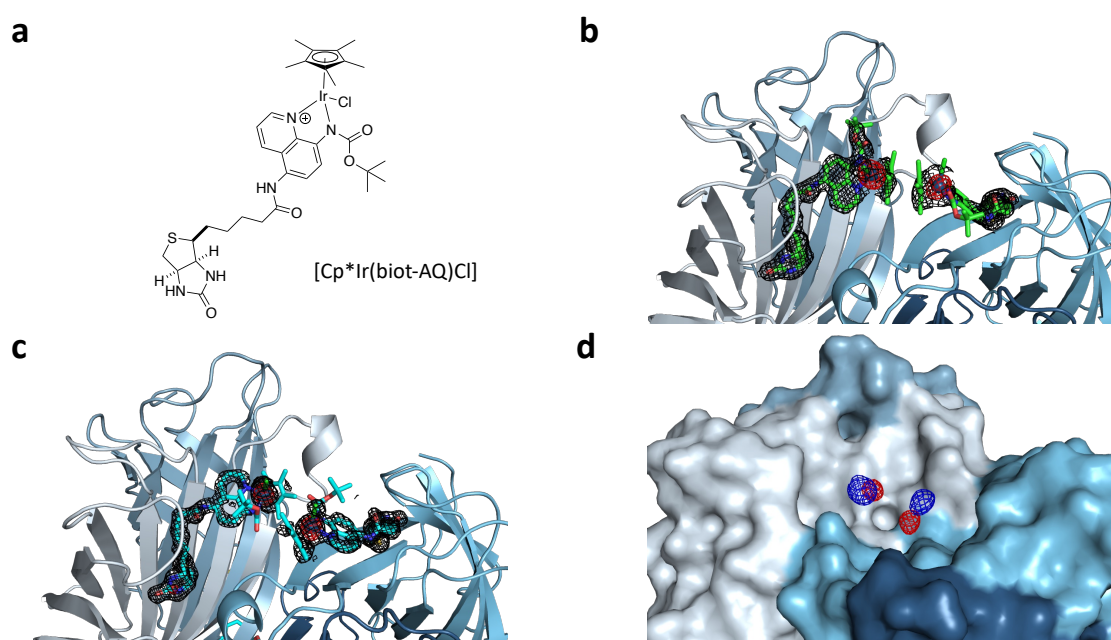
*Figure 156 **a** Structure of [Cp*Ir(biot-AQ)Cl] **b** Close-up view of the biotin binding vestibule from the crystal structure of [Cp*Ir(biot-AQ)Cl]·SavS112I (pdb 8aqx). **c** Close-up view of the biotin binding vestibule from crystal structure of [Cp*Ir(biot-AQ)Cl]·SavS112A (pdb 8aqy).**d** Close-up view of the biotin binding vestibule with the anomalous density maps from both crystal structures. (blue = 8aqx, red = 8aqy) The protein is represented as a cartoon model (**b, c**) and surface model (**d**) with different shades of blue for each monomer. [Cp*Ir(biot-AQ)Cl] is represented as stick model in green (8aqx) and cyan (8aqy) (atoms are color coded; nitrogen = blue, oxygen = red, iron = orange and sulfur = yellow). The 2Fo-Fc difference map is displayed as black mesh (1σ). The anomalous electron density highlighted as red mesh (8σ) (**b** and **c**).*

Comparing both structures, one apparent difference arises; the tert-butyloxycarbonyl of the complex points in different directions depending on the mutation of Sav S112. In the case of the Sav S112I mutant, the tert-butyloxycarbonyl is pointing towards the S112I residue, whereas for Sav S112A, the group point towards the solvent-exposed space. Therefore, two different enantiomers of the complex might bind to Sav depending on the mutation of S112. In case of the [Cp*Ir(biot-AQ)Cl]·SavS112A a $Ir_{(R)}$ configuration was found and for [Cp*Ir(biot-AQ)Cl]·SavS112I a $Ir_{(s)}$ was detected. This could be (part of) the reason why a pronounced difference in enantioselectivity in the reaction was observed, depending on the nature of the residue at position S112. Nevertheless, this observation cannot be considered as the only evidence for selectivity. It needs to be supported by further experiments. An additional difference of the structures is the slightly changed position of the iridium (Figure 156d). In case the metal center is more deeply embedded into the protein, the influence of the secondary coordination sphere would be more prominent on the reactivity.

5.5.7 Experimental

General information. Chemicals and crystallization screens were purchased from Sigma Aldrich, Jena Biosience, Molecular Dimensions and Hampton Research.

Protein crystallization. Sav and the mutants thereof were expressed and purified as described previously.[104] Lyophilized Sav was dissolved in ultrapure water (18.2 MΩ·cm, MilliQ, Millipore Corporation, Burlington, USA). For sitting-drop vapor diffusion in a 24-well crystallization plate Sav (5 µL of a 20 mg/mL stock solution) was mixed with the precipitation buffer (5 µL). The drop was equilibrated against a reservoir of the precipitation buffer (700 µL at 20 °C). Crystals of Sav grew within three days. For soaking, Sav crystals were individually transferred into sitting-drop depressions containing the precipitation buffer (1.5 µL) and the ligand solution (0.15 µL, 10 mM in DMSO). After soaking (20 °C) the crystals were flash- frozen in liquid nitrogen. The crystallization procedure for anaerobic crystals is adapted as follows; For sitting drop vapor diffusion in a 96-well 2-drop plate (SwissSci) 1 µL reservoir solution of crystallization screen was combined with 1 µL protein (20 mg/ml) solution under strict

exclusion of O_2 in an inert gas vinyl tent (Coy Laboratory Products Inc.) containing 95/5 % mixture of N_2/H_2.

High-throughput protein crystallization. The 96 well crystallization screens were set up mixing 0.20 μL mother liquor with 0.20 μL protein solution (20 mg/mL) with a robot (Crystal Gryphon, Art Robins Instruments, USA; MRC 3-well plates, Jena Bioscience, Germany) and stored at 20 °C in an automated imaging system (Formulatrix Rock Imager SONICC). Crystals from the screens were picked and directly flash-frozen in liquid nitrogen prior to data collection.

Co-crystallization. Protein crystals were obtained using a sitting drop vapor diffusion experiment. The lyophilized protein mutants Sav S112A and Sav S112L were dissolved in a 20 mM Tris-HCl solution (pH = 7.0) to a concentration of 2.5 mg/mL. A 10 mM solution of [Cp*Ir(biot-AQ)Cl] in dimethylsulfoxide (DMSO) (5 μL) was added to the solution and the mixture was incubated at room temperature overnight. After concentration to 25 mg/mL by ultracentrifugation (Amicon® ultra centrifugal filters, Merck; cut-off of 10 kDa), crystals were grown by sitting drop vapor diffusion using 0.1 M HEPES 7.5 pH, 10 % w/v PEG 8K and 8 % v/v EG or 0.2 M NaCl, 0.1 M BIS-TRIS 6.5 pH and 25 % w/v PEG 3350 as crystallization buffer. Crystals grew within 21 days and were flash frozen in liquid nitrogen without further cryoprotection.

Data collection and processing. Protein-crystal diffraction data were collected (at 100 K) at the Swiss Light Source beam line PSI and PSIII at a wavelength of 1.0 Å. Crystal indexing, integration and scaling[106,107] were carried out with the program XDS[108] and AIMLESS[109] using the graphical interface CCP4i2[110] of the CCP4 suite.[334]

Structure-solution and refinement. The ligand structure files were generated with Phenix and eLBOW.[335] The structure was solved by molecular replacement using the program PHASER MR[111] and the structure pdb 3bc2 and 3pk2, devoid of the Ir-cofactor and water molecules. For the structure refinement, REFMAC[112,113] of the CCP4 Suite was used. For structure modelling, water picking and electron-density visualization the software COOT[114] was used.

Figures were generated with PyMOL (the PyMOL Molecular Graphics System, Version 1.821, Schrödinger, LLC). Amino acid residues 1-14 and 172-196 are not resolved in the electron density, presumably due to disorder.

6. Conclusion and Outlook

In the past, our group mainly relied on the chemical optimization of the catalysts by tuning their electronic and steric properties by changing the ligands' chemical properties. Furthermore, they could use some degree of genetic optimization by changing the amino acid residues of the protein host. However, this optimization was limited by the number of residues influencing the reactivity and selectivity. Also, the metal catalyst remained, in many cases, rather solvent-exposed. Nevertheless, with chemo-genetic optimization, new-to-nature reactions with remarkable selectivities and reactivities could be catalyzed by the ArMs. Herein, we engineered a chimeric version of streptavidin, with a shielded biotin binding vestibule, which mimics an active site of an enzyme more closely. This chimeric host significantly broadens the possibilities for genetic optimization of the scaffold for the optimization of ArMs. Furthermore, inspired by the scdSav, we generated the chimeric scdSav-SOD, which may allow to mutate independently aminoacid residues that line the biotin-binding vestibule

The chimeric Sav-SOD was evaluated for two reactions; the Noyori-type catalyst-based artificial transfer hydrogenation and the newly introduced gold(I) biotinylated N-heterocyclic carbene complex-based hydroamination reaction. In both cases, the influence of the loop on the reactivity and selectivity could be highlighted.

We furthermore extended the biorthogonal gold(I) catalysis beyond ArM technology by functionalizing somatostatin analogs LM3 and TATE with gold(I) N-heterocyclic carbene complexes. These gold peptides were shown to be stable and catalytically active for the hydroarylation of coumarin analogs without requiring base or additives. The same peptides have been shown to be efficient for the release of an alcohol and furan by an intramolecular cyclization followed by 1,2-elimination of pro-furan moieties.

Native mass spectrometry was introduced to the lab to investigate the ArMs and was shown to be a powerful tool. Furthermore, different ArMs were crystallized in order gain information concerning: i) the stoichiometry of the complexes as well as ii) the details of the molecular structure. Most importantly, the crystal structure of the Fe4S4-cluster was obtained.

Having revealed the influence of the SOD-loop, in future different loops from naturally occurring proteins or *de-novo* designed ones could be implemented to the chimeric Sav

approach and used as novel scaffolds for the generation of ArMs. Also, the scdSavSOD could be used to study the influence of cofactor:Sav ratios, as well as the symmetry-related amino acid residues on ArMs.

Native mass spectrometry experiments could be further developed in future in order to obtain not only information of the ArMs formation and stability, but also to screen for the best binding ligands and evaluate the association/dissociation parameters of the ArMs.

X-ray crystallography is still one of the most powerful approaches concerning the atomic structure of ArMs and will be used whenever possible to inspect the ArMs generated.

7. Curriculum Vitae

Nico Valerio Igareta

From Möhlin (AG), Switzerland

Personal information:

Date of birth:	03.03.1992
Place of birth:	Liestal (BL) Switzerland
Nationality:	Swiss/Spanish

Education:

2018-2022	**PhD in Chemistry, Ward Lab, University of Basel**
2017-2018	**Master of Sciences in Chemistry, Seebeck Lab, University of Basel**
2013-2017	**Bachelor of Sciences in Chemistry, University of Basel**

Academic Supervision and Teaching:

2018–2021	**General Chemistry Practicals for Biologists**

Workshop:

2018/09	**Hands-on-Pichia, TU Graz**

Scientific Publications:

A) Christoffel, F.*; Igareta, N. V.*; Pellizzoni, M. M.; Tiessler-Sala, L.; Lozhkin, B.; Spiess, D. C.; Lledós, A.; Maréchal, J.-D. D.; Peterson, R. L.; Ward, T. R. Design and Evolution of Chimeric Streptavidin for Protein-Enabled Dual Gold Catalysis. *Nat. Catal.* **2021**, *4* , 643–653. https://doi.org/10.1038/s41929-021-00651-9. (* = equal contribution)

B) Stein, A.; Chen, D.; Igareta, N. V.; Cotelle, Y.; Rebelein, J. G.; Ward, T. R. A Dual Anchoring Strategy for the Directed Evolution of Improved Artificial Transfer Hydrogenases Based on Carbonic Anhydrase. *ACS Cent. Sci.* **2021**, *7*, 1874–1884. https://doi.org/10.1021/acscentsci.1c00825.

C) Wu, S.; Zhou, Y.; Rebelein, J. G.; Kuhn, M.; Mallin, H.; Zhao, J.; Igareta, N. V.; Ward, T. R. Breaking Symmetry: Engineering Single-Chain Dimeric Streptavidin as Host for Artificial Metalloenzymes. *J. Am. Chem. Soc.* **2019**, 141, 15869–15878. https://doi.org/10.1021/jacs.9b06923.

D) Igareta, N. V.; Tachibana, R.; Spiess, D. C.; Peterson, R. L.; Ward, T. R. Spiers Memorial Lecture: Shielding the active site: a streptavidin superoxide-dismutase chimera as a host protein for asymmetric transfer hydrogenation., *Faraday Disscusions, Advanced Article* **2023**.

8. Bibliography

(1)   Robertus, J. D.; Kraut, J.; Alden, R. A.; Birktoft, J. J. Subtilisin; a Stereochemical Mechanism Involving Transition-State Stabilization. *Biochemistry* **1972**, *11* (23), 4293–4303. https://doi.org/10.1021/bi00773a016.

(2)   Åqvist, J.; Kazemi, M.; Isaksen, G. V.; Brandsdal, B. O. Entropy and Enzyme Catalysis. *Acc. Chem. Res.* **2017**, *50* (2), 199–207. https://doi.org/10.1021/acs.accounts.6b00321.

(3)   Richard, J. P. Enzymatic Rate Enhancements: A Review and Perspective. *Biochemistry* **2013**, *52* (12), 2009–2011. https://doi.org/10.1021/bi301515j.

(4)   Haldane, J. *Enzymes*; Longmans, Green and co: New York, 1930.

(5)   Pauling, L. NATURE OF FORCES BETWEEN LARGE MOLECULES OF BIOLOGICAL INTEREST. *Nature* **1948**, *161* (4097), 707–709.

(6)   Wolfenden, R. Thermodynamic and Extrathermodynamic Requirements of Enzyme Catalysis. *Biophys. Chem.* **2003**, *105* (2–3), 559–572. https://doi.org/10.1016/S0301-4622(03)00066-8.

(7)   Schramm, V. L. Enzymatic Transition States, Transition-State Analogs, Dynamics, Thermodynamics, and Lifetimes. *Annu. Rev. Biochem.* **2011**, *80*, 703–732. https://doi.org/10.1146/annurev-biochem-061809-100742.

(8)   Miyashita, A.; Yasuda, A.; Takaya, H.; Toriumi, K.; Ito, T.; Souchi, T.; Noyori, R. Synthesis of 2,2/-Bis(Diphenylphosphino)-l,l/-Binaphthyl (BINAP), an Atropisomeric Chiral Bis(Triaryl) Phosphine, and Its Use in the Rhodium(I)-Catalyzed Asymmetric Hydrogenation of a-(Acylamino)Acrylic Acids. *J. Am. Chem. Soc.* **1980**, *102*, 7932–7934.

(9)   Knowles, W. S. Asymmetric Hydrogenations (Nobel Lecture 2001). *Adv. Synth. Catal.* **2003**, *345* (1–2), 3–13. https://doi.org/10.1002/adsc.200390028.

(10)   Noyori, R. Asymmetric Catalysis: Science and Opportunities (Nobel Lecture 2001). *Adv. Synth. Catal.* **2003**, *345* (1–2), 15–32. https://doi.org/10.1002/adsc.200390002.

(11)   Kolb, H. C.; VanNieuwenhze, M. S.; Sharpless, K. B. Catalytic Asymmetric Dihydroxylation. *Chem. Rev.* **1994**, *94*, 2483–2547.

(12)   Yamamura, K.; Kaiser, E. T. Studies on the Oxidase Activity of Copper(II) Carboxypeptidase A. *J. Chem. Soc. Chem. Commun.* **1976**, No. 20, 830–831. https://doi.org/10.1039/C39760000830.

(13) Wilson, M. E.; Whitesides, G. M. Conversion of a Protein to a Homogeneous Asymmetric Hydrogenation Catalyst by Site-Specific Modification with a Diphosphinerhodium(I) Moiety. *J. Am. Chem. Soc.* **1978**, *100* (1), 306–307. https://doi.org/10.1021/ja00469a064.

(14) Akabori, S.; Izumi, Y.; Fujii, Y.; Sakurai, S. An Asymmetric Catalyst. *Nature* **1956**, *178*, 323–324. https://doi.org/10.1246/nikkashi1948.77.1374.

(15) Hyster, T. K.; Ward, T. R. Genetische Optimierung von Metalloenzymen: Weiterentwicklung von Enzymen Für Nichtnatürliche Reaktionen. *Angew. Chemie* **2016**. https://doi.org/10.1002/ange.201508816.

(16) Renata, H.; Wang, Z. J.; Arnold, F. H. Expanding the Enzyme Universe: Accessing Non-Natural Reactions by Mechanism-Guided Directed Evolution. *Angew. Chemie - Int. Ed.* **2015**, *54* (11), 3351–3367. https://doi.org/10.1002/anie.201409470.

(17) Reetz, M. T. Directed Evolution of Artificial Metalloenzymes: A Universal Means to Tune the Selectivity of Transition Metal Catalysts? *Acc. Chem. Res.* **2019**, *52*, 336–344. https://doi.org/10.1021/acs.accounts.8b00582.

(18) Schwizer, F.; Okamoto, Y.; Heinisch, T.; Gu, Y.; Pellizzoni, M. M.; Lebrun, V.; Reuter, R.; Köhler, V.; Lewis, J. C.; Ward, T. R. Artificial Metalloenzymes: Reaction Scope and Optimization Strategies. *Chem. Rev.* **2018**, *118* (1), 142–231. https://doi.org/10.1021/acs.chemrev.7b00014.

(19) Dydio, P.; Key, H. M.; Nazarenko, A.; Rha, J. Y.-E.; Seyedkazemi, V.; Clark, D. S.; Hartwig, J. F. An Artificial Metalloenzyme with the Kinetics of Native Enzymes. *Science (80-.).* **2016**, *354* (6308), 102–106.

(20) Basler, S.; Studer, S.; Zou, Y.; Mori, T.; Ota, Y.; Camus, A.; Bunzel, H. A.; Helgeson, R. C.; Houk, K. N.; Jiménez-Osés, G.; Hilvert, D. Efficient Lewis Acid Catalysis of an Abiological Reaction in a de Novo Protein Scaffold. *Nat. Chem.* **2021**, *13* (3), 231–235. https://doi.org/10.1038/s41557-020-00628-4.

(21) Oohora, K.; Onoda, A.; Hayashi, T. Hemoproteins Reconstituted with Artificial Metal Complexes as Biohybrid Catalysts. *Acc. Chem. Res.* **2019**, *52* (4), 945–954. https://doi.org/10.1021/acs.accounts.8b00676.

(22) Rosati, F.; Roelfes, G. Artificial Metalloenzymes. *ChemCatChem* **2010**, *2* (8), 916–927. https://doi.org/10.1002/cctc.201000011.

(23) Cangelosi, V. M.; Deb, A.; Penner-Hahn, J. E.; Pecoraro, V. L. A de Novo Designed Metalloenzyme for the Hydration of CO2. *Angew. Chemie - Int. Ed.* **2014**, *53* (30), 7900–7903. https://doi.org/10.1002/anie.201404925.

(24) Regan, L.; DeGrado, W. F. Characterization of a Helical Protein Designed from First Principles. *Sience* **1988**, *241* (4868), 976–978.

(25) Lombardi, A.; Pirro, F.; Maglio, O.; Chino, M.; DeGrado, W. F. De Novo Design of Four-Helix Bundle Metalloproteins: One Scaffold, Diverse Reactivities. *Acc. Chem. Res.* **2019**, *52* (5), 1148–1159. https://doi.org/10.1021/acs.accounts.8b00674.

(26) Song, W. J.; Tezcan, F. A. A Designed Supramolecular Protein Assembly with in Vivo Enzymatic Activity. *Science (80-.).* **2014**, *346* (6216), 1525–1528. https://doi.org/10.1126/science.1259680.

(27) Zastrow, M. L.; Peacock, A. F. A.; Stuckey, J. A.; Pecoraro, V. L. Hydrolytic Catalysis and Structural Stabilization in a Designed Metalloprotein. *Nat. Chem.* **2012**, *4* (2), 118–123. https://doi.org/10.1038/nchem.1201.

(28) Huang, P. S.; Boyken, S. E.; Baker, D. The Coming of Age of de Novo Protein Design. *Nature* **2016**, *537* (7620), 320–327. https://doi.org/10.1038/nature19946.

(29) Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Žídek, A.; Potapenko, A.; Bridgland, A.; Meyer, C.; Kohl, S. A. A.; Ballard, A. J.; Cowie, A.; Romera-Paredes, B.; Nikolov, S.; Jain, R.; Adler, J.; Back, T.; Petersen, S.; Reiman, D.; Clancy, E.; Zielinski, M.; Steinegger, M.; Pacholska, M.; Berghammer, T.; Bodenstein, S.; Silver, D.; Vinyals, O.; Senior, A. W.; Kavukcuoglu, K.; Kohli, P.; Hassabis, D. Highly Accurate Protein Structure Prediction with AlphaFold. *Nature* **2021**, *596* (7873), 583–589. https://doi.org/10.1038/s41586-021-03819-2.

(30) Deuss, P. J.; Denheeten, R.; Laan, W.; Kamer, P. C. J. Bioinspired Catalyst Design and Artificial Metalloenzymes. *Chem. - A Eur. J.* **2011**, *17* (17), 4680–4698. https://doi.org/10.1002/chem.201003646.

(31) Xie, J.; Schultz, P. G. A Chemical Toolkit for Proteins--an Expanded Genetic Code. *Nat. Rev. Mol. cell Biol.* **2006**, *7* (10), 775–782. https://doi.org/10.1038/nrm2005.

(32) Nanda, V.; Rosenblatt, M. M.; Osyczka, A.; Kono, H.; Getahun, Z.; Dutton, P. L.; Saven, J. G.; DeGrado, W. F. De Novo Design of a Redox-Active Minimal Rubredoxin Mimic. *J. Am. Chem. Soc.* **2005**, *127* (16), 5804–5805. https://doi.org/10.1021/ja050553f.

(33)   Key, H. M.; Dydio, P.; Liu, Z.; Rha, J. Y. E.; Nazarenko, A.; Seyedkazemi, V.; Clark, D. S.; Hartwig, J. F. Beyond Iron: Iridium-Containing P450 Enzymes for Selective Cyclopropanations of Structurally Diverse Alkenes. *ACS Cent. Sci.* **2017**, *3* (4), 302–308. https://doi.org/10.1021/acscentsci.6b00391.

(34)   Wu, Z.-P.; Hilvert, D. Conversion of a Protease into an Acyl Transferase: Selenolsubtilisin. *J. Am. Chem. Soc* **1989**, *111* (7), 4513–4514.

(35)   Bos, J.; Fusetti, F.; Driessen, A. J. M.; Roelfes, G. Enantioselective Artificial Metalloenzymes by Creation of a Novel Active Site at the Protein Dimer Interface. *Angew. Chemie - Int. Ed.* **2012**, *51* (30), 7472–7475. https://doi.org/10.1002/anie.201202070.

(36)   Christoffel, F.; Igareta, N. V.; Pellizzoni, M. M.; Tiessler-Sala, L.; Lozhkin, B.; Spiess, D. C.; Lledós, A.; Maréchal, J.-D. D.; Peterson, R. L.; Ward, T. R. Design and Evolution of Chimeric Streptavidin for Protein-Enabled Dual Gold Catalysis. *Nat. Catal.* **2021**, *4* (August), 643–653. https://doi.org/10.1038/s41929-021-00651-9.

(37)   Jeschek, M.; Reuter, R.; Heinisch, T.; Trindler, C.; Klehr, J.; Panke, S.; Ward, T. R. Directed Evolution of Artificial Metalloenzymes for in Vivo Metathesis. *Nature* **2016**, *537* (7622), 661–665. https://doi.org/10.1038/nature19114.

(38)   Serrano-Plana, J.; Rumo, C.; Rebelein, J. G.; Peterson, R. L.; Barnet, M.; Ward, T. R. Enantioselective Hydroxylation of Benzylic C(Sp3)-H Bonds by an Artificial Iron Hydroxylase Based on the Biotin-Streptavidin Technology. *J. Am. Chem. Soc.* **2020**, *142* (24), 10617–10623. https://doi.org/10.1021/jacs.0c02788.

(39)   Pordea, A.; Creus, M.; Letondor, C.; Ivanova, A.; Ward, T. R. Improving the Enantioselectivity of Artificial Transfer Hydrogenases Based on the Biotin-Streptavidin Technology by Combinations of Point Mutations. *Inorganica Chim. Acta* **2010**, *363* (3), 601–604. https://doi.org/10.1016/j.ica.2009.02.001.

(40)   Letondor, C.; Pordea, A.; Humbert, N.; Ivanova, A.; Mazurek, S.; Novic, M.; Ward, T. R. Artificial Transfer Hydrogenases Based on the Biotin-(Strept)Avidin Technology: Fine Tuning the Selectivity by Saturation Mutagenesis of the Host Protein. *J. Am. Chem. Soc.* **2006**, *128* (25), 8320–8328. https://doi.org/10.1021/ja061580o.

(41)   Baiyoumy, A.; Vallapurackal, J.; Schwizer, F.; Heinisch, T.; Kardashliev, T.; Held, M.; Panke, S.; Ward, T. R. Directed Evolution of a Surface-Displayed Artificial Allylic

Deallylase Relying on a GFP Reporter Protein. *ACS Catal.* **2021**, *11* (17), 10705–10712. https://doi.org/10.1021/acscatal.1c02405.

(42) Chatterjee, A.; Mallin, H.; Klehr, J.; Vallapurackal, J.; Finke, A. D.; Vera, L.; Marsh, M.; Ward, T. R. An Enantioselective Artificial Suzukiase Based on the Biotin-Streptavidin Technology. *Chem. Sci.* **2016**, *7* (1), 673–677. https://doi.org/10.1039/c5sc03116h.

(43) Köhler, V.; Mao, J.; Heinisch, T.; Pordea, A.; Sardo, A.; Wilson, Y. M.; Knörr, L.; Creus, M.; Prost, J. C.; Schirmer, T.; Ward, T. R. OsO4•streptavidin: A Tunable Hybrid Catalyst for the Enantioselective Cis-Dihydroxylation of Olefins. *Angew. Chemie - Int. Ed.* **2011**, *50* (46), 10863–10866. https://doi.org/10.1002/anie.201103632.

(44) Boaz, M. A. Avidin. *Biochem. J* **1924**, *18* (422).

(45) György, P.; Rose, C. S.; Eakin, R. E.; Snell, E. E.; Williams, R. J. Egg-White Injury As The Result Of Nonabsorbtion Or Inactivation Of Biotin. *Science (80-.).* **1941**, *93*, 477–478. https://doi.org/10.1016/S0140-6736(58)90176-4.

(46) Miller, A. K.; Tausig, F. Biotin-Binding by Parenterally-Administered Streptavidin or Avidin. *Biochem. Biophys. Res. Commun.* **1964**, *14* (3), 210–214. https://doi.org/10.1016/0006-291X(64)90437-1.

(47) Laitinen, O. H.; Hytönen, V. P.; Nordlund, H. R.; Kulomaa, M. S. Genetically Engineered Avidins and Streptavidins. *Cell. Mol. Life Sci.* **2006**, *63* (24), 2992–3017. https://doi.org/10.1007/s00018-006-6288-z.

(48) Wilchek, M.; Bayer, E. A.; Livnah, O. Essentials of Biorecognition: The (Strept)Avidin-Biotin System as a Model for Protein-Protein and Protein-Ligand Interaction. *Immunol. Lett.* **2006**, *103* (1), 27–32. https://doi.org/10.1016/j.imlet.2005.10.022.

(49) Michael Green, N. Avidin and Streptavidin. *Methods Enzymol.* **1990**, *184* (C), 51–67. https://doi.org/10.1016/0076-6879(90)84259-J.

(50) Livnah, O.; Bayer, E. A.; Wilchek, M.; Sussman, J. L. Three-Dimensional Structures of Avidin and the Avidin-Biotin Complex. *Proc. Natl. Acad. Sci. U. S. A.* **1993**, *90* (11), 5076–5080. https://doi.org/10.1073/pnas.90.11.5076.

(51) Jones, M. L.; Kurzban, G. P. Noncooperativity of Biotin Binding to Tetrameric Streptavidin. *Biochemistry* **1995**, *34* (37), 11750–11756. https://doi.org/10.1021/bi00037a012.

(52) Le Trong, I.; Wang, Z.; Hyre, D. E.; Lybrand, T. P.; Stayton, P. S.; Stenkamp, R. E.

Streptavidin and Its Biotin Complex at Atomic Resolution. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2011**, *67* (9), 813–821. https://doi.org/10.1107/S0907444911027806.

(53)   Green, M. T. *Avidin*; 1975; Vol. 29. https://doi.org/10.1007/978-3-540-37885-3_11.

(54)   González, M.; Argaraña, C. E.; Fidelio, G. D. Extremely High Thermal Stability of Streptavidin and Avidin upon Biotin Binding. *Biomol. Eng.* **1999**, *16* (1–4), 67–72. https://doi.org/10.1016/S1050-3862(99)00041-8.

(55)   Gallizia, A.; De Lalla, C.; Nardone, E.; Santambrogio, P.; Brandazza, A.; Sidoli, A.; Arosio, P. Production of a Soluble and Functional Recombinant Streptavidin in Escherichia Coil. *Protein Expr. Purif.* **1998**, *14* (2), 192–196. https://doi.org/10.1006/prep.1998.0930.

(56)   Jeschek, M.; Panke, S.; Marlière, P.; Schneider, V.; Bahls, M. O.; Ward, T. R.; Schneider, V.; Marlière, P.; Ward, T. R.; Panke, S. Biotin-Independent Strains of Escherichia Coli for Enhanced Streptavidin Production. *Metab. Eng.* **2017**, *40* (December 2016), 33–40. https://doi.org/10.1016/j.ymben.2016.12.013.

(57)   Määttä, J. A. E.; Eisenberg-Domovich, Y.; Nordlund, H. R.; Hayouka, R.; Kulomaa, M. S.; Livnah, O.; Hytönen, V. P. Chimeric Avidin Shows Stability against Harsh Chemical Conditions-Biochemical Analysis and 3D Structure. *Biotechnol. Bioeng.* **2011**, *108* (3), 481–490. https://doi.org/10.1002/bit.22962.

(58)   Sano, T.; Pandori, M. W.; Chen, X.; Smith, C. L.; Cantor, C. R. Recombinant Core Streptavidins: A Minimum-Sized Core Streptavidin Has Enhanced Structural Stability and Higher Accessibility to Biotinylated Macromolecules. *J. Biol. Chem.* **1995**, *270* (47), 28204–28209. https://doi.org/10.1074/jbc.270.47.28204.

(59)   Bayer, E. A.; Ben-Hur, H.; Gitlin, G.; Wilchek, M. An Improved Method for the Single-Step Purification of Streptavidin. *J. Biochem. Biophys. Methods* **1986**, *13* (2), 103–112. https://doi.org/10.1016/0165-022X(86)90022-9.

(60)   Stenkamp, R. E.; Trong, I. Le; Klumb, L.; Stayton, P. S.; Freitag, S. Structural Studies of the Streptavidin Binding Loop. *Protein Sci.* **1997**, *6* (6), 1157–1166. https://doi.org/10.1002/pro.5560060604.

(61)   Schmidt, T. G. M. M.; Eichinger, A.; Schneider, M.; Bonet, L.; Carl, U.; Karthaus, D.; Theobald, I.; Skerra, A. The Role of Changing Loop Conformations in Streptavidin Versions Engineered for High-Affinity Binding of the Strep-Tag II Peptide. *J. Mol. Biol.* **2021**, *433* (9), 166893. https://doi.org/10.1016/j.jmb.2021.166893.

(62) Le Trong, I.; Chu, V.; Xing, Y.; Lybrand, T. P.; Stayton, P. S.; Stenkamp, R. E. Structural Consequences of Cutting a Binding Loop: Two Circularly Permuted Variants of Streptavidin. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2013**, *69* (6), 968–977. https://doi.org/10.1107/S0907444913003855.

(63) McConnell, D. B. Biotin's Lessons in Drug Design. *J. Med. Chem.* **2021**, *64* (22), 16319–16327. https://doi.org/10.1021/acs.jmedchem.1c00975.

(64) Weber, P. C.; Ohlendorf, D. H.; Wendoloski, J. J.; Salemme, F. R. Structural Origins of High-Affinity Biotin Binding to Streptavidin. *Science (80-.).* **1989**, *243* (4887), 85–88. https://doi.org/10.1126/science.2911722.

(65) Le Trong, I.; Humbert, N.; Ward, T. R.; Stenkamp, R. E. Crystallographic Analysis of a Full-Length Streptavidin with Its C-Terminal Polypeptide Bound in the Biotin Binding Site. *J. Mol. Biol.* **2006**, *356* (3), 738–745. https://doi.org/10.1016/j.jmb.2005.11.086.

(66) Creus, M.; Pordea, A.; Rossel, T.; Sardo, A.; Letondor, C.; Ivanova, A.; LeTrong, I.; Stenkamp, R. E.; Ward, T. R.; LeTrong, I.; Stenkamp, R. E.; Creus, M.; Pordea, A.; Rossel, T.; Sardo, A.; Letondor, C.; Ivanova, A.; Ward, T. R. X-Ray Structure and Designed Evolution of an Artificial Transfer Hydrogenase. *Angew. Chemie - Int. Ed.* **2008**, *47* (8), 1400–1404. https://doi.org/10.1002/anie.200704865.

(67) Dürrenberger, M.; Heinisch, T.; Wilson, Y. M.; Rossel, T.; Nogueira, E.; Knörr, L.; Mutschler, A.; Kersten, K.; Zimbron, M. J.; Pierron, J.; Schirmer, T.; Ward, T. R. Artificial Transfer Hydrogenases for the Enantioselective Reduction of Cyclic Imines. *Angew. Chemie - Int. Ed.* **2011**, *50* (13), 3026–3029. https://doi.org/10.1002/anie.201007820.

(68) Hestericová, M.; Heinisch, T.; Alonso-Cotchico, L.; Maréchal, J. D.; Vidossich, P.; Ward, T. R. Directed Evolution of an Artificial Imine Reductase. *Angew. Chemie - Int. Ed.* **2018**, *57* (7), 1863–1868. https://doi.org/doi.org/10.1002/anie.201711016.

(69) Wu, S.; Zhou, Y.; Rebelein, J. G.; Kuhn, M.; Mallin, H.; Zhao, J.; Igareta, N. V.; Ward, T. R. Breaking Symmetry: Engineering Single-Chain Dimeric Streptavidin as Host for Artificial Metalloenzymes. *J. Am. Chem. Soc.* **2019**, jacs.9b06923. https://doi.org/10.1021/jacs.9b06923.

(70) Hyster, T. K.; Knörr, L.; Ward, T. R.; Rovis, T. Biotinylated Rh(III) Complexes in Engineered Streptavidin for Accelerated Asymmetric C–H Activation. *Science (80-.).* **2012**, *338* (October), 500–503. https://doi.org/10.1126/science. 1226132.

(71)   Zimbron, J. M.; Heinisch, T.; Schmid, M.; Hamels, D.; Nogueira, E. S.; Schirmer, T.; Ward, T. R. A Dual Anchoring Strategy for the Localization and Activation of Artificial Metalloenzymes Based on the Biotin-Streptavidin Technology. *J. Am. Chem. Soc.* **2013**, *135* (14), 5384–5388. https://doi.org/10.1021/ja309974s.

(72)   Pordea, A.; Creus, M.; Panek, J.; Duboc, C.; Mathis, D.; Novic, M.; Ward, T. R.; Creusa, M.; Panek, J.; Duboc, C.; Mathis, D.; Novic, M.; Ward, T. R. Artificial Metalloenzyme for Enantioselective Sulfoxidation Based on Vanadyl-Loaded Streptavidin. *J. Am. Chem. Soc.* **2008**, *130* (25), 8085–8088. https://doi.org/10.1021/ja8017219.

(73)   Pellizzoni, M. M.; Schwizer, F.; Wood, C. W.; Sabatino, V.; Cotelle, Y.; Matile, S.; Woolfson, D. N.; Ward, T. R. Chimeric Streptavidins as Host Proteins for Artificial Metalloenzymes. *ACS Catal.* **2018**, *8* (2), 1476–1484. https://doi.org/10.1021/acscatal.7b03773.

(74)   Chen, K.; Arnold, F. H. Engineering New Catalytic Activities in Enzymes. *Nature Catalysis*. Nature Research March 1, 2020, pp 203–213. https://doi.org/10.1038/s41929-019-0385-5.

(75)   Mallin, H.; Hestericová, M.; Reuter, R.; Ward, T. R. Library Design and Screening Protocol for Artificial Metalloenzymes Based on the Biotin-Streptavidin Technology. *Nat. Protoc.* **2016**, *11* (5), 835–852. https://doi.org/10.1038/nprot.2016.019.

(76)   Robles, V. M.; Dürrenberger, M.; Heinisch, T.; Lledós, A.; Schirmer, T.; Ward, T. R.; Maréchal, J. D. Structural, Kinetic, and Docking Studies of Artificial Imine Reductases Based on Biotin-Streptavidin Technology: An Induced Lock-and-Key Hypothesis. *J. Am. Chem. Soc.* **2014**, *136* (44), 15676–15683. https://doi.org/10.1021/ja508258t.

(77)   Chilkoti, A.; Schwartz, B. L.; Smith, R. D.; Long, C. J.; Stayton, P. S. Engineered Chimeric Streptavidin Tetramers as Novel Tools for Bioseparations and Drug Delivery. *Bio/Technology* **1995**, *13* (11), 1198–1204. https://doi.org/10.1038/nbt1195-1198.

(78)   Reznik, G.; Vajda, S.; Smith, C. L.; Cantor, C. R.; Sano, T. Streptavidins with Intersubunit Crosslinks Have Enhanced Stability. *Nat. Biotechnol.* **1996**, *14*, 1007–1011. https://doi.org/10.1038/nbt0896-1007.

(79)   Nordlund, H. R.; Laitinen, O. H.; Hytönen, V. P.; Uotila, S. T. H.; Porkka, E.; Kulomaa, M. S. Construction of a Dual Chain Pseudotetrameric Chicken Avidin by Combining Two Circularly Permuted Avidins. *J. Biol. Chem.* **2004**, *279* (35), 36715–36719.

https://doi.org/10.1074/jbc.M403496200.

(80)   Nordlund, H. R.; Hyt, V. P.; E M, J. A.; White, D. J.; Halling, K.; Porkka, E. J.; Peter SLOTTE, J.; Laitinen, O. H.; Kulomaa, M. S. Tetravalent Single-Chain Avidin: From Subunits to Protein Domains via Circularly Permuted Avidins. *Biochem. J* **2005**, *392*, 485–491. https://doi.org/10.1042/BJ20051038.

(81)   Hytönen, V. P.; Hörhä, J.; Airenne, T. T.; Niskanen, E. A.; Helttunen, K. J.; Johnson, M. S.; Salminen, T. A.; Kulomaa, M. S.; Nordlund, H. R. Controlling Quaternary Structure Assembly: Subunit Interface Engineering and Crystal Structure of Dual Chain Avidin. *J. Mol. Biol.* **2006**, *359* (5), 1352–1363. https://doi.org/10.1016/j.jmb.2006.04.044.

(82)   Reznik, G. O.; Vajda, S.; Sano, T.; Cantor, C. R. A Streptavidin Mutant with Altered Ligand-Binding Specificity. *Proc. Natl. Acad. Sci. U. S. A.* **1998**, *95* (23), 13525–13530. https://doi.org/10.1073/pnas.95.23.13525.

(83)   Spagnolo, L.; Töro, I.; D'Orazio, M.; O'Neil, P.; Pedersen, J. Z.; Carugo, O.; Rotilio, G.; Battistoni, A.; Djinović-Carugo, K. Unique Features of the SodC-Encoded Superoxide Dismutase from Mycobacterium Tuberculosis, a Fully Functional Copper-Containing Enzyme Lacking Zinc in the Active Site. *J. Biol. Chem.* **2004**, *279* (32), 33447–33455. https://doi.org/10.1074/jbc.M404699200.

(84)   Lee, C.-D.; Sun, H.-C.; Hu, S.-M.; Chiu, C.-F.; Homhuan, A.; Liang, S.-M.; Leng, C.-H.; Wang, T.-F. An Improved SUMO Fusion Protein System for Effective Production of Native Proteins. *Protein Sci.* **2008**, *17* (7), 1241–1248. https://doi.org/10.1110/ps.035188.108.

(85)   Mikolajczyk, J.; Drag, M.; Békés, M.; Cao, J. T.; Ronai, Z.; Salvesen, G. S. Small Ubiquitin-Related Modifier (SUMO)-Specific Proteases: Profiling the Specificities and Activities of Human SENPs. *J. Biol. Chem.* **2007**, *282* (36), 26217–26224. https://doi.org/10.1074/jbc.M702444200.

(86)   Mossessova, E.; Lima, C. D. Ulp1-SUMO Crystal Structure and Genetic Analysis Reveal Conserved Interactions and a Regulatory Element Essential for Cell Growth in Yeast. *Mol. Cell* **2000**, *5* (5), 865–876. https://doi.org/10.1016/S1097-2765(00)80326-3.

(87)   Studier, F. W. Protein Production by Auto-Induction in High Density Shaking Cultures. *Protein Expr. Purif.* **2005**, *41* (1), 207–234. https://doi.org/10.1016/j.pep.2005.01.016.

(88)   Waner, M. J.; Mascotti, D. P. A Simple Spectrophotometric Streptavidin-Biotin Binding

Assay Utilizing Biotin-4-Fluorescein. *J. Biochem. Biophys. Methods* **2008**, *70* (6), 873–877. https://doi.org/10.1016/j.jbbm.2007.06.001.

(89)   Bessette, P. H.; Åslund, F.; Beckwith, J.; Georgiou, G. Efficient Folding of Proteins with Multiple Disulfide Bonds in the Escherichia Coli Cytoplasm. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96* (24), 13703–13708. https://doi.org/10.1073/pnas.96.24.13703.

(90)   Rebelein, J. G.; Ward, T. R. In Vivo Catalyzed New-to-Nature Reactions. *Curr. Opin. Biotechnol.* **2018**, *53* (ii), 106–114. https://doi.org/10.1016/j.copbio.2017.12.008.

(91)   Vidal, C.; Tomás-Gamasa, M.; Destito, P.; López, F.; Mascareñas, J. L. Concurrent and Orthogonal Gold(I) and Ruthenium(II) Catalysis inside Living Cells. *Nat. Commun.* **2018**, *9* (1), 1–9. https://doi.org/10.1038/s41467-018-04314-5.

(92)   Wittwer, M.; Markel, U.; Schiffels, J.; Okuda, J.; Sauer, D. F.; Schwaneberg, U. Engineering and Emerging Applications of Artificial Metalloenzymes with Whole Cells. **2021**, *4* (October). https://doi.org/10.1038/s41929-021-00673-3.

(93)   Himiyama, T.; Okamoto, Y. Artificial Metalloenzymes: From Selective Chemical Transformations to Biochemical Applications. *Molecules* **2020**, *25* (13). https://doi.org/10.3390/molecules25132989.

(94)   Gu, Y.; Bloomer, B. J.; Liu, Z.; Chen, R.; Clark, D. S.; Hartwig, J. F. Directed Evolution of Artificial Metalloenzymes in Whole Cells. *Angew. Chemie Int. Ed.* **2022**, *61* (5). https://doi.org/10.1002/anie.202110519.

(95)   Meister, A.; Anderson, M. E. Glutathione. *Annu. Rev. Biochem* **1983**, *52*, 711–760.

(96)   Tsubokura, K.; Vong, K. K. H. H.; Pradipta, A. R.; Ogura, A.; Urano, S.; Tahara, T.; Nozaki, S.; Onoe, H.; Nakao, Y.; Sibgatullina, R.; Kurbangalieva, A.; Watanabe, Y.; Tanaka, K. In Vivo Gold Complex Catalysis within Live Mice. *Angew. Chemie - Int. Ed.* **2017**, *56* (13), 3579–3584. https://doi.org/10.1002/anie.201610273.

(97)   Okamoto, Y.; Kojima, R.; Schwizer, F.; Bartolami, E.; Heinisch, T.; Matile, S.; Fussenegger, M.; Ward, T. R. A Cell-Penetrating Artificial Metalloenzyme Regulates a Gene Switch in a Designer Mammalian Cell. *Nat. Commun.* **2018**, *9* (1), 1–7. https://doi.org/10.1038/s41467-018-04440-0.

(98)   Chordia, S.; Narasimhan, S.; Lucini Paioni, A.; Baldus, M.; Roelfes, G. In Vivo Assembly of Artificial Metalloenzymes and Application in Whole-Cell Biocatalysis**. *Angew. Chemie - Int. Ed.* **2021**, *60* (11), 5913–5920. https://doi.org/10.1002/anie.202014771.

(99) Wilson, Y. M.; Dürrenberger, M.; Nogueira, E. S.; Ward, T. R. Neutralizing the Detrimental Effect of Glutathione on Precious Metal Catalysts. *J. Am. Chem. Soc.* **2014**, *136* (25), 8928–8932. https://doi.org/10.1021/ja500613n.

(100) Jeschek, M.; Panke, S.; Ward, T. R. *Periplasmic Screening for Artificial Metalloenzymes*, 1st ed.; Elsevier Inc., 2016; Vol. 580. https://doi.org/10.1016/bs.mie.2016.05.037.

(101) Zhao, J.; Rebelein, J. G.; Mallin, H.; Trindler, C.; Pellizzoni, M. M.; Ward, T. R. Genetic Engineering of an Artificial Metalloenzyme for Transfer Hydrogenation of a Self-Immolative Substrate in Escherichia Coli's Periplasm. *J. Am. Chem. Soc.* **2018**, *140* (41), 13171–13175. https://doi.org/10.1021/jacs.8b07189.

(102) Rebelein, J. G.; Cotelle, Y.; Garabedian, B.; Ward, T. R. Chemical Optimization of Whole-Cell Transfer Hydrogenation Using Carbonic Anhydrase as Host Protein. *ACS Catal.* **2019**, *9* (5), 4173–4178. https://doi.org/10.1021/acscatal.9b01006.

(103) Heinisch, T.; Schwizer, F.; Garabedian, B.; Csibra, E.; Jeschek, M.; Vallapurackal, J.; Pinheiro, V. B.; Marlière, P.; Panke, S.; Ward, T. R. E. Coli Surface Display of Streptavidin for Directed Evolution of an Allylic Deallylase. *Chem. Sci.* **2018**, *9* (24), 5383–5388. https://doi.org/10.1039/c8sc00484f.

(104) Klehr, J.; Zhao, J.; Kron, A. S.; Ward, T. R.; Köhler, V. Streptavidin (Sav)-Based Artificial Metalloenzymes: Cofactor Design Considerations and Large-Scale Expression of Host Protein Variants. In *Peptide and Protein Engineering*; 2020; pp 213–235. https://doi.org/10.1007/978-1-0716-0720-6_12.

(105) Mallin, H.; Hestericová, M.; Reuter, R.; Ward, T. R. Library Design and Screening Protocol for Artificial Metalloenzymes Based on the Biotin-Streptavidin Technology. *Nat. Protoc.* **2016**, *11* (5), 835–852. https://doi.org/10.1038/nprot.2016.019.

(106) Evans, P. R.; Murshudov, G. N. How Good Are My Data and What Is the Resolution? *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2013**, *69* (7), 1204–1214. https://doi.org/10.1107/S0907444913000061.

(107) Evans, P. Scaling and Assessment of Data Quality. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2006**, *62* (1), 72–82. https://doi.org/10.1107/S0907444905036693.

(108) Kabsch, W. XDS. *Acta Crystallogr. - Sect. D Biol. Crystallogr.* **2010**, No. D66, 125–132. https://doi.org/10.1107/S0907444909047337.

(109) Evans, P. R. An Introduction to Data Reduction: Space-Group Determination, Scaling

and Intensity Statistics. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2011**, *67* (4), 282–292. https://doi.org/10.1107/S090744491003982X.

(110) Potterton, L.; Agirre, J.; Ballard, C.; Cowtan, K.; Dodson, E.; Evans, P. R.; Jenkins, H. T.; Keegan, R.; Krissinel, E.; Stevenson, K.; Lebedev, A.; McNicholas, S. J.; Nicholls, R. A.; Noble, M.; Pannu, N. S.; Roth, C.; Sheldrick, G.; Skubak, P.; Turkenburg, J.; Uski, V.; Von Delft, F.; Waterman, D.; Wilson, K.; Winn, M.; Wojdyr, M. CCP 4 i 2: The New Graphical User Interface to the CCP 4 Program Suite. *Acta Crystallogr. Sect. D Struct. Biol.* **2018**, *74* (August 2017), 68–84. https://doi.org/10.1107/S2059798317016035.

(111) McCoy, A. J.; Grosse-Kunstleve, R. W.; Adams, P. D.; Winn, M. D.; Storoni, L. C.; Read, R. J. Phaser Crystallographic Software. *J. Appl. Crystallogr.* **2007**, *40* (4), 658–674. https://doi.org/10.1107/S0021889807021206.

(112) Kovalevskiy, O.; Nicholls, R. A.; Long, F.; Carlon, A.; Murshudov, G. N. Overview of Refinement Procedures within REFMAC 5: Utilizing Data from Different Sources. *Acta Crystallogr. Sect. D Struct. Biol.* **2018**, *74*, 215–227. https://doi.org/10.1107/S2059798318000979.

(113) Murshudov, G. N.; Skubák, P.; Lebedev, A. A.; Pannu, N. S.; Steiner, R. A.; Nicholls, R. A.; Winn, M. D.; Long, F.; Vagin, A. A. REFMAC5 for the Refinement of Macromolecular Crystal Structures. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2011**, *67* (4), 355–367. https://doi.org/10.1107/S0907444911001314.

(114) Emsley, P.; Lohkamp, B.; Scott, W. G.; Cowtan, K. Features and Development of Coot. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2010**, *66* (4), 486–501. https://doi.org/10.1107/S0907444910007493.

(115) Eckart, K.; Spiess, J.; Sp, J.; Spiess, J. Electrospray Ionization Mass Spectrometry of Biotin Binding to Streptavidin. *J. Am. Soc. Mass Spectrom.* **1995**, *6* (10), 912–919. https://doi.org/10.1016/1044-0305(95)00480-2.

(116) Robinson, P. J.; Holbrook, K. A. Why Gold Is the Noblest of All the Metals. *Nature* **1995**, *376* (July), 238–240.

(117) Bond, G. C. The Catalytic Properties of Gold - Potential Applications in the Chemical Industry. *Gold Bull.* **1972**, *5* (1), 11–13. https://doi.org/10.1007/BF03215149.

(118) Bond, G. C.; Sermon, P. A.; Webb, G.; Buchanan, D. A.; Wells, P. B. Hydrogenation over Supported Gold Catalysts. *J. Chem. Soc. Chem. Commun.* **1973**, No. 13, 444–445.

https://doi.org/10.1039/C3973000444b.

(119) Haruta, M.; Kobayashi, T.; Sano, H.; Yamada, N. Novel Gold Catalysts for the Oxidation of Carbon Monoxide at a Temperature below 0°C. *Chem. Lett.* **1987**, *16* (16), 405–408.

(120) Hutchings, G. J. Vapor Phase Hydrochlorination of Acetylene: Correlation of Catalytic Activity of Supported Metal Chloride Catalysts. *J. Catal.* **1985**, *96* (1), 292–295. https://doi.org/10.1016/0021-9517(85)90383-5.

(121) Ito, Y.; Sawamura, M.; Hayashi, T. Catalytic Asymmetric Aldol Reaction: Reaction of Aldehydes with Isocyanoacetate Catalyzed by a Chiral Ferrocenylphosphine-Gold(I) Complex. *J. Am. Chem. Soc.* **1986**, *108* (20), 6405–6406. https://doi.org/10.1021/ja00280a056.

(122) Hashmi, A. S. K.; Hutchings, G. J. Gold Catalysis. *Angew. Chemie - Int. Ed.* **2006**, *45* (47), 7896–7936. https://doi.org/10.1002/anie.200602454.

(123) Gorin, D. J.; Toste, F. D. Relativistic Effects in Homogeneous Gold Catalysis. *Nature* **2007**, *446* (7134), 395–403. https://doi.org/10.1038/nature05592.

(124) Hashmi, A. S. K. Gold-Catalyzed Organic Reactions. *Chem. Rev.* **2007**, *107*, 3180–3211. https://doi.org/10.1021/cr000436x.

(125) Sadler, P. J.; Sue, R. E. The Chemistry of Gold Drugs. *Chem. Gold Drugs* **1994**, *1*, 108–144.

(126) Murphy, C. J.; Gole, A. M.; Stone, J. W.; Sisco, P. N.; Alkilany, A. M.; Goldsmith, E. C.; Baxter, S. C. Gold Nanoparticles in Biology: Beyond Toxicity to Cellular Imaging. *Acc. Chem. Res.* **2008**, *41* (12), 1721–1730. https://doi.org/10.1021/ar800035u.

(127) Vong, K.; Eda, S.; Kadota, Y.; Nasibullin, I.; Wakatake, T.; Yokoshima, S.; Shirasu, K.; Tanaka, K. An Artificial Metalloenzyme Biosensor Can Detect Ethylene Gas in Fruits and Arabidopsis Leaves. *Nat. Commun.* **2019**, *10* (1), 2–12. https://doi.org/10.1038/s41467-019-13758-2.

(128) Vong, K.; Yamamoto, T.; Tanaka, K. Artificial Glycoproteins as a Scaffold for Targeted Drug Therapy. *Small* **2020**, *16* (27), 1–16. https://doi.org/10.1002/smll.201906890.

(129) Eda, S.; Nasibullin, I.; Vong, K.; Kudo, N.; Yoshida, M.; Kurbangalieva, A.; Tanaka, K. Biocompatibility and Therapeutic Potential of Glycosylated Albumin Artificial Metalloenzymes. *Nat. Catal.* **2019**, *2* (9), 780–792. https://doi.org/10.1038/s41929-019-0317-4.

(130) McLean, E. B.; Lee, A. L. Golden Potential. *Nat. Chem.* **2019**, *11* (9), 760–761. https://doi.org/10.1038/s41557-019-0315-9.

(131) Dorel, R.; Echavarren, A. M. Gold(I)-Catalyzed Activation of Alkynes for the Construction of Molecular Complexity. *Chem. Rev.* **2015**, *115* (17), 9028–9072. https://doi.org/10.1021/cr500691k.

(132) Pflästerer, D.; Hashmi, A. S. K. Gold Catalysis in Total Synthesis - Recent Achievements. *Chemical Society Reviews*. Royal Society of Chemistry March 7, 2016, pp 1331–1367. https://doi.org/10.1039/c5cs00721f.

(133) Stephen, A.; Hashmi, K.; Braun, I.; Rudolph, M.; Rominger, F. The Role of Gold Acetylides as a Selectivity Trigger and the Importance of Gem-Diaurated Species in the Gold-Catalyzed Hydroarylating-Aromatization of Arene-Diynes. *Organometallics* **2012**, *31* (2), 644–661. https://doi.org/10.1021/om200946m.

(134) Gimeno, A.; Medio-Simón, M.; De Arellano, C. R.; Asensio, G.; Cuenca, A. B. NHC-Stabilized Gold(I) Complexes: Suitable Catalysts for 6-Exo-Dig Heterocyclization of 1-(o-Ethynylaryl)Ureas. *Org. Lett.* **2010**, *12* (9), 1900–1903. https://doi.org/10.1021/ol100595s.

(135) Ye, L.; Wang, Y.; Aue, D. H.; Zhang, L. Experimental and Computational Evidence for Gold Vinylidenes: Generation from Terminal Alkynes via a Bifurcation Pathway and Facile C-H Insertions. *J. Am. Chem. Soc.* **2012**, *134* (1), 31–34. https://doi.org/10.1021/ja2091992.

(136) Zhao, X.; Rudolph, M.; Hashmi, A. S. K. Dual Gold Catalysis-an Update. *Chem. Commun.* **2019**, *55* (81), 12127–12135. https://doi.org/10.1039/c9cc06078b.

(137) Wang, W.; Ji, C.-L.; Liu, K.; Zhao, C.-G.; Li, W.; Xie, J. Dinuclear Gold Catalysis. *Chem. Soc. Rev.* **2021**, *50*, 1874–1912. https://doi.org/10.1039/d0cs00254b.

(138) Cheong, P. H. Y.; Morganelli, P.; Luzung, M. R.; Houk, K. N.; Toste, F. D. Gold-Catalyzed Cycloisomerization of 1,5-Allenynes via Dual Activation of an Ene Reaction. *J. Am. Chem. Soc.* **2008**, *130* (13), 4517–4526. https://doi.org/10.1021/ja711058f.

(139) Bertini, I.; Gray, H. B.; Valentine, J. S.; Stiefel, E. I. *Biological Inorganic Chemistry : Structure and Reactivity*, 2nd ed.; CA: University Science Books, 2007.

(140) Jewett, J. C.; Bertozzi, C. R. Cu-Free Click Cycloaddition Reactions in Chemical Biology. *Chem. Soc. Rev.* **2010**, *39* (4), 1272–1279. https://doi.org/10.1039/b901970g.

(141) Devaraj, N. K. The Future of Bioorthogonal Chemistry. *ACS Cent. Sci.* **2018**, *4* (8), 952–959. https://doi.org/10.1021/acscentsci.8b00251.

(142) Pickens, C. J.; Johnson, S. N.; Pressnall, M. M.; Leon, M. A.; Berkland, C. J. Practical Considerations, Challenges, and Limitations of Bioconjugation via Azide-Alkyne Cycloaddition. *Bioconjug. Chem.* **2018**, *29* (3), 686–701. https://doi.org/10.1021/acs.bioconjchem.7b00633.

(143) Boren, B. C.; Narayan, S.; Rasmussen, L. K.; Zhang, L.; Zhao, H.; Lin, Z.; Jia, G.; Fokin, V. V. Ruthenium-Catalyzed Azide-Alkyne Cycloaddition: Scope and Mechanism. *J. Am. Chem. Soc.* **2008**, *130* (28), 8923–8930. https://doi.org/10.1021/ja0749993.

(144) Tsubokura, K.; Vong, K. K. H.; Pradipta, A. R.; Ogura, A.; Urano, S.; Tahara, T.; Nozaki, S.; Onoe, H.; Nakao, Y.; Sibgatullina, R.; Kurbangalieva, A.; Watanabe, Y.; Tanaka, K. In Vivo Gold Complex Catalysis within Live Mice. *Angew. Chemie - Int. Ed.* **2017**, *56* (13), 3579–3584. https://doi.org/10.1002/anie.201610273.

(145) Pérez-López, A. M.; Rubio-Ruiz, B.; Sebastián, V.; Hamilton, L.; Adam, C.; Bray, T. L.; Irusta, S.; Brennan, P. M.; Lloyd-Jones, G. C.; Sieger, D.; Santamaría, J.; Unciti-Broceta, A. Gold-Triggered Uncaging Chemistry in Living Systems. *Angew. Chem. Int. Ed.* **2017**, *56* (41), 12548–12552. https://doi.org/10.1002/anie.201705609.

(146) Jung Jou, M.; Chen, X.; Swamy, K. M. K.; Na Kim, H.; Kim, H. J.; Lee, S. G.; Yoon, J. Highly Selective Fluorescent Probe for Au(III) Based on Cyclization of Propargylamide. *Chem. Commun.* **2009**, *46*, 7218–7220. https://doi.org/10.1039/b917832e.

(147) Wang, J. B.; Wu, Q. Q.; Min, Y. Z.; Liu, Y. Z.; Song, Q. H. A Novel Fluorescent Probe for Au(III)/Au(I) Ions Based on an Intramolecular Hydroamination of a Bodipy Derivative and Its Application to Bioimaging. *Chem. Commun.* **2012**, *48* (5), 744–746. https://doi.org/10.1039/c1cc16128h.

(148) Studer, S.; Hansen, D. A.; Pianowski, Z. L.; Mittl, P. R. E. E.; Debon, A.; Guffy, S. L.; Der, B. S.; Kuhlman, B.; Hilvert, D. Evolution of a Highly Active and Enantiospecific Metalloenzyme from Short Peptides. *Science (80-.).* **2018**, *362* (6420), 1285–1288. https://doi.org/10.1126/SCIENCE.AAU3744.

(149) Gu, Y.; Natoli, S. N.; Liu, Z.; Clark, D. S.; Hartwig, J. F. Site-Selective Functionalization of (Sp3)C–H Bonds Catalyzed by Artificial Metalloenzymes Containing an Iridium-Porphyrin Cofactor. *Angew. Chemie - Int. Ed.* **2019**, *58* (39), 13954–13960.

https://doi.org/10.1002/anie.201907460.

(150) Roelfes, G. LmrR : A Privileged Scaffold for Artificial Metalloenzymes. *Acc. Chem. Res.* **2019**, *52*, 545–556. https://doi.org/10.1021/acs.accounts.9b00004.

(151) Zhou, Z.; Roelfes, G. Synergistic Catalysis in an Artificial Enzyme by Simultaneous Action of Two Abiological Catalytic Sites. *Nat. Catal.* **2020**, *3* (3), 289–294. https://doi.org/10.1038/s41929-019-0420-6.

(152) Alonso, S.; Santiago, G.; Cea-Rama, I.; Fernandez-Lopez, L.; Coscolín, C.; Modregger, J.; Ressmann, A. K.; Martínez-Martínez, M.; Marrero, H.; Bargiela, R.; Pita, M.; Gonzalez-Alfonso, J. L.; Briand, M. L.; Rojo, D.; Barbas, C.; Plou, F. J.; Golyshin, P. N.; Shahgaldian, P.; Sanz-Aparicio, J.; Guallar, V.; Ferrer, M. Genetically Engineered Proteins with Two Active Sites for Enhanced Biocatalysis and Synergistic Chemo- and Biocatalysis. *Nat. Catal.* **2020**, *3* (3), 319–328. https://doi.org/10.1038/s41929-019-0394-4.

(153) Martínez-Calvo, M.; Couceiro, J. R.; Destito, P.; Rodríguez, J.; Mosquera, J.; Mascareñas, J. L. Intracellular Deprotection Reactions Mediated by Palladium Complexes Equipped with Designed Phosphine Ligands. *ACS Catal.* **2018**, *8* (7), 6055–6061. https://doi.org/10.1021/acscatal.8b01606.

(154) Monnard, F. W.; Nogueira, E. S.; Heinisch, T.; Schirmer, T.; Ward, T. R. Human Carbonic Anhydrase II as Host Protein for the Creation of Artificial Metalloenzymes: The Asymmetric Transfer Hydrogenation of Imines. *Chem. Sci.* **2013**, *4* (8), 3269–3274. https://doi.org/10.1039/c3sc51065d.

(155) Mirts, E. N.; Petrik, I. D.; Hosseinzadeh, P.; Nilges, M. J.; Lu, Y. A Designed Heme-[4Fe-4S] Metalloenzyme Catalyzes Sulfite Reduction like the Native Enzyme. *Science (80-.).* **2018**, *361* (6407), 1098–1101. https://doi.org/10.1126/science.aat8474.

(156) Lewis, J. C. Beyond the Second Coordination Sphere: Engineering Dirhodium Artificial Metalloenzymes to Enable Protein Control of Transition Metal Catalysis. *Acc. Chem. Res.* **2019**, *52* (3), 576–584. https://doi.org/10.1021/acs.accounts.8b00625.

(157) Chino, M.; Maglio, O.; Nastri, F.; Pavone, V.; Degrado, W. F.; Lombardi, A. Artificial Diiron Enzymes with a de Novo Designed Four-Helix Bundle Structure. *European Journal of Inorganic Chemistry*. 2015, pp 3371–3390. https://doi.org/10.1002/ejic.201500470.

(158) Grimm, A. R.; Sauer, D. F.; Polen, T.; Zhu, L.; Hayashi, T.; Okuda, J.; Schwaneberg, U. A

Whole Cell E. Coli Display Platform for Artificial Metalloenzymes: Poly(Phenylacetylene) Production with a Rhodium-Nitrobindin Metalloprotein. *ACS Catal.* **2018**, *8* (3), 2611–2614. https://doi.org/10.1021/acscatal.7b04369.

(159) Liang, A. D.; Serrano-Plana, J.; Peterson, R. L.; Ward, T. R. Artificial Metalloenzymes Based on the Biotin–Streptavidin Technology: Enzymatic Cascades and Directed Evolution. *Acc. Chem. Res.* **2019**, *52* (i), acs.accounts.8b00618. https://doi.org/10.1021/acs.accounts.8b00618.

(160) Heinisch, T.; Ward, T. R. Artificial Metalloenzymes Based on the Biotin-Streptavidin Technology: Challenges and Opportunities. *Acc. Chem. Res.* **2016**, *49* (9), 1711–1721. https://doi.org/10.1021/acs.accounts.6b00235.

(161) Vornholt, T.; Christoffel, F.; Pellizzoni, M. M.; Panke, S.; Ward, T. R.; Jeschek, M. Systematic Engineering of Artificial Metalloenzymes for New-to-Nature Reactions. *Sci. Adv.* **2021**, *7* (4), 1–12. https://doi.org/10.1126/sciadv.abe4208.

(162) Gimeno, A.; Cuenca, A. B.; Suárez-Pantiga, S.; De Arellano, C. R.; Medio-Simõn, M.; Asensio, G. Competitive Gold-Activation Modes in Terminal Alkynes: An Experimental and Mechanistic Study. *Chem. - A Eur. J.* **2014**, *20* (3), 683–688. https://doi.org/10.1002/chem.201304087.

(163) Vreeken, V.; Broere, D. L. J.; Jans, A. C. H.; Lankelma, M.; Reek, J. N. H.; Siegler, M. A.; van der Vlugt, J. I. Well-Defined Dinuclear Gold Complexes for Preorganization-Induced Selective Dual Gold Catalysis. *Angew. Chem. Int. Ed.* **2016**, *55* (34), 10042–10046. https://doi.org/10.1002/anie.201603938.

(164) Ye, D.; Wang, J.; Zhang, X.; Zhou, Y.; Ding, X.; Feng, E.; Sun, H.; Liu, G.; Jiang, H.; Liu, H. Gold-Catalyzed Intramolecular Hydroamination of Terminal Alkynes in Aqueous Media: Efficient and Regioselective Synthesis of Indole-1-Carboxamides. *Green Chem.* **2009**, *11* (8), 1201–1208. https://doi.org/10.1039/b904044g.

(165) Breker, V.; Sak, H.; Baracchi-Krause, G.; Krause, N. Synthesis and Properties of a Biotin-Tagged NHC-Gold Complex. *Tetrahedron Lett.* **2015**, *56* (23), 3390–3392. https://doi.org/10.1016/j.tetlet.2015.01.023.

(166) Collado, A.; Gómez-Suárez, A.; Martin, A. R.; Slawin, A. M. Z.; Nolan, S. P. Straightforward Synthesis of [Au(NHC)X] (NHC = N-Heterocyclic Carbene, X = Cl, Br, I) Complexes. *Chem. Commun.* **2013**, *49* (49), 5541–5543.

https://doi.org/10.1039/c3cc43076f.

(167) Kajetanowicz, A.; Chatterjee, A.; Reuter, R.; Ward, T. R. Biotinylated Metathesis Catalysts: Synthesis and Performance in Ring Closing Metathesis. *Catal. Letters* **2014**, *144* (3), 373–379. https://doi.org/10.1007/s10562-013-1179-z.

(168) Anhäuser, L.; Teders, M.; Rentmeister, A.; Glorius, F. Bio-Additive-Based Screening: Toward Evaluation of the Biocompatibility of Chemical Reactions. *Nat. Protoc.* **2019**, *14* (9), 2599–2626. https://doi.org/10.1038/s41596-019-0190-2.

(169) Loving, G.; Imperiali, B. Thiol-Reactive Derivatives of the Solvatochromic 4-N,N-Dimethylamino-1,8- Naphthalimide Fluorophore: A Highly Sensitive Toolset for the Detection of Biomolecular Interactions. *Bioconjug. Chem.* **2009**, *20* (11), 2133–2141. https://doi.org/10.1021/bc900319z.

(170) Alonso-Cotchico, L.; Rodr͠íguez-Guerra, J.; Lledós, A.; Maréchal, J. D. Molecular Modeling for Artificial Metalloenzyme Design and Optimization. *Acc. Chem. Res.* **2020**, *53* (4), 896–905. https://doi.org/10.1021/acs.accounts.0c00031.

(171) Marenich, A. V.; Cramer, C. J.; Truhlar, D. G. Universal Solvation Model Based on Solute Electron Density and on a Continuum Model of the Solvent Defined by the Bulk Dielectric Constant and Atomic Surface Tensions. *J. Phys. Chem. B* **2009**, *113* (18), 6378–6396. https://doi.org/10.1021/jp810292n.

(172) Kiss, P. T.; Baranyai, A. A Systematic Development of a Polarizable Potential of Water. *J. Chem. Phys.* **2013**, *138* (20), 204507. https://doi.org/10.1063/1.4807600.

(173) Webb, Benjamin, A. S.; Webb, B.; Sali, A. Comparative Protein Structure Modeling Using MODELLER. *Curr. Protoc. Bioinforma.* **2016**, *2016* (June), 5.6.1-5.6.37. https://doi.org/10.1002/cpbi.3.

(174) D.A. Case, I.Y. Ben-Shalom, S.R. Brozell, D.S. Cerutti, T.E. Cheatham, III, V.W.D. Cruzeiro, T.A. Darden, R.E. Duke, D. Ghoreishi, M.K. Gilson, H. Gohlke, A.W. Goetz, D. Greene, R Harris, N. Homeyer, Y. Huang, S. Izadi, A. Kovalenko, T. Kurtzman, T.S. Lee, D. M. Y. and P. A. K.; Case, D. A.; Walker, R. C.; Cheatham, T. E.; Simmerling, C.; Roitberg, A.; Merz, K. M.; Luo, R.; Darden, T.; D.A. Case, I.Y. Ben-Shalom, S.R. Brozell, D.S. Cerutti, T.E. Cheatham, III, V.W.D. Cruzeiro, T.A. Darden, R.E. Duke, D. Ghoreishi, M.K. Gilson, H. Gohlke, A.W. Goetz, D. Greene, R Harris, N. Homeyer, Y. Huang, S. Izadi, A. Kovalenko, T. Kurtzman, T.S. Lee, D. M. Y. and P. A. K.; Case, D. A.; Walker, R. C.; Cheatham, T. E.;

Simmerling, C.; Roitberg, A.; Merz, K. M.; Luo, R.; Darden, T. Amber 2018. *Univ. California, San Fr. 2018* **2018**, 1–923.

(175) Jones, G.; Willett, P.; Glen, R. C.; Leach, A. R.; Taylor, R.; Gareth Jones, Peter Willett, Robert C. Glen , Andrew R. Leach, R. T. Development and Validation of a Genetic Algorithm for Flexible Docking Gareth. *J. Mol. Biol.* **1997**, *267* (3), 727–748. https://doi.org/10.4000/formationemploi.3853.

(176) Nanda, V.; Koder, R. L. Designing Artificial Enzymes by Intuition and Computation. *Nat. Chem.* **2010**, *2* (1), 15–24. https://doi.org/10.1038/nchem.473.

(177) Morris, J. H.; Huang, C. C.; Babbitt, P. C.; Ferrin, T. E. StructureViz: Linking Cytoscape and UCSF Chimera. *Bioinformatics* **2007**, *23* (17), 2345–2347. https://doi.org/10.1093/bioinformatics/btm329.

(178) Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E. UCSF Chimera - A Visualization System for Exploratory Research and Analysis. *J. Comput. Chem.* **2004**, *25* (13), 1605–1612. https://doi.org/10.1002/jcc.20084.

(179) Qu, G.; Li, A.; Acevedo-Rocha, C. G.; Sun, Z.; Reetz, M. T. The Crucial Role of Methodology Development in Directed Evolution of Selective Enzymes. *Angew. Chem. Int. Ed.* Wiley-VCH Verlag August 3, 2020, pp 13204–13231. https://doi.org/10.1002/anie.201901491.

(180) Wen, J.; Xia, P.; Zheng, Z.; Xu, Y.; Li, H.; Liu, F.; Sun, S. Naphthalimide-Rhodamine Based Fluorescent Probe for Ratiometric Sensing of Cellular PH. *Chinese Chem. Lett.* **2017**, *28* (10), 2005–2008. https://doi.org/10.1016/j.cclet.2017.09.014.

(181) Chu, V.; Stayton, P. S.; Freitag, S.; Le Trong, I.; Stenkamp, R. E. Thermodynamic and Structural Consequences of Flexible Loop Deletion by Circular Permutation in the Streptavidin-Biotin System. *Protein Sci.* **1998**, *7* (4), 848–859. https://doi.org/10.1002/pro.5560070403.

(182) Sletten, E. M.; Bertozzi, C. R. Bioorthogonal Chemistry: Fishing for Selectivity in a Sea of Functionality. *Angew. Chemie - Int. Ed.* **2009**, *48* (38), 6974–6998. https://doi.org/10.1002/anie.200900942.

(183) Latocheski, E.; Dal Forno, G. M.; Ferreira, T. M.; Oliveira, B. L.; Bernardes, G. J. L.; Domingos, J. B. Mechanistic Insights into Transition Metal-Mediated Bioorthogonal

Uncaging Reactions. *Chem. Soc. Rev.* **2020**, *49* (21), 7710–7729. https://doi.org/10.1039/d0cs00630k.

(184) Christoffel, F.; Ward, T. R. Palladium-Catalyzed Heck Cross-Coupling Reactions in Water: A Comprehensive Review. *Catal. Letters* **2018**, *148* (2), 489–511. https://doi.org/10.1007/s10562-017-2285-0.

(185) Sabatino, V.; Rebelein, J. G.; Ward, T. R. "Close-to-Release": Spontaneous Bioorthogonal Uncaging Resulting from Ring-Closing Metathesis. *J. Am. Chem. Soc.* **2019**, *141* (43), 17048–17052. https://doi.org/10.1021/jacs.9b07193.

(186) Lozhkin, B.; Ward, T. R. Bioorthogonal Strategies for the in Vivo Synthesis or Release of Drugs. *Bioorganic Med. Chem.* **2021**, *45*, 116310. https://doi.org/10.1016/j.bmc.2021.116310.

(187) Martínez-Calvo, M.; Mascareñas, J. L. Organometallic Catalysis in Biological Media and Living Settings. *Coord. Chem. Rev.* **2018**, *359*, 57–79. https://doi.org/10.1016/j.ccr.2018.01.011.

(188) Tomás-Gamasa, M.; Martínez-Calvo, M.; Couceiro, J. R.; Mascarenãs, J. L. Transition Metal Catalysis in the Mitochondria of Living Cells. *Nat. Commun.* **2016**, *7*. https://doi.org/10.1038/ncomms12538.

(189) Gupta, A.; Das, R.; Yesilbag Tonga, G.; Mizuhara, T.; Rotello, V. M. Charge-Switchable Nanozymes for Bioorthogonal Imaging of Biofilm-Associated Infections. *ACS Nano* **2018**, *12* (1), 89–94. https://doi.org/10.1021/acsnano.7b07496.

(190) Sánchez, M. I.; Penas, C.; Vázquez, M. E.; Mascareñas, J. L. Metal-Catalyzed Uncaging of DNA-Binding Agents in Living Cells. *Chem. Sci.* **2014**, *5* (5), 1901–1907. https://doi.org/10.1039/c3sc53317d.

(191) Yusop, R. M.; Unciti-Broceta, A.; Johansson, E. M. V. V.; Sánchez-Martín, R. M.; Bradley, M. Palladium-Mediated Intracellular Chemistry. *Nat. Chem.* **2011**, *3* (3), 241–245. https://doi.org/10.1038/nchem.981.

(192) Wang, J.; Zheng, S.; Liu, Y.; Zhang, Z.; Lin, Z.; Li, J. J.; Zhang, G.; Wang, X.; Li, J. J.; Chen, P. R. Palladium-Triggered Chemical Rescue of Intracellular Proteins via Genetically Encoded Allene-Caged Tyrosine. *J. Am. Chem. Soc.* **2016**, *138* (46), 15118–15121. https://doi.org/10.1021/jacs.6b08933.

(193) Sasmal, P. K.; Carregal-Romero, S.; Han, A. A.; Streu, C. N.; Lin, Z.; Namikawa, K.; Elliott,

S. L.; Köster, R. W.; Parak, W. J.; Meggers, E. Catalytic Azide Reduction in Biological Environments. *ChemBioChem* **2012**, *13* (8), 1116–1120. https://doi.org/10.1002/cbic.201100719.

(194) Sancho-Albero, M.; Rubio-Ruiz, B.; Pérez-López, A. M.; Sebastián, V.; Martín-Duque, P.; Arruebo, M.; Santamaría, J.; Unciti-Broceta, A. Cancer-Derived Exosomes Loaded with Ultrathin Palladium Nanosheets for Targeted Bioorthogonal Catalysis. *Nat. Catal.* **2019**, *2* (10), 864–872. https://doi.org/10.1038/s41929-019-0333-4.

(195) Bray, T. L.; Salji, M.; Brombin, A.; Pérez-López, A. M.; Rubio-Ruiz, B.; Galbraith, L. C. A.; Patton, E. E.; Leung, H. Y.; Unciti-Broceta, A. Bright Insights into Palladium-Triggered Local Chemotherapy. *Chem. Sci.* **2018**, *9* (37), 7354–7361. https://doi.org/10.1039/c8sc02291g.

(196) Clore, A. A New Route to Synthetic DNA. *Nat. Biotechnol.* **2018**, *36* (7), 593–595. https://doi.org/10.1038/nbt.4185.

(197) Clavadetscher, J.; Indrigo, E.; Chankeshwara, S. V.; Lilienkampf, A.; Bradley, M. In-Cell Dual Drug Synthesis by Cancer-Targeting Palladium Catalysts. *Angew. Chemie - Int. Ed.* **2017**, *56* (24), 6864–6868. https://doi.org/10.1002/anie.201702404.

(198) Indrigo, E.; Clavadetscher, J.; Chankeshwara, S. V.; Lilienkampf, A.; Bradley, M. Palladium-Mediated in Situ Synthesis of an Anticancer Agent. *Chem. Commun.* **2016**, *52* (99), 14212–14214. https://doi.org/10.1039/C6CC08666G.

(199) Stenton, B. J.; Oliveira, B. L.; Matos, M. J.; Sinatra, L.; Bernardes, G. J. L. A Thioether-Directed Palladium-Cleavable Linker for Targeted Bioorthogonal Drug Decaging. *Chem. Sci.* **2018**, *9* (17), 4185–4189. https://doi.org/10.1039/c8sc00256h.

(200) Weiss, J. T.; Dawson, J. C.; Fraser, C.; Rybski, W.; Torres-Sánchez, C.; Bradley, M.; Patton, E. E.; Carragher, N. O.; Unciti-Broceta, A. Development and Bioorthogonal Activation of Palladium-Labile Prodrugs of Gemcitabine. *J. Med. Chem.* **2014**, *57* (12), 5395–5404. https://doi.org/10.1021/jm500531z.

(201) Streu, C.; Meggers, E. Ruthenium-Induced Allylcarbamate Cleavage in Living Cells. *Angew. Chemie - Int. Ed.* **2006**, *45* (34), 5645–5648. https://doi.org/10.1002/anie.200601752.

(202) Völker, T.; Meggers, E. Chemical Activation in Blood Serum and Human Cell Culture: Improved Ruthenium Complex for Catalytic Uncaging of Alloc-Protected Amines.

ChemBioChem **2017**, *18* (12), 1083–1086. https://doi.org/10.1002/cbic.201700168.

(203) Unciti-Broceta, A.; Johansson, E. M. V.; Yusop, R. M.; Sánchez-Martín, R. M.; Bradley, M. Synthesis of Polystyrene Microspheres and Functionalization with Pd0 Nanoparticles to Perform Bioorthogonal Organometallic Chemistry in Living Cells. *Nat. Protoc.* **2012**, *7* (6), 1207–1218. https://doi.org/10.1038/nprot.2012.052.

(204) Miller, M. A.; Mikula, H.; Luthria, G.; Li, R.; Kronister, S.; Prytyskach, M.; Kohler, R. H.; Mitchison, T.; Weissleder, R. Modular Nanoparticulate Prodrug Design Enables Efficient Treatment of Solid Tumors Using Bioorthogonal Activation. *ACS Nano* **2018**, *12* (12), 12814–12826. https://doi.org/10.1021/acsnano.8b07954.

(205) Pérez-López, A. M.; Rubio-Ruiz, B.; Sebastián, V.; Hamilton, L.; Adam, C.; Bray, T. L.; Irusta, S.; Brennan, P. M.; Lloyd-Jones, G. C.; Sieger, D.; Santamaría, J.; Unciti-Broceta, A. Gold-Triggered Uncaging Chemistry in Living Systems. *Angew. Chemie - Int. Ed.* **2017**, *56* (41), 12548–12552. https://doi.org/10.1002/anie.201705609.

(206) Bethge, N.; Diel, F.; Rosick, M.; Holz, J. Somatostatin Half-Life: A Case Report in One Healthy Volunteer and a Three Month Follow-Up. *Horm. Metab. Res.* **1981**, *13* (13), 709–710.

(207) Van Bist, G.; Tourwé, D. Backbone Modifications in Somatostatin Analogues: Relation between Conformation and Activity. *Pept. Res.* **1992**, *5* (1), 8–13.

(208) Testa, C.; D'addona, D.; Scrima, M.; Tedeschi, A. M.; D'ursi, A. M.; Bernhard, C.; Denat, F.; Bello, C.; Rovero, P.; Chorev, M.; Papini, A. M. Design, Synthesis, and Conformational Studies of [DOTA]-Octreotide Analogs Containing [1,2,3]Triazolyl as a Disulfide Mimetic. *Pept. Sci.* **2018**, *110* (5). https://doi.org/10.1002/pep2.24071.

(209) Bass, R. T.; Buckwalter, B. L.; Patel, B. P.; Pausch, M. H.; Price, L. A.; Strnad, J.; Hadcock, J. R. Identification and Characterization of Novel Somatostatin Antagonists. *Mol. Pharmacol.* **1996**, *50* (4), 709–715.

(210) Hruby, V. J. Designing Peptide Receptor Agonists and Antagonists. *Nat. Rev. Drug Discov.* **2002**, *1* (11), 847–858. https://doi.org/10.1038/nrd939.

(211) Fani, M.; Nicolas, G. P.; Wild, D. Somatostatin Receptor Antagonists for Imaging and Therapy. *J. Nucl. Med.* **2017**, *58*, 61S-66S. https://doi.org/10.2967/jnumed.116.186783.

(212) Fani, M.; Del Pozzo, L.; Abiraj, K.; Mansi, R.; Tamma, M. L.; Cescato, R.; Waser, B.;

Weber, W. A.; Reubi, J. C.; Maecke, H. R. PET of Somatostatin Receptor-Positive Tumors Using64Cu- and 68Ga-Somatostatin Antagonists: The Chelate Makes the Difference. *J. Nucl. Med.* **2011**, *52* (7), 1110–1118. https://doi.org/10.2967/jnumed.111.087999.

(213) Dong, C.; Liu, Z.; Wang, F. Peptide-Based Radiopharmaceuticals for Targeted Tumor Therapy. *Curr. Med. Chem.* **2013**, *21* (1), 139–152. https://doi.org/10.2174/09298673113209990219.

(214) Rylova, S. N.; Stoykow, C.; Del Pozzo, L.; Abiraj, K.; Tamma, M. L.; Kiefer, Y.; Fani, M.; Maecke, H. R. The Somatostatin Receptor 2 Antagonist 64Cu-NODAGA-JR11 Outperforms 64Cu-DOTA-TATE in a Mouse Xenograft Model. *PLoS One* **2018**, *13* (4), 1–16. https://doi.org/10.1371/journal.pone.0195802.

(215) Vong, K.; Yamamoto, T.; Chang, T. C.; Tanaka, K. Bioorthogonal Release of Anticancer Drugs: Via Gold-Triggered 2-Alkynylbenzamide Cyclization. *Chem. Sci.* **2020**, *11* (40), 10928–10933. https://doi.org/10.1039/d0sc04329j.

(216) Yada, Y.; Miyake, Y.; Nishibayashi, Y. Ruthenium-Catalyzed Intramolecular Cyclization of 3-Butyne-1,2-Diols into Furans. *Organometallics* **2008**, *27* (14), 3614–3617. https://doi.org/10.1021/om800298r.

(217) Richard, J. A.; Meyer, Y.; Jolivel, V.; Massonneau, M.; Dumeunier, R.; Vaudry, D.; Vaudry, H.; Renard, P. Y.; Romieu, A. Latent Fluorophores Based on a Self-Immolative Linker Strategy and Suitable for Protease Sensing. *Bioconjug. Chem.* **2008**, *19* (8), 1707–1718. https://doi.org/10.1021/bc8001997.

(218) de Souza, L. G.; Rennã, M. N.; Figueroa-Villar, J. D. Coumarins as Cholinesterase Inhibitors: A Review. *Chem. Biol. Interact.* **2016**, *254*, 11–23. https://doi.org/10.1016/j.cbi.2016.05.001.

(219) Barnard, D. L.; Xu, Z. Q.; Stowell, V. D.; Yuan, H.; Smee, D. F.; Samy, R.; Sidwell, R. W.; Nielsen, M. K.; Sun, L.; Cao, H.; Li, A.; Quint, C.; Deignan, J.; Crabb, J.; Flavin, M. T. Coumarins and Pyranocoumarins, Potential Novel Pharmacophores for Inhibition of Measles Virus Replication. *Antivir. Chem. Chemother.* **2002**, *13* (1), 39–59. https://doi.org/10.1177/095632020201300104.

(220) Saeed, A.; Mahesar, P. A.; Channar, P. A.; Larik, F. A.; Abbas, Q.; Hassan, M.; Raza, H.; Seo, S. Y. Hybrid Pharmacophoric Approach in the Design and Synthesis of Coumarin Linked Pyrazolinyl as Urease Inhibitors, Kinetic Mechanism and Molecular Docking.

Chem. Biodivers. **2017**, *14* (8). https://doi.org/10.1002/cbdv.201700035.

(221) Mishra, N.; Maurya, A. K.; Mulpuru, V. Discovery of Novel Coumarin Analogs against the α-Glucosidase Protein Target of Diabetes Mellitus: Pharmacophore-Based QSAR, Docking, and Molecular Dynamics Simulation Studies. *ACS Omega* **2020**, *5* (50), 32234–32249. https://doi.org/10.1021/acsomega.0c03871.

(222) Christoffel, F. Directed Evolution of Gold-Based Artificial Metalloenzymes and Design of Gold-Triggered Drug-Release Systems. **2021**.

(223) Wild, D.; Fani, M.; Behe, M.; Brink, I.; Rivier, J. E. F.; Reubi, J. C.; Maecke, H. R.; Weber, W. A. First Clinical Evidence That Imaging with Somatostatin Receptor Antagonists Is Feasible. *J. Nucl. Med.* **2011**, *52* (9), 1412–1417. https://doi.org/10.2967/jnumed.111.088922.

(224) Laznicek, M.; Laznickova, A.; Mäcke, H. R.; Eisenwiener, K.; Reubi, J. C.; Wenger, S. Octreotide and Octreotate Derivatives Radiolabeled with Yttrium: Pharmacokinetics in Rats. *Cancer Biother. Radiopharm.* **2002**, *17* (5), 527–533. https://doi.org/10.1089/108497802760804754.

(225) Kaiser, E.; Colescott, R. L.; Bossinger, C. D.; Cook, P. I. Color Test for Detection of Free Terminal Amino Groups in the Solid-Phase Synthesis of Peptides. *Anal. Biochem.* **1970**, *34* (2), 595–598. https://doi.org/10.1016/0003-2697(70)90146-6.

(226) Wild, D.; Fani, M.; Fischer, R.; Pozzo, L. Del; Kaul, F.; Krebs, S.; Rivier, J. E. F.; Reubi, J. C.; Maecke, H. R.; Weber, W. A. Comparison of Somatostatin Receptor Agonist and Antagonist for Peptide Receptor Radionuclide Therapy: A Pilot Study. *J. Nucl. Med.* **2014**, *55* (8), 1248–1252. https://doi.org/10.2967/jnumed.114.138834.

(227) Emil Thomas Kaiser, Robert L. Colescott, Charles D. Bossinger, P. I. C. Color Test for Detection of Free Terminal Amino Groups in the Solid-Phase Synthesis of Peptides. *Anal. Biochem.* **1970**, *34* (2), 595–598.

(228) Alhassan, M.; Al Musaimi, O.; Collins, J. M.; Albericio, F.; De La Torre, B. G. Cleaving Protected Peptides from 2-Chlorotrityl Chloride Resin. Moving Away from Dichloromethane. *Green Chem.* **2020**, *22* (9), 2840–2845. https://doi.org/10.1039/d0gc00834f.

(229) Chrzanowska, M.; Grajewska, A.; Rozwadowska, M. D. Asymmetric Synthesis of Isoquinoline Alkaloids: 2004-2015. *Chem. Rev.* **2016**, *116* (19), 12369–12465.

https://doi.org/10.1021/acs.chemrev.6b00315.

(230) Breuer, M.; Ditrich, K.; Habicher, T.; Hauer, B.; Keßeler, M.; Stürmer, R.; Zelinski, T. Industrial Methods for the Production of Optically Active Intermediates. *Angew. Chemie - Int. Ed.* **2004**, *43* (7), 788–824. https://doi.org/10.1002/anie.200300599.

(231) Cosgrove, S. C.; Brzezniak, A.; France, S. P.; Ramsden, J. I.; Mangas-Sanchez, J.; Montgomery, S. L.; Heath, R. S.; Turner, N. J. Imine Reductases, Reductive Aminases, and Amine Oxidases for the Synthesis of Chiral Amines: Discovery, Characterization, and Synthetic Applications. In *Methods in Enzymology*; Academic Press Inc., 2018; Vol. 608, pp 131–149. https://doi.org/10.1016/bs.mie.2018.04.022.

(232) Heberling, M. M.; Wu, B.; Bartsch, S.; Janssen, D. B. Priming Ammonia Lyases and Aminomutases for Industrial and Therapeutic Applications. *Curr. Opin. Chem. Biol.* **2013**, *17* (2), 250–260. https://doi.org/10.1016/j.cbpa.2013.02.013.

(233) Grogan, G. Synthesis of Chiral Amines Using Redox Biocatalysis. *Curr. Opin. Chem. Biol.* **2018**, *43*, 15–22. https://doi.org/10.1016/j.cbpa.2017.09.008.

(234) Uematsu, N.; Fujii, A.; Hashiguchi, S.; Ikariya, T.; Noyori, R. Asymmetric Transfer Hydrogenation of Imines. *J. Am. Chem. Soc.* **1996**, *118*, 4916–4917.

(235) Muñoz Robles, V.; Vidossich, P.; Lledós, A.; Ward, T. R.; Maréchal, J. D. Computational Insights on an Artificial Imine Reductase Based on the Biotin-Streptavidin Technology. *ACS Catal.* **2014**, *4* (3), 833–842. https://doi.org/10.1021/cs400921n.

(236) Koike, T.; Ikariya, T. Mechanistic Aspects of Formation of Chiral Ruthenium Hydride Complexes from 16-Electron Ruthenium Amide Complexes and Formic Acid: Facile Reversible Decarboxylation and Carboxylation. *Adv. Synth. Catal.* **2004**, *346* (1), 37–41. https://doi.org/10.1002/adsc.200303152.

(237) Quinto, T.; Schwizer, F.; Zimbron, J. M.; Morina, A.; Köhler, V.; Ward, T. R. Expanding the Chemical Diversity in Artificial Imine Reductases Based on the Biotin-Streptavidin Technology. *ChemCatChem* **2014**, *6* (4), 1010–1014. https://doi.org/10.1002/cctc.201300825.

(238) Okamoto, Y.; Köhler, V.; Ward, T. R. An NAD(P)H-Dependent Artificial Transfer Hydrogenase for Multienzymatic Cascades. *J. Am. Chem. Soc.* **2016**, *138* (18), 5781–5784. https://doi.org/10.1021/jacs.6b02470.

(239) Quinto, T.; Häussinger, D.; Köhler, V.; Ward, T. R. Artificial Metalloenzymes for the

Diasterecselective Reduction of NAD+ to NAD2H. *Org. Biomol. Chem.* **2015**, *13* (2), 357–360. https://doi.org/10.1039/c4ob02071e.

(240) Köhler, V.; Wilson, Y. M.; Dürrenberger, M.; Ghislieri, D.; Churakova, E.; Quinto, T.; Knörr, L.; Häussinger, D.; Hollmann, F.; Turner, N. J.; Ward, T. R. Synthetic Cascades Are Enabled by Combining Biocatalysts with Artificial Metalloenzymes. *Nat. Chem.* **2013**, *5* (2), 93–99. https://doi.org/10.1038/nchem.1498.

(241) Michaelis, L.; Menten, M. L.; Goody, R. S.; Johnson, K. A. Die Kinetik Der Invertinwirkung/ The Kinetics of Invertase Action. *Biochemistry* **1913**, *49* (2), 333–369.

(242) Briggs, G. E.; Haldane, J. A Note On The Kinetics Of Enzyme Action. *Biochem. J.* **1925**, *19*, 338–339.

(243) Oohora, K.; Onoda, A.; Hayashi, T. Hemoproteins Reconstituted with Artificial Metal Complexes as Biohybrid Catalysts. *Acc. Chem. Res.* **2019**, *52* (4), 945–954. https://doi.org/10.1021/acs.accounts.8b00676.

(244) Hestericová, M.; Heinisch, T.; Lenz, M.; Ward, T. R. Ferritin Encapsulation of Artificial Metalloenzymes: Engineering a Tertiary Coordination Sphere for an Artificial Transfer Hydrogenase. *Dalt. Trans.* **2018**, *47* (32), 10837–10841. https://doi.org/10.1039/c8dt02224k.

(245) Hestericová, M.; Correro, M. R.; Lenz, M.; Corvini, P. F. X.; Shahgaldian, P.; Ward, T. R. Immobilization of an Artificial Imine Reductase within Silica Nanoparticles Improves Its Performance. *Chem. Commun.* **2016**, *52* (60), 9462–9465. https://doi.org/10.1039/c6cc04604e.

(246) Chu, V. V.; Freitag, S.; Le Trong, I. I.; Stenkamp, R. E.; Stayton ', P. S.; Stayton, P. S.; Freitag, S.; Le Trong, I. I.; Stenkamp, R. E. Thermodynamic and Structural Consequences of Flexible Loop Deletion by Circular Permutation in the Streptavidin-Biotin System. *Protein Sci.* **1998**, *7* (4), 7848–7859. https://doi.org/10.1002/pro.5560070403.

(247) Chivers, C. E.; Koner, A. L.; Lowe, E. D.; Howarth, M. How the Biotin-Streptavidin Interaction Was Made Even Stronger: Investigation via Crystallography and a Chimaeric Tetramer. *Biochem. J.* **2011**, *435* (1), 55–63. https://doi.org/10.1042/BJ20101593.

(248) Schwizer, F.; Köhler, V.; Dürrenberger, M.; Knörr, L.; Ward, T. R. Genetic Optimization of the Catalytic Efficiency of Artificial Imine Reductases Based on Biotin-Streptavidin Technology. *ACS Catal.* **2013**, *3* (8), 1752–1755. https://doi.org/10.1021/cs400428r.

(249) Martins, J. E. D.; Contreras Redondo, M. A.; Wills, M. Applications of N'-Alkylated Derivatives of TsDPEN in the Asymmetric Transfer Hydrogenation of C=O and C=N Bonds. *Tetrahedron Asymmetry* **2010**, *21* (18), 2258–2264. https://doi.org/10.1016/j.tetasy.2010.07.013.

(250) Martins, J. E. D.; Clarkson, G. J.; Wills, M. Ru(II) Complexes of N-Alkylated TsDPEN Ligands in Asymmetric Transfer Hydrogenation of Ketones and Imines. *Org. Lett.* **2009**, *11* (4), 847–850. https://doi.org/10.1021/ol802801p.

(251) McPherson, A.; Gavira, J. A. Introduction to Protein Crystallization. *Acta Crystallogr. Sect. FStructural Biol. Commun.* **2014**, *70* (1), 2–20. https://doi.org/10.1107/S2053230X13033141.

(252) Ladd, M.; Palmer, R. *Structure Determination by X-Ray Crystallography: Analysis by X-Rays and Neutrons*; 2013. https://doi.org/10.1007/978-1-4614-3954-7.

(253) Heck, A. J. R. Native Mass Spectrometry: A Bridge between Interactomics and Structural Biology. *Nat. Methods* **2008**, *5* (11), 927–933. https://doi.org/10.1038/nmeth.1265.

(254) Lorenzen, K.; Duijn, E. Van; van Duijn, E.; Duijn, E. Van; van Duijn, E. Native Mass Spectrometry as a Tool in Structural Biology. *Curr. Protoc. Protein Sci.* **2010**, No. SUPPL.62, 1–17. https://doi.org/10.1002/0471140864.ps1712s62.

(255) Tanaka, K.; Waki, H.; Ido, Y.; Akita, S.; Yoshida, Y.; Yoshida, T.; Matsuo, T. Protein and Polymer Analyses up to m/z 100 000 by Laser Ionization Time-of-flight Mass Spectrometry. *Rapid Commun. Mass Spectrom.* **1988**, *2* (8), 151–153. https://doi.org/10.1002/rcm.1290020802.

(256) Fenn, J. B.; Mann, M.; Meng, C. K.; Wong, S. F.; Whitehouse, C. M. Electrospray Ionization for Mass Spectrometry of Large Biomolecules. *Science (80-.).* **1989**, *246* (4926), 64–71. https://doi.org/10.1126/science.2675315.

(257) Ganem, B. Detection of Noncovalent Receptor-Ligand Complexes by Mass Spectrometry Bruce. *J. Am. Chem. Soc* **1991**, *113*, 6294–6296.

(258) Light-Wahl, K. J.; Winger, B. E.; Smith, R. D. Observation of the Multimeric Forms of Concanavalin A by Electrospray Ionization Mass Spectrometry. *J. Am. Chem. Soc* **1993**, *115*, 5869–5870.

(259) Schwartz, B. L.; Light-wahl, K. J.; Smith, R. D. Observation of Noncovalent Complexes. *J.*

Am. Soc. Mass Spectrom. **1994**, *5*, 201–204.

(260) Van Den Heuvel, R. H. H.; Heck, A. J. R. Native Protein Mass Spectrometry: From Intact Oligomers to Functional Machineries. *Curr. Opin. Chem. Biol.* **2004**, *8* (5), 519–526. https://doi.org/10.1016/j.cbpa.2004.08.006.

(261) Morris, H. R.; Paxton, T.; Dell, A.; Langhorne, J.; Berg, M.; Bordoli, R. S.; Hoyes, J.; Bateman, R. H. High Sensitivity Collisionally-Activated Decomposition Tandem Mass Spectrometry on a Novel Quadrupole/Orthogonal-Acceleration Time-of-Flight Mass Spectrometer. *Rapid Commun. Mass Spectrom.* **1996**, *10* (8), 889–896. https://doi.org/10.1002/(SICI)1097-0231(19960610)10:8<889::AID-RCM615>3.0.CO;2-F.

(262) Chernushevich, I. V.; Loboda, A. V.; Thomson, B. A. An Introduction to Quadrupole-Time-of-Flight Mass Spectrometry. *J. Mass Spectrom.* **2001**, *36* (8), 849–865. https://doi.org/10.1002/jms.207.

(263) Collings, B. A.; Douglas, D. J. An Extended Mass Range Quadrupole for Electrospray Mass Spectrometry. *Int. J. Mass Spectrom. Ion Process.* **1997**, *162* (1–3), 121–127. https://doi.org/10.1016/s0168-1176(96)04478-3.

(264) Konermann, L. Addressing a Common Misconception: Ammonium Acetate as Neutral PH "Buffer" for Native Electrospray Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* **2017**, *28* (9), 1827–1835. https://doi.org/10.1007/s13361-017-1739-3.

(265) Peschke, M.; Verkerk, U. H.; Kebarle, P. Prediction of the Charge States of Folded Proteins in Electrospray Ionization. *Eur. J. Mass Spectrom.* **2004**, *10* (6), 993–1002. https://doi.org/10.1255/ejms.692.

(266) Erba, E. B.; Zenobi, R. Mass Spectrometric Studies of Dissociation Constants of Noncovalent Complexes. *Annu. Reports Prog. Chem. - Sect. C* **2011**, *107*, 199–228. https://doi.org/10.1039/c1pc90006d.

(267) Cubrilovic, D.; Biela, A.; Sielaff, F.; Steinmetzer, T.; Klebe, G.; Zenobi, R. Quantifying Protein-Ligand Binding Constants Using Electrospray Ionization Mass Spectrometry: A Systematic Binding Affinity Study of a Series of Hydrophobically Modified Trypsin Inhibitors. *J. Am. Soc. Mass Spectrom.* **2012**, *23* (10), 1768–1777. https://doi.org/10.1007/s13361-012-0451-6.

(268) Arno Wortmann, Matthias C. Jecklin, David Touboul, M. B. and R. Z. Binding Constant

Determination of High-Affinity Protein–Ligand Complexes by Electrospray Ionization Mass Spectrometry and Ligand Competition. *J. Mass Spectrom.* **2008**, *43* (6), 600–608. https://doi.org/10.1002/jms.1355.

(269) Schwartz, B. L.; Gale, D. C.; Smith, R. D.; Chilkoti, A.; Stayton, P. S. Investigation of Non-covalent Ligand Binding to the Intact Streptavidin Tetramer by Electrospray Ionization Mass Spectrometry. *J. Mass Spectrom.* **1995**, *30* (8), 1095–1102. https://doi.org/10.1002/jms.1190300806.

(270) Bornscheuer, U. T. The Fourth Wave of Biocatalysis Is Approaching. *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.* **2018**, *376* (2110), 1–7. https://doi.org/10.1098/rsta.2017.0063.

(271) Hammer, S. C.; Knight, A. M.; Arnold, F. H. Design and Evolution of Enzymes for Non-Natural Chemistry. *Curr. Opin. Green Sustain. Chem.* **2017**, *7*, 23–30. https://doi.org/10.1016/j.cogsc.2017.06.002.

(272) Wu, S.; Zhou, Y.; Gerngross, D.; Jeschek, M.; Ward, T. R. Chemo-Enzymatic Cascades to Produce Cycloalkenes from Bio-Based Resources. *Nat. Commun.* **2019**, *10* (1), 1–10. https://doi.org/10.1038/s41467-019-13071-y.

(273) Booth, R. L.; Grogan, G.; Wilson, K. S.; Duhme-Klair, A.-K. Artificial Imine Reductases: Developments and Future Directions. *RSC Chem. Biol.* **2020**, *1* (5), 369–378. https://doi.org/10.1039/d0cb00113a.

(274) Leveson-Gower, R. B.; Zhou, Z.; Drienovská, I.; Roelfes, G. Unlocking Iminium Catalysis in Artificial Enzymes to Create a Friedel-Crafts Alkylase. *ACS Catal.* **2021**, *11* (12), 6763–6770. https://doi.org/10.1021/acscatal.1c00996.

(275) Markel, U.; Sauer, D. F.; Schiffels, J.; Okuda, J.; Schwaneberg, U. Towards the Evolution of Artificial Metalloenzymes—A Protein Engineer's Perspective. *Angew. Chemie - Int. Ed.* **2019**, *58* (14), 4454–4464. https://doi.org/10.1002/anie.201811042.

(276) Raines, D. J.; Clarke, J. E.; Blagova, E. V.; Dodson, E. J.; Wilson, K. S.; Duhme-Klair, A. K. Redox-Switchable Siderophore Anchor Enables Reversible Artificial Metalloenzyme Assembly. *Nat. Catal.* **2018**, *1* (9), 680–688. https://doi.org/10.1038/s41929-018-0124-3.

(277) Large, B.; Baranska, N. G.; Booth, R. L.; Wilson, K. S.; Duhme-Klair, A. K. Artificial Metalloenzymes: The Powerful Alliance between Protein Scaffolds and Organometallic

Catalysts. *Curr. Opin. Green Sustain. Chem.* **2021**, *28*, 100420. https://doi.org/10.1016/j.cogsc.2020.100420.

(278) Morra, S.; Pordea, A. Biocatalyst-Artificial Metalloenzyme Cascade Based on Alcohol Dehydrogenase. *Chem. Sci.* **2018**, *9* (38), 7447–7454. https://doi.org/10.1039/C8SC02371A.

(279) Thiel, A.; Sauer, D. F.; Markel, U.; Mertens, M. A. S.; Polen, T.; Schwaneberg, U.; Okuda, J. An Artificial Ruthenium-Containing β-Barrel Protein for Alkene-Alkyne Coupling Reaction. *Org. Biomol. Chem.* **2021**, *19* (13), 2912–2916. https://doi.org/10.1039/d1ob00279a.

(280) Markel, U.; Sauer, D. F.; Wittwer, M.; Schiffels, J.; Cui, H.; Davari, M. D.; Kröckert, K. W.; Herres-Pawlis, S.; Okuda, J.; Schwaneberg, U. Chemogenetic Evolution of a Peroxidase-like Artificial Metalloenzyme. *ACS Catal.* **2021**, *11* (9), 5079–5087. https://doi.org/10.1021/acscatal.1c00134.

(281) Doble, M. V.; Obrecht, L.; Joosten, H. J.; Lee, M.; Rozeboom, H. J.; Branigan, E.; Naismith, J. H.; Janssen, D. B.; Jarvis, A. G.; Kamer, P. C. J. Engineering Thermostability in Artificial Metalloenzymes to Increase Catalytic Activity. *ACS Catal.* **2021**, *11* (6), 3620–3627. https://doi.org/10.1021/acscatal.0c05413.

(282) Yang, H.; Srivastava, P.; Zhang, C.; Lewis, J. C. A General Method for Artificial Metalloenzyme Formation through Strain-Promoted Azide-Alkyne Cycloaddition. *ChemBioChem* **2014**, *15* (2), 223–227. https://doi.org/10.1002/cbic.201300661.

(283) Sauer, D. F.; Himiyama, T.; Tachikawa, K.; Fukumoto, K.; Onoda, A.; Mizohata, E.; Inoue, T.; Bocola, M.; Schwaneberg, U.; Hayashi, T.; Okuda, J. A Highly Active Biohybrid Catalyst for Olefin Metathesis in Water: Impact of a Hydrophobic Cavity in a β-Barrel Protein. *ACS Catal.* **2015**, *5* (12), 7519–7522. https://doi.org/10.1021/acscatal.5b01792.

(284) Haquette, P.; Salmain, M.; Svedlunq, K.; Martel, A.; Rudolf, B.; Zakrzewski, J.; Cordier, S.; Roisnel, T.; Fosse, C.; Jaouen, G. Cysteine-Specific, Covalent Anchoring of Transition Organometallic Complexes to the Protein Papain from Carica Papaya. *ChemBioChem* **2007**, *8* (2), 224–231. https://doi.org/10.1002/cbic.200600387.

(285) Ricoux, R.; Dubuc, R.; Dupont, C.; Marechal, J.; Martin, A.; Sellier, M. Hemozymes Peroxidase Activity Of Artificial Hemoproteins Constructed From the Streptomyces

LiWidans Xylanase A and Iron(III)-Carboxy-Substituted Porphyrins. *Bioconjug Chem* **2008**, *19* (Iii), 899–910.

(286) Rondot, L.; Girgenti, E.; Oddon, F.; Marchi-Delapierre, C.; Jorge-Robin, A.; Ménage, S. Catalysis without a Headache: Modification of Ibuprofen for the Design of Artificial Metalloenzyme for Sulfide Oxidation. *J. Mol. Catal. A Chem.* **2016**, *416*, 20–28. https://doi.org/10.1016/j.molcata.2016.02.015.

(287) Sano, Y.; Onoda, A.; Hayashi, T. A Hydrogenase Model System Based on the Sequence of Cytochrome c: Photochemical Hydrogen Evolution in Aqueous Media. *Chem. Commun.* **2011**, *47* (29), 8229–8231. https://doi.org/10.1039/c1cc11157d.

(288) Joseph Sommer, D.; David Vaughn, M.; Ghirlanda, G. Protein Secondary-Shell Interactions Enhance the Photoinduced Hydrogen Production of Cobalt Protoporphyrin IX. *Chem. Commun.* **2014**, *50* (100), 15852–15855. https://doi.org/10.1039/c4cc06700b.

(289) Soltau, S. R.; Dahlberg, P. D.; Niklas, J.; Poluektov, O. G.; Mulfort, K. L.; Utschig, L. M. Ru-Protein-Co Biohybrids Designed for Solar Hydrogen Production: Understanding Electron Transfer Pathways Related to Photocatalytic Function. *Chem. Sci.* **2016**, *7* (12), 7068–7078. https://doi.org/10.1039/C6SC03121H.

(290) Jing, Q.; Okrasa, K.; Kazlauskas, R. J. Manganese-Substituted α-Carbonic Anhydrase as an Enantioselective Peroxidase. In *Bio-inspired Catalysts*; 2008; pp 45–61. https://doi.org/10.1007/3418_2008_1.

(291) Coleman, J. E. Metal Lon Dependent Binding of Sulphonamide to Carbonic Anhydrase. *Nature* **1967**, *214* (3), 193–194.

(292) Carey, J. R.; Ma, S. K.; Pfister, T. D.; Garne, D. K.; Kim, H. K.; Abramite, J. A.; Wang, Z.; Guo, Z.; Lu, Y. A Site-Selective Dual Anchoring Strategy for Artificial Metalloprotein Design. *J. Am. Chem. Soc.* **2004**, *126* (35), 10812–10813. https://doi.org/10.1021/ja046908x.

(293) Garner, D. K.; Liang, L.; Barrios, D. A.; Zhang, J. L.; Lu, Y. The Important Role of Covalent Anchor Positions in Tuning Catalytic Properties of a Rationally Designed Mnsalen-Containing Metalloenzyme. *ACS Catal.* **2011**, *1* (9), 1083–1089. https://doi.org/10.1021/cs200258e.

(294) Laureanti, J. A.; Ginovska, B.; Buchko, G. W.; Schenter, G. K.; Hebert, M.; Zadvornyy, O.

A.; Peters, J. W.; Shaw, W. J. A Positive Charge in the Outer Coordination Sphere of an Artificial Enzyme Increases CO2 Hydrogenation. *Organometallics* **2020**, *39* (9), 1532–1544. https://doi.org/10.1021/acs.organomet.9b00843.

(295) Drienovská, I.; Rioz-Martínez, A.; Draksharapu, A.; Roelfes, G. Novel Artificial Metalloenzymes by in Vivo Incorporation of Metal-Binding Unnatural Amino Acids. *Chem. Sci.* **2015**, *6* (1), 770–776. https://doi.org/10.1039/c4sc01525h.

(296) Heinisch, T.; Pellizzoni, M.; Dürrenberger, M.; Tinberg, C. E.; Köhler, V.; Klehr, J.; Häussinger, D.; Baker, D.; Ward, T. R.; Dü, M.; Tinberg, C. E.; Kö, V.; Klehr, J.; Hä, D.; Baker, D.; Ward, T. R. Improving the Catalytic Performance of an Artificial Metalloenzyme by Computational Design. *J. Am. Chem. Soc.* **2015**, *137* (32), 10414–10419. https://doi.org/10.1021/jacs.5b06622.

(297) Zhao, J.; Kajetanowicz, A.; Ward, T. R. Carbonic Anhydrase II as Host Protein for the Creation of a Biocompatible Artificial Metathesase. *Org. Biomol. Chem.* **2015**, *13* (20), 5652–5655. https://doi.org/10.1039/c5ob00428d.

(298) Krishnamurthy, V. M.; Kaufman, G. K.; Urbach, A. R.; Gitlin, I.; Gudiksen, K. L.; Weibel, D. B.; Whitesides, G. M. Carbonic Anhydrase as a Model for Biophysical and Physical-Organic Studies of Proteins and Protein-Ligand Binding. *Chem. Rev.* **2008**, *108* (3), 946–1051. https://doi.org/10.1021/cr050262p.

(299) DiMario, R. J.; Machingura, M. C.; Waldrop, G. L.; Moroney, J. V. The Many Types of Carbonic Anhydrases in Photosynthetic Organisms. *Plant Sci.* **2018**, *268* (December 2017), 11–17. https://doi.org/10.1016/j.plantsci.2017.12.002.

(300) Hou, G. H.; Xie, J. H.; Yan, P. C.; Zhou, Q. L. Iridium-Catalyzed Asymmetric Hydrogenation of Cyclic Enamines. *J. Am. Chem. Soc.* **2009**, *131* (4), 1366–1367. https://doi.org/10.1021/ja808358r.

(301) Leipold, F.; Hussain, S.; Ghislieri, D.; Turner, N. J. Asymmetric Reduction of Cyclic Imines Catalyzed by a Whole-Cell Biocatalyst Containing an (S)-Imine Reductase. *ChemCatChem* **2013**, *5* (12), 3505–3508. https://doi.org/10.1002/cctc.201300539.

(302) Schrittwieser, J. H.; Velikogne, S.; Kroutil, W. Biocatalytic Imine Reduction and Reductive Amination of Ketones. *Adv. Synth. Catal.* **2015**, *357* (8), 1655–1685. https://doi.org/10.1002/adsc.201500213.

(303) Marshall, J. R.; Yao, P.; Montgomery, S. L.; Finnigan, J. D.; Thorpe, T. W.; Palmer, R. B.;

Mangas-Sanchez, J.; Duncan, R. A. M.; Heath, R. S.; Graham, K. M.; Cook, D. J.; Charnock, S. J.; Turner, N. J. Screening and Characterization of a Diverse Panel of Metagenomic Imine Reductases for Biocatalytic Reductive Amination. *Nat. Chem.* **2020**, *13* (February). https://doi.org/10.1038/s41557-020-00606-w.

(304) Mitsukura, K.; Suzuki, M.; Shinoda, S.; Kuramoto, T.; Yoshida, T.; Nagasawa, T. Purification and Characterization of a Novel (R)-Imine Reductase from Streptomyces Sp. GF3587. *Biosci. Biotechnol. Biochem.* **2011**, *75* (9), 1778–1782. https://doi.org/10.1271/bbb.110303.

(305) Aleku, G. A.; France, S. P.; Man, H.; Mangas-Sanchez, J.; Montgomery, S. L.; Sharma, M.; Leipold, F.; Hussain, S.; Grogan, G.; Turner, N. J. A Reductive Aminase from Aspergillus Oryzae. *Nat. Chem.* **2017**, *9* (10), 961–969. https://doi.org/10.1038/nchem.2782.

(306) Abrahamson, M. J.; Vázquez-Figueroa, E.; Woodall, N. B.; Moore, J. C.; Bommarius, A. S. Development of an Amine Dehydrogenase for Synthesis of Chiral Amines. *Angew. Chemie - Int. Ed.* **2012**, *51* (16), 3969–3972. https://doi.org/10.1002/anie.201107813.

(307) France, S. P.; Howard, R. M.; Steflik, J.; Weise, N. J.; Mangas-Sanchez, J.; Montgomery, S. L.; Crook, R.; Kumar, R.; Turner, N. J. Identification of Novel Bacterial Members of the Imine Reductase Enzyme Family That Perform Reductive Amination. *ChemCatChem* **2018**, *10* (3), 510–514. https://doi.org/10.1002/cctc.201701408.

(308) Roiban, G. D.; Kern, M.; Liu, Z.; Hyslop, J.; Tey, P. L.; Levine, M. S.; Jordan, L. S.; Brown, K. K.; Hadi, T.; Ihnken, L. A. F.; Brown, M. J. B. Efficient Biocatalytic Reductive Aminations by Extending the Imine Reductase Toolbox. *ChemCatChem* **2017**, *9* (24), 4475–4479. https://doi.org/10.1002/cctc.201701379.

(309) Cabrero-Antonino, J. R.; Adam, R.; Papa, V.; Beller, M. Homogeneous and Heterogeneous Catalytic Reduction of Amides and Related Compounds Using Molecular Hydrogen. *Nat. Commun.* **2020**, *11* (1), 1–18. https://doi.org/10.1038/s41467-020-17588-5.

(310) Tin, S.; Fanjul, T.; Clarke, M. L. Hydrogenation of Unactivated Enamines to Tertiary Amines: Rhodium Complexes of Fluorinated Phosphines Give Marked Improvements in Catalytic Activity. *Beilstein J. Org. Chem.* **2015**, *11*, 622–627. https://doi.org/10.3762/bjoc.11.70.

(311) Liu, Y.; Yue, X.; Luo, C.; Zhang, L.; Lei, M. Mechanisms of Ketone/Imine Hydrogenation Catalyzed by Transition-Metal Complexes. *Energy Environ. Mater.* **2019**, *2* (4), 292–312. https://doi.org/10.1002/eem2.12050.

(312) Agouridas, V.; El Mahdi, O.; Diemer, V.; Cargoët, M.; Monbaliu, J. C. M.; Melnyk, O. Native Chemical Ligation and Extended Methods: Mechanisms, Catalysis, Scope, and Limitations. *Chem. Rev.* **2019**, *119* (12). https://doi.org/10.1021/acs.chemrev.8b00712.

(313) Friedmann, E.; Marrian, D. H.; Simon-Reuss, I. ANTIMITOTIC ACTION OF MALEIMIDE AND RELATED SUBSTANCES. *Br. J. Pharmacol.* **1949**, *4*, 105–108.

(314) Gempf, K. L.; Butler, S. J.; Funk, A. M.; Parker, D. Direct and Selective Tagging of Cysteine Residues in Peptides and Proteins with 4-Nitropyridyl Lanthanide Complexes. *Chem. Commun.* **2013**, *49* (80), 9104–9106. https://doi.org/10.1039/c3cc45875j.

(315) Moore, E. J.; Zorine, D.; Hansen, W. A.; Khare, S. D.; Fasan, R. Enzyme Stabilization via Computationally Guided Protein Stapling. *Proc. Natl. Acad. Sci. U. S. A.* **2017**, *114* (47), 12472–12477. https://doi.org/10.1073/pnas.1708907114.

(316) Goddard, D. R.; Michaelis, L. Derivatives of Keratin. *J. Biol. Chem.* **1935**, *112* (1), 361–371. https://doi.org/10.1016/s0021-9258(18)74993-4.

(317) Chalker, J. M.; Gunnoo, S. B.; Boutureira, O.; Gerstberger, S. C.; Fernández-González, M.; Bernardes, G. J. L.; Griffin, L.; Hailu, H.; Schofield, C. J.; Davis, B. G. Methods for Converting Cysteine to Dehydroalanine on Peptides and Proteins. *Chem. Sci.* **2011**, *2* (9), 1666–1676. https://doi.org/10.1039/c1sc00185j.

(318) Toda, N.; Asano, S.; Barbas, C. F. Rapid, Stable, Chemoselective Labeling of Thiols with Julia- Kocieński-like Reagents: A Serum-Stable Alternative to Maleimide-Based Protein Conjugation. *Angew. Chemie - Int. Ed.* **2013**, *52* (48), 12592–12596. https://doi.org/10.1002/anie.201306241.

(319) Ngo, A. H.; Ibañez, M.; Do, L. H. Catalytic Hydrogenation of Cytotoxic Aldehydes Using Nicotinamide Adenine Dinucleotide (NADH) in Cell Growth Media. *ACS Catal.* **2016**, *6* (4), 2637–2641. https://doi.org/10.1021/acscatal.6b00395.

(320) Ngo, A. H.; Do, L. H. Structure-Activity Relationship Study of Half-Sandwich Metal Complexes in Aqueous Transfer Hydrogenation Catalysis. *Inorg. Chem. Front.* **2020**, *7* (3), 583–591. https://doi.org/10.1039/c9qi01310e.

(321) Arnold, F. H. Innovation by Evolution: Bringing New Chemistry to Life (Nobel Lecture)**. *Angew. Chemie - Int. Ed.* **2018**, *58* (41), 14420–14426. https://doi.org/10.1002/ange.201907729.

(322) Cho, I.; Arnold, F. H. Site-Selective Enzymatic C–H Amidation for Synthesis of Diverse Lactams. *Science (80-.).* **2019**, *346*, 575–578.

(323) Tang, L.; Gao, H.; Zhu, X.; Wang, X.; Zhou, M.; Jiang, R. Construction of "Small-Intelligent" Focused Mutagenesis Libraries Using Well-Designed Combinatorial Degenerate Primers. *Biotechniques* **2012**, *52* (3), 149–158. https://doi.org/10.2144/000113820.

(324) Kachur, A. V.; Koch, C. J.; Biaglow, J. E. Mechanism of Copper-Catalyzed Oxidation of Glutathione. *Free Radic. Res.* **1998**, *28* (3), 259–269. https://doi.org/10.3109/10715769809069278.

(325) Ngamchuea, K.; Batchelor-mcauley, C.; Compton, R. G. The Copper (II) -Catalyzed Oxidation of Glutathione. *Chem. - A Eur. J.* **2016**, *22* (15937–157944).

(326) Harpaz, Y.; Gerstein, M.; Chothia, C. Volume Changes on Protein Folding. *Structure* **1994**, *2* (7), 641–649. https://doi.org/10.1016/S0969-2126(00)00065-4.

(327) Fleming, S. A.; Rawlins, D. B.; Samano, V.; Robins, M. J. Photochemistry of Nucleoside Transport Inhibitor 6-S-Benzylated Thiopurine Ribonucleosides. Implications for a New Class of Photoaffinity Labels. *J. Chem. Soc. Chem. Commun.* **1992**, *57* (22), 5968–5976.

(328) Barrios-Rivera, J.; Xu, Y.; Wills, M. Asymmetric Transfer Hydrogenation of Unhindered and Non-Electron-Rich 1-Aryl Dihydroisoquinolines with High Enantioselectivity. *Org. Lett.* **2020**, *22* (16), 6283–6287. https://doi.org/10.1021/acs.orglett.0c02034.

(329) Zoppi, C.; Nocentini, A.; Supuran, C. T.; Pratesi, A.; Messori, L. Native Mass Spectrometry of Human Carbonic Anhydrase I and Its Inhibitor Complexes. *J. Biol. Inorg. Chem.* **2020**, *25* (7), 979–993. https://doi.org/10.1007/s00775-020-01818-8.

(330) Stierand, K.; Rarey, M. PoseView -- Molecular Interaction Patterns at a Glance. *J. Cheminform.* **2010**, *2* (S1), 20146. https://doi.org/10.1186/1758-2946-2-s1-p50.

(331) Abdalla, M. Important Factors Influencing Protein Crystallization. *Glob. J. Biotechnol. Biomater. Sci.* **2016**, *2*, 025–028. https://doi.org/10.17352/gjbbs.000008.

(332) Capasso, S.; Mazzarella, L.; Sica, F.; Zagari, A.; Salvadori, S. Spontaneous Cyclization of the Aspartic Acid Side Chain to the Succinimide Derivative. *J. Chem. Soc. Chem.*

Commun. **1992**, No. 12, 919–921. https://doi.org/10.1039/C39920000919.

(333) Plumeré, N.; Henig, J.; Campbell, W. H. Enzyme-Catalyzed O 2 Removal System for Electrochemical Analysis under Ambient Air: Application in an Amperometric Nitrate Biosensor. *Anal. Chem.* **2012**, *84* (5), 2141–2146. https://doi.org/10.1021/ac2020883.

(334) Winn, M. D.; Ballard, C. C.; Cowtan, K. D.; Dodson, E. J.; Emsley, P.; Evans, P. R.; Keegan, R. M.; Krissinel, E. B.; Leslie, A. G. W.; McCoy, A.; McNicholas, S. J.; Murshudov, G. N.; Pannu, N. S.; Potterton, E. A.; Powell, H. R.; Read, R. J.; Vagin, A.; Wilson, K. S. Overview of the CCP4 Suite and Current Developments. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2011**, *67* (4), 235–242. https://doi.org/10.1107/S0907444910045749.

(335) Moriarty, N. W.; Grosse-Kunstleve, R. W.; Adams, P. D. Electronic Ligand Builder and Optimization Workbench (ELBOW): A Tool for Ligand Coordinate and Restraint Generation. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2009**, *65* (10), 1074–1080. https://doi.org/10.1107/S0907444909029436.